Proceedings of the Challenges for Analysis of the Economy, the Businesses, and Social Progress

International Scientific Conference

Szeged, November 19-21, 2009

ISBN 978-963-06-9558-9

Edited by
Péter Kovács
Katalin Szép
Tamás Katona

Published by
Unidocument Kft.
www.e-document.hu

Szeged, 2010

UNIVERSITY OF SZEGED
FACULTY OF ECONOMICS AND
BUSINESS ADMINISTRATION

HUNGARIAN CENTRAL
STATISTICAL OFFICE

Preface

This proceeding contains reviewed papers from Challenges for Analysis of the Economy, the Businesses, and Social Progress International Scientific Conference, which was organized by the Faculty of Economics and Business Administration, University of Szeged (Hungary) and the Hungarian Central Statistical Office from November 19 to 21, 2009 in Szeged.

More than 220 colleagues took part in this conference form 35 countries (from Mexico to Japan and Australia). 190 presentations were taken from various areas of statistics and methodology and applications of those in 2 Plenary and 16 Parallel Sessions.

The reviewed and accepted papers are available, free of charge, from http://www.e-doc.hu/conferences/statconf2009

You can read the papers online and furthermore separately download them.

Editorial Correspondence should be addressed to:

Peter Kovacs
Faculty of Economics and Business Administration, University of Szeged,
Kálvária sgt. 1.,
6722 Szeged,
Hungary

Email: pepe@eco.u-szeged.hu

The Editors

Table of contents

1 - Plenary session

2-Economic & Business Statistics

3 - Agribusiness

4 - Financies - Empirical Issues, Models & Simulations

5 - Consumer Behaviour, Marketing and Management

6 - Labour Markets

7 - Educational Issues

8 - Social Statistics

9 - Regional & Spatial Analysis

10 - Measurement Issues

11 - Statistical Theory

New directions in the measurement of social progress

Eszter Bagó[1]

The assessment of the measure of social progress (development, well-being) and its changes has been on the agenda of science (and politics) for many decades. Despite numerous initiatives, proposals and tentatives, the concepts themselves have not been clarified yet. At the same time, the demand to measure progress is stronger than ever before.

In the past decades GDP and its value measured at purchasing power parity have played the role of the principal indicator of development. Today, as a consequence of the growing importance of environmental considerations and the sharpening of social inequalities, it has become clear for science and politics that the multidimensional notion of social progress must be examined and managed in a complex way, and its measurement requires also new initiatives.

The current initiatives aimed to measure social progress develop in four directions:

- *The „correction" of GDP calculations with environmental aspects*
- *Establishment of indicator sets*
- *Development of composite indicators*
- *Measurement of well-being (happiness) with subjective indicators*

In 2007, the OECD launched an international project to measure the progress of societies. The Hungarian Central Statistical Office has also joined the project. In September of this year, the European Commission issued a Communication on the tasks of measuring social progress. A Commission comprising prominent Nobel Prize laureates and renowned experts was set up to clarify the notion of progress and elaborate proposals for its measurement. In October, a world conference will deal with the subject.

The lecture gives an overview of the initiatives aimed at measuring social progress. It presents statistical methodological issues, dilemmas and expected tasks related to the creation and development of the measurement.

Social progress- is a concept that in the different periods and different nations covers changing content. The several experts, explaining this subject do not reach a consensus concerning the notion of social progress. Most of them are agreed that social progress does not refer only to material well-being. The general opinion is that it would be good to have an exact picture on the degree of development of our country compared to other countries, and on the progress reached in the past period. Today, statistics are not able to give a totally adequate answer to these questions of common interest.

[1] Eszter Bagó, Deputy president of the Hungarian Central Statistical Office

In autumn of this year, a new, promising chapter was opened in the history of the measurement of social progress. Last September, it was in the ceremonial hall of the Sorbonne that the president of the French Republic presented in the presence of five hundred prominent guests the report of the Commission chaired by the Nobel Prize economist Joseph Stiglitz (Stiglitz 2009) on the measurement of economic performance and social progress.

The European Commission signaled also its clear commitment to measure social progress and issued a Communication on the subject (Comission of the European Communities 2009). The paper sets: The aim is to provide indicators that do what people really want them to do, namely measure progress in delivering social, economic and environmental goals in a sustainable manner.

The newest initiatives aimed at measuring social progress are led since 2007 by the OECD. The Global Project (Measuring the Progress of the Society) established by the OECD gives a framework to international initiatives; it advocates the necessity of new ways to measure social progress and organizes the works aimed at methodological development. The OECD has already organized three world forums dealing with the measurement of the progress of societies.

In the following part of my lecture, I would like to present the background of these initiatives.

For more than a half century, the most widely accepted measure of a country's progress has been the GDP. The system of national accounts, the creation of the GDP are linked to the name of Simon Kuznetz, American Nobel Prize winning economist. The use of GDP globally as a measure of economic progress began after the Second World War. In that time the growth of the economy was seen as the improvement of economic well-being. The system of national accounts and its headline indicator, the GDP fit perfectly into this concept.

Later, the indicator taking into account the purchasing power parity of GDP was also created for the sake of international comparison. The indicators which adjust the GDP for the relations with the rest of the world, the gross national income and gross disposable income became also widespread.

Since the introduction of the GDP, many economists have warned: that the GDP was a specialized tool. Gross domestic product is, by definition an aggregate measure of production. The notion of economic well being is a broader concept.

The most widespread criticism (Constanza 2009) against GDP is for example that the reconstruction after a natural catastrophe induces the increase of production and consequently of the GDP. Another popular example is that traffic jams may increase GDP as a result of the increased use of gasoline, but increase also air pollution and deteriorate our quality of life.

The GDP mainly measures market production, though it has often been treated as if it was a measure of economic well-being.

Material living standards are more closely associated with measures of real income and consumption. Production can expand while income decreases or vise versa. Because –for example- of income flows into and out of a country.

In the next part I will give a quick overview of the main tools that have been proposed to better measure socio-economic well-being.

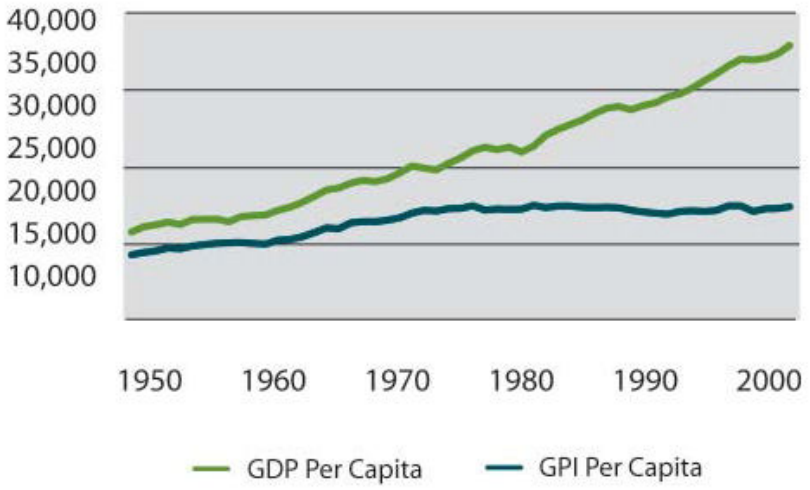
The first group of indicators corrects the existing GDP.

In the nineties there have been several attempts to develop alternative national income accountings to eliminate the deficiencies of the GDP accounting. The major objectives of these so-called „green GDP-s" are to provide a more accurate measure of welfare and to assess whether the development of an economy is sustainable or not.

The computation of these indices begins with the estimation of consumption expenditures, weighted by inequality in the distribution of income. Accounts for the non market benefits are added (socially productive time). Deductions are made to account the defensive expenditures (pollution) and costs associated with the use of natural capital.

These indices imply a decreasing proportion of economic benefits registered by the growth of the GDP, because such benefits are increasingly offset by the costs associated with growing inequality and deteriorating social and environmental condition.

Table 1. Real GDP and GPI Per Capita 1950-2004 in $ 2000



Source: Talberth, J. et al.(2007)

The most well-known among these indices is the Genuine Progress Indicator. On the figure, the clearer green line is the GDP, and the darker shows the values of GPI concerning the United States. It shows the significant differences of the two indices.
The trend found in many of the GPI calculations completed over the past years has put forward the evidence of a „threshold effect".

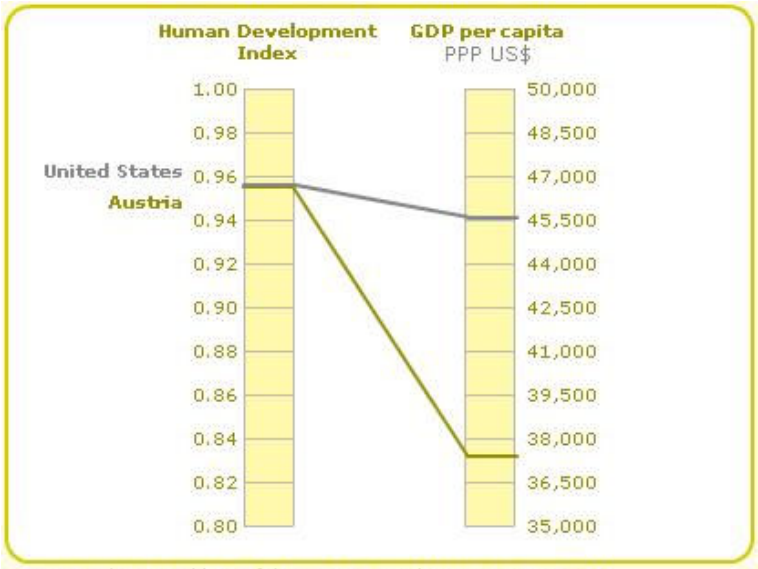
For every society there seems to be a period in which economic growth brings about an improvement in the quality of life, but only up to a point. This point is the threshold point. Beyond this point, if there is more economic growth, quality of life remains unchanging, or may begin to deteriorate.

The most important critic to these alternative indices is that they are arbitrary in the choice of the variables. Many methodological questions are open: the valuation of non renewable resources, the cost of environmental damages.
The other approach to measure social progress is to build a composite index. Composite indices can be integrated measure of complex issues, they are easy to interpret, but they can be misinterpreted, because of arbitrary weighting of their components

The Human Development Index, published by UNDP is the archetype of such composite indicator. The human development index is a summary composite index that measures a country's average achievements in three basic aspects of human development: health, knowledge, and standard of living

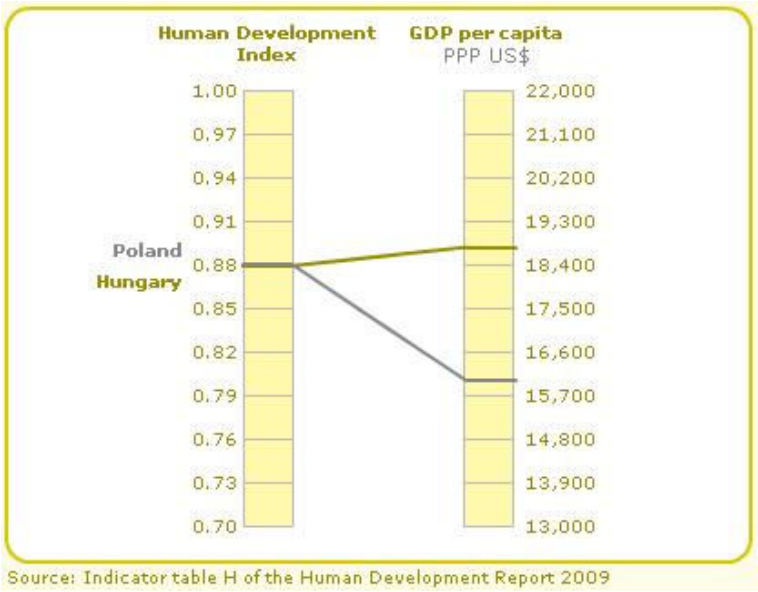
The Human Development Index combines indicators of life expectancy, educational attainment and income. The HDI sets a minimum and a maximum for each dimension, called goalposts, and then shows where each country stands in relation to these goalposts, expressed as a value between 0 and 1.
The next figures are examples to compare GDP and HDI value per capita.

Table 2. Human Development Reports- United Nations Development Programme



Source: Indicator table H of the Human Development Report 2009

Table 3. Human Development Reports- United Nations Development Programme



Source: Indicator table H of the Human Development Report 2009

We can see that in the case of the United States the value of the two indices of development is at the same level, while in the case of Japan, the value of the HDI is considerably higher than the level of development measured by the GDP.

In the case of Hungary, the level of development measured on the basis of the GDP is slightly higher than the value of the HDI.

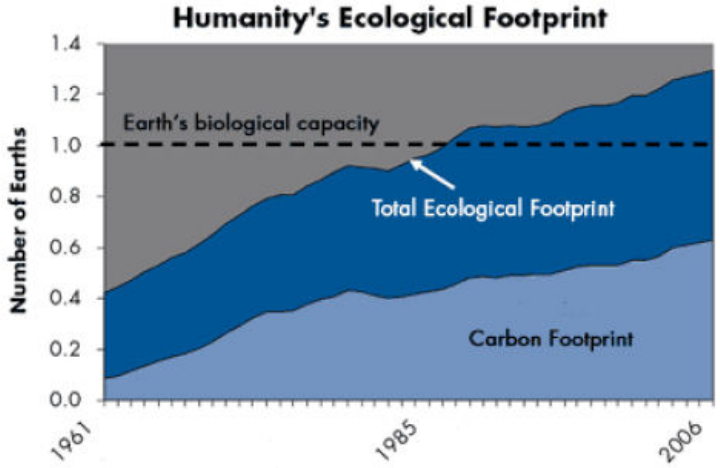
The danger of the climate change has highlighted the environmental aspects of development. The approach analyzing the sustainability of development only from the point of view of environmental sustainability has become popular. The ecological footprint index reflects this approach.

The ecological footprint is a measure of human demand on the Earth's ecosystem. While the term ecological footprint is widely used, methods of measurement differ. For 2005, humanity's total ecological footprint was estimated at 1.3 planet Earths – in other words, humanity uses ecological services 1.3 times as fast as Earth can renew them.

Another calculation: the total world Ecological Footprint is 2.7 hectare per person, while the world average biocapacity is 2.1 per person. This leads to an ecological deficit of 0,6 hectares per person. The countries which don't have an ecological deficit are called ecological creditor countries. They are first of all the underdeveloped countries: Congo, Gabon, and Mongolia. The countries with the greatest deficit, the ecological debtor countries are: Japan, Kuwait, United Arab Emirates, the USA. The data for Hungary is about the world average.

For practical reason, it is easier to measure the carbon footprint. Carbon dioxide emission is well measurable.

Table 4. Ecological Footprint, carbon component and earth's biocapacity



Source: Constanza et al. (2009)

The figure shows that environmental pressure deriving from economic growth exceeded already at the end of the past century the capacity of the environment. The biosphere does not have sufficient capacity to supply the needs of economic growth.

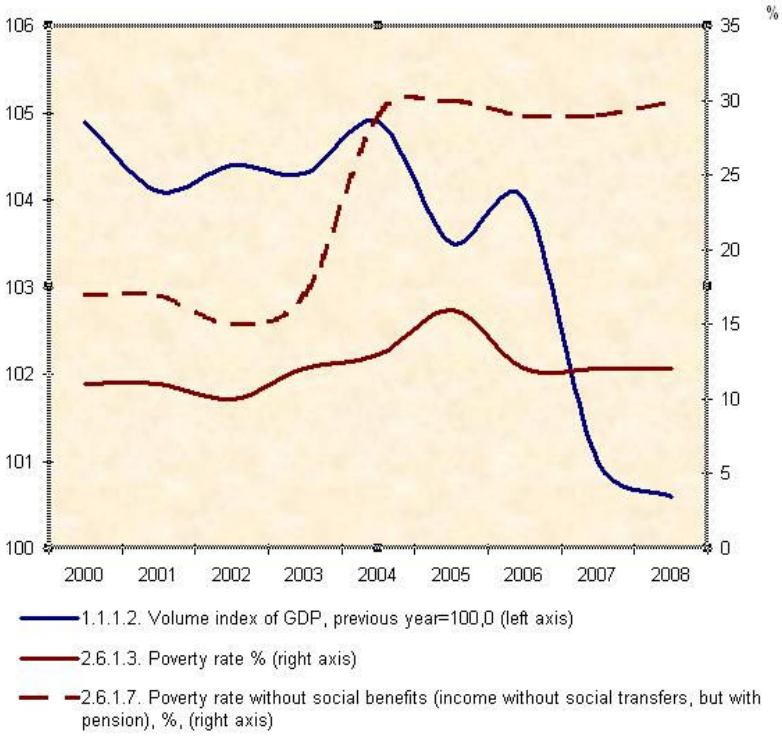
The use of indicators sets is a very widespread way of measuring progress. The indicator set collects the indicators characterizing the key aspects of progress. The users can evaluate and choose themselves the indicators they consider relevant. A lot of very good usable indicator sets are available. For international comparison, the most popular is the sustainable development indicator set of the European Union. As part of the OECD Global project, several indicator sets are available, which have been established in several countries to measure the progress of society. The Australian and Irish Statistical Offices have been publishing for several years their indicator sets, that politics consider also as a reliable assessment of the situation of the country.

The Hungarian Central Statistical Office published on its website at the beginning of 2009 its own developed indicator system. The selection of the indicators had been preceded by broad scientific consultation.

The indicator system has a hierarchical structure. The factors of social progress are split into three modules: indicators describing changes in the economy, society and the state of the environment. Within the three modules, we determined the most important factors of development to which we have assigned headline indicators.

Altogether 23 headline indicators were established. Headline indicators rely on secondary indicators, which explain and describe the developments of headline indicators. Detailed meta-descriptions have been enclosed to each indicator, describing the exact definition and sources of the data.

Table 5. Indicators to measure progress in Hungary



Source: www.ksh.hu

Note: The figure shows an example of the Hungarian indicator system. The blue line presents the evolution of the GDP, the red line shows the poverty rate. We can see that the evolution of the poverty rate does not move together with the GDP. The dotted line shows also that without taking into account social benefits, poverty rate

The figure shows an example of the Hungarian indicator system. The blue line presents the evolution of the GDP, the red line shows the poverty rate. We can see that the evolution of the poverty rate does not move together with the GDP. The dotted line shows also that without taking into account social benefits, poverty rate follows the decrease of the GDP.

An important advantage of indicator sets is that they highlight the complexity of developments and the opposite moving of the different processes. The limits of the dashboard of indicators is that there is no consensus about what are the main elements of good life, the selection of the indicators included in the set is ad hoc. There is no method to weight the different indicators, or the method is subjective, depends on individual value judgment.

Which are the factors that after the decades of several attempts have speeded up today the need of measuring social progress?

First of all there is the well-known fact that the increase of production results in the un-sustainable increase of environmental pressure. It is necessary to measure the use of non renewable natural resources. The decrease of environmental pressure can be considered as a key question from the point of view of sustainable development.

The other current factor is the recognition, that the rapid growth of developed economies of the second half of the past century seems to come to an end. In the past years, the quickest growth was characteristic of a group of less developing countries, let's think about China. Politicians in the developed countries obviously ask themselves the question: Is the speeding up of economic growth *the* only way to increase the well-being of society?

Moreover, several researches have shown that in the developed countries, well-being perceived by the population does not increase parallel with economic growth. According to calculations done with long time series, there is a turning point in the change of social well-being: Above a certain level of satisfaction of the needs, it becomes difficult to further increase the perception of well-being. Social well-being cannot any more increase in the same rhythm as measurable economic growth. In some cases, the relationship is opposite: rapid economic growth lowers community cohesion and the sense of setting goals. Growing income goes together with rising rates of alcoholism, depression, divorce, etc. The attempts to measure happiness show clearly that the feeling of happiness of people changes very slowly even in the conditions of rapid economic growth. This finding is named the Easterlin paradox (Easterlin 1974).

Easterlin pointed out that in spite of a 30 % increase of the GDP per head in the United States, the share of individuals who declare to be very happy did not increase in the eighties.

Last, but not least, there is a factor that is in close connection with statistics. It can be observed that while assessing their situation people tend to neglect statistical indices, and neglect those phenomena which are measured by statistics.

I would also like to use the example frequently cited by Enrico Giovanninin, head of the OECD Global Project: When we think about the future of our children, we hope that they will be healthy and will have a happy, peaceful, secure existence and family life. This is not the increase of the GDP that we wish, when we think about the future.

People don't know what to do with the great aggregates of macrostatistical indices. For individuals, there is no link between macro-aggregates and micro perception. Politics cannot successfully refer to the data provided by statistics. In many developed countries there is mistrust in official statistics.

The Stiglitz Commission mentioned in the introduction of my lecture has issued 12 recommendations concerning the development of the measurement of

social progress. In the following part of my lecture I would like to summarize briefly these proposals.

The first task is to overcome the shortcomings in the method of GDP calculations. In the evaluation of material well-being, income and consumption are more important than production. The available national accounts data show that in a number of OECD countries real household income has grown quite differently from the real GDP per capita, and typically at a lower rate. Measures of wealth are also central to measuring sustainability. A household that spends its wealth on consumption goods increases its current well-being but at the expense of its future well-being.

There is a need to give more prominence to the distribution of income, consumption and wealth. A rise in average income could be unequally shared across groups, leaving some households relatively worse-off than others.

Median consumption provides a better measure of what is happening to the „typical" individual or household than average consumption (income or wealth).

There is a need to broaden the measures of non-market activities.

Focusing on non-market activities, the question of leisure arises. Consuming the same bundle of goods and services but working for 1500 hours a year instead of 2000 hours a year implies an increase in one's standard of living.

Objective and subjective dimensions of well being are both important. Research has shown that it is possible to collect meaningful and reliable data on subjective as well as objective well-being. Despite the persistence of many unresolved issues, these subjective measures provide important information about quality of life. Those types of questions that have proved their value within small-scale and unofficial surveys should be included in larger-scale surveys undertaken by official statistical offices.

There is a need of a pragmatic approach towards measuring sustainability. An important notice of the Stiglitz Commission is that confusion may arise when one tries to combine current well-being and sustainability into a single indicator.

The environmental aspects of sustainability deserve a separate follow-up based on a well-chosen set of physical indicators. These separate sets of physical indicators will be needed to monitor the state of the environment.

It seems that the factors previously mentioned, urging the renewal of the measurement of social progress have led to a consensus between science and politics.

Are science and statistics prepared to find in a short time solutions to satisfy political needs? The answer cannot be a definite yes.

Critics of the new measures of social progress argue that data and methodological issues are barriers to the new solutions.

Having a long history, the SNA calculation developed an infrastructure and know-how which ensure a good quality of the GDP. The new indicator, if based on the GDP and SNA data, will lead to accurate measure. When the measure is based

on environmental or social data, the indicator may be less accurate. There will always be differences in data quality between indicators. That is, what made difficult to analyze the several aspects together.

But the main problem comes from the different scope and scale of the statistics characterizing social progress. Some data are collected at the level of individuals, others at the level of businesses, others at national level. It is also obvious that there is no solution to convert all the determinants of societal progress into one monetary equivalent.

Data availability varies in different statistical domains. The GDP figures are published quarterly. But, for example, the statistics about the income distribution are published worldwide with two years delay.

Last but not least the creation of new indicators must be preceded by the choice of the main values and goals of the society. The determination of the values and goals which the societies would like to pursue is not of the competence of statistics but is a political issue.

It seems that political determination has strengthened at international level. The recommendations of the Stiglitz Commission and the framework to measure the progress of societies worked out by the OECD this autumn could provide good starting point for the next phase of the work.

The creation of the new method of the measurement of social progress supported by international consensus can only begin. There have been a multitude of attempts in this respect. Today there are only some elements on which a new system can be built.

Obviously the GDP alone does not measure progress in relation to sustainability and well-being. Progress has to be measured in its complexity. But we must recognize that there are elements of subjectivity in evaluating progress. There is a need of a global dialogue for defining global goals and the unified approach is difficult to be found.

The chance of reaching international consensus is demonstrated by the process leading to the creation of the Millennium Development Goals by the United Nations.

It is quite clear that all indicators are proxies and limited in scope. It is sure that none can measure all significant aspects of progress. The several scopes and scales of determinants require the plurality of indicators. The indicator sets, like a dashboard, give information about the state of the important dimensions of society. So, the statistical offices should provide the information needed to measure all dimensions of progress, allowing the construction of different indexes and aggregates. We can also be sure that better indicators can be made and data quality can be improved.

We can read in the report of the Stiglitz Commission: „What we measure affects what we do, and if our measurements are flawed, decisions may be distorted". I am sure that we are before a paradigm shift. The sole objective of economic growth will be replaced by the complex goal of social progress. Political

debates concerning the components of social progress will continue. But the demands towards statistics will not become simpler. We have to prepare for the measurement of social-economic processes with an even broader scope, in an even better quality and timeliness.

References

Commission of the European Communities 2009: *Communication to the Council and the European Parliament. GDP and beyond. Measuring progress in a changing world.* http://www.beyond-gdp.eu/EUroadmap.html#communication

Constanza, R.- Hart, M.- Posner, S.- Talbrecht, J. 2009: Beyond GDP: The need for New Measures of Progress. *Boston University, The Pardee Papers* No. 4, January 2009. www.sustainable-economy.org/main/send_client_files?f...pdf

Deutsche Bank Research 2008: *The broad basis of societal progress.* www.db**research**.de/PROD/CIB...EN.../PROD0000000000232129.pdf

Easterlin, Richard A. 1974: Does Economic Growth Improve the Human Lot? In Paul A. David and Melvin W. Reder, eds., *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz.* Academic Press Inc, New York

European Parliament, Policy Department 2007: *Alternative progress indicators to Gross Domestic Product (GDP) as a means towards sustainable development* www.beyond-**gdp**.eu/download/b**gdp**-bp-goossens.pdf

Giovannini, E.- Hall, J.- Morrone, A. 2009: *A Framework to Measure the Progress of Societies.* OECD Working Paper, September 2009.

Measuring the Progress of Society. The website of the Global Project. http://www.oecd.org/pages/0,3417,en_40033426_40033828_1_1_1_1_1,00.html

OECD, INSEE 2009: *Survey of existing approaches to measuring socio-economic progress.* http://www.stiglitz-sen-fitoussi.fr/en/documents.htm

Stiglitz, J.- Sen, A.- Fitoussi, J. 2009: *Report by the commission on the Measurement of Economic Performance and Social Progress.* http://www.stiglitz-sen-fitoussi.fr/en/index.htm

Talberth, j.- Cobb, C.- Slattery, N. 2007: *The Genuine Progress Indicator 2006.* Redefining Progress, Oakland http://www.rprogress.org/sustainability_indicators/genuine_progress_indicator.htm

UNDP 2009: *Human Development Report 2009.* http://hdr.undp.org/en/

Emergence of new branches of statistics (Science, technology and innovation statistics)

Dr. Annamária Inzelt[1]

Statistical information and its analysis are key factors in any kind of decision-making process. In response to modern society's increasing demand for information, new branches of statistics must be developed in order to provide decision-makers with detailed and timely information. One of the relatively young branches of statistics, only one century old, is the branch of science, technology and innovation (STI) statistics.

This paper focuses on the emergence and development of science, technology and innovation (STI) statistics that has resulted in internationally harmonised norms, classification and comparable time series.

STI indicator development is an ongoing process. In the 21st century, it is critical to improve measures for the internationalisation of STI in order to provide new tools for policymaking and evaluation. This process requires additional internationally comparable databanks as well as a better understanding of currently unmeasured factors in the STI internationalisation process. The first section of this paper gives a short overview of the background leading up to the emergence of STI statistics. The second section focuses on the new epoch: the post-war period when demand, actors and speed of development in STI statistics changed significantly. These changes resulted from the recognition of the importance of scientific policy, which created the need for research and development indicators. The third section gives a detailed account of international comparability, an area which gained importance as competition between nations as well as the internationalisation of research and development (R&D) activities both created a strong demand for internationally comparable indicators. The OECD played an important role in this process that lead to the creation of the Frascati Family manuals and internationally comparable time-series. The adoption of OECD standards in a transition economy, namely in Hungary, is described in the fourth section, while the final section gives some concluding notes.

Keywords: science, technology and innovation statistics, Frascati Family manuals, transition economies

[1] Annamária Inzelt, Director of IKU Innovation Research Centre, Financial Research Corp., core member of Doctoral School of Szeged University (Hungary)

1. Background

As modern society's demand for information increases, new branches of statistics must be developed in order to provide decision-makers with detailed and timely information. Statistical information and its analysis are key factors in any kind of decision-making processes. It was because of the needs of states to collect data on their people and economies, and to administer these data that modern statistics emerged in the 18th century.

One of the relatively young branches of statistics is the branch of *science, technology and innovation* (STI) statistics. Dramatic changes in the socio-economic environment resulted in the need for systematic information on research and development activities, and on the output of scientific efforts.

During the era of the second industrial revolution, in the late 19th century, the role of science and research activities had significantly transformed the economic life of forerunner countries. The emergence of industrial research and development (R&D) transformed the mode of operations for innovative work and the attitude toward the cost and benefit of scientific work. In the majority of industries, links between science and technological innovations grew closer, as this era was marked by changes in various areas: great inventions, organisational innovations, and the expansion of educated people.

One of the important organisational innovations was the creation of industrial laboratories. At the same time, the growing need for an educated workforce affected the education system, including colleges and universities. The growing role of profit-oriented funders in science and the increasing cost of research made investors in science much more interested in the input, output, and outcome of research and development activities. *All of these changes created a demand for STI statistics that has developed gradually over the last century.* Scientists themselves also became interested in S&T statistics.

The history of STI indicators reaches back more than just 100 years. As identified by Benoit Godin (2007), an historian of STI indicators, the first systematic STI publication was American Men of Science, compiled regularly by the American James McKeen Cattell and published between 1906 and 1944. Cattell had edited the still prestigious Science journal for decades. This journal published short scientific CVs on the authors and accumulated thousands of CVs. Cattell exploited the information on authors thus gathered in order to create a repertory of American scientists. Its first publication contained demographic, geographic and scientific performance indicators on 4000 American scientists. This publication provided

information on the relative strength of individual scientific fields per geographic region, and on who are the most successful scientists by fields and age cohorts.[2]

There were *various users* of this series of publications: for example, universities considering whether to appoint an applicant to a tenured position. Researchers used it when seeking collaboration partners in their own or another field of science, as did various clients to know which science is strong in a given region or which region is strong in a particular science.

One of the first attempts to focus on R&D activity from the *policy point of view* was in the 1930s. It was at that time that policymakers first appeared among the users, even if only from a distance. There was an attempt to measure input and output of research and development activities in the Soviet Union's centralised planned economy, where everything was approached as important macroeconomic growth factors. Lundvall and Borrás (2005 p 604) mentioned in Western Europe, 'according to Chritopher Freeman science policy was recognised as a policy area through the pioneering work by Bernal (1939) Bernal was a pioneer in measuring the R&D effort at the nationa level in England.' 'In the 1930s Bernal made the first attempt to measure the effort made in science by relating R&D expenditure to the national income of the UK.' (616)[3]

Until the end of World War II, there were only few countries that prepared S&T statistics, the majority of which existed merely as research products and focused on researchers as the most important assets of science. Demand for S&T statistics and the involvement of stakeholders changed in the early post-war years.

The shift in the concept of S&T statistics is usually linked to the Vannevar Bush report (1945, Science, The Endless Frontier) prepared by V. Bush as the first presidential advisor to President Roosevelt right after World War II. He proposed to create a peacetime government research and development agency. In 1950, the National Science Foundation was created in the US. This organisation made the first systematic collection of data employing surveys and administrative data and analysis of data and indicators.

In the early post-war years, S&T policy became an immanent and independent part of governmental policies. (The emergence of science policy was influenced by the experiences of World War II and by the start of the Cold War.) When the

---

[2] Similar data sources were available for at least few fields of science in several other countries. (For example: a biologist compiled data on biologic researchers in Belgium in the mid-19th century.) This publication remained as a matter of special interest. French experts prepared a repertory based on a systematic datasheet on S&T personnel but did not exploit this source statistically to support S&T policymaking. They employed only for identifying relevant knowledgeable people for military purposes.

[3] Bernal J. D. (crystallography and molecular biology) worked on these indicators following his visit to the Soviet Union. He focused on social function of science, attempts to measure progress in science by relating R&D expenditure to the national income in Great Britain.

importance of science policy as a new policy area was recognised (Bernal 1939, Bush 1945) the need for R&D indicators for policymaking came into the spotlight. (Lundvall-Borrás 2005 pp. 599-631) The post-war era opened a new epoch in STI statistics.

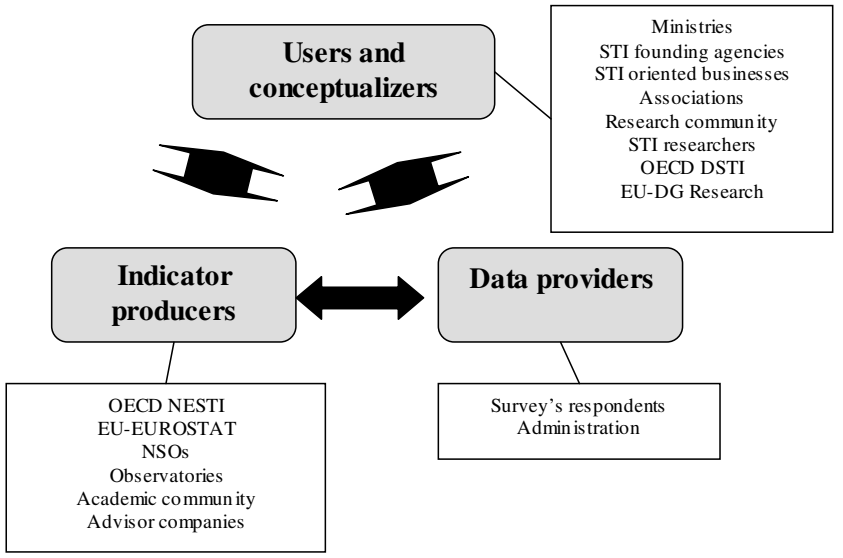
2. New epoch in STI statistics: changing demand, actors and development in STI statistic

The main difference between pre- and post-war S&T statistics was in their *conceptual framework*, as the issues covered by STI statistics and the actors producing the statistics changed significantly after the war. The measurement concept for RDI became economic in character. The result is a collection of economic indicators that are compatible with other economic datasets. (Many dimensions of RDI activities remained out of the measurable field.)

Conceptual foundations are crucial in the development of STI statistics even if they are rarely considered when indicators are used, as S&T (and innovation) statistics always rest on some kind – explicit or implicit – of conceptual foundation. In the 1950s, the conceptualisation and construction of STI indicators, as well as the collection and analysis of internationally comparable STI data and indicators, started. In the 1950s, British researchers developed the conceptual framework, definitions and classifications for measuring R&D. One of the main problems was to define research in a way that allows to measure research activities in different fields of science in a comparable way. During the preparation of the 'Green Book' for government R&D policy, the House of Lords discussed a study that backed up a unified definition of R&D for national policy. This study presented more than 40 definitions of research that was used previously in ad hoc measures. (Lord Rothschild 1972)

STI statistic work is an interactive process and the statistic is a joint product of various actors in the process. Figure 1 shows the schema of actors and their relations.

Figure 1. Actors and Linkages in Preparing Indicators



Source: own creation

The following lists some breakthrough concepts that led new streams of STI indicators. These examples illustrate that research community was also an important initiator of and actor in the development of STI indicators.

- *Polanyi's concept* on the codified and tacit dimensions of knowledge is a great challenge to measure diffusion of knowledge. Developments of scientometrics and patent indicators are a good illustration of designing indicators to measure codified knowledge. Indicators for measuring the diffusion of tacit knowledge are still in the blue-sky or development stages. Pilot indicators on the mobility of science and technology personnel are promising indicators in this field.
- *Rosenberg and Kline's work* (1986) relates to the conceptual foundation of innovation indicators. That work had an explicit effect on the OECD's (Organisation for Economic Co-operation and Development) Innovation Manual (Oslo Manual 1992, 1997, 2005). As Smith summarised (2005, p. 150), the chain-link innovation model (Rosenberg and Kline 1986) has important implications for indicator developments: (a) innovation indicators should pick up small-scale changes that originated from the learning process and that may lead to important technological and economic outcomes; (b) innovation input indicators should cover non-R&D input (design, engineering developments, experimentation, training,

exploration of markets for new products. Measuring the networking dimension of innovations is still in the blue-sky stage.

3. Demand of Users

Demand of users has always had a crucial impact on the measuring, disaggregation and frequency of data collection. Beside the conceptual development, the emergence of new policy issues played an important role in broadening the STI statistical field. Both important scientific studies on STI-policy related issues and new policy challenges (European sclerosis, emergence of new technologies, globalisation, global warming, and ageing population) have created a demand for more fact-based analysis on science, technological development and on the innovation process. Heightened demand for S&T statistics also increased the involvement of stakeholders.

In the 1980s, policymakers had no reliable relevant indicators to support them in better understanding the changing world and to back up strategy-making. However, there was a recognition in international policy circles that technological development and innovation are crucial factors influencing economic growth, efficiency and employment.

Since the 1990s, the demand for indicators has been increasing. Policymakers as well as economic actors seek an accurate portrayal of the relationship between technological development and economic performance. This increasing demand has lead to the development of information that allows for the identification of the economic importance of high-tech industries, in particular the role of information and communication technology, their contribution to national performance in global competition. Detailed information on R&D personnel, on R&D expenditures, and on the effects of public investment in R&D is becoming important for policymakers. Other stakeholders such as leading industrialists are becoming more and more interested in S&T statistical information to back up their strategic decisions, since R&D investments play an important role in competition. The academic community also showed an interest in some types of information.

A good example of demand-led indicator development is the appearance of innovation indicators. The emergence of innovation in scientific work and as policy issue was another breakthrough in the second half of the 20th century in the development of this new branch of statistics. Beside science and technology policy, innovation policy also emerged in the 1950s and 1960s and created a new demand for statistics. Gradually developed indicators and more detailed disaggregation by large and small businesses and by manufacturing and technology intensive service sectors improved our knowledge on the innovation process, and provided support for further development of concepts and measures. (Smith 2005)

Innovation statistics gained further impetus in the 1990s. Several deployments have affected the innovation systems of world leading economies, such as the globalisation of STI. (OECD 1997) This also created a new demand for indicators. Decision-makers are now seeking information on issues that were not even of peripheral interest to them a decade ago. They need information on different RDI and competition performances by global regions; globalisation of RDI; cross-border RDI collaborations; cross-border mobility of highly skilled workers, and its impact on flow of knowledge. In recent years, new indicators have emerged on innovation input and output combined with economic data. Economy-wide measures have some degree of international comparability. Beside sectoral disaggregation, improving data coverage allows for regional disaggregation as well.

These mutual developments of quantitative information, scientific advance and policy needs have initiated a new track for STI indicator developments in the second half of the 20th century and are also the locomotives for 21st century RDI statistics.

4. Key producers and developers

The key producers of R&D statistics changed significantly after World War II. The first systematic collection of data was carried out in the US by NSF. Since its establishment in 1950, the National Science Foundation has been organising surveys and analyses of data and indicators. Other regions of the world also developed their national STI statistics. Nowadays the leading role rests on official statisticians (institutions vary by national settings). Scientists and hobby indicator developers are being gradually replaced by statisticians.

In several countries, STI observatories and STI platforms are important producers of indicators and analytical reports. Researchers have remained important and visible figures in designing indicators, in developing academic databanks, carrying feasibility studies, and in identifying emerging needs.

Beside official statistics, individuals or research teams have developed important classes of indicators with related databases as research tools. *Even spot data and short time series can help to put old questions of science in a new light.*

As was illustrated by this little detour through conceptual work, policy needs and indicator development, conceptualisers, users, indicator producers and data providers are crucial actors in the development of STI statistics and their interactions have an important influence on the availability and quality of information.

5. International comparability

Competition among nations, the internationalisation of R&D activities and the diffusion of international collaboration have all served to create a stronger demand for the international comparison of relevant indicators. Thus, the search for the international comparability of STI indicators is an important chapter in the development of this branch of statistics. The differences between countries in their knowledge producing and accumulating capabilities, the role of these capabilities, and their impact on economic competitiveness can only be analysed using internationally comparable indicators.

The collection and analysis of internationally comparable STI indicators and data started in the 1950s along with the conceptualisation and construction of STI indicators. The exchange of knowledge between nations and active collaboration between countries were crucial in the international harmonisation procedure. (Sirilli 2005, 2006) Work on international comparability started in the 1950s at the predecessor of OECD.[4] 'The OECD played not a single but a unique role among international organisations in STI policy and conceptual debates, in the development of instruments used for measuring and producing an internationally comparable databank and indicators.' 'In the late 1950s and early 1960s Christopher Freeman played a key role in developing the analytical basis of science policy and it is significant that he also was one of the architects behind the Frascati manual that in 1963 gave the OECD and national authorities methods to measure R&D compare the effort across countries.' (Lundvall-Borrás 2005)

The OECD member states in 1960s were very active players in developing internationally comparable databanks and indicators. The Nordic countries were among the forerunners of international harmonisation. (Young-Westholm 2006) Parallel to the conceptualisation and construction of STI indicators, the collection and analysis of internationally comparable STI data and indicators were organised.

---

[4] The forerunner of the OECD was the Organisation for European Economic Co-operation (OEEC), which was formed to administer American and Canadian aid under the Marshall Plan for the reconstruction of Europe after World War II. Since it took over from the OEEC in 1961, the OECD's vocation has been to build strong economies in its member countries, improve efficiency, home market systems, expand free trade and contribute to development in industrialised as well as developing countries.

OECD was established on 30 September 1961. The founding members are: Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States. Latter other countries became members Japan (1964), Finland (1969), Australia (1971), New Zealand (1973), Mexico (1994), the Czech Republic (1995), Hungary, Poland and Korea (1996), the Slovak Republic (2000).

Since the early 1960s, collected data coverage has broadened significantly. More and more issues were measured and various types of input-output data became available for the majority of countries and time series have been gradually developing. Nowadays several throughput (process) and impact data also occur.

In the field of STI indicators, the OECD manuals, called the *Frascati Manual Family*, became international standards. (Sirilli 2006, Gault 2009, OECD 2002)

Following the first milestone in the development of internationally comparable indicator methods, the publication of the Frascati Manual – focusing on measuring research and experimental development, financing issues and human resources devoted to R&D activities – other manuals have been prepared reflecting broadening needs for STI information. Table 1 summarises the Frascati Manual Family by the first appearance of new topics covered by new manuals.

Table 1. Frascati Manual Family: International standards for measuring STI

Year of publication		Title of the Manual	Type of Data	Availability in Hungarian
First	Revision			
1963	1988 1993 2002	***Frascati Manual:*** Proposed Standard Practice for Surveys of R&D	Research and Experimental Development	✓
1989	Since 1993 in the FM	*R&D Statistics and Output Measurement in the HE Sector.*	Higher education R&D	
1990	-	***TBP Manual*** : for the Measurement and Interpretation of Technology Balance of Payments Data	Technology balance of payments	✓
1992	1997 2005	***Oslo Manual:*** Proposed guidelines for Collecting and Interpreting Technological Innovation Data	Innovation	✓
1994	2009	***Patent Statistics Manual***	Patent data	(1994)✓
1995		***Canberra Manual***: The Measurement of Human Resources Devoted to S&T	S&T personnel	✓

Source: own creation on the basis of the Frascati Manual 2002, p. 16.

Notes: * since 1997, OECD and Eurostat joint publication

** TRIAD: United States, EU and Japan

The manuals provide internationally harmonised definitions and such tools that are vital in order to speak in the same language when comparing indicators internationally.

These manuals are basically technical, methodological documents that were written by experts for experts. The preparation of each manual took a few years with the involvement of many experts working on the conceptualisation, feasibility and pilot surveys at national and international levels.

Regular revision dates in Table 1 show how methodological development is an ongoing interactive process. Manuals support development in surveying and analysing processes, while accumulated data collection and analyses likewise encourage the revision of manuals from time to time, as illustrated in the second column of Table 1.

At the OECD, there are some experimental methodological works that have yet not resulted in internationally accepted, harmonised manuals but which are contributing to the development of STI measures in several fields. Table 2 summarises these manuals.

Table 2. OECD STI Manuals besides the Frascati Family

Type of data	Title	**Availability in Hungarian**
High technology	Revision of High-technology Sector and Product Classification (OECD, STI Working Paper 1997/2)	-
Bibliometrics	Bibliometric Indicators and Analysis of Research Systems, Methods and Examples, by Yoshiko Okubo (OECD, STI Working Paper 1997/1)	✓/-
Globalisation	Manual of Economic Globalisation Indicators	-

Source: own creation on the basis of the Frascati Manual 2002, p. 16.

Beside so-called STI manuals, there are some other relevant internationally harmonised statistical frameworks prepared primarily for other measuring purposes that are regularly employed in the preparation of STI statistics. (Table 3) These borrowed methodologies are important in combining various branches of statistics for analysing complex systems such as innovation system or higher education system.

Table 3. Other relevant OECD statistical frameworks

Type of data	**Title**	**Availability in Hungarian**
Education statistics	OECD Manual for Comparative Education Statistics	✓
Education classification	Classifying Educational Programmes, Manual for ISCED-97 Implementation in OECD countries (OECD 1999)	✓
Training statistics	Manual for Better Training Statistics – Conceptual, Measurement an Survey Issues (OECD 1997b)	
International standard of Industrial classification	ISIC Rev 3. (NACE)	✓
Classification of Occupations	ISCO (International Labour Organization, 1990)	✓
Field of Science Classification	FOS (OECD)	✓
Classification by field of research	ISI (classification of journals covered by Web of Science ISI) CWTS	✓
Classification of R&D activities by functions	COFOG SNA/OECD	✓

Source: own creation on the basis of the Frascati Manual 2002, p. 16.

The manuals and availability of data and indicators contribute considerably to the better understanding of the role and importance of science, technology and innovation, the importance of codified and tacit knowledge. We can understand better how the science system works, how the system of innovation is changing, what are the links between innovation activities, sectors and size of companies and so on.

The OECD at a global level and the EU as a regional organisation are playing important roles either as initiator and/or coordinator in developing novel RDI indicators to respond to new challenges. To mention only a few new activities: measuring R&D outsourced abroad inside or outside corporations; handling immaterial R&D assets in SNA (System of National Accounts); measuring new

emerging fields such as ICT, biotechnology and nanotechnology; measuring various types of RDI collaborations; measuring the diffusion of knowledge; measuring the impact of globalisation (or Europeanisation) of RDI activities.

Appropriate indicators and time-series can be used not only in analyses but also to support other tools of intelligent policymaking such as evaluation, assessment and foresight exercises. Today national R&D and innovation statistics are quite detailed and quite a significant part thereof is internationally comparable.

6. Adaptation of OECD standards in a transition economy

Before the transition period Hungary, similarly to other former socialist countries, employed different standards (if any) to measure STI activities. As part of their accession to the OECD (1996) and to the European Union (2004), it was crucial for Hungary and for other transition economies to adopt the international standards that were employed by democratic market economies.

As Hungary is a full member of the OECD and EU, it had to accept their standards and organise its data collection in an internationally comparable way. EU laws are compulsory for Hungary as a member state. In addition to compulsory EU tasks, national demand for RDI time-series is certainly important as well.

The adoptation of these standards was not a simple exercise, as the OECD countries which had developed the Frascati Family manuals all shared the characteristic of being advanced economies. It worth emphasising this feature, as it has an important influence on the demand for information. Countries such as transition economies that joined as latecomers to the club could not use everything they got as readymade.

Since the beginning of the transition period, Hungary has done a lot to revise, modernise and adjust its STI system to market economy demand and international standards. (Hüttl et al. 1997, Inzelt 1994, 2002, 2003, Szunyogh-Varga 2004,)

The adaptation procedure and dissemination of STI indicator knowledge in Hungary are summarised in Table 4.

Table 4. Adaptation of Frascati Family and other STI manuals in Hungary

Hungarian publication of manuals (translations or summaries in Hungarian)		**Pilot surveys**		
OECD Manuals	Year	Type	Year	Prepared by
Oslo Manual	1993 & up-dated	Feasibility and pilot surveys on innovation - manufacturing sector (large and medium firms) - selected service sectors - manufacturing sector (large and medium firms) - small and micro firms - manufacturing sector (large sample)	 1994 1999 2000 2001 2000	 IKU IKU HCSO-IKU IKU-PTE HCSO
R&D at small businesses (OECD working document)	1994	Testing the journal publication based method	1994	IKU
Technology Balance of Payments (Summary) (TBP)	1995	Hungarian National Bank takes into account technology payments of its information system Introduction of R&D export/import survey	1996- 2004	MNB KSH
Frascati Manual	1996 up-dated	-inserting some elements into the regular economic survey - pilot survey for revising regular R&D survey by FM - revised R&D survey	1993 '96/7 1998	HCSO IKU-HCSO HCSO
Patent Statistic Manual (Summary)	1999	- research work has started	2003	
Canberra Manual	2000	- feasibility studies - Blue Sky projects (EU funded ENMOB, ERAWATCH)	'98/9 2000-	IKU IKU-HCSO IKU

Source: own creation

Notes: The first translations or summaries of the manuals were prepared by IKU except for the Oslo manual. The manuals were published by the OMFB (predecessor of the National Office of Research and Technology). In the pre-OECD membership period, the translations were supported by OECD. Pilot and feasibility studies were used as samples for other transition economies and developing countries.

IKU: Innovation Research Centre, Hungary (now belongs to Financial Research Ltd.)

HCSO: Hungarian Central Statistical Office

MNB: Hungarian National Bank

An important step of knowledge dissemination was the publication of a Hungarian translation or summary of the manuals. (See 1st and 2nd columns of Table 4) The publication of the Hungarian versions of manuals was accompanied by feasibility or pilot surveys and their analyses (listed in column 3). The regular R&D statistic survey was revised by the Frascati Manual, and previous time-series were made comparable with methodological bridges. Novel indicators and their survey methods on innovation were introduced based on the Oslo Manual and through an adoptation of EU-Eurostat CIS (Community Innovation Survey).

Hungary has at its disposal more than 10 years of time-series of many RDI data and indicators, but some important indicators (GBOARD, financial data on R&D programs) are still missing. Further revision is needed on higher education expenditures (HERD) data, while sectoral mobility and the international mobility of highly skilled workers are hardly measured.

A detailed overview on the availability of Hungarian RDI data and indicators by international standards can be found in Inzelt et al., 2008. Demands of national users are summarised in Inzelt et al., 2009. These studies identified the strengths and weaknesses of Hungarian RDI indicators, surveying methods and also discussed how the shortcomings of the STI information system can be overcome.

Besides developing its own system, Hungary can participate with its capacities in the revision of survey methods, existing international standards and in the international development of novel indicators that attempt to respond to new STI policy challenges. Hungary has to make its own decisions regarding which topics are important for its stakeholders in forthcoming years.

7. Concluding notes

Accumulated quantitative information on the availability of RDI data supports fact-based policymaking, business strategy formulation and further research. The fact-finding approach has improved our understanding of the innovation system and initiated a new track for STI indicator developments in the second half of the 20th century. All of the relevant indicators and their analysis helped to understand the mechanisms that influence scientific and innovative performance, as well as how policy can strengthen or diminish their roles. At the same time, STI statistical measuring gained a more important role in intelligent policymaking.

At the beginning of the 21st century, evidence-based policymaking requires indicators to monitor, assess and evaluate research programs and STI policies. The quality of information depends on improving the availability of statistical data and the development of indicators that reflect the complexity of the STI process.

Today, as this relatively young branch of statistics is becoming an adult member among the various branches of statistics, the country specific development

of STI statistics strongly depends on the national culture of policymaking that influences national demand.

References

Bernal, J. D. 1939: *The Social Function of Science*. George Routledge & Sons Ltd., London.

Bush, V. 1945: *Science: The Endless Frontier.* United States Government Printing Office, Washington.

Gault, F. 2009: The OECD Innovation Strategy: Delivering Value. *Foresight,* 1,9, 16-28. p.

Godin, B. 2007: From Eugenics to Scientometrics: Galton, Cattell, and Men of Science. *Social Studies of Science*, 37, 5, 691-728. p.

Hüttl, A. - Inzelt, A. - Varga, E. 1997: A kutatási-fejlesztési statisztika megújítása. (Revision of S&T Statistical System). *Statisztikai Szemle*, 75, 11, 893-907. p.

Inzelt, A. 1994: Preliminary lessons of Hungarian Pilot Innovation Survey. OECD Working Paper, Paris.

Inzelt, A. 2002: Attempts to Survey Innovation in the Hungarian Service Sector. *Science and Public Policy* 29, 5, 367-383. p.

Inzelt, A. 2003: Innovation Process in Hungary. In Shavinina, L. (szerk.) *The International Handbook on Innovation*, Elsevier, 859-872. p.

Inzelt, A. - Goldperger, I. - Szunyogh, Zs. 2008: The Information base of science, technology and innovation policy. Decisions: on what basis? *Pénzügyi Szemle,* 53, 2, 214-228. p.

Inzelt, A. – Csonka, L. – Nyiri, L. – Varga, Gy. 2009: Mindenki másképp, de csinálja! A tudomány, technológia és az innovációpolitika statisztikai elemzési bázisa. *Európai Tükör*, 15, 1, 5-33. p.

Kline, S. J. - Rosenberg, N. 1986: An Overview of Innovation. In Landau, R. – Rosenberg, N. (eds): *The Positive Sum Strategy. Harnessing Technology for Economic Growth*. National Academy Press, Washington.

Lord Rothschild 1972: Forty-five Varieties of Research (and Development). *Nature,* 239, 373-378. p.

Lundvall, B. A. - Borrás, S. 2005: Science, technology and innovation policy. In Fagerberg, J. - Mowery, D. – Nelson, R. (eds): *The Oxford Handbook of Innovation*. Oxford University Press, Oxford-New York, 599-631. p.

OECD 1997: *National Innovation Systems*. OECD Publishing, Paris.

OECD 2002: *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development.* OECD Publishing, Paris.

OECD 2005: *Oslo Manual.* OECD Publishing, Paris.

Sirilli, G. 2005: *Ricerca & Sviluppo*. Il Mulino, Bologna.

Sirilli, G. 2006: Developing science and technology indicators at the OECD: the NESTI network, 16 p. Presented at the ENID / PRIME International

Conference "Indicators on Science, Technology and Innovation: History and New Perspectives", Lugano, Switzerland, 16-17 November 2006.

Smith, K. 2005: Measuring Innovation. In Fagerberg, J. - Mowery, D. - Nelson, R. (eds): *The Oxford Handbook of Innovation*. Oxford University Press, Oxford, UK, 148-177. p.

Szunyogh, Zs. - Varga, A. 2004: A K+F statisztika módszertana. *KSH módszertani füzetek*, 42. Központi Statisztikai Hivatal, Budapest.

Young, A. - Westholm, G. 2006: The Participation of the Nordic Countries in the Early Development of OECD Science and Technology Statistics, 9. p. Presented at the ENID / PRIME International Conference "Indicators on Science, Technology and Innovation: History and New Perspectives", Lugano, Switzerland, 16-17 November 2006.

Re-engineering of eNAQ - An electronic National Accounts Questionnaire Based on 1993 & 2008 SNA

Jens Chorus – Hans – J. Lenz[1]

1. Introduction

The prototype "eNAQ" - an electronic National Accounts Questionnaire - was developed for the electronic capturing and administration of National Accounts Questionnaires (NAQ) at the United Nations National Account Section (UN NAS).

In 2005 ENAQ was first implemented by Eichler (2004)) based on a requirement analysis and design specification of Müller (2003)). Recently, a complete re-engineering was undertaken by the first author, paying special regard to the additional requirements derived from the pending switch to the 2008 SNA version.

The task of eNAQ is to support all users (experts and administrators) at national and international organizations to efficiently enter data in a user friendly way by a personalized user interface (GUI). The systems novelty is to allow users to enter data worldwide, and to assist the NA experts in validating the data at the point of entry. eNAQ is not meant to pre-describe, produce or even dominate how to produce statistical annual reports on any national level, nor to compute automatically statistical data. eNAQ is a software tool supporting data entry backed up by an integrated validation rules checker. The rule system is based on internationally agreed rule sets, yet, allowing for national particularities of national accounts methodologies.

This paper only summarizes main features of the prototype, and the revised requirements according to the 2008 System of National Accounts (SNA). The features presented are based on the NAQ and the system currently used at the United

[1] Institute of Statistics and Econometrics, Freie Universität Berlin

Acknowledgment: The authors express their gratitude to Jörn Eichler, Darmstadt, for his contributions to clarify and improve parts of this paper as far as the implementation of eNAQ is concerned. They are very thankful to Patrick Müller, Washington, who always helped us to keep track of 2008 SNA new directions and changes. We feel free but are not authorized to call him "The Third Man".

Nations. The prototype and requirements analysis include visionary elements and features, which are not necessary part of current mission and strategy of the UN National Accounts Section. Major parts of the 2003 specifications where implemented with the new system at the United Nations, other essential parts were implemented later via satellite systems. The core part of eNAQ, the local data entry and validation by the country reporting to the United Nations, was never implemented. A lot of the work with respect to data entry and validation at National Statistical Offices (NSOs) and the UN NAS is still done manually and could be efficiently supported by eNAQ.

The envisioned benefits from using eNAQ are a simplified data transmission between NSOs and the UN NAS that don't require double manual data entry or editing for the UN. Data can be easily transferred with eNAQ via Internet, and, where Internet is not available, import and export functions guarantee a flawless data transfer. Only one single data entry effort needs to be done by the country office providing the data. In real time an automatic electronic validation is performed. Data submissions and entry using other sources (paper publications, pdf-files, or individualized excel tables) that are time consuming, costly and vulnerable to mistakes should become the exception. The main goal is the acceleration and simplification of workflows, reducing data handling cost while improving data quality.

2. National Accounts Compilation

Some figures are displayed to illustrate the complexity of the annual UN NAS report. The UN NAQ for the 1993 SNA consists of 31 tables containing in total 1046 items. The current rule set for the 1993 SNA contains of 814 rules that need to be validated. Note that additionally, revised country data beyond the coverage of the standard NAQ years, have to be updated and validated in the UN NAS database when it becomes available. Let us assemble some facts about national accounts compilation. NA seems to play a role in economics and national statistics similar to the role mathematics plays in science: National accounts and, of course, input-output analysis is the "Queen". NA delivers information on main statistics as economic indicators for Government, Parliament, Press, and Science. It is an ensemble of flows, stocks and values per units (VpUs). It assumes that the fundamental balance equations of an economic equilibrium are fulfilled: Supply ≡ Demand (EX-Post Identities). The main economic indicator system includes many indicators of which the Gross Domestic Product (GDP) is not the only but the most prominent one. For instance, it helps to determine per capita GDP using the population size PS and GDP.

The current world wide standardized and fairly harmonized system of national accounts (SNA) - with respect to definitions, measurements and quality - is a

distributed, loosely federated information system. This is true in the sense that national data is collected locally, autonomously on heterogeneous hard- and software platforms owned by very divergent countries, and validated using a national set of validation rules of type "edits" or "if…then" rules. Furthermore, there are supra-national rule sets for each pair of countries used for
validation on the international level. A very simple, nearly trivial example of a rule is the "Production in a Sector" - Rule:

$$Intermediate\ Consumption_{Sector} + Value\ Added_{Sector} = Output_{Sector}$$

NA computing is in its core 'Estimation', a term precisely coined in Mathematical Statistics. Missing units like companies or citizens, missing values ("null values" termed by IT specialists), outliers, non-linearity of relationships like sales=quantity*price/unit, non-Gaussian distributions, or national economic policy, the human behavior in general, or even strikes of labor-unions or economic sectors influence its timeliness, quality (precision, certainty) and cost of computing. Moreover, think of the existence and size of black
markets around the world.

Finally, it is worthwhile mentioning that various Excel-like tables which represent the annual NA reports are mandatory and necessary for any input-output-analysis, both on the national as well as international level. Take for instance Germany as a country, the EU (the European Union) as one out of many political or economical groups or even the world (=ALL) representing the world's economy status per toto at the UNO, IMF or World Bank.

3. The System of National Accounts

The System of National Accounts is a set of standard recommendations regarding definitions, concepts, classifications and foremost, accounting rules for the measurement and compilation of economic activities. Harmonization is a major problem in NA. For example, there is a lot of skepticism within the European Union regarding what is referred to as an over-standardization. The argument is that some national differences are vital to European diversity, as expressed by the second author in a meeting of EuroStat a couple of years ago, in the term the "Standardized European Tomato". A well designed "European Federated
Statistical Information System", as firstly proposed by Lenz and Shoshani (1996) could provide a viable solution for the compilation of NA data. While the technical prerequisites in form of software, hardware and middleware have been available for more than 15 years, political power to implement such a solution remains questionable.

Reasons for divergent statistical systems are as follows:

- Divergent efforts in analysis of statistical data
- Lacking autonomy of statistical institutions
- Budget restrictions
- Varying intellectual potentials and traditions in the classic fields of economics and statistics
- Divergent measures of services and financial assets
- Heterogeneous handling of insurance premiums
- Income from shares and other equities
- Different statistical definitions and approaches
- Measuring economic activities varies across membership countries
- Size of black markets
- Degree of illiteracy
- Degree of mobility of people and assets
- Data capturing facilities.

A further discussing of the problems of harmonization of statistical systems is out of the scope of this paper, and would drive us far away from transmitting the main ideas about an electronic questionnaire. Instead, below we will elaborate the current standard of NA computing, illustrating that it is viewed mathematically, rather then statistically, while using simple arithmetic operations. Let us remark, that many economic and IT experts underestimate the role and influence of applied statistics in conjunction with modern IT technology for improving data quality while reducing cost of data production. As stated by B. Efron, Stanford University: *"Those who ignore Statistics are condemned to reinvent it".*

The SNA offers three different approaches for the computation of the main economic indicator Gross Domestic Product (GDP). While various approaches to compute indicators are quite common for economists and other "soft domain scientists", civil engineers using precise measurements will stumble over this fact, as their metering is usually dependent upon one and only one measurement technique. The three approaches for computing NA are:

- Production Approach
    - Aggregation of Production data
- Expenditure Approach
    - Aggregation of Demand data
- Income Approach
    - Aggregation of Income data.

Below, the GDP computation of each approach is listed introducing a number of intermediate variables. As all these examples equate to the same indicator, GDP, the reader can deduce that all these variables form an equation system of main economic indicators based on relations given by given by definitions. According to

our main topic and the space limitations of this contribution the main indicator system will remain "behind the scene"[2].

Utilizing the *Production Approach* the GDP can be compiled as follows:

+ Output value (at basic prices)
- Intermediate consumption
= Gross value added (GVA)
+ Taxes on production
- Subsidies on production
= Gross Domestic Product (GDP)

The GDP can also be compiled by the *Expenditure Approach*

+ Final consumption expenditures (households, government and NPISH)
+ Gross capital formation (corporations, households, government and NPISH)
+ Exports
- Imports
= Gross Domestic Product (GDP)

Finally, the GDP can be computed using the *Income Approach*

+ Compensation of employees (wages and salaries, employers social contributions)
+ Other taxes on production
- Other subsidies on production
+ Taxes on products
- Subsidies on producs
+ Operating surplus
+ Mixed income (unincorporated enterprises owned by households)
= Gross Domestic Product (GDP)

[2] The interested reader is referred to the Pubns (2009), Nations et al. (2008), Nations et al. (1993),
or Wikipedia

At this point the reader should remind himself that using these approaches for computing GDP will most likely lead to statistical discrepancies. Hence, any questionnaire must allow an entry of statistical discrepancies for the variables. It will be shown later how eNAQ handles those deviances.

4. System Analysis of the SNA and the UN NAS

There are various aspects concerning UN NAS activities, i.e. the present situation at the UN NAS, the variants of questionnaires, and drawbacks of the working process:

- Situation at the UN Statistics Division
    - Tasks
    - Workflow
    - Budget
- National Accounts Questionnaire
    - Questionnaire for the 1993 SNA
    - Future questionnaire for the 2008 SNA
- Weak points and bottlenecks in the workflow

Task at the UN NAS

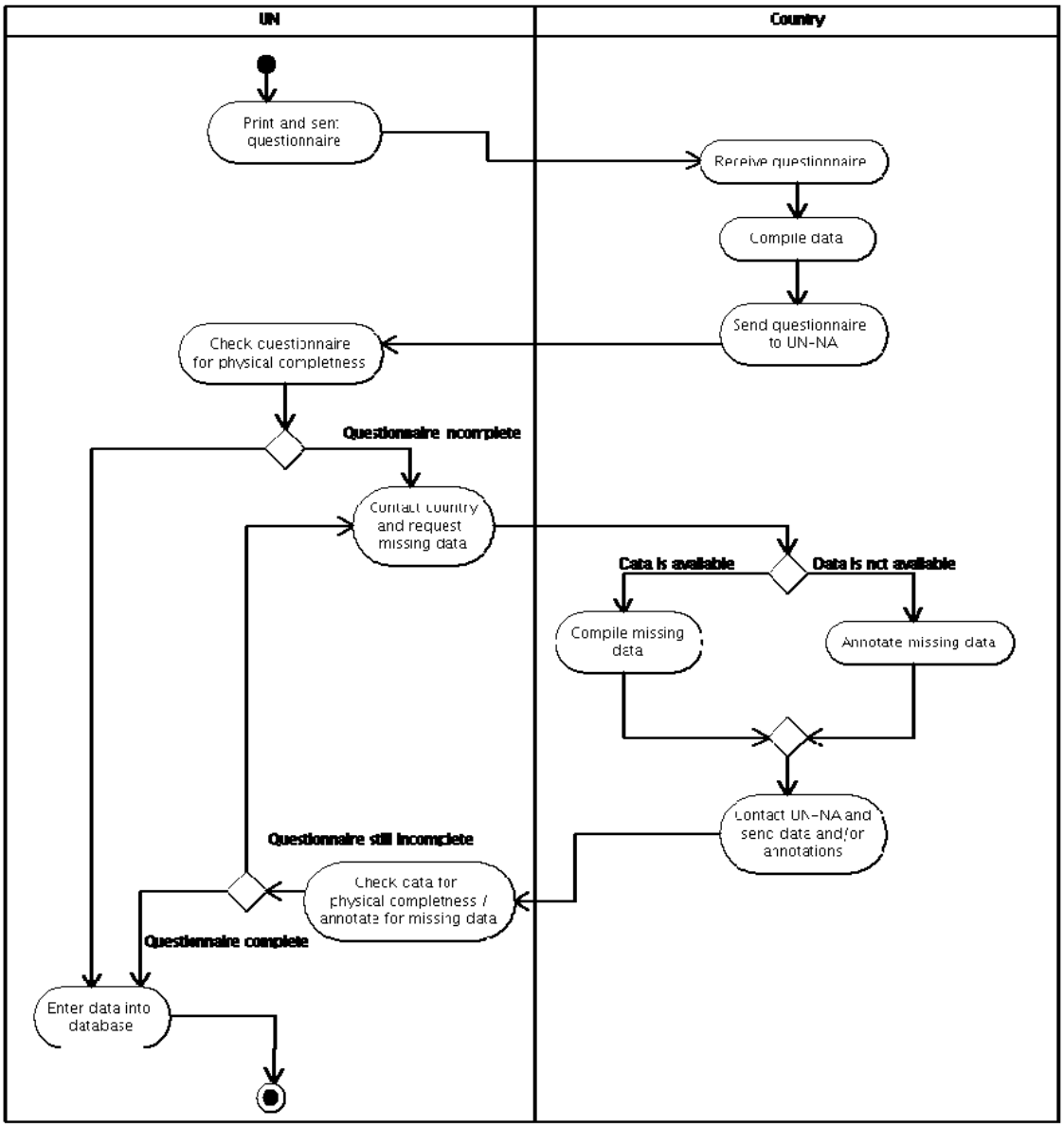
Tasks of the National Accounts Section include:

- Collection and compilation of data received from member countries of the UN
- Contribution to the statistical Yearbook published by the UN
- Publication of the NA Yearbook: National Accounts Statistics: Main Aggregates & Detailed Tables.
- Publication of National Accounts Statistics: Analysis of Main Aggregates.
- Provision of data to governments, the private sector, or Universities for scientific purposes etc.
- Provide support for the member countries in data compilation and adherence to the guidelines developed in the SNA

Workflow

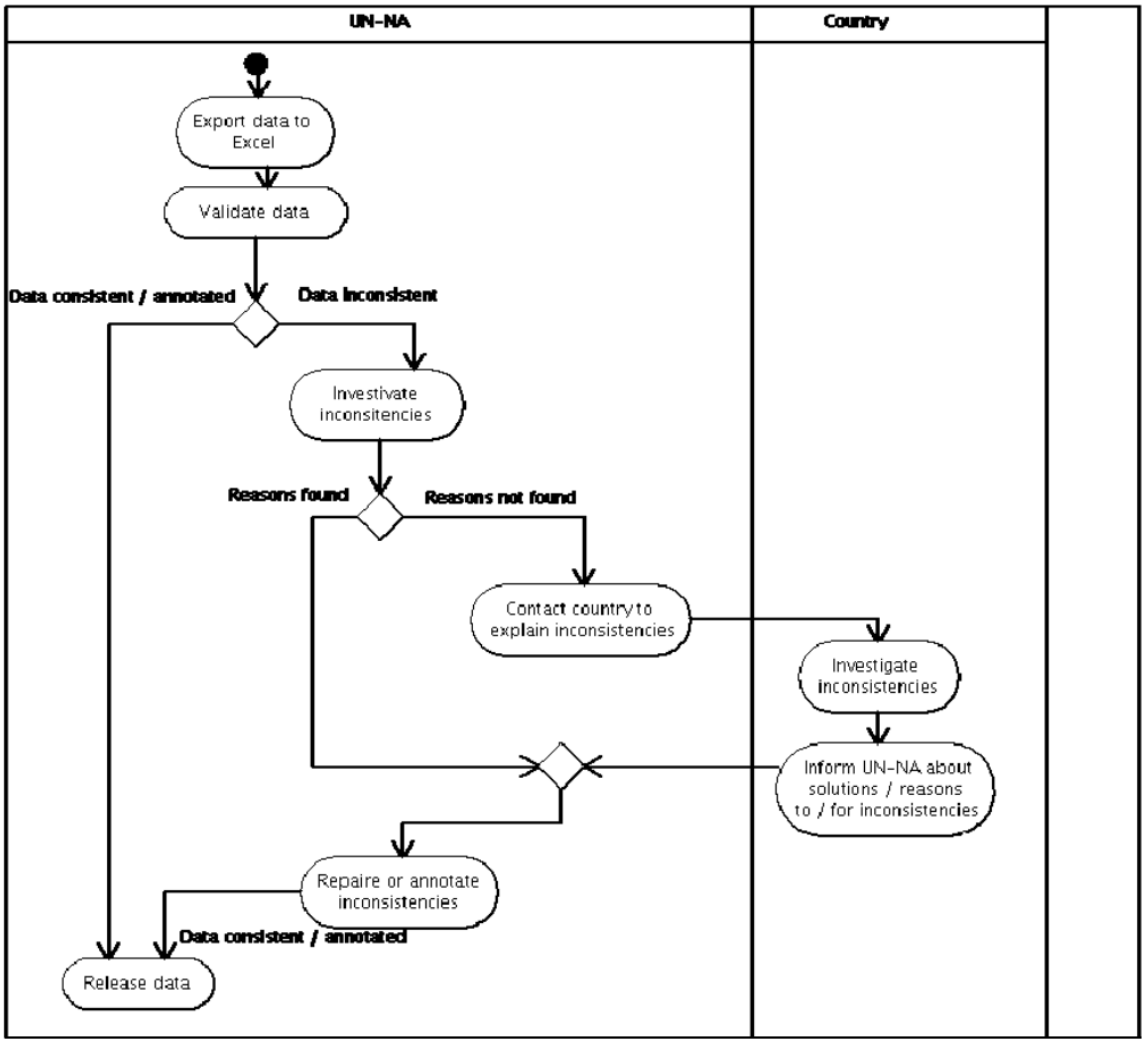
The illustrations 1 and 2 represent the data capturing process and its validation as currently done by the UN.

Illustration 1. Current Workflow "Enter data of a country's questionnaire"



The diagrams are self explaining, but let us stress the high number of data transfers between the UN NAS and breaks which require manual data input.

Illustration 2. Current Workflow "Validate and release country's data"



4.1. The National Accounts Questionnaire (NAQ)

A NAQ consists of the following elements:

Country Notes
The *country notes* provide information regarding the national institution that compiled the data, a general note regarding the kinds of tables provided, the method

of computation, special remarks concerning the countries' data and a specification of the method used for GDP computation, cf. Nations et al. (1993)).

Conceptual Table

The *conceptual table* is an addition to the NAQ and was used to collect information on the progress of SNA 1993 implementation.

National Accounts Data

The *national accounts data* is comprised of various tables within the categories briefly summarized below.

- Main Aggregates
  - In this part of the NAQ, three tables are requested that summarize information on gross domestic product by expenditure in current and constant prices (Tables 1.1 and 1.2), as well as information regarding the relations among the main aggregates: production, income, saving, and net lending aggregates (Table 1.3).
- Domestic Production by Industries
  - The tables 2.1 and 2.2 compile the GDP in current and constant prices summing the individual GDP value of industries as defined by the ISIC Rev.3.1[3]. Table 2.3 computes the contribution of these industries to various main aggregates, such as gross value added, gross capital formation, and others.
- Classification of Expenditure According to Purpose
  - Table 3.1 requests data on the government consumption expenditure, classified according to the 10 divisions of COFOG[4]. Table 3.2 requests similar information for individual household consumption expenditure, individual consumption expenditure for NPISH[5] and individual consumption expenditure of general government, classified according to COICOP[6].
- Integrated Economic Accounts
  - The integrated accounts assemble information for the institutional sector[7] accounts, the rest of the world accounts, and asset and liabilities accounts. They show principal economic relations as well as the main aggregates. In theory, they form an accounting framework in which the sum of resources and uses of the total economy should balance to zero with those of the rest of the world.

[3] International Standard Industrial Classification

[4] Classification of the Functions of Government

[5] Non Profit Institutions Serving Households

[6] Classification of Individual Consumption According to Purpose

[7] Sectors are the Total economy, Rest of the world, Non-financial corporations, Financial corporations, General government, Households, and Non-profit institutions serving households

- Cross Classification of Gross Value Added by Industries and Institutional Sectors
  - The table 5.1 requests the individual contributions of sectors to the gross value added by industries, as classified by the ISIC Rev.3.1.

4.2. Changes Introduced by the 2008 Revision of the SNA

To date the UN NAS has not yet published a revised version of the NAQ for the 2008 SNA. The revised NAQ will also reflect the new ISIC Rev. 4. The study of the new framework suggests that no new tables will be introduced. However, the tables themselves will be changed, introducing new items, removing and renaming items. This will affect the structure of the tables as well as the relations between their items[8], and will approximately double the numbers of 1046 items and 814 rules mentioned above.

For illustration purposes a data and a conceptual table is presented below. Note, that entries for twelve successive years are requested within the data table. This provides an opportunity for the countries to update already submitted data. However if data were revised for more than the years shown in the NAQ, this data should be provided to UNSD, too., in order to maintain the latest available data for all the countries in the UN National Accounts database.

Table 1. Gross domestic product by expenditures at current prices - NAQ: 1.1

Item Description	Item Code	S	1995	1996	...	2006
EXPENDITURES OF THE GROSS DOMESTIC PRODUCT						
P.3 Final consumption expenditure	10107					
P.3 Household final consumption expenditure	10108					
P.3 NPISHs final consumption expenditure	10109					
P.3 General government final consumption expenditure	10110					
P.31 Individual consumption expenditure	10111					
P.32 Collective consumption expenditure	10112					
P.5 Gross capital formation	10113					
P.51 Gross fixed capital formation	10114					
P.52 Changes in inventories	10115					
P.53 Acquisitions less disposals of valuables	10116					
P.6 Exports of goods and services	10117					
P.61 Exports of goods	10118					
P.62 Exports of services	10119					
P.7 Less: Imports of goods and services	10120					
P.71 Imports of goods	10121					
P.72 Imports of services	10122					
Plus: Statistical discrepancy	10123					

[8] The relations between the items are also referred to as *rules*. They will be elaborated below.

Table 2. Conceptual Table 1993 SNA

	Yes	No	Partly	Not Applicable
Is government defence expenditure on fixed assets that can be used for civilian Purposes included in GCF?				
Is consumption of fixed capital included on all government fixed assets (airfields, roads, hospitals, docks, dams and breakwaters and other forms of construction)?				
Is all successful mineral exploration capitalised?				
Is all unsuccessful mineral exploration capitalised?				
Are purchases of computer software included in GCF?				
Is expenditure on software development on own account and for sale included in GCF and output?				
Is expenditure on entertainment, literary or artistic originals included in GCF and on Their development included in output?				
Is expenditure on valuables included in GCF?				
Is the natural growth of cultivated forests included in output and GCF?				
Is financial intermediation services indirectly measured (FISIM) allocated to users?				
1993 SNA extends the production boundary of households to include goods that are not made from primary goods? Are these goods included In output?				
1993 SNA extends the production boundary of households to include all goods that are produced by households whether for sale or not ? Are These goods included in output?				
Are volumes estimated using a chaining procedure at least on an annual Basis?				
Are un-funded social contributions (for sickness, unemployment, retirement etc) by enterprises imputed as compensation of employees and included as contributions to social insurance?				
Do non-life insurance estimates include premium supplements rather than being based just on premiums less claims?				
Do life insurance estimates include premium supplements rather than being based just on premiums less claims?				
Are reinvested earnings estimates included In the rest of the world Account?				
Are foreign workers? remittances excluded from GNI?				
Is output and value added measured at basic prices?				
Is output and value added measured at producers? prices?				
Is value added measured at factor cost?				
Is government final consumption expenditure broken down into individual and collective consumption?				

4.3. Weak Points and Bottlenecks in the Workflow

Currently the UN NAS provides Excel tables, but submissions by coutries in other formats, electronic as well as paper, are accepted, too. Not all countries have real time and on-line Internet connections, creating obstacles for the communication with national offices, especially for the dispatch of the Excel NAQs, and follow ups. Tough nearly all Excel NAQs are received via email and in the desired format, some countries provide their data in their own excel spreadsheets or formats like pdf. This situation causes three negative main effects: *(1)* diminished data quality due to media breaks, *(2)* much time spent on double data entries,

and *(3)* high cost of data manipulations.
Some of the problems are that:

- Multiple changes of media cause high costs and ample opportunity for errors.
- The validation process is tedious and time consuming, requiring 814 rules to be validated. Typical problems of the validation include avoidable errors like switched numbers, balance equations contradicting data and missing data.
- Ex-post validation of country data at the UN NAS, New York, causes a high need for communication in order to correct invalid data and is politically risky, since it is not necessarily the country submitting the data finally published.

5. Requirement Engineering for the eNAQ System

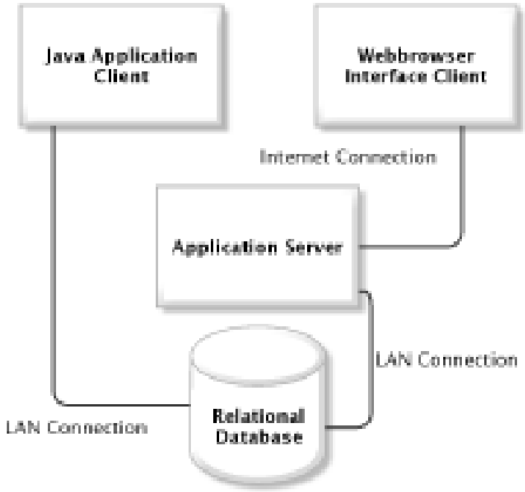
5.1. Objectives

The main goal for an eNAQ is to improve the workflow at the United Nations Statistics Division. This is to be achieved by a framework that supports direct data entry into the national electronic questionnaire. The system should assist the national offices by providing help functions, and deliver feedback about the validation of the entered data, thus strongly raising data quality. Of course, such a system must support multiple languages. Furthermore, it should support communication between the national offices and the United Nations Statistics Division (UNSD) about the data. When data is accepted by the UNSD, automatic data storage within the database is required. Stored and transmitted data has to be protected against unauthorized modifications, cf. Müller, Eichler and Lenz (2004)). For countries with no reliable Internet connection and for importing delivered data, import and export functions are required. Additionally, the generation of data for reports for scientific or public purposes and the production of a statistical yearbook are to be considered.
eNAQ's main function is to offer an electronic version of the *System of National Accounts Questionnaire* as a framework for the whole process of entering data sets, checking data sets for completeness, checking data with respect to the validation rule set, and generating reports. Note, that eNAQ is explicitly not meant to compute (new) variables in order to generate derived data.

eNAQ may be used by the national statistical offices of all UNO membership countries and UNSD. Their *user groups* include: UN staff members, country specialists, and system administrators. Data access for other users (scientific purposes etc.) requires an additional user group with different access rights, to be managed by the UNSD staff.

eNAQ has to support two different user modes. It should be usable as a local Java application client within a local area network connecting to a local database via JDBC[9] by a TCP/IP[10] connection. The second scenario

Illustration 3. Three tier architecture of eNAQ



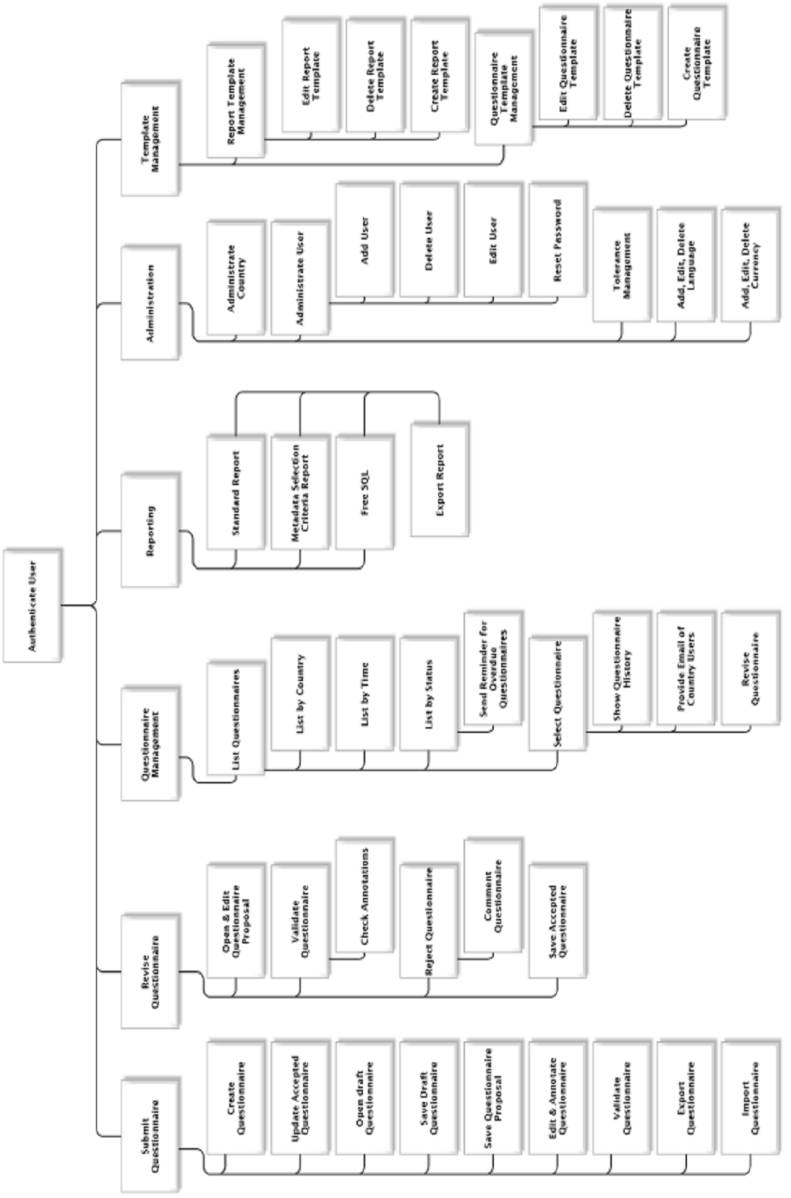
uses an Internet connection. The client connects to the server using SOAP[11], while on the server site eNAQ is run as a web service, locally accessing the database via JDBC.This results in a three-tier client-server architecture (displayed in illustration 3) where the presentation layer (top tier) provides the user interface. The interface can be implemented using either a browser interface or an application program written for example in Java, cf. Eichler (2004). The middle tier is an application server which provides security checks, identity verification and may implement part of the business logic[12]. The bottom tier is a database server running a relational Database Management System (DBMS). During the implementation specialregard should be paid to the software that will be necessary to use eNAQ. Countries using the client must not be forced to buy expensive software licenses from international vendors. In order to allow countries to set up their own database system and application servers, embedded commercial licenses must be minimized. The user interface should be similar to the paper version of NAQ, so experienced users feel familiar with the all tables and intuitively know where to enter data.

[9] Java Database Connectivity.
[10] Transmission Control Protocol - Internet Protocol.
[11] Simple Object Access Protocol
[12] Business logic may also be implemented by the client or the database.

5.2. Functions

The function tree (displayed in illustration 4) provides an overview of the required functions that can be deduced from the objectives specified above.

Illustration 4. Function tree



5.3. Validation Rules and Meta Data

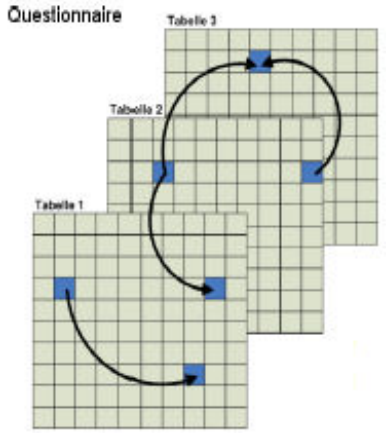
Validation rules are of a great importance for any system. How validation rules look like and how do they work? Remember that data tables contain data items. Items are a synonym for variables or economic indicators which are the basics of SNA. Hence, relations among items are defined as rules, according to economic accounting principles. Some rules define relations between items within only one table; other rules define relations between items of various tables. Table 3 illustrates a rule defining the relations within table 1.1 GDP *by expenditure in current prices*. It contains all the information a user should recognize when he validates (aggregated) data of a NA questionnaire. The table is to be read as an equation, where the elements on the right hand side must equate the element on the left hand side (the "caller"of the rule). SNA codes identify economic variables and the item codes uniquely identify any item within the NAQ. A description provides interpretation of a given variable. Special regard should be paid to the entry *Statistical Discrepancy*, which designates how much the sum of values of the right hand side may deviate from the left hand side.

Table 3. SNA validation rule of eNAQ

Rule Group				**101**	**Accounting Rule**			
Rule #	SNA Code	Item Code	Name	,	+/-	SNA Code	Item Code	Name
1010001	B.1*g	1012400	GDP by expenditure	=		P.3	1010700	Final consumption expenditure
					+	P.5	1011300	Gross capital formation
					+	P.6	1011700	Exports of goods and services
					-	P.7	1012000	Imports of goods and services
Applies to SNA				1993 ; 2008	+		1012300	Statistical discrepancy

Now remind yourself that the various tables are stored as spreadsheets with relations among their items defined in rules. One can imagine a complex rule set based on arithmetic relationships, cf. illustration 5. Keep in mind the complexity of the data structure with 31 tables, 1046 items and 814 rules, for the 1993 SNA, only.

Illustration 5. Multi relationship between items and tables

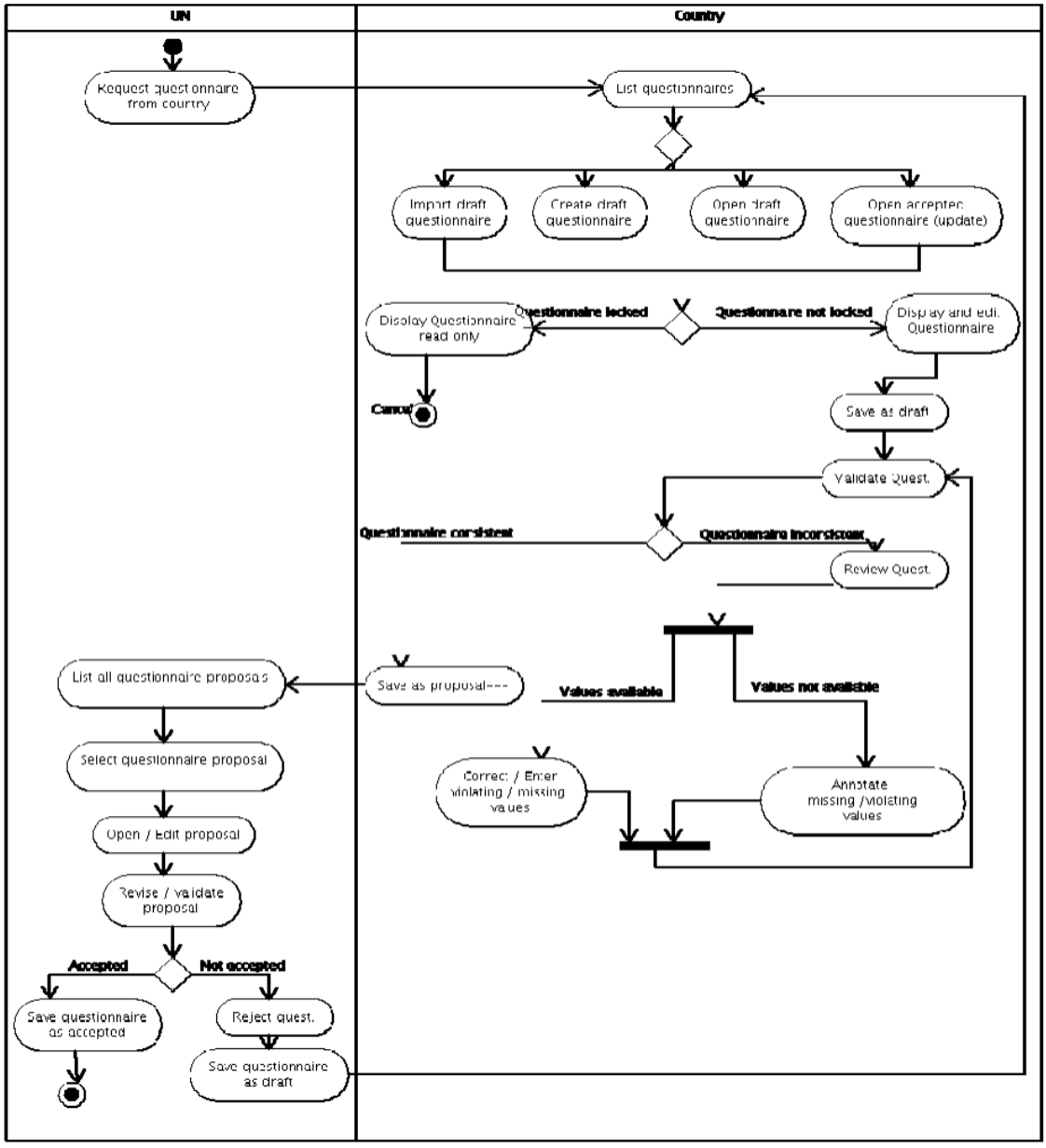


Note, that not only variables have to be provided by the system, but also meta data about the variables themselves. Such meta data deliver information for the user like "To which table does a certain item belongs?", "What is its semantic?", and "What is its relation to all ancestral variables?". Moreover, remember that this information has to be provided for multiple languages. Clearly, the complexity of data structure stresses the need for a sound database design. But before going a step ahead, let us take a short look at the streamlined workflow of eNAQ.

5.4. Workflow required

The diagram below represents the required workflow. It encapsulates the same functionality as the two diagrams displayed above. Note that the major improvement lies in shifting the validation process to the countries and forcing country users to validate data before submitting it to the UNSD. Comparing only the length of the graph, it becomes clear that the new workflow should be more efficient.

Illustration 6. Data compilation and validation of eNAQ



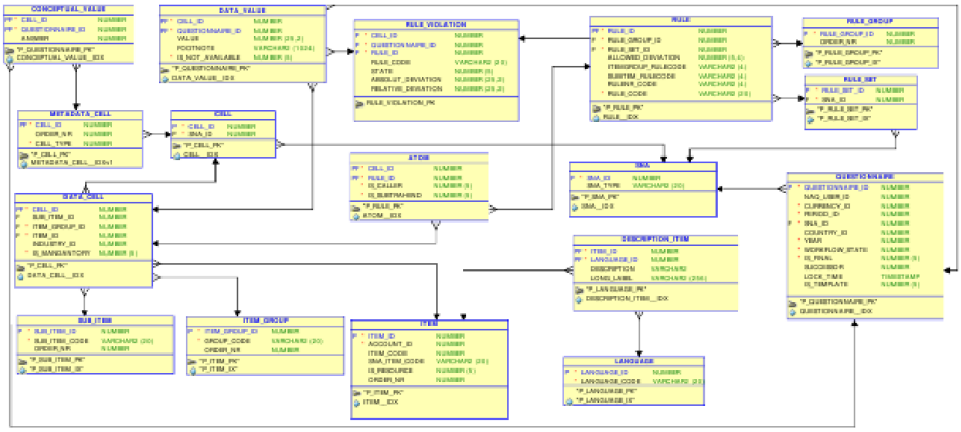
6. System Design

6.1. Conceptual Database Schemata

In this section some highlights of the conceptual database design are provided. It is based upon the relational model. Note, that this presentation is far from comprehensive, concentrating on the elements essential to the display of questionnaires, the storage of data entered within questionnaires and the validation process. Within the relation QUESTIONNAIRE information regarding a specific questionnaire is stored, such as the country, currency, year, and the SNA version. The relation is linked with the relation SNA, which provides connection to a specific rule set for any version of the SNA as well as a connection to the relation CELL.
Cells are used to store all elements of a questionnaire. Using the relations DATA_CELL and METADATA_CELL provides a necessary distinction between conceptual data and economic variables. The relation DATA_CELL is used to construct tables containing the economic variables. Tables are represented in the relation ITEM_GROUP, table items are stored in the relation ITEM, and the sub items of table 2.3 are stored in the relation SUBITEM. Note, that the relation ITEM has a link to the relation DESCRIPTION_ITEM. This relation contains meta data for every single item, such as a label or a description which can be used, for example, by the help system to provide context information regarding single variables. Using a composite primary key consisting of the ITEM_ID and the LANGUAGE_ID this provides the necessary basis for storing meta data in multiple languages[13].

[13] Note that this construct is used in other relations as well, but for reasons of clearness they are removed.

llustration 7. Questionnaire and Validation Rules (simplified) of eNAQ



A composite primary key is used in the relations DATA_VALUE and CONCEPTUAL VALUE, too. It is composed of the CELL_ID and the QUESTIONNAIRE_ID. In these relations specific values entered by users for table items, and respectively the questions in the conceptual table as well as the country notes, sources, special remarks and explanations regarding the chosen method of GDP computation, are assigned to a specific questionnaire.

Concerning the validation process, the relation ATOM is used to assign further attributes to each combination of a data cell and a rule applying to this cell. This relation specifies whether the value of a specific cell is either positive or negative in the context of a given rule, and also whether it is found on the left hand side, or the right hand side of the rule equation, as illustrated in table 3. The relation RULE in turn, belongs to a rule set, which is assigned to a specific version of the SNA, as can be seen in the relation RULE_SET.

Furthermore, on each update of the relation DATA_VALUE a combination of stored procedures and triggers uses this information to validate the new values against the stored rule information. Rules may be satisfied, violated but with a feasible statistical discrepancy, violated, with an infeasible statistical discrepancy, or simply refer to missing values. Depending upon the single case, information is stored within the relation RULE VIOLATION. Further triggers ensure that questionnaires can't be sent to the UN NAS for revision until at least all rules are satisfied inside the statistical discrepancy or violations have been annotated. This greatly facilitates client development as it implements essential parts
of the business logic inside the database, thus decoupling business and application logic, allowing for the development of light clients that only display views of the database.

The database stores much more information, regarding the various countries, currencies, the different users and user groups within the system as well as templates, but their presentation is outside the scope of this paper. The same is true for *external views* and the *internal schema*.

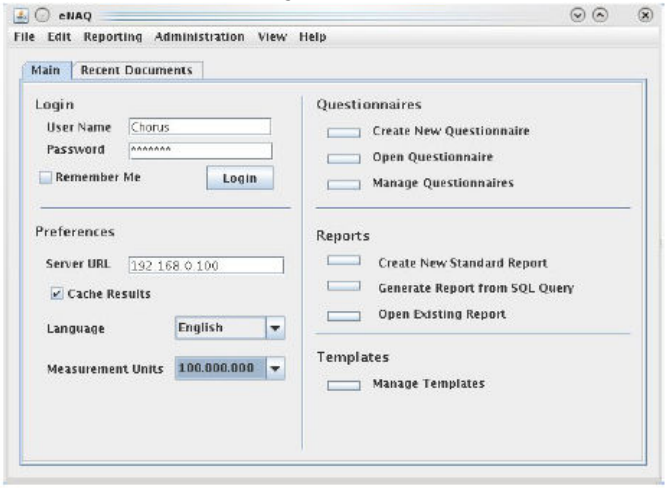
6.2. Functionalities embedded into the GUI

Finally, we have a look at the GUI. Its design includes calls of functions for editing and revising questionnaires, a help function, functions for questionnaire management, user, country and currency administration, functions for the generation of statistical reports as well as functions for using questionnaire and report templates. Due to the space limitations we will concentrate on presenting the main windows of the eNAQ client, as well as the functions for editing questionnaires and questionnaire management, as they are considered core functionalities. For better oversight of all available functions, the reader is referred to the function tree as displayed in illustration 4.

6.3. The Main Window

The main window, see illustration 8, provides an overview of the core functions to be invoked by eNAQ's GUI. It presents a login field, the required information for connecting to the server, as well as the selected language and the chosen measurement units. Furthermore, it provides an overview of core functions available, such as creating, opening and managing questionnaires as well as reports and templates.

Illustration 8. eNAQ Main Window, Start Screen

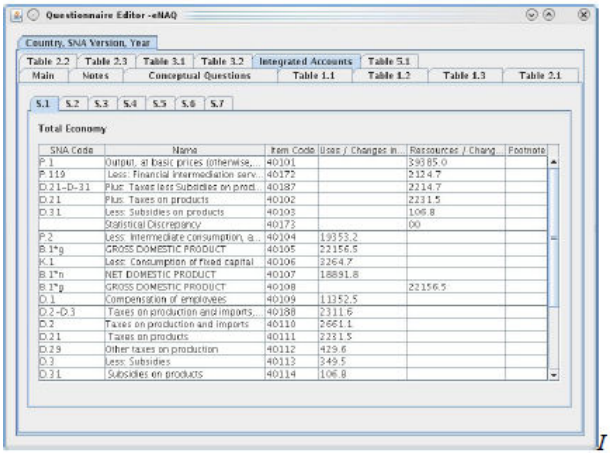


All these functions, as well as additional functions are also available through the menus below the title bar. The *Recent Documents* tab provides easy access to the last opened questionnaires and reports.

6.4. The Questionnaire Editor

Selecting the country and year, activates the questionnaire editor (illustration 9). In the top region of the window, the available tables are displayed, including the country notes, and the conceptual table.

llustration 9. Questionnaire Editor - Table 4.1



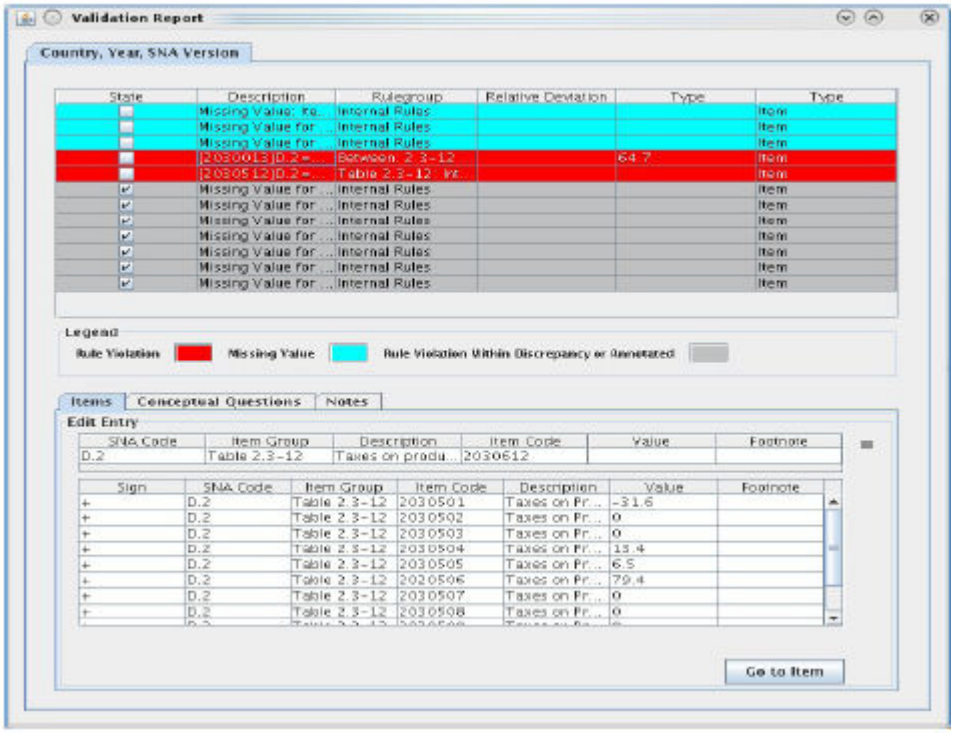
The main tab contains general information regarding the opened questionnaire, such as a questionnaire history, and the questionnaire status. Directly below more tabs are used to display a selection of sub tables for the selected category. For example, the integrated accounts tables contain seven tables, one for each sector. The table itself contains the SNA codes, a short description of the corresponding items, the item codes and a column for values to be entered by the user. An additional column is available for entering footnotes to annotate discrepancies or missing values.

6.5. The Validation Report

According to the requirement engineering any country expert is expected to validate his questionnaire before he can send it to the UN NAS. The validation function is invoked by a menu click. The validation form is displayed in illustration 10. The top table provides an overview of all violated rules. The coloring indicates different degrees of violations. Red is assigned to rule violations beyond the statistical discrepancy allowed. Cyan indicates fields with missing data. Light Grey is used for

rules that are violated, but within the statistical discrepancy or violations that have been annotated with footnotes. The checkboxes within the state column are automatically checked when a rule violation is resolved to the degree that a questionnaire can be saved as a proposal. Violations that are completely resolved are removed from the table. The description provides a short explanation regarding the cause of the violation. If the cause is a missing value, the item code is provided so that the user can identify the missing item within the questionnaire. If rules are breached the rule formula is provided in SNA Codes. The table also provides information about the relative and the absolute deviation, as well as the type of the violating entry. Types are items, conceptual questions and notes. The table can be sorted by any column.

llustration 10. Validation Report, based on Eichler (2004)



The validation report provides another table within the panel *Data Entry*. It contains tabs for the different types of violations (violations regarding relations between the various items, missing entries in the conceptual table or missing country notes). The tabs are automatically selected and filled with the corresponding data according to the type of the row selected in the top table. For items, on the top of the item tab the

item causing the violation[14] and its properties are displayed, including a field for its value and a footnote, in case a discrepancy cannot be resolved. If the violation concerns an unanswered question within the conceptual table, the value field is displayed as a combo box, containing the allowed answers. The equality operator indicates the rule to the user. The bottom table displays the SNA codes, item codes, the item group, a description containing the item names and their values. Thus, all elements designated by the violated rule are displayed with their properties, allowing for their identification within the NAQ and their interpretation regarding economic significance within the SNA. Users can directly edit or annotate these values, either by correcting or annotating the item calling the rule, or by correcting or annotating the rule elements. For missing values the bottom table remains empty and only the missing item and its properties are displayed.

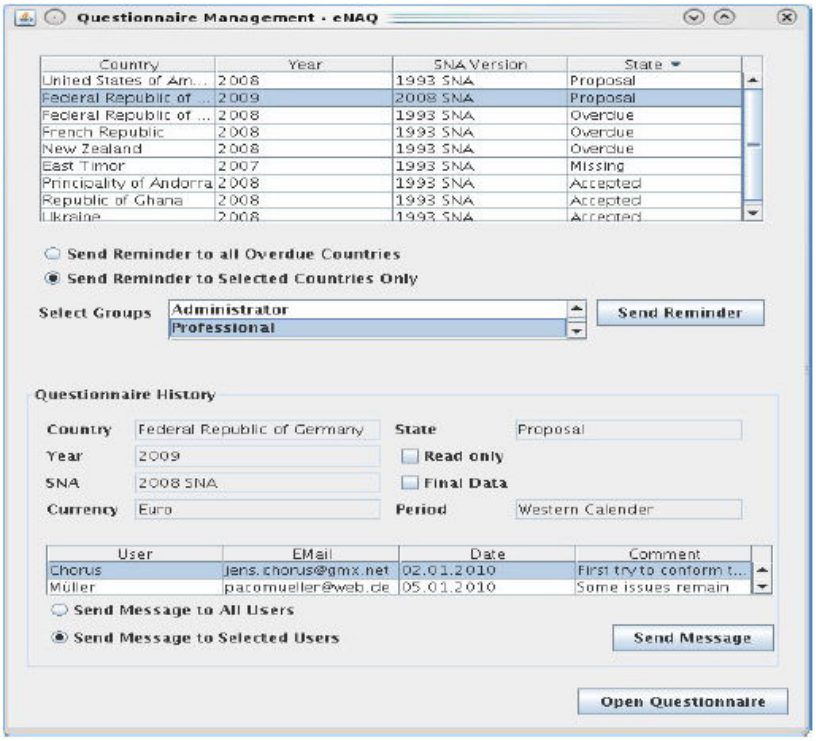
Note, that entries within the questionnaire editor are colored according to the violations detected within the validation report. This provides additional oversight, and allows users to use the questionnaire editor to correct values should they choose to do so. Entries for which violations have been detected, now also allow for editing footnotes.

6.6. Questionnaire Management

This function provides an overview of all questionnaires within the database, allows for contacting the

[14] In terms of the database design this item is the caller of the rule.

Illustration 11. Questionnaire Management, based on Eichler (2004)



countries of selected questionnaires, displays the questionnaire history of a selection, and provides an alternative path to access a questionnaire for further editing or revision. The top table of this form (illustration 11) can be used to sort all questionnaires within the database by country, year, SNA version and status. The value set (domain) for the status column is *Accepted, Proposal, Overdue and Missing*. The missing property was introduced to allow UN employees to flag countries for which no questionnaire is expected to be submitted given a fixed year[15]. The radio buttons *Send Reminder to all Overdue Countries* and *Send Reminder to Selected Countries only* allow users to choose the corresponding action for the *Send Reminder* button. A message is sent to all users selected in the corresponding list field. Selecting multiple user groups is allowed. The table also allows for the selection of multiple rows.

[15] For example a severe breakdown of government structures due to war, natural catastrophes or other events typically cause data insufficiency or lack of participation.

The *Questionnaire History* provides detailed information for any single questionnaire selected in the top table. If none or multiple tables are selected this should be deactivated. Users can contact other users who edited the questionnaire, using the radio buttons and the corresponding *Send Message* button. They can also directly edit the selected questionnaire by using the *Open Questionnaire* button. This action will activate the questionnaire editor.

7. Conclusions

Let us summarize the major improvements that can be achieved by implementing eNAQ at the UN NAS and the various NSIs around the world. First of all, the workflows are simplified and streamlined. Shifting the validation process to the countries themselves which submit the data, an increase in data quality will arise in most cases while simultaneously decreasing data transfers and communications between the UN NAS and the various NSIs around the world.

Once a questionnaire has been filled in, no double manual data entry is required at any stage of the workflow. The support facilities for annotations of missing values or data which contradict rules together with the integrated communication support, the intuitive GUI and the context sensitive help function guarantee a high level of user friendliness for all parties.

All components needed for the implementation of the presented design can be implemented using all major operating systems and do not require the purchase of expensive licenses. Even a cost reduction is achieved in comparison to the software currently required, as the tables provided by the UN NAS require Microsoft Excel for usage. However, this is negligible relative to cost reduction for data production and management gained by an improved workflow and no further need of a manual data validation process.

The system design allows an easy integration of new rules or rule sets, as will be required when 2008 SNA will be rolled out. Furthermore, statistical discrepancies can not only be set for each desired indicator, but even for indicators in the context of a specific rule. Once implemented in the UN NAS database, a rule change could be propagated to all countries with only minor efforts.

eNAQ can easily be adapted to fit other requirements such as scientific, government, and press appliances. As all data is stored in a relational database, derived data can easily be generated using simple SQL aggregation functions like *min, max, sum* and *count,* and utilizing meta information from the database.

The solid start-of-the-art design and the chosen technologies provide a good basis for future developments. Relational database management systems are available from many vendors and the GUI can be implemented, choosing from a

wide range of options. This ensures that no vendor lock in can occur, and thus guarantees future independency.

eNAQ promises to open the door for a cost reduced computing of NA questionnaires at an improved data quality level. The authors hope that this approach will close the gap between an advanced NA methodology due to the 2008 SNA and modern software engineering based either on web services and browsers, or application programs backed up by a user friendly GUI.

References

Chorus, J. 2010: Reengineering of eNAQ - electronic National Accounts Questionnaire - 1993 & 2008 SNA - Requirements and Design Specification. Msc Thesis, Freie Universität Berlin, Berlin.

Eichler, J. 2004: Entwurf und Implementierung eines Systems zur Validierung von Daten des 'System of National Accounts' der Vereinten Nationen. MSc thesis, Freie Universität Berlin, Berlin.

Lenz, H. J. - Eichler, J. - Müller, P. 2008: A Prototype for an electronic version of an UN National Accounts Questionnaire - or A Pledge for using Advanced IT and Applied Statistics. Revista Romana de Statistica, 11, 7–16. p.

Müller, P. 2003: Analysis and Requirement Engineering of the UN National Accounts. Diplomarbeit, Freie Universität Berlin, Berlin.

Müller, P. - Eichler, J. - Lenz, H. J. 2004: eNAQ: A Prototype for an electronic Version of the UN National Accounts Questionnaire, Technical Report, Freie Universität Berlin, Berlin.

Pubns, United Nations 2009: System of national accounts 1993.

URL: http://unstats.un.org/unsd/sna1993/toctop.asp

United Nations et al. 1993: Handbook of National Accounting - Links between Business and National Accounting. United Nations.

United Nations et al. (2008): System of National Accounts 2008 -Pre-edited version of Volume 1. United Nations Pubns.

Decomposition of the Gray-Williams "tau" in main and interaction effects by ANOVA in three-way contingency table

Luigi D'Ambra[1] - Antonello D'Ambra[2] - Pasquale Sarnacchiaro[3]

The identification of meaningful relationships between two or more categorical variables is an important, and ongoing, element to the analysis of contingency tables. It involves detecting categories that are similar and/or different to other categories. Correspondence analysis can be used to detect such relationships by providing a graphical interpretation of the association between the variables, and it is especially useful when it is known that this association is of a symmetric nature. (Greenacre 1984), (Lebart et al. 1984).

In this paper, we will explore the Gray-Williams index when used as the measure of association in non-symmetrical correspondence analysis (NSCA). It will be shown that, by concatenating a predictor variable of a three-way contingency table, the two measures are equivalent. The paper will analyse the sum of squares for nominal data partitioning the Sum of squares for main effects and the interaction in the sense of analysis of variance giving an orthogonal decomposition of Gray Williams index .

Keywords: Three-way contingency table, The Gray-Williams measure of association, Catanova, Main effects Interaction

1. Introduction

The identification of meaningful relationships between two or more categorical variables is an important, and ongoing, element to the analysis of contingency tables. It involves detecting categories that are similar and/or different to other categories. Correspondence analysis can be used to detect such relationships by providing a graphical interpretation of the association between the variables, and it is especially useful when it is known that this association is of a symmetric nature. (Greenacre 1984), (Lebart et al. 1984).

There are many real-life applications where it is not appropriate to perform classical correspondence analysis because of the obvious asymmetry of the association between the variables. In these cases non-symmetrical correspondence

[1] Luigi D'Ambra , Dept. Biological Sciences, Università di Napoli Federico II (Italy)
[2] Antonello D'Ambra, Department of Management and Accounting (Italy)
[3] Pasquale Sarnacchiaro, University of Naples Federico II (Italy)

analysis can be considered. (D'Ambra-Lauro 1989, 1992), (Gimaret et al. 1998) and (Kroonenberg–Lombardo 1999).

The key difference between the symmetrical and non-symmetrical versions of correspondence analysis rests in the measure of association used to quantify the relationship between the variables. For a two-way, or multi-way, contingency table, the Pearson chi-squared statistic is commonly used when it can be assumed that the categorical variables are symmetrically related. However, for a two-way table, it may be that one variable can be treated as a predictor variable and the second variable can be considered a response variable. For such a variable structure, the Pearson chi-squared statistic is not an appropriate measure of association. Instead one may consider the Goodman-Kruskal tau index. Where there are more than two cross-classified variables, multivariate versions of the Goodman-Kruskal tau index can be considered. These include Marcotorchino's index (Marcotorchino 1985) and Gray-Williams' indices (Gray-Williams 1975), (Anderson-Landis 1980)

In this paper, we will explore the Gray-Williams index when used as the measure of association in non-symmetrical correspondence analysis (NSCA). It will be shown that, by concatenating a predictor variable of a three-way contingency table, the two measures are equivalent. The paper will analyse the sum of squares for nominal data partitioning the Sum of squares for main effects and the interaction in the sense of analysis of variance giving an orthogonal decomposition of Gray Williams index .

This paper is divided into six further sections. In Section 2 we consider the measure of association for two asymmetric cross-classified categorical variables. In Section 3 we provide a description of NSCA where the Goodman-Kruskal tau index is used as a measure of asymmetric association. This section also offers two tools that can be used to delve deeper into the source of association using this index. One is the C-statistic based on the work of Light-Margolin (1971), and the other is confidence circles. This latter tool was discussed in some detail for symmetrical, or classical, correspondence analysis of nominal variables by Lebart et al. (1984).
The Gray-Williams measure of complete association and its link to the Goodman Kruskal tau index when concatenating a predictor variable is discussed in section 4. In section 5 we analyse the interaction between the predictor variables and we present an orthogonal decomposition of Gray-Williams "Multiple" τ in which we have the part of main effects and the part of interaction.
A case study we present in the section 6, some final consideration ended the paper.

2. Measuring Non-SymmetricAssociation

Suppose we consider the cross classification of n individuals/units according to two categorical variables, X_1 and Y, that form a two-way contingency table, N. Let X_1 be the column variable that consists of c categories, and Y be the row variable

consisting of r categories. Denote the *(i, j)*th cell entry by n_{ij} for i = 1, 2, . . . , r and j = 1, 2, . . . , c, and the *(i, j)*th joint proportion by $p_{ij} = n_{ij} / n$ so that $\sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij} = 1$. Define the *i*th row marginal proportion by $p_{i\bullet} = \sum_{j=1}^{c} p_{ij}$ and define the *j*th column marginal proportion by $p_{\bullet j} = \sum_{i=1}^{r} p_{ij}$.

The chi-squared statistics commonly used as a means of formally measuring the departure from independence between X_1 and Y. By considering this statistic, it is assumed that there is a symmetric relationship between the two variables. However, there are many situations where the association between two categorical variables is not symmetric.

Suppose there exist an asymmetric association between two categorical variables such that X_1 is treated as a predictor variable and Y is the response variable. Therefore, a more appropriate measure of their association is to adjust the chi-squared statistic and consider instead

$$n\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(p_{ij} - p_{i\bullet}p_{\bullet j}\right)^2}{p_{\bullet j}} \qquad (1)$$

This measure was proposed by Goodman-Kruskal (1954) as a means of measuring the proportional reduction in error (PRE) in the prediction of the response variable given a predictor variable. Mirkin (2001,) also discussed these measures for nominal variables, as did Light-Margolin (1971) in the context of ANOVA for contingency tables. Therefore, suppose we let $\pi_{ij} = p_{ij} / p_{\bullet j} - p_{i\bullet}$ be the difference between the unconditional prediction of the *i*th row category, $p_{i\bullet}$, and the conditional prediction of that, given the *j*th column category, $p_{ij} / p_{\bullet j}$. Given the *j*th column category, if it does not contribute to the predictability of the *i*th row category, then $\pi_{ij} = 0$. Formal procedures can be adopted to measure the predictability of the row response categories given the information in the column categories by considering the Goodman - Kruskal (1954) tau index

$$\tau = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{\bullet j}\left(\frac{p_{ij}}{p_{\bullet j}} - p_{i\bullet}\right)^2 \Bigg/ \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2\right) = \tau_{num} \Bigg/ \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2\right) \qquad (2)$$

Light-Margolin (1971) also considered such a measure and described it as "the proportion of total variation in the response variable which is accounted for by knowledge of the grouping [predictor] variable".

The numerator of (2), τ_{num}, can be alternatively expressed as (1) divided by the sample size n, and is bounded by the interval [0, 1]. When the distribution of each of the response (row) categories across each of the columns is identical to the overall marginal proportion, such that $p_{ij} / p_{\bullet j} = p_{i\bullet}$, there is no relative increase in predicability of the row variable and thus τ is zero. Note that zero predictability also implies no association (ie independence) between the two categorical variables. When $\tau = 1$, there is perfect predictability of the response categories (rows) given the predictor categories (columns).

3. Non-Symmetrical Correspondence Analysis: testing and confidence circles

The measure of the departure from independence of the *(i, j)*th cell of the two-way contingency table, N, when there is an asymmetric association between two categorical variables, can be quantified by the π_{ij} that is defined in Section 2. To obtain characteristics and low-dimensional summaries of the structure of this association, NSCA involves applying a singular value decomposition (SVD) to π_{ij} so that

$$\pi_{ij} = \frac{p_{ij}}{p_{\bullet j}} - p_{i\bullet} = \sum_{m=1}^{M} a_{im}\lambda_m b_{jm} \qquad (3)$$

where $M = \min(r, c) - 1$ and λ_m is the *m*th singular value of π_{ij} for $m = 1, \ldots, M$. The quantities a_{im} and b_{jm} are, respectively, the elements of the singular vectors $\mathbf{a}_m$ and $\mathbf{b}_m$ associated with the *i*th row and *j*th column categories and have the property

$$\sum_{i=1}^{r} a_{im}a_{im'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases} \qquad \sum_{j=1}^{c} p_{\bullet j} b_{jm} b_{jm'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases}.$$

By considering the decomposition (3), the numerator of the Goodman-Kruskal tau index can be decomposed so that

$$\tau_{num} = \sum_{m=1}^{M} \lambda_m^2 .$$

When performing NSCA (Beh-D'Ambra 2010), we can graphically depict the association between the row and column categories by plotting along the *m*th dimension of the non-symmetrical correspondence plot the row and column profile coordinates

$$f_{im} = a_{im}\lambda_m \qquad \text{and} \qquad g_{jm} = b_{jm}\lambda_m .$$

If one considers these coordinates, then it must be kept in mind that they are not guaranteed to be centred about the origin of the correspondence plot (a useful property underlying the coordinates from classical, symmetrical, correspondence analysis). However, with respect to the unit metric and $p_{\bullet j}$ metric, the row and column coordinates are closely related to the numerator of the tau index through

$$\tau_{num} = \sum_{i=1}^{r}\sum_{m=1}^{M} f_{im}^2 = \sum_{j=1}^{c}\sum_{m=1}^{M} p_{\bullet j} g_{jm}^2 .$$

Therefore, points lying at a distance from the origin of the plot indicate that these categories contribute more to τ_{num} than those points that lie near the origin. Also, if column points lie close to the origin, these categories do not contribute to the predictability of the response variable. If predictor (row) points lie close to the origin, these categories are not affected by any variation in the predictor variable.

The Goodman-Kruskal tau index is a good measure for determining the predictability of the rows given the columns. However, as Agresti (1990) indicated, a low value of τ does not mean that there is a "low" association between the two variables. While τ is an appropriate measure of the predictability, the statistic cannot, in its current form, be used to formally test for association. Instead such tests are carried out using the C-statistic of Light-Margolin (1971)

$$C = (n-1)(r-1)\tau = (n-1)(r-1)\sum_{m=1}^{M}\lambda_m^2 \Big/ \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2\right) \qquad (4)$$

Under the hypothesis of zero predictability ($H_0 : \Pi_{ij} = 0$), Light - Margolin (1971) showed that the C-statistic is asymptotically chi-squared distributed with $(r-1)(c-1)$ degrees of freedom. These authors introduced this statistic when deriving an analysis of variance procedure for contingency tables, commonly referred to as CATANOVA (Categorical Analysis of Variance).

When the variables of a two-way contingency table are considered to be symmetrically related, as is in the case for classical correspondence analysis, Lebart et al. (1984) presented the idea of confidence circles to identify those categories that contribute to the hypothesis of independence and those that do not. These circles are similar to the regions that Mardia et al. (1982, p. 346) derived for canonical analysis. Ringrose (1992, 1996) also explored the use of these types of circles for correspondence analysis, although a bootstrap procedure was employed for their construction. When categorical variables are ordinal in nature showed that the radii of these circles are identical to those of Lebart et al. (1984). However the confidence circles derived for use in symmetrical correspondence analysis are not applicable for NSCA. Here we present the radii lengths of confidence circles for NSCA.

Suppose that a two-way contingency table consists of row and column variables asymmetrically structured in the manner described in Section 2. The C–statistic of (4) can be expressed in terms of the predictor (row) coordinates such that

$$C = (n-1)(r-1)\sum_{j=1}^{c}\sum_{m=1}^{M} p_{\bullet j} g_{jm}^2 \Bigg/ \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2\right) \sim \chi^2 .$$

For the *j*th column (predictor) coordinate,

$$(n-1)(r-1)\sum_{m=1}^{M} p_{\bullet j} g_{jm}^2 \Bigg/ \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2\right) \sim \chi^2_{(r-1)} .$$

Therefore

$$\sum_{m=1}^{M} p_{\bullet j} g_{jm}^2 \sim \chi^2_{(r-1)} \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2\right) \Big/ \left((n-1)(r-1)\right).$$

Since, for higher dimensions, the coordinates will be close to zero (as the singular values associated with these dimensions are generally relatively close to zero), the relationship between the *j*th column coordinates for the first two dimensions of a two-dimensional non-symmetrical correspondence plot is

$$p_{\bullet j} g_{j1}^2 + p_{\bullet j} g_{j2}^2 = \chi_{(2)}^2 \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2 \right) \Big/ (n-1)(r-1).$$

At the 5% level of significance, this can be expressed as

$$g_{j1}^2 + g_{j2}^2 = \left[\sqrt{5.99 \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2 \right) \Big/ p_{\bullet j} (n-1)(r-1)} \right]^2 .$$

Therefore, the 95% confidence circle for the *j*th column coordinate in the two-dimensional non-symmetrical correspondence plot has a radius of length

$$o_j = \sqrt{5.99 \left(1 - \sum_{i=1}^{r} p_{i\bullet}^2 \right) \Big/ \left(p_{\bullet j} (n-1)(r-1) \right)} . \qquad (5)$$

Note that (5) depends on the *j*th marginal proportion classified into that category. Thus, if there is a very small number of classifications made in the *j*th predictor category, its radius length associated with this category will be relatively large. Similarly, for a relatively large classification, the radius length will be relatively small. Since we are interested in the predictability of the row categories given the column categories, confidence circles will only be constructed for the predictor variable.

Careful attention must be given to the interpretation of these regions. They do not suggest that a point has any significant link with an axis, since the axes have no direct interpretation (other than to graphically depict the proportion of the association between the variables it reflects). Overlapping regions may provide some indication as to the level of association between intra-variable categories but they do not provide formal evidence that such an association exists, although employing the *uncertainty circles* of Gabriel (1995) can provide such insight. The real strength of the confidence circles described here lies in their ability to reflect the significance of a particular predictor category in accounting for the level of predictability on a response variable. If the origin is enclosed within the confidence circle of a predictor category, then that category does not contribute to the predictability of the response variable. Similarly, if the origin falls outside of a confidence circle, then that particular predictor category does contribute to the predictability of the response variable. Such conclusions can be made keeping in mind the level of significance used to construct these circular regions.

4. Multiple NSCA – The Gray-Williams index

Suppose we consider the cross classification of n individuals/units according to three categorical variables X_1, X_2 and Y, that form a three-way contingency table, N. Let X_1 be the second (column) variable that consists of c categories, X_2 be the third (tube) variable that consists of t categories, and Y be the first (row) variable consisting of r categories. The terminology "tube" is used to be consistent with much of the discussion that has been made on multiple categorical data analysis; for example, Kroonenberg (1989) uses the expression. One may also consider X_2 to be a stratifying variable. The resulting contingency table is therefore of size $r \times c \times t$. Here we consider the relationship between the three variables to be asymmetric, in that Y is the response variable and depends on the two predictor variables X_1 and X_2.

To measure the asymmetric association of the three variables, one may consider multivariate extensions of the Goodman-Kruskal tau index.(Anderson-Landis 1980) Two examples include Marcotorchino's index and Lombardo's index. Another measure, with which we will concern ourselves here, is the Gray-Williams index (Gray-Williams 1975).

Let $\pi_{ijk} = p_{ijk} / p_{\bullet jk} - p_{i\bullet\bullet}$, for i = 1, 2, . . . r, j = 1, 2, . . . c and k = 1, 2, . . . , t, be the difference between the unconditional marginal proportion of the *i*th response category, $p_{i\bullet\bullet}$, and the (conditional) prediction of the *i*th response given the joint proportion of the two predictor variables, $p_{ijk} / p_{\bullet jk}$. Gray-Williams (1975) proposed an extension of the Goodman-Kruskal tau index for three categorical variables where the proportional reduction in error for the prediction of the response (row) variable can be measured by considering

$$\tau_{GW} = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{t} p_{\bullet jk}\left(\frac{p_{ijk}}{p_{\bullet jk}} - p_{i\bullet\bullet}\right)^2 \Bigg/ \left(1 - \sum_{i=1}^{r} p_{i\bullet\bullet}^2\right) = \tau_{GWnum} \Bigg/ \left(1 - \sum_{i=1}^{r} p_{i\bullet\bullet}^2\right) \qquad (6)$$

Just as was done for the Goodman-Kruskal tau index of (2), the numerator of the Gray-Williams index, τ_{GWnum}, will be the focus of our discussion here, since the denominator is independent of the any of the joint cell proportions of the table N.

To determine the structure of the dependence between three categorical variables (one criterion variable and two predictor variables), one may consider a three-way extension of the SVD of (3):

$$\pi_{ijk} = \sum_{m=1}^{M} a_{im}\lambda_m b_{jm} c_{km}$$

where

$$\sum_{i=1}^{r} a_{im}a_{im'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases}, \qquad \sum_{j=1}^{c} p_{\bullet j \bullet} b_{jm} b_{jm'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases}$$

and

$$\sum_{k=1}^{t} p_{\bullet\bullet k} c_{km} c_{km'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases}.$$

This approach is analogous to the PARAFAC/CANDECOMP models independently considered by Harshman (1970) and Carroll-Chang (1970) (Faber et al 2003). An alternative approach was considered by Lombardo-Carlier-D'Ambra (1996). For their approach, π_{ijk} (constructed to reflect the variation in predictability as measured by the Marcotorchino index) is decomposed using the Tucker3 decomposition (Tucker 1966).

Another method of decomposition that can be considered is

$$\pi_{ijk} = \frac{p_{ijk}}{p_{\bullet jk}} - p_{i\bullet\bullet} = \sum_{m=1}^{M} a_{im}\lambda_m b_{jkm} \tag{7}$$

where $\mathrm{M} = \min(\mathrm{r}, \mathrm{c}+\mathrm{t}) - 1$ and a_{im} and $\mathrm{b}_{\mathrm{jkm}}$ are subject to the constraints

$$\sum_{i=1}^{r} a_{im}a_{im'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases} \quad \text{and} \quad \sum_{j=1}^{c}\sum_{k=1}^{t} p_{\bullet jk} b_{jkm} b_{jkm'} = \begin{cases} 1, & m = m' \\ 0, & m \neq m' \end{cases}$$

respectively. This approach is called Multiple Non Symmetrical Correspondence Analysis (MNSCA).

The generalised singular values, λ_m, are again arranged in descending order such that $1 > \lambda_1 > \lambda_2 > \cdots > \lambda_M > 0$. The value a_{im} is an element of the singular vector $\mathbf{a}_{\mathrm{m}}$ and is associated with the *i*th row response category. Similarly the value b_{jkm} is an element of the joint singular vector $\mathbf{b}_{\mathrm{m}}$ of length *ct* and is associated with the joint association between the two predictor variables. The calculation of these quantities can be easily performed, not through any modification of the SVD

procedure of (3), but by simply concatenating one of the predictor variables to form a two-way table.

To demonstrate this point, suppose we transform the $r \times c \times t$ contingency table N in such a way that the tube predictor variable is concatenated so that N is of size $\mathrm{r} \times \mathrm{ct}$.

For the *k*th *(k = 1, 2, . . . , t)* $r \times c$ submatrix, the Goodman-Kruskal tau numerator, $\tau_{num|k}$, is

$$\tau_{num|k} = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{\bullet jk}\left(\frac{p_{ijk}}{p_{\bullet jk}} - p_{i\bullet\bullet}\right)^2 .$$

Aggregating each of these t measures of asymmetry yields

$$\begin{aligned}\sum_{k=1}^{K}\tau_{num|k} &= \sum_{i=1}^{r}\sum_{j=1}^{c} p_{\bullet j1}\left(\frac{p_{ij1}}{p_{\bullet j1}} - p_{i\bullet\bullet}\right)^2 + \cdots + \sum_{i=1}^{r}\sum_{j=1}^{c} p_{\bullet jt}\left(\frac{p_{ijt}}{p_{\bullet jt}} - p_{i\bullet\bullet}\right)^2 \\ &= \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{t} p_{\bullet jk}\left(\frac{p_{ijk}}{p_{\bullet jk}} - p_{i\bullet\bullet}\right)^2\end{aligned} \tag{8}$$

Which is the numerator of the Gray-Williams index, τ_{GWnum}, defined by (6).

For a concatenated three-way contingency table with response (row) marginal proportions $\{\mathrm{p_{1\bullet}}, \ldots, \mathrm{p_{r\bullet}}\}$ and predictor (column) marginal proportions $\{\mathrm{p_{\bullet 11}}, \mathrm{p_{\bullet 21}}, \ldots, \mathrm{p_{\bullet ct}}\}$, equation (8) is equivalent to the Goodman-Kruskal tau index. This is apparent since

$$\begin{aligned}\tau_{num} &= \sum_{i=1}^{r}\left[p_{\bullet 11}\left(\frac{p_{i11}}{p_{\bullet 11}} - p_{i\bullet\bullet}\right)^2 + p_{\bullet 21}\left(\frac{p_{i21}}{p_{\bullet 21}} - p_{i\bullet\bullet}\right)^2 + \cdots + p_{\bullet ct}\left(\frac{p_{ict}}{p_{\bullet ct}} - p_{i\bullet\bullet}\right)^2\right] \\ &= \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{t} p_{\bullet jk}\left(\frac{p_{ijk}}{p_{\bullet jk}} - p_{i\bullet\bullet}\right)^2 \\ &= \tau_{GWnum} .\end{aligned}$$

Therefore, the Gray-Williams index for the $\mathrm{r} \times \mathrm{c} \times \mathrm{t}$ contingency table, N, is equivalent to the Goodman-Kruskal tau index when concatenating a predictor variable. When performing a NSCA for a three-way contingency table, the influence of the predictor variables $\mathrm{X_1}$ and $\mathrm{X_2}$ on the response variable Y may therefore be

made by considering the SVD of the concatenated data. By considering the concatenated contingency table, NSCA can be applied to obtain profile coordinates

$$f_{im}^{*} = a_{im}^{*}\lambda_{m}^{*} \qquad \text{and} \qquad g_{(jk)m}^{*} = b_{(jk)m}^{*}\lambda_{m}^{*} \tag{9}$$

Here, a_{im}^{*} is the *i*th element of the *m*th singular vector associated with the rows of the concatenated table. Similarly $b_{(jk)m}^{*}$ is the *(j, k)*th element of the *m*th singular vector associated with the columns of the concatenated table, and λ_{m}^{*} is the *m*th singular value.

By considering (9) for the concatenated NSCA, the numerator of the Gray-Williams index may be expressed as the weighted sum of squares of these coordinates so that

$$\tau_{GWnum} = \sum_{i=1}^{r}\sum_{m=1}^{M}\left(f_{im}^{*}\right)^{2} = \sum_{j=1}^{c}\sum_{k=1}^{t}\sum_{m=1}^{M} p_{\bullet jk}\left(g_{(jk)m}^{*}\right)^{2} = \sum_{m=1}^{M}\left(\lambda_{m}^{*}\right)^{2} .$$

For the application of confidence circles we consider Gray-Williams τ and the C-statistic of Anderson –Landis

$$C = (n-1)(r-1)\tau = (n-1)(r-1)\sum_{m=1}^{M}\lambda_{m}^{2} \Big/ \left(1 - \sum_{i=1}^{r} p_{i\bullet}^{2}\right).$$

Under the hypothesis of zero predictability ($H_0 : \Pi_{ijk} = 0$), Anderson Landis (1980) showed that the C-statistic is asymptotically chi-squared distributed with $(r-1)(ct-1)$ degrees of freedom. Therefore the radius can be computed as shown in formula (5).

5. Analysis of interaction term

Interaction effects represent the combined effects of predictor variables on the response variable. When interaction effects are present, the impact of one predictor variable depends on the level of the other predictor, in other words it means that interpretation of the main effects is incomplete or misleading.

In case of no interaction effect, a difference in level between the two lines would indicate a main effect of predictor variable.

Many texts stipulate that you should interpret the interaction first. If the interaction is not significant, you can then examine the main effects without needing

to qualify the main effects because of the interaction. If the interaction is significant, you cannot examine the main effects because the main effects do not tell the complete story. It seems that it makes more sense to tell the simple story first and then the more complex story. In the two-way case, we prefer to examine each of the main effects first and then the interaction.

Regarding MNSCA, in order to consider the different effects of the predictor variables (main and interaction effects) on response variable, our approach starts from the exact reconstruction formula of the contingency table using eigen values and coordinates, particularly

$$p_{ijk} = p_{.jk}\left[p_{i.} + \sum_{m=1}^{M}\left(\frac{1}{\sqrt{\lambda_m}}\right) a_{im}\, g_{jkm} \right] \tag{10}$$

The coordinates g_{jkm} (for $j = 1, 2, \ldots c$ and $k = 1, 2, \ldots, t$) include the main effects (j,k) and the interaction (jXk), the a_{im} (for $i = 1, 2, \ldots r$) are the row coordinates (response variable) and λ_m is the eigenvalue with $m = \min\left[(r-1);(ct-1)\right]$ equal to the rank of matrix.

We replace in formula (10) the coordinates g_{jkm} ($m=1\ldots.M$) with the functions h_{jm} and w_{km} obtained by two way analysis of variance without interaction.

$$\hat{p}_{ijk} = p_{.jk}\left[p_{i.} + \sum_{m=1}^{M}\left(\frac{1}{\sqrt{\lambda_m}}\right) f_{im}(h_{jm} + w_{km}) \right] \tag{11}$$

This new matrix $\hat{P}$ represents the dependence between categories of rows and columns after the elimination of the interaction effect. Performing a MNSCA on $\hat{P}$ we improve the interpretation of the main effect, to be more precise we represent only the effect of the prediction variables on the response variable purified to the interaction between predictors.

The choice of this functions yields the following orthogonal decomposition:

$$\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{t}\left[\left(\frac{P_{ijk}}{P_{\bullet jk}}-P_{i\bullet\bullet}\right)^2 p_{\cdot jk}\right]=\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{t}\left[\left(\frac{\hat{p}_{ij,k}}{P_{\bullet jk}}-P_{i\bullet\bullet}\right)^2 p_{\cdot jk}\right]+$$

$$+\sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{t}\left[\left(\frac{p_{ijk}}{p_{\bullet jk}}-\frac{\hat{p}_{ij,k}}{p_{\bullet jk}}\right)^2 p_{\bullet jk}\right]$$

The ***Gray Williams "Multiple" numerator = Main effects + Interaction***

It easy to verify that the matrix $\hat{P}$, with general terms $\hat{p}_{ij,k}$, and P have the same column and row marginals. Moreover considering the matrix $\overline{P}=P-\hat{P}$ and performing a MNSCA, we compute the analysis of the interaction between predictors without the main effects.

It is possible to show that if in our approach we use one-way analysis of variance instead of two way analysis of variance without interaction, we get the solution proposed by Takane-Jung (2009) based on linear constraints on the predictor categories (Takane-Shibayama 1991). In this last case the Gray-Williams multiple τ is decomposed in two components: the formes gives the Goodman-Kruskal numerator and the other gives partial τ_{num} Gray-Williams.

6. Case study

In this section, we present a detailed application of the proposed method. The case study pertains to the analysis of a 5 x 6 x 4 contingency table obtained cross-classifying subjects by mathematical score at University, teaching method used and final grade at school (independence variables). The data collected are placed in a bivariate table (table 1).

Table 1. Cross classification of students in term of mathematical score at University (criterion variable), teaching method used and final grade at school (predictor variables)

	Method	A	A	A	A	A	A
	School score	School very low score (---)	School low score (--)	School middle low score (-)	School middle high score (+)	School high score (++)	School very high score (+++)
Mathematic score very low	(--)	12	10	9	8	6	5
Mathematic score low	(-)	7	18	8	3	6	4
Mathematic score middle	(±)	1	8	10	5	1	1
Mathematic score high	(+)	2	8	6	12	16	5
Mathematic score very high	(++)	5	2	9	6	12	25

	Method	B	B	B	B	B	B
	School score	School very low score (---)	School low score (--)	School middle low score (-)	School middle high score (+)	School high score (++)	School very high score (+++)
Mathematic score very low	(--)	12	10	5	5	4	4
Mathematic score low	(-)	5	7	7	4	2	2
Mathematic score middle	(±)	1	6	6	4	3	2
Mathematic score high	(+)	5	4	3	17	11	3
Mathematic score very high	(++)	3	3	2	5	9	28

	Method	C	C	C	C	C	C
	School score	School very low score (---)	School low score (--)	School middle low score (-)	School middle high score (+)	School high score (++)	School very high score (+++)
Mathematic score very low	(--)	2	3	4	1	3	2
Mathematic score low	(-)	4	5	8	2	4	5
Mathematic score middle	(±)	5	6	9	6	7	8
Mathematic score high	(+)	5	8	13	13	22	23
Mathematic score very high	(++)	20	22	30	35	40	45

	Method	D	D	D	D	D	D
	School score	School very low score (---)	School low score (--)	School middle low score (-)	School middle high score (+)	School high score (++)	School very high score (+++)
Mathematic score very low	(--)	1	1	3	2	1	1
Mathematic score low	(-)	2	2	5	4	2	2
Mathematic score middle	(±)	2	2	7	7	8	9
Mathematic score high	(+)	5	6	16	12	18	21
Mathematic score very high	(++)	12	20	24	28	38	44

Source: own creation

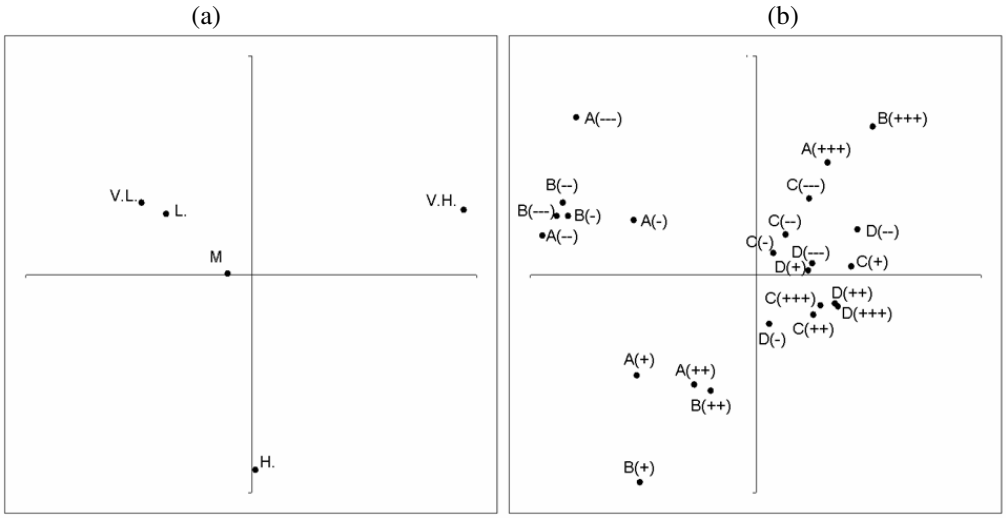
The first variable is a response or criterion variable and it has five classes of ordered categories: 18-20 (VL), 21-23 (L), 24-26 (M), 27-29 (H), 30 (VH). The second and third variables are predictor variables. The variable Final grade at Italian school are grouped in six categories: 60-64 (---), 65-69 (--), 70-79 (-), 80-89 (+), 90-94 (++), 95-100 (+++). Teaching methods has four categories: Technological tools - Projector/video/slide (A), Problem solving Brainstorming (B), Direct Teaching (C) and Lecture (D). The symbol in parentheses are the label in graphic representations.

In order to analyze the statistical dependence of mathematical score at University from teaching method used and final grade at school we perform a

MNSCA. We represent the two dimensional configuration as the best solution in Figure1, particularly in the left side we plot the criterion categories, in the right side we project the modalities of the predictor variables.

In order to selection the dimension we use permutation test. It works in the following way: first, on compute the singular value (SV) from the original data set. Then, the columns of predictor variables are randomly permuted, and SV's are computed from the permuted data set. The largest SV from the permuted data set is compared with that from the original data set. To test the statistical significance of the SV from the original data set, we repeat the same procedure K time (with K very great) and count how many times the former is larger than the latter. If this count is smaller than Kα (where α is the prescribed significance level), the largest SV being tested is significantly different from 0. Each subsequent SV can be tested in the same way after eliminating the effect of the preceding SV's. In our case we found two axis significative.

Figure 1. Classical MNSCA (a) row coordinates (b) column coordinates (total inertia explained 91,51%)



Source: own creation

The predictive power of a particular predictor category on a particular criterion category can be evaluated by the magnitude of the inner product between the two vectors representing the two categories. For example overlapping the two plots we can remark that A (+++) and B(+++) are closest to VH (the highest mathematical score). This means that the students having the highest final grade at school and used as teaching method A and B have achieved highest mathematical score.

In table 2, for each category has been computed the radius of confidence circle and the distance from the origin of the axes. The decision rule is: if the radius is greater than the distance then the category is significant

Table 2. Radius of the Confidence Circle and Distance from the origin of the axes (* Category statistical significant at the 5% level).

	Distance from origin	**Radius**
A(---)	0,453	0,200*
A(--)	0,476	0,153*
A(-)	0,279	0,160*
A(+)	0,299	0,178*
A(++)	0,204	0,162*
A(+++)	0,223	0,164*
B(---)	0,447	0,204*
B(--)	0,438	0,190*
B(-)	0,424	0,217*
B(+)	0,384	0,176*
B(++)	0,189	0,193
B(+++)	0,330	0,166*
C(---)	0,158	0,173
C(--)	0,087	0,157
C(-)	0,049	0,130
C(+)	0,214	0,138*
C(++)	0,140	0,119*
C(+++)	0,151	0,114*
D(---)	0,128	0,221
D(--)	0,235	0,187*
D(-)	0,074	0,140
D(+)	0,118	0,143
D(++)	0,182	0,127*
D(+++)	0,188	0,118*

Source: own creation

The classical MNSCA is based on the decomposition of Gray Williams "Multiple" τ including together main effects and interaction term. In order to know the statistical significance of the single main effect and of the interaction, we can use the factorial representation analysis of variance of nominal data (Onukogu 1984). The results are summarized in table 2.

Table 3. CATANOVA table

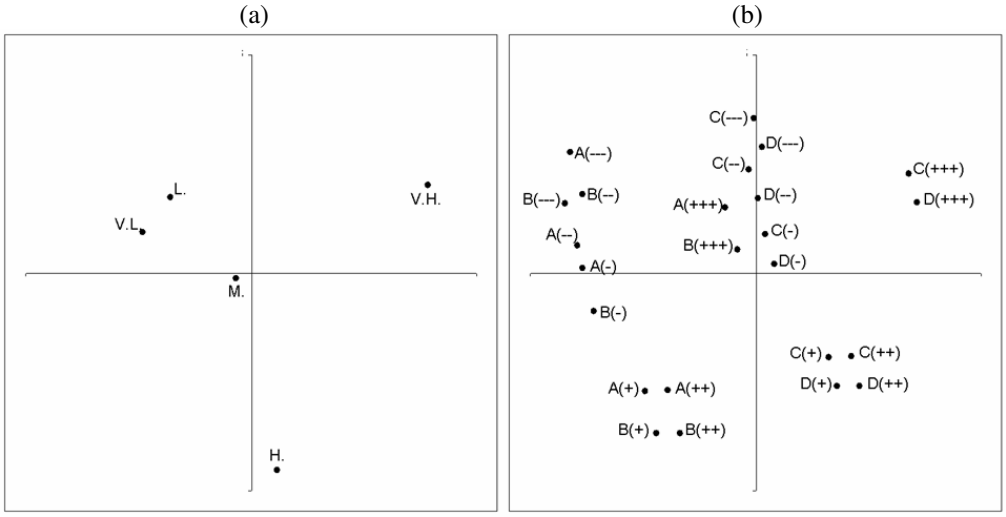
Source	SS	C-statistic	dof	p-value
Final grade at school (1)	0,021	22,75	5	0,0002
Teaching methods (2)	0,029	31,26	3	0,0000
Interaction (3) Final grade at school *Teaching methods	0,018	19,98	15	0,0447
Between (1) + (2) + (3)	0,068	73,99	23	0,0000
Within	0,652		1053	
Total	0,720		1076	

Source: own creation

All the sources of variation are statistically significant, therefore the levels of Matematical Score depends on the final grade at school, on teaching methods and on their interaction. In classical MNSCA the effect of interaction could make unclear the interpretation of the axis.

Following the approach proposed in section 5, we can separate the effect of the main sources of variation from the interaction. Particularly in figure 2 and 3 we represent the main effects and the interaction term respectively. Following the procedure presented previously, in both cases the dimension selected is composed by two axis significative.

Figure 2. MNSCA only main effects (a) row coordinates (b) column coordinates (total inertia explained 94,83%)



Source: own creation

In figure 2, we note that overlapping the two plots the predictor categories closest to VH (the highest mathematical score) are C (+++) and D (+++), moreover it seems that the teaching methods more effective are C and D because all categories of Final grade at school are closest to the criterion categories VH and H.

Considering only the main effects, in table 4 for each category has been computed the radius of confidence circle and the distance from the origin of the axes. The decision rule is: if the radius is greater than the distance then the category is significant

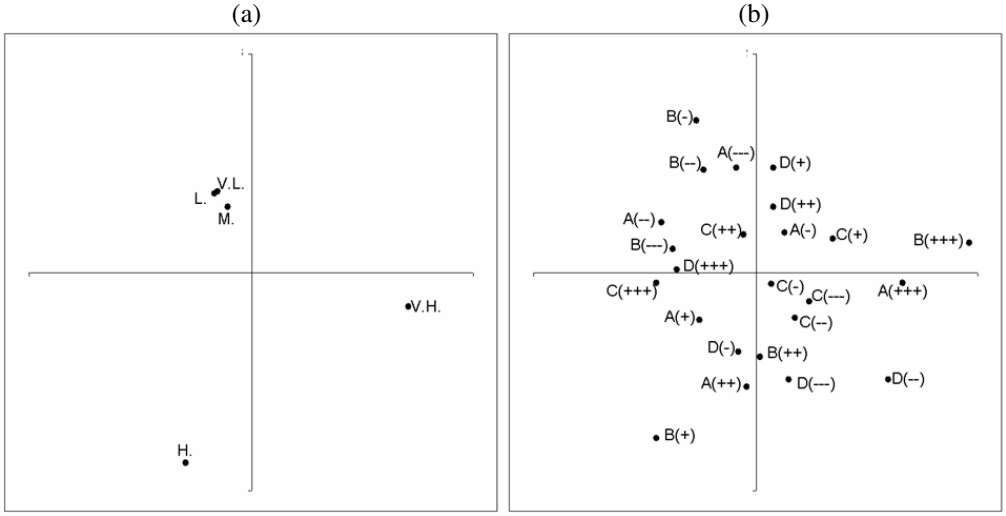
Table 4. Radius of the Confidence Circle and Distance from the origin of the axes (* Category statistical significant at the 5% level) for main effects

	Distance from origin	**Radius**
A(---)	0,339	0,200*
A(--)	0,341	0,153*
A(-)	0,308	0,160*
A(+)	0,213	0,178*
A(++)	0,176	0,162*
A(+++)	0,070	0,164
B(---)	0,312	0,204*
B(--)	0,317	0,190*
B(-)	0,288	0,217*
B(+)	0,207	0,176*
B(++)	0,174	0,193
B(+++)	0,036	0,166
C(---)	0,106	0,173
C(--)	0,072	0,157
C(-)	0,032	0,130
C(+)	0,142	0,138*
C(++)	0,180	0,119*
C(+++)	0,281	0,114*
D(---)	0,087	0,221
D(--)	0,051	0,187
D(-)	0,034	0,140
D(+)	0,164	0,143*
D(++)	0,201	0,127*
D(+++)	0,292	0,118*

Source: own creation

These results are supported by the plot of interaction in which we can observe how the students with the highest score in mathematic are those that had the best Final grade at school and used A and B teaching methods.

Figure 3. MNSCA only interaction effect (a) row coordinates (b) column coordinates (total inertia explained 86,34%)



Source: own creation

In this paper we have presented a method of analysing two complementary parts of the predictive relationships between the columns and rows of a three way contingency table, one part can be explained by main effect of predictive categories on criterion variable and the other represents the effect of interaction between the two predictive variables on the criterion variable. The usefulness of the method is shown by a study regarding the statistical dependence of mathematical score at University from teaching method used and final grade at school.

References

Agresti, A. 1990: *Analysis of Categorical Data*. Wiley, New York.

Anderson, R. J. - Landis, J. R. 1980: CATANOVA for multidimensional contingency tables: Nominal-scale response. *Comm. Statist. Theory Methods,* 9, 1191-1206. p.

Beh, E. J. - D'Ambra, L. 2009: Non symmetrical correspondence analysis with concatenation and linear constraints. *Australian and New Zealand Journal of statistics.* (to appear 2010)

Caroll, J. D. - Chang, J.-J. 1970: Analysis of individual differences in multi-dimensional scaling via an N-way generalisation of "Eckart-Young" decomposition. *Psychometrika* 35, 283-319. p.

D'Ambra, L. – Lauro, N. 1989: Non symmetrical analysis of three way contingency tables. In Coppi, R. - Bolasco, S. (eds): *Multiway Data Analysis*. North Holland, 301–315. p.

D'Ambra, L. - Lauro, N. 1992: Exploratory non symmetric data analysis. *Italian J. Appl. Stats.* 4, 511-529. p.

Faber, N. M. - Bro, R. - Hopke, P. K. 2003: Recent developments in CANDECOMP/PARAFAC algorithms: a critical review. *Chemo. Intell. Lab. Syst.,* 65, 119-137. p.

Gabriel, K. R. 1995: MANOVA biplots for two-way contingency tables. In Krzanowski, W. J. (eds): *Recent Advances in Descriptive Multivariate Analysis*. Oxford University Press, 227 – 268. p.

Gimaret-Carpentier, C. - Chessel, D. - Pascal, J.-P. (1998): Non-symmetric correspondence analysis: an alternative for species occurrences data. *Plant Ecology* 138, 97-112. p.

Goodman, L. A. - Kruskal, W. H. 1954: Measures of association for cross-classifications. *J. Amer. Statist. Assoc.,* 49, 732-764. p.

Gray, L. N. - Williams, J. S. 1975: Goodman and Kruskals tau b: Multiple and partial analogs. *American Statistical Association (Proceedings of the Social Statistics Section)*, 444-448. p.

Greenacre, M. J. 1984: *Theory and Applications of Correspondence Analysis*. Academic Press, London.

Harshman, R. A. 1970: Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multi-modal factor analysis. *UCLA Working Paper in Phonetics,* 16, 1-84. p.

Kroonenberg, P. 1989: Singular value decompositions of interactions in three-way contingency tables. In Coppi, R. - Bolasco, S. (eds): *Multiway Data Analysis*. Amsterdam, North-Holland, 169 – 184. p.

Kroonenberg, P. - Lombardo, R. 1999: Nonsymmetric correspondence analysis: A tool for analysing contingency table with a dependence structure. *Multiv. Beh. Res.,* 34, 367–397. p.

Lebart, L. - Morineau, A. - Warwick, K. M. 1984: *Multivariate Descriptive Statistical Analysis*. Wiley, New York .

Light, R. J. - Margolin, B. H. 1971: An analysis of variance for categorical data. *J. Amer. Statist. Assoc.,* 66, 534-544. p.

Lombardo, R. - Carlier, A. - D'Ambra, L. 1996 : Nonsymmetric correspondence analysis for three-way contingency tables. *Methodologica,* 4, 59-80. p.

Marcotorchino, F. 1985: *Utilisation des Comparaisons par Paires en Statistique de Contingences: Parte III*. Report # F 071, Etude du Centre Scientifique, IBM, France.

Mardia, K. V. - Kent, J. T. - Bibby, J. M. 1982: *Multivariate Analysis*. Academic Press, London.

Mirkin, B. 2001: Eleven ways to look at the chi-squared coefficient for contingency tables. *Amer. Statist.,* 55, 111-120. p.

Onukogu .I. B. 1984: An Analysis of variance of nominal data . Statistica anno XLIV n.

Ringrose, T. J. 1992: Bootstrapping and correspondence analysis in archaeology. *J. Arch. Sci.,* 19, 615-629. p.

Ringrose, T. J. 1996: Alternative confidence regions for canonical variate analysis. *Biometrika,* 83, 575-587. p.

Takane, Y. - JUNG, S. 2009: Tests of ignoring and eliminating in non symmetric correspondence analysis. *Advances in data analysis and classification* Vol. 3, Issue 3, 315-340. p.

Takane, Y. - Shibayama, T. 1991: Principal component analysis with external information on both subjects and variables. *Psychometrika,* 56, 97-120. p.

Tucker, L. R. 1966: Some mathematical notes on three mode factor analysis. *Psychometrika*, 31, 279 - 311. p.

Is Free Market Access Enough for Development? – Lessons of the Lomé Conventions

Beáta Udvari[1]

Nowadays, problems of the developing countries belong to the global issues. More and more initiatives and financial assistance are to promote their development, but their effectiveness is at a low level. Perhaps as a result of it, the role of trade becomes more important and international organizations emphasize its role in development. Despite its current relevance, there were also attempts on this field in the past, the European Union's activity is outstanding. The EU granted significant trade and other economic preferences for the African, Caribbean, Pacific (ACP-) countries on a non-reciprocal basis for a long time in the framework of the Lomé Conventions.

In our study – using multivariate statistical methods – we aim to investigate the Conventions' impact on the economic development of the beneficiary countries. Our before-after comparison is based on cluster analysis investigating the relative situation of countries within and outside of, and before and after the Lomé-system. Our empirical research shows that the Lomé Conventions do not have significant impact on the economic development of the beneficiary countries.

Keywords: Lomé Conventions, free trade, trade and development, cluster analysis

1. Introduction

Nowadays trade plays an essential role, although some countries are unable to take part in the global trade effectively. It is accepted that trade could contribute to the eradication of poverty, and several international organizations accept it. They represent the interest to enable extremely favourable or free market access to the developing, mainly the least developed countries, and this would enable these countries to develop. Among these institutions, the European Union has been paying also great attention on trade in its development policy, as the EU has been providing favourable market access to several developing countries to the European markets.
The EU differentiates among the developing countries, and it has built up a special relationship with the African, Pacific, Caribbean (ACP) countries, and the Lomé

[1] Beáta Udvari, PhD student, University of Szeged, Faculty of Economics and Business Administration, Hungary.
The author would like to express her gratitude to Péter Kovács for the regular and useful meetings throughout the empirical research.

Conventions provide the framework of this relationship. 46 ACP countries and 9 EC member states signed the first Lomé Convention in 1975, and three other Conventions followed this providing free market access and other economic preferences for the ACP countries. The objective of the conventions was to promote the economic and social development of the contracting developing countries.

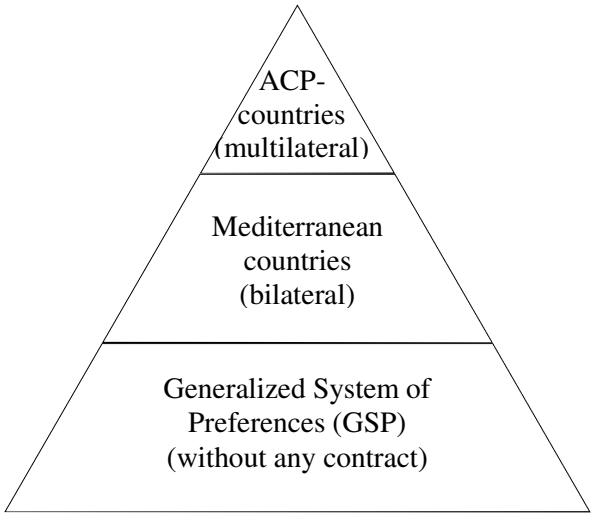
All these mean that there have already been initiatives which tried to integrate developing countries into the world trade and promote development through trade. Consequently, it is worth analysing how effective they were. Therefore *the objective of this paper is to investigate whether the Lomé Conventions as a unique initiative could contribute to the economic development of the contracting African, Pacific and Caribbean (ACP) countries*. To achieve this objective, we hammered out a new statistical methodology: the empirical analysis is based on a *before-after comparison using cluster analysis*. The analysis does cover not only the ACP countries, but other developing countries were also involved into the empirical research, playing as a control group. The paper first discusses the EU's relationship with the developing countries, then it introduces the Lomé Conventions as the greatest tool of the EU's development policy. In the next section we detail the methodology of the statistical analysis and present the results of the cluster analysis.

2. The EU's relationship with the developing countries

The European Union has been playing a great role in promoting development in the less-advanced countries. To achieve this aim, the EU grants not only financial assistance, but trade preferences to the developing countries in its development policy (Udvari 2008). The EU has established relationship with almost all the developing countries, but the preferences it grants for them are different. According to the obtained trade preferences we can analyse the relationship and the EU's influence on the developing countries (Figure 1).

The pyramid of beneficiaries shows that *the ACP countries are staying at the top level*, since they receive the greatest preferences and the most support from the European Union (Balázs 2002, Persson-Wilhelmsson 2006), that is these countries are the main beneficiaries of the EU's development policy. The EU-ACP relationship is based on contracts which ensure high level of security for the contracting parties. Besides, the contracts have a multilateral form as they are not signed with only one country but with several ones (with a country group). The preferences granted by the contracts cover a wide range of products, and there are only few restrictions and exemptions (Persson-Wilhelmsson 2006). Nowadays, 79 countries belong to the ACP-group (EC 2009a), out of which 48 stays in Africa. The framework of the EU-ACP relationship is based on the Lomé Conventions, which provided different preferences for the ACP countries. As the Lomé Conventions are the key point of our analysis, we detail them in the next section.

Figure 1 Pyramid of beneficiaries in the aspect of trade preferences



Source: own construction based on Persson-Wilhelmsson (2006)

At the *second level of the pyramid the Mediterranean countries are staying* – they receive similar preferences to the ACP countries, although the contracts are bilateral and there are more exemptions and several restricting rules than in the case of the ACP countries (Persson-Wilhelmsson 2006). Among the Mediterranean countries we can mention 16 countries with which the EU re-launched the Euro-Mediterranean Partnership in 2008 (EC 2009b).

At the *lowest level* of the pyramid those countries are standing which obtain only preferences *under the GSP*[2]*-rules*. Although this system is the oldest tool in the EU's trade preferences, the GSP-system provides the less favourable preferences and the EU grants them in an autonomous way: there are several restrictions, the rules on country of origin are extremely strict, besides, and here are several exemptions (Horváth 2005, Persson-Wilhelmsson 2006). Nowadays almost all the developing countries enjoy preferences under this system (EC 2009c).

This pyramid is essential for our analysis, since it gives us the opportunity to choose out the potential countries for our analysis. The pyramid shows us that in the aspect of trade preferences the EU provides for developing countries the ACP countries are the main beneficiaries, but other countries are also in connection with the European Union in a closer or a looser way. Consequently, this pyramid gives the basis for determining the group of control countries.

[2] Generalised System of Preferences

3. The Lomé Conventions

As the Lomé Conventions and their interventional areas provide the basic points to our empirical research, it is important to know their key elements. As the EU's development activity contains not only financial assistance but trade preferences, we accept that the Lomé Conventions belong to the tools of the EU's development policy (Horváth 2005). They granted preferences only to the ACP countries – but the preferences contained not only trade preferences, but interventions on other economic areas (diversification, investment), too (Babarinde-Faber 2004, Bjornskov-Krivonos 2001, Nunn-Price 2004).

The first convention was signed in 1975 and three other followed it. They expired after five years, excluding the fourth one with its 10-years. Consequently, between 1975 and 2000 the ACP countries could enjoy great preferences in the framework of these agreements.[3] The number of the ACP countries changed during this time: 46 ACP countries signed the first Lomé Convention, while 68 countries the fourth one (Bilal-te Velde 2003, Udvari 2008). All of the conventions involve the following areas (Babarinde 1994, Udvari 2008):

- *Trade preferences* ensured the ACP-countries free market access to the European markets in a non-reciprocal way.
- *Industrial cooperation* aimed at the industrial and technical development of the beneficiaries and at promoting the technology transfer.
- *Financial cooperation* enabled the ACP-countries to partake in financial support under the framework of the European Development Fund, European Investment Bank and the EU-budget.
- *Foreign direct investment* prescribed that no restriction may hinder the movement of capital between the contracting parties.
- *Technical cooperation* aimed at the diversification and modernization of the beneficiaries' economy.

All of these areas can be found in every convention, the difference is only the growing number of the articles in the conventions. It was supposed that the interventions would contribute to the economic and social development of the contracting ACP countries. Although most of the relevant literature emphasizes the trade preferences in connection with the Conventions, the importance of these interventions is larger.

3.1. The importance of the Lomé Conventions

These Conventions were unique in the era when they were signed. The concept seems to be successful as the number of the ACP countries signed the Conventions grew from 46 to 68 (and there was no left). It is unambiguous that *the key point of the Conventions is the trade preferences and long-term free access* to the European

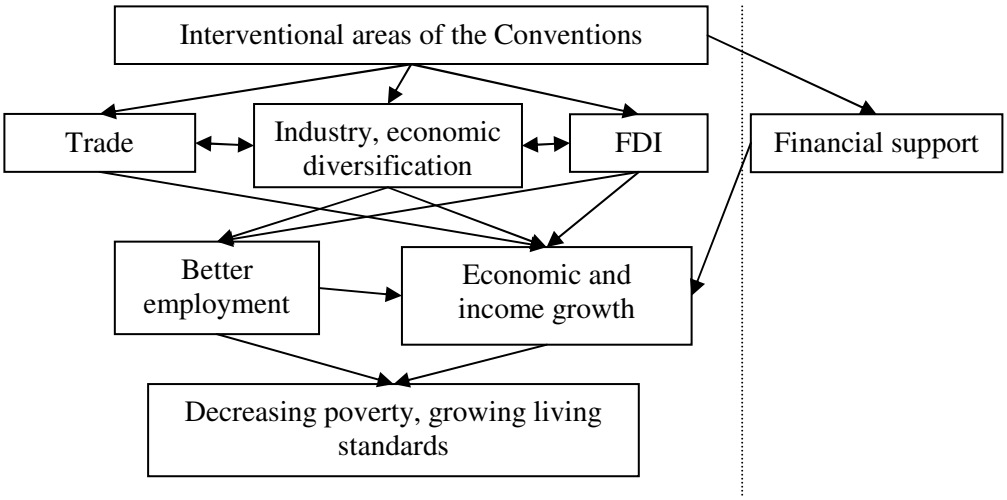
[3] Since 2000 the Cotonou Agreement plays this role for the ACP countries (Babarinde-Faber 2004).

market granted in contractual form to the recipient countries (Babarinde-Faber 2004, EC 1996).

The granted trade preferences are important because it covered 90% of the ACP export products (Babarinde-Faber 2004, Dezséri 2003). Besides, more than half of the ACP countries were contracting parties of the GATT[4], which dealt with tariff cuts of only industrial products and not of agricultural ones – which meant a large disadvantage for the developing countries (Stiglitz 2003). Furthermore, non-tariff barriers were imposed on goods produced by the ACP countries (Somai 1997), so these countries had a non-favourable position in the international trade.

The literature about the Lomé Conventions emphasizes the trade preferences only but we believe they were more than simple trade conventions, since they contained elements in connection with industrialization and foreign direct investments, as well as elements about financial assistance (Figure 2). All of these had the aim of economic diversification and to hinder that the ACP countries be independent from one (generally agricultural) export product and be able to decrease the risk coming from the fluctuating price of the products they can export.

Figure 2. Scheme of the Lomé Conventions



Source: own construction

All these facts suggest that the intervention of the EU should be favourable for achieving real economic development of the beneficiary ACP countries. Considering this, we try to analyse the economic effects of these Conventions using multivariate statistical methods. We have to emphasize that the interventional areas

[4] General Agreement on Tariffs and Trade, signed in 1947. It is the former institution of the World Trade Organization.

provide the basis for our empirical research, but in our analysis we investigate only the economic side, since, on one hand, because of lack data we had to leave out the effects of financial support from our analysis; on the other hand, lack data hindered us to analyse social indicators and social development. In the next section we detail the empirical research we made for analysing the impacts of the Lomé Conventions.

4. The empirical analysis

The four Lomé Conventions concentrated on the same areas, consequently most of the ACP countries could enjoy trade and other economic preferences for 25 years. This enables the empirical analysis of their impacts on the economy of the ACP countries. *The analysis is based on a before-after comparison* to investigate the changes happened after the Conventions. To get the most relevant results, control countries were involved into the investigation: the ACP-development was compared to that of the control states. In this section we first introduce the relevant countries and the indicators chosen out for the analysis. Later, we detail the methodology we hammered out for the investigation, and at the end the results will be presented.

4.1. Relevant countries and indicators

For choosing out the potential countries and indicators, we considered two things: on one hand, we concentrated on the interventional areas of the Conventions, on the other hand, we paid attention on the indicators in related literature. The literature review included not only studies on the empirical analysis of the Lomé Conventions, but on the relationship between trade and economic growth (e.g. UNCTAD 2007, Yanikkaya 2003) and on competitiveness (e.g. Lengyel 2000, Lukovics 2008), as well. In our analysis we studied only the direct economic effects of the Conventions, therefore we left out social indicators. Finally, we have determined *14 – exclusively economic – indicators to analyse the direct impacts of the Conventions*, they are as follows:

1. Trade
    a. EU-share from the country's export
    b. EU-share from the country's import
    c. Country's share from the EU's extra-export
    d. Country's share from the EU's extra-import
    e. Share of export from the GDP
    f. Share of total trade from the GDP
2. Industry, economic diversification
    a. Share of agriculture from the GDP
    b. Share of manufacture from the GDP
    c. Share of services from the GDP
    d. Agricultural employed within the population

3. Foreign direct investment
    a. Inward FDI per capita
    b. Inward FDI in proportion of the GDP
4. Economic and income growth
    a. Household consumption expenditure in proportion of the GDP
    b. GDP per capita (in logarithm way)

As for the trade performance, we found that we should analyse the trade from two sides: on one hand, we have to investigate whether the EU is an important partner for the country (*1a* and *1b* indicators); on the other hand, we have to analyse whether the country is important for the EU (*1c* and *1d* indicators) and this could show the effects of the Lomé-system. The *1a* and *1b* indicators showed that the EU played great role in the trade activity of the investigated countries[5], therefore we assumed the interventions of the EU could have a relatively strong impact on these economies.

The data were collected in case of all the indicators for the years of 1970-1975 and 2000-2005, and to avoid the outlier data we counted averages for these years. The data are from the on-line database of the UNCTAD Handbook of Statistics (UNCTAD 2009a), UNCTAD FDI (UNCTAD 2009b), and the United Nations Statistics Division (UN 2009). To eliminate the different measures of the indicators, we used standardized variables.

For determining the relevant countries to the analysis, the pyramid of the beneficiaries provided the basis. Beside the ACP countries, the Mediterranean countries and economies under the GSP (as a control group) were involved into the investigation. The relevant countries are from the study of Persson-Wilhelmsson (2006). Originally, 112 countries were involved into the analysis (as was in the study mentioned above), but throughout the investigation process we had to leave some countries out of the analysis because of three reasons (so-called country-filter): first, countries were left out which were outliers in the aspect of trade (e.g. China and India); secondly, we missed countries where data for most of the indicators were not available. Consequently, the sample contained 82 countries and the analysis was based on 14 indicators. But we had to leave out another two countries (the Bahamas and Singapore) as we experienced their distorting effects on the results of the cluster analysis. Finally, *80 countries composed the sample of our analysis,* out of which 50 countries belong to the ACP-block, while 30 countries (7 Mediterranean[6] and 23 GSP-countries) belong to the control group.

[5] The mean of this indicator was 45%, and 42%.

[6] Although nowadays more countries belong to the Mediterranean countries, we left out the 'newer' countries and concentrated on the countries that Persson-Wilhelmsson (2006) had determined.

4.2. *Methodology*

For the analysis *we hammered out a new methodology*. In our assumption, the Lomé Conventions granted remarkable preferences for the ACP countries for long-term, therefore some development should be experienced in these countries after the Lomé-system expired. To gain the most appropriate result, we involved countries outside of the ACP-group to the analysis as a control group. Consequently, we compared the results of the ACP-countries and that of the control group.

Our methodology is based on a *before-after comparison*. The investigation periods are the pre- (1970-75) and the after-period (2000-05) of the Lomé-system. Analysing the pre-period gives the opportunity to get a picture of the relative situation of the ACP- and the non-ACP countries before the Lomé Conventions appeared. Comparing this situation with the after-period, we can state the changes and investigate the economic results of the Lomé Conventions. *For the comparison we used cluster analysis*, as the aim of this multivariate statistical technique is to organize cases into homogeneous groups along the involved variables, and as a result, cases in a cluster are nearer (similar) to each other than to any member in other clusters (Sajtos-Mitev 2007, Székelyi-Barna 2005).

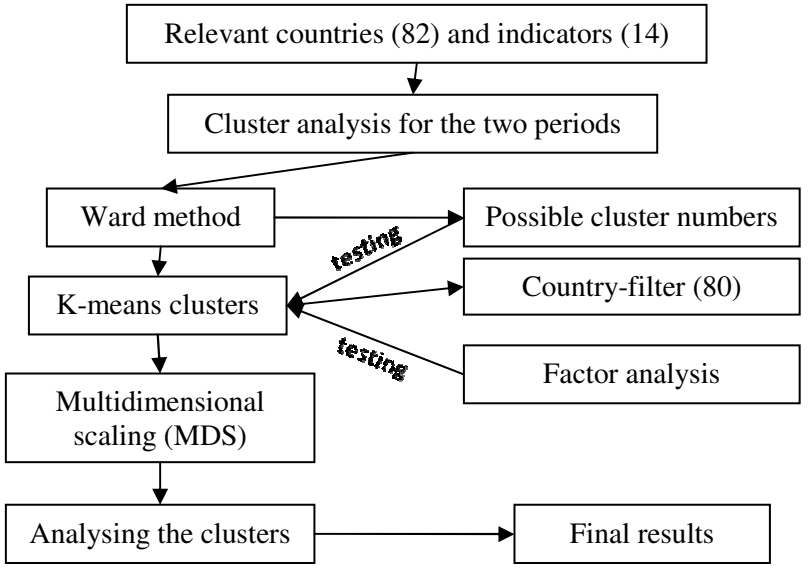
In our analysis *the clusters enabled us to investigate the economic changes between the two periods*. The *clusters of the pre-period show the relative situation of the involved countries before the EU's interventions*, as countries with similar economic results were organized into the same group. Consequently, we could determine which countries are more developed than the others and which lag behind along the indicators we involved into the analysis. Besides, the *clusters of the after-period show the results of the Lomé-system*, and we assumed the Lomé Conventions could contribute to the economic development of the beneficiaries; therefore changes should happen among the clusters. Consequently, during our analysis, *we compared two things in connection with the clusters*: on one hand, we paid great attention on the changing meaning of the clusters between the two periods; on the other hand, we analysed the changes of the cluster membership, as well.

Although the cluster analysis is the core point of our methodology, we tested and *checked the primary results with several other multivariate statistical methods* (Figure 3)[7]. These additional statistical techniques were necessary, since there were no references from the earlier studies how many clusters were assumed to gain. Therefore to determine the appropriate number of the clusters, we used two methods for the grouping and we followed the practice as preparing a hierarchical cluster analysis at first and then a non-hierarchical one (Sajtos-Mitev 2007). Consequently, the *hierarchical cluster analysis* (with Ward method) was the starting point *to determine the potential number of the clusters* with analysing the difference between the coefficients, as great difference between the coefficients shows that clusters are far from each other, therefore it is no worth uniting them (Sajtos-Mitev 2007).

[7] The process of the analysis was the same for both periods.

The hierarchical cluster analysis showed only uncertain results[8], therefore we tested the potential cases with *K-means clusters* (non-hierarchical method). This method showed us that two countries (the Bahamas and Singapore) appeared in one single cluster during the tests of all the cases, therefore we found they had a distorting effect and we should leave these two countries out of the analysis. Consequently, the number of the relevant countries decreased to 80.

Figure 3. Steps of the empirical analysis



Source: own construction

By K-means clusters method we could test the potential solutions of the cluster numbers we received by the hierarchical analysis. To choose out the most appropriate one, the *multidimensional scaling* (MDS) was taken. MDS is such a multivariate statistical technique, which decreases the data and enables the geometric representation of cases according to their distances in a decreased dimension area (Székelyi-Barna 2005). In our case, the MDS was prepared for two dimensions on all the indicators we involved into the analysis.[9] The two axis (dimensions) could be named as economic and trade performance. The two-dimension graphic appearance contributed to analyse the distances between the cases (countries) and to decide which cluster-number could give the most

[8] The potential number of the clusters was between four and eight.

[9] The value of the S-stress indicator shows whether the MDS is appropriate for the analysis (Kovács-Petres-Tóth 2006). The value of this indicator was 0.071 (good) for the pre-period, and 0.091 (good) for the after period.

appropriate solution. Using the coordinates we gained by the MDS, we represented the countries in a two-dimension area where we analysed their positions according to their cluster membership. We rejected those solutions for the potential number of clusters where the countries' appearance according to their cluster membership was not obvious.

Beside the graphic representation, we used *factor (principal component) analysis for testing the results of the K-means cluster analysis*, as well. Principal component analysis is a technique to reduce the number of the variables by getting a new variable which cannot be investigated directly (Hajdu 2003, Sajtos-Mitev 2007). We decreased the number of the variables according to the scheme and interventional areas of the Lomé Conventions, and as a result we received six factors.[10] The cluster analysis was prepared again (beginning with the hierarchical method) along these new factors as variables. The result was persuasive: *the difference between the two cluster analyses was not significant* – the most appropriate number of clusters was the same, besides, most of the countries belonged to the same cluster as originally. As there was no significant difference and as there was no contradiction with the testing methods, *we could accept the results of the K-means cluster analysis*.

4.3. Results of the cluster analysis

In this chapter we will present only the results of the K-mean cluster analysis for both periods. As there was no contradiction between the testing methods, we do not detail either the pre-calculations or the results of the testing methods. Because of the relatively large number of the countries, we cannot introduce the membership of the clusters, we will present only some examples, the final results of both periods can be found in the appendix.

4.3.1. Cluster analysis for the pre-period of the Lomé Conventions

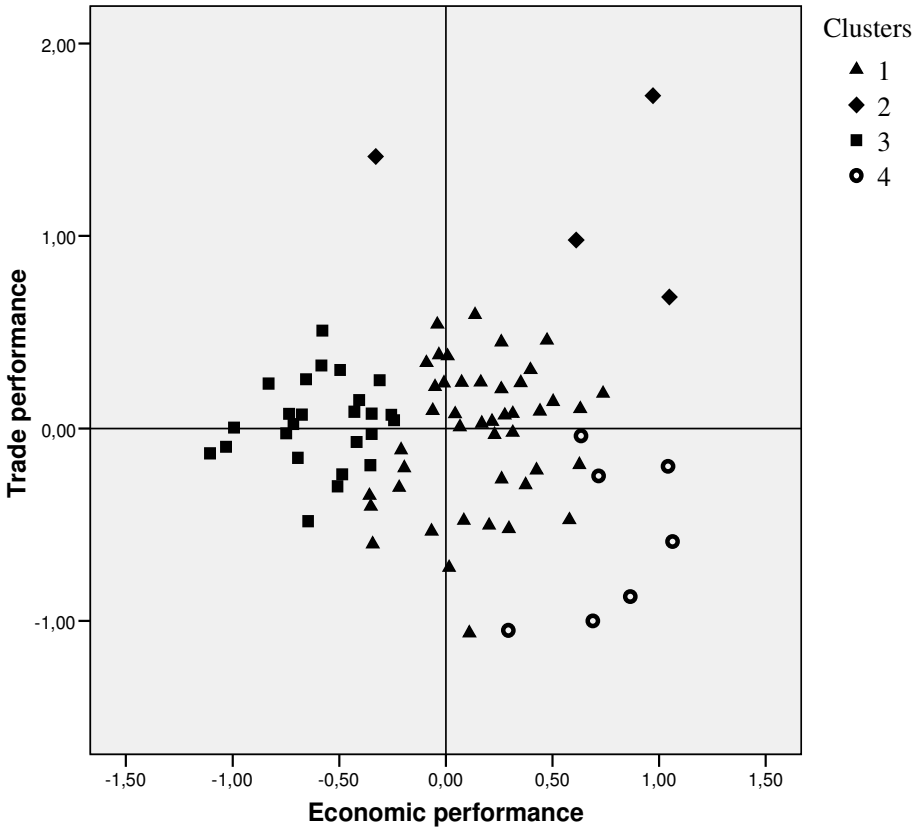
The analysis for the pre-period was prepared along the above-mentioned 14 indicators and for 80 countries. The MDS and the K-means cluster analysis strengthened us (complying with the hierarchical cluster analysis) that the most appropriate number for the clusters are four (Figure 4).

Analysing the final cluster centres and the means of all the clusters, we could give meanings for them and make economic distinguish between them. Using cross

[10] All the principal components preserved more than 80% of the information of the original indicators. These components are as follows: (1) openness, (2) EU's importance in the country's trade, (3) country's importance in the EU's trade, (4) FDI, (5) economic activity outside the industry, (6) economic income.

tables we could also analyse other features of these clusters with other variables which were not directly involved into the analysis. [11]

Figure 4. Clusters before the Conventions



Source: own construction

The cluster names and their features are as follows:

- *Middle opened, averagely developed countries (cluster 1)*, containing 44 countries, where we can find African and Latin-American countries, as well (e.g. Angola, Bolivia, Egypt, Nicaragua, Uruguay).

- *Relatively developed industrial economies with strong EU-relations (cluster 2)*, containing only 4 countries (Argentina, Brazil, Israel, Nigeria). The

[11] In these cases it is no worth speaking about ACP countries, as the ACP-group was created a bit later, and we speak about them in connection only with the Lomé Conventions.

strong EU-relations mean that the countries are important partners for the European Union, but the EU's role in the countries' trade is under the average.

- *Relatively closed and underdeveloped, agricultural countries (cluster 3)*, containing 25 economies, out of which 84% are African countries (e.g. Chad, Ghana, Burundi, Uganda).
- *Relatively developed, opened economies independent from the EU (cluster 4)*, containing only 7 countries. The openness refers to the openness for both the foreign investments and trade. Half of these countries belong to the Caribbean countries (e.g. Barbados, Jamaica).

These results show that before the Lomé-system there were two smaller groups which were relatively developed than the other countries. Furthermore, these relatively developed countries are outside of the ACP-group. It is important to emphasize that *the relatively underdeveloped, agricultural cluster contains mostly African countries*. As the African countries are the greatest beneficiaries of the Lomé-system, and they obtained more favourable preferences than any other countries, we have to pay attention on this cluster while analysing the after-period. Consequently, we should experience development and changes in this cluster and in its membership.

4.3.2. Cluster analysis for the after-period of the Lomé Conventions

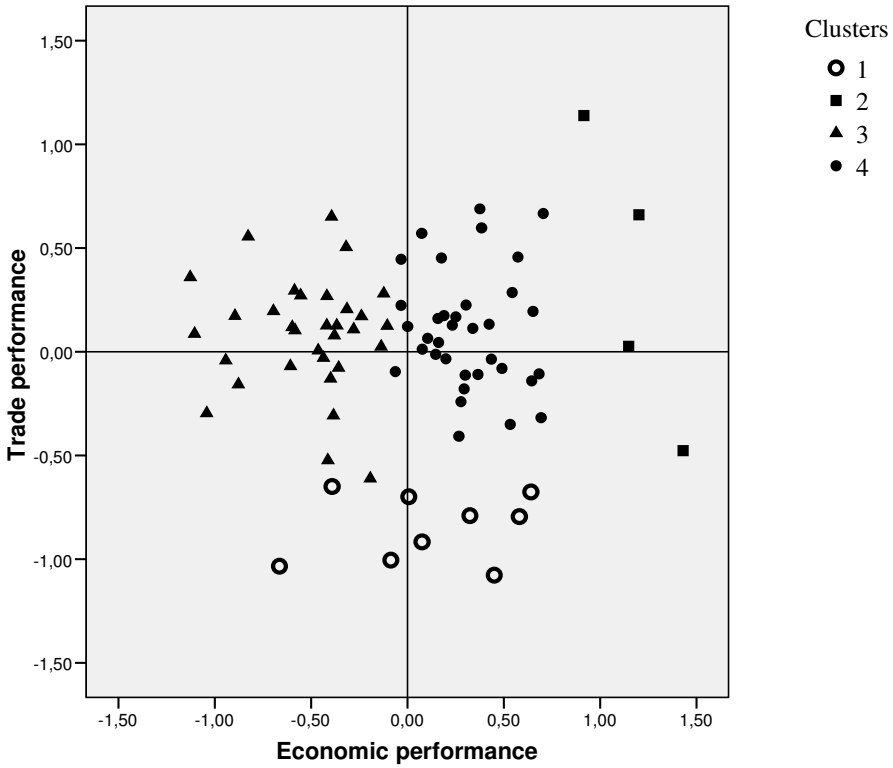
For the after-period similar analysis was prepared as happened in the case of the pre-period. Because of the interventional areas of the Lomé Conventions we assumed that there would be great changes not only in the number of the clusters, but in their meanings, as well. Despite, our analysis gave a surprising result. The K-means cluster analysis together with the MDS shows that the four-cluster solution seems to be the most appropriate (Figure 5).

Beside the similar number of clusters, the meaning of the clusters remained similar to that of the earlier period. The clusters' names are as follows and their features are analysed by using cross tables:

- *Opened economies with strong EU-relations (cluster 1)*, containing only 9 countries. They all are ACP countries and they all signed the first Lomé Convention. The EU is an important trading partner for the countries, but it is not true from the other point of view.
- *Relatively developed, industrial economies with strong EU-relations (cluster 2)*, containing only four countries, but there is a little change in comparison to the earlier period: Brazil and Israel remained in the relatively developed cluster as were in the first period, but two new countries joined them (Thailand, Malaysia). None of the cluster members signed any of the Lomé Conventions, and three of them are under the GSP-system.

- *Relatively underdeveloped, agricultural economies (cluster 3)*, containing 32 countries. 78% of the cluster members signed the first Lomé Convention (1975), and most of the cluster members (94%) are ACP countries. Besides, 60% of the ACP countries belong to this cluster.
- *Averagely developed economies without any important EU-relations (cluster 4)*, containing 35 countries. 80% of the countries which did not sign any of the Conventions became member of this cluster; consequently most of the control countries are in this cluster.

Figure 5. Clusters after the Conventions



Source: own construction

Although we expected significant changes in the meanings and features of the clusters, our results show the opposite. The clusters are similar to the pre-period ones in the aspect of their meaning. Looking at the features of the cluster members, it is unambiguous that most of the ACP countries could not gain from the preferences of the Lomé Conventions, since most of these states belong to the relatively underdeveloped cluster with an agricultural feature. Consequently, the

Lomé Conventions seem to be unsuccessful in the objective of economic development and diversification in the beneficiary countries. The features of the clusters' membership present that countries out of the Lomé-system could perform better in the aspect of the investigated variables. It is important to compare the two periods' results, and we can do it as the clusters are extremely similar to each other.

4.3.3. Comparing the two periods

The results of the two periods are similar to each other: the clusters of the after-period could be corresponded to those of the pre-period. This provides the basis for comparing the two periods. Since the sample of our before-after comparison was equal in the two periods (there was no change), that is the data of the two periods belong to the same sample, we can control these two related samples' homogeneity. And homogeneity of two related samples can be checked with nonparametric tests (Ketskeméty-Izsó 2005, Vargha 2000). Consequently, for comparing the results of the two periods in our analysis *we used a nonparametric test*[12] *(Wilcoxon Signed Ranks Test)*, with which we investigated two things:

- whether there is *significant difference between the cluster memberships* of the two periods;
- whether *signing one of the Lomé Conventions contributed to the changes of the cluster membership*.

In our analysis, the Wilcoxon-test shows that there is *no significant difference between the cluster memberships* of the two periods (p-value: 0.152). This means that countries belonging one cluster of the pre-period became members mostly the same cluster in the after-period. Besides, as for the effects of the Lomé conventions on the cluster membership, the result of the non-parametric test presents that there is *no importance whether the countries signed the Lomé Conventions or not*, the contracts do not have determining role in the changes of the cluster membership (p-value varies between 0.102 and 0.336 along the Lomé Conventions).

Consequently, *the Lomé Conventions* – though they seemed to be effective intervention by the European Union – *seem to perform without any spectacular results*. Countries which were out of the Lomé system could achieve better results and they are more developed than the ACP countries getting significant preferences. Our analysis shows that the most unfavourable result occurs in case of the countries, which belonged to the underdeveloped, agricultural economies in the pre-period: although they became beneficiaries of the Lomé-system, they could not perform much better and they remained in the same cluster. The objective of economic diversification was not successful, either – as the most underdeveloped cluster

[12] At first we had to recode the number of two clusters to correspond to each other in the two periods. As a matter of fact, cluster *4* in the after-period corresponds to cluster *1* in the pre-period (the new number is 1), while cluster *1* in the after-period to cluster *4* in the pre-period.

contains mostly ACP countries. Analysing the cluster membership in the after-period we could see that there are some success countries – but unfortunately they are little and happened to countries mostly which are outside the ACP-group (e.g. Thailand, Malaysia).

5. Conclusions

The aim of this paper was to study the impacts of the Lomé Conventions on the economic development of the ACP countries. This paper and the empirical analysis show us that *the Lomé Conventions do not have important impact on the economic development of the contracting countries*.

The paper introduced that the ACP countries became the greatest beneficiaries of the EU's development policy, the preferences they can enjoy are at the highest level. These show that a special relationship exists between the EU and the ACP countries. The framework of their relationship is regulated by the Lomé Conventions which granted trade and other economic preferences for the contracting ACP-group. The preferences were on non-reciprocal basis, meaning the ACP countries could enjoy free market access to the European markets without any compensation. As the Conventions granted long-term preferences on the same areas (trade, economic diversification, foreign direct investments) in contractual form, the basis for an empirical analysis is given to investigate their direct impact on the development of the beneficiary countries.

The analysis was based on a before-after comparison: the pre-period of the Lomé Conventions was compared to the after-period. Besides, our empirical analysis contained countries outside of the ACP-group playing as a control group. Consequently, we compared the results of the ACP countries to those of the control group, stating the relative situation of the countries within and outside of the Lomé-system. For the analysis we used multivariate statistical techniques and we hammered out a new methodology. The empirical research was based on cluster analysis using other multivariate statistical methods for testing the results.

As 25 years passed between the two periods, changes and some economic development were expected. But our analysis does not show significant changes. Interestingly, the clusters' number remained the same with the same meaning in the after period as they were in the pre-period. Comparing the memberships of the two periods we could state that there is no significant change. Moreover, most of the African ACP countries remained members in the least developed cluster. This result is more important if we remember that the Lomé Conventions were not only about trade preferences, but contained elements in connection with economic diversification and foreign investments.

Consequently, free market access and non-reciprocal trade preferences themselves are not enough for achieving economic development. This kind of

initiative of the international organizations should be complemented with other development assistance to achieve more favourable results.

References

Babarinde, O. A. 1994: *The Lomé Conventions and Development. An empirical assessment.* Aveburg Ashgate Publishing Limited, Aldershot.

Babarinde, O. – Faber, G. 2004: From Lomé to Cotonou: Business as usual? *European Foreign Affairs Review*, 9, 27-47. o.

Balázs, P. 2002: *Az Európai Unió külpolitikája és a magyar-EU kapcsolatok fejlődése*. KJK-Kerszöv Jogi és Üzleti Kiadó Kft, Budapest.

Bilal, S. – te Velde, D. W. 2003: *Foreign direct investment and home country measures in the Lomé Conventions and Cotonou Agreement*. Study, UNCTAD, New York.

Bjornskov, C. – Krivonos, E. 2001: *From Lomé to Cotonou. The new EU-ACP Agreement*. Working Paper, 14, Danish Institute of Agricultural and Fisheries Economics (SJFI).

Dezséri, K. 2003: *Az EU fejlesztési politikája a fejlődő országok megsegítésére.* Tanulmány, MTA – Világgazdasági Kutatóintézet, Budapest.

EC 1996: *Green Paper on relations between the European Union and the ACP countries on the eve of the 21st century. Challenges and options for a new partnership*. European Commission, Brussels.

EC 2009a: *African, Caribbean, and Pacific states*. http://europa.eu/legislation_summaries/development/african_caribbean_pacific_states/index_en.htm, downloaded: 25 August 2009.

EC 2009b: *The Euro-Mediterranean Partnership*. http://ec.europa.eu/external_relations/euromed/index_en.htm, downloaded: 27 December 2009

EC 2009c: *Generalised System of Preferences (GSP)*. http://ec.europa.eu/trade/wider-agenda/development/generalised-system-of-preferences/, downloaded: 27 December 2009

Hajdu, O. 2003: *Többváltozós statisztikai számítások*. KSH, Budapest.

Horváth, Z. 2005: *Kézikönyv az Európai Unióról*. HVG-ORAC Lap- és Könyvkiadó Kft., Budapest.

Ketskeméty, L. – Izsó, L. 2005: *Bevezetés az SPSS programrendszerbe*. ELTE Eötvös Kiadó, Budapest.

Kovács, P. – Petres, T. – Tóth, L. 2006: *Válogatott fejezetek statisztikából: többváltozós statisztikai módszerek.* JATEPress, Szeged.

Lengyel, I. 2000: A regionális versenyképességről. *Közgazdasági Szemle*, 12, 962-987. o.

Lukovics, M. 2008: *Térségek versenyképességének mérése*. JATEPress, Szeged.

Nunn, A. – Price, S. 2004: Managing development: EU and African relations through the evolution of the Lomé and Cotonou Agreements. *Historical Materialism*, 4, 203-230. o.

Persson, M. – Wilhelmsson, F. 2006: *Assessing the effects of EU trade preferences for developing countries*. Working Paper, 4, Lund University, Department of Economics, Sweeden.

Sajtos, L. – Mitev, A. 2007: *SPSS kutatási és adatelemzési kézikönyv.* Alinea Kiadó, Budapest.

Somai, M. 1997: A GATT-ból a WTO-ba: az Uruguayi forduló előzményei és eredményei. *Külgazdaság*, 10, 4-20. o.

Stiglitz, J. E. 2003: *A globalizáció és visszásságai*. Napvilág Kiadó, Budapest.

Székelyi, M. – Barna, I. 2005: *Túlélőkészlet az SPSS-hez. Többváltozós elemzési technikákról társadalomkutatók számára*. Typotex, Budapest.

Udvari, B. 2008: A Loméi Egyezmények gazdasági hatásai Fekete-Afrikában. *Kül-Világ*, 3-4, 39-60. o.

UN 2009: *Statistics Division, on-line.* http://unstats.un.org/unsd/snaama/selbasicFast.asp, downloaded: 2 July 2009.

UNCTAD 2007: *Trade and Development Report, 2007*. United Nations Conference on Trade and Development, New York and Geneva.

UNCTAD 2009a: *Handbook of Statistics, on-line*. http://stats.unctad.org/handbook/ReportFolders/ReportFolders.aspx?CS_referer=&CS_ChosenLang=en, downloaded: 18 July 2009.

UNCTAD 2009b: *FDI on-line*. http://stats.unctad.org/fdi/, downloaded: 24 July 2009.

Yanikkaya, H. 2003: Trade openness and economic growth: a cross-country empirical investigation. *Journal of Development Economics*, 1, 57-89. o.

Vargha, A. 2000: *Matematikai statisztika pszichológiai, nyelvészeti és biológiai alkalmazásokkal*. Pólya Kiadó, Budapest.

Appendix

Cluster membership in both of the periods

Country	Cluster (pre)	Distance (pre)	Cluster (after)	Distance (after)
Angola	*1*	2,453	*1*	2,329
Argentina	*2*	2,522	*4*	2,171
Bangladesh	*3*	1,903	*3*	2,372
Barbados	*4*	2,382	*4*	2,522
Belize	*4*	3,17	*4*	1,719
Benin	*1*	2,388	*3*	1,674
Bolivia	*1*	1,663	*4*	1,679
Brazil	*2*	3,459	*2*	2,92
Burkina Faso	*3*	2,395	*3*	1,718
Burundi	*3*	2,166	*3*	1,758
Cameroon	*1*	2,296	*4*	1,998
Cape Verde	*1*	3,128	*4*	2,845
Central African R.	*3*	2,646	*3*	3,637
Chad	*3*	2,027	*1*	3,911
Chile	*1*	2,938	*4*	2,547
Colombia	*1*	1,925	*4*	1,178
Comoros	*3*	2,273	*3*	1,714
Congo	*1*	3,478	*1*	2,587
Costa Rica	*1*	2,835	*4*	1,339
Cote d'Ivoire	*1*	2,055	*4*	2,343
Dem. Rep. Congo	*1*	2,374	*3*	2,608
Ecuador	*1*	2,166	*4*	1,351
Egypt	*1*	2,289	*4*	2,481
El Salvador	*1*	1,811	*4*	2,22
Ethiopia	*3*	2,109	*3*	3,128
Fiji	*1*	2,243	*4*	1,93
Gabon	*1*	4,736	*1*	3,306
Gambia	*3*	3,099	*3*	2,421
Ghana	*3*	1,453	*3*	1,649
Grenada	*1*	2,753	*1*	3,976

Guatemala	*1*	2,191	*4*	1,97
Guinea-Bissau	*3*	2,908	*3*	2,875
Guyana	*1*	2,5	*1*	2,931
Haiti	*3*	1,611	*3*	2,809
Honduras	*1*	2,128	*4*	1,871
Israel	*2*	3,66	*2*	2,962
Jamaica	*4*	2,36	*4*	1,312
Jordan	*1*	3,126	*4*	1,832
Kenya	*3*	1,656	*3*	1,228
Lao P. D. R.	*3*	3,004	*3*	2,241
Lebanon	*1*	3,431	*4*	2,949
Liberia	*4*	3,596	*3*	3,587
Madagascar	*3*	1,815	*3*	1,389
Malawi	*3*	2,051	*3*	1,243
Malaysia	*1*	2,441	*2*	3,134
Mali	*3*	2,882	*3*	1,238
Mauritania	*1*	3,407	*1*	2,886
Mauritius	*1*	2,135	*4*	2,64
Morocco	*1*	2,769	*4*	3,658
Mozambique	*3*	2,828	*3*	1,979
New Caledonia	*1*	3,327	*4*	2,999
Nicaragua	*1*	2,473	*4*	1,811
Niger	*3*	1,609	*3*	1,335
Nigeria	*2*	4,233	*3*	3,255
Pakistan	*3*	1,912	*4*	2,091
Panama	*4*	2,106	*4*	2,524
Papua New Guinea	*1*	3,234	*3*	3,716
Paraguay	*1*	1,625	*4*	1,706
Peru	*1*	1,939	*4*	1,224
Philippines	*1*	3,091	*4*	2,329
Rwanda	*3*	1,157	*3*	1,658
Samoa	*1*	2,465	*4*	2,325
Sao Tome and P.	*1*	3,112	*3*	3,788
Senegal	*1*	2,316	*3*	2,264

Seychelles	*4*	3,172	*1*	4,451
Sierra Leone	*3*	2,46	*3*	3,994
Somalia	*3*	1,631	*3*	3,319
Sri Lanka	*1*	2,429	*4*	1,73
Sudan	*3*	1,773	*3*	2,226
Suriname	*1*	2,829	*1*	2,711
Syrian Arab Rep.	*1*	1,365	*4*	2,922
Thailand	*1*	2,413	*2*	2,673
Togo	*1*	2,881	*3*	1,71
Trinidad & Tobago	*4*	2,523	*4*	2,767
Tunisia	*1*	2,123	*4*	4,405
Uganda	*3*	1,475	*3*	1,145
Tanzania	*3*	1,425	*3*	1,148
Uruguay	*1*	2,511	*4*	1,228
Vanuatu	*1*	2,964	*4*	2,756
Zambia	*1*	3,047	*3*	1,777

Source: own calculation

Economic growth and convergence in the world economies: an econometric analysis

Attila Gaspar[1]

This paper investigates, whether convergence or divergence can be observed among world economies between 1992 and 2008. The catching-up process is calculated for GDP per capita, as this indicator is a relative good quantitative proxy for economic growth. However, the factors behind economic growth were not analysed, only the existence and rate of (actual and expected) catching-up.

Both econometric and purely statistical methods have been applied and I have also created an indicator called omega to analyse the convergence process by a new approach. Convergence can be confirmed by all of the indicators: economies are converging to their own level of steady-states but only very slow catching-up can be measured for low income economies (and also for developed countries). Long-run catching-up process is also expected to take place but this primary applies for middle income countries, which mainly confirms the existence of club convergence.

Keywords: Conditional convergence, club convergence, economic growth

1. Introduction

Economic growth and convergence is one of the most discussed fields in economics as the long-run growth basically determines the welfare of countries. Actually, it is assumed that countries with lower GDP per capita tend to grow faster, than the richer ones. This process is called catching-up. This is of great importance as convergence assumes long-run common (or group-specific / own) steady-state levels.

Developing countries might be able to converge towards high income countries, as they can e.g. adopt new technologies of high income countries. These flows might lead to higher rate of economic growth exceeding the growth rate of developed countries. However, the convergence process of lower income countries is not guaranteed. There are many factors leading to divergence: e.g. high level of debts and net lending.

The analysis of economic growth is widely applied in the literature: both the determinants of economic growth and the rate of convergence are measured in vari-

[1] Attila Gáspár, Statistician, Hungarian Central Statistical Office, Sector Accounts Department, Hungary

ous ways by many researchers. In this paper only the rate of convergence is analysed.

On the basis of economic, statistical and econometric models several studies found empirical evidence of convergence among countries. However, mostly conditional convergence can be confirmed. This means that economies are converging but the steady-state is not common, countries are converging towards different steady-states. Therefore, convergence can be measured when control variables (which determine the steady state) are also integrated into the models (e.g. investment/depreciation rate).

Convergence is especially found among homogenous groups of countries (e.g. EU regions, US States, Australian States, Japanese prefectures, Canadian provinces), while generally divergence can be measured among heterogeneous economies and among world economies on the whole.

A catching-up rate of 2% can be estimated very often in various country groups and countries, which might be interpreted as a general rate of convergence. However, this rate might also be defined as a 'statistical artifact' (Johnson et al. 2004).

In this paper an indicator is also presented (which is based on cluster analysis) to analyse convergence by a new approach. In growth econometrics cluster analysis is primarily applied for determining convergence clubs. Researchers (Hobijn – Franses 2000, Corrado et al. 2004) identified groups of converging countries or pattern of differences among world economies. However, in this paper an indicator is primarily applied to calculate convergence.

In the following sections different approaches of convergence are presented. Both econometric and purely statistical methods are calculated: new and some modified classical approaches are also applied.

2. Methodology

Table 1 shows the main economic indicators of the world economies:

Table 1. Economic outlook (2000-2007)

Country group	Year	GDP (current billion USD)	GDP deflator (annual, %)	Agriculture, (% of GDP)	Industry, (% of GDP)	Services, (% of GDP)	Exports (goods, services; % of GDP)	GCF (% of GDP)	FDI net inflows (current billion USD)
World	2000	31 969	5	4	29	67	25	22	1 519
	2005	45 179	5	3	28	69	27	22	1 117
	2006	48 863	6	3	28	69	28	22	1 457
	2007	54 584	5	N/A	N/A	N/A	N/A	N/A	2 139
OECD	2000	24 727	2	2	28	70	22	22	1 273
	2005	33 506	2	2	26	73	23	21	771
	2006	35 223	2	1	26	73	24	21	993
	2007	38 278	2	N/A	N/A	N/A	N/A	N/A	1 515
Middle income	2000	5 660	7	11	36	54	27	24	155
	2005	9 413	5	10	37	53	33	27	270
	2006	11 103	7	9	38	53	33	28	342
	2007	13 490	6	9	37	53	33	29	495
Low income	2000	371	6	30	24	45	28	20	6
	2005	602	9	27	28	45	31	23	12
	2006	701	9	26	29	45	32	24	21
	2007	801	7	25	30	46	32	25	32

Source: Data of World Bank 2009

Substantial differences can be observed among the country groups. The GDP growth is fast, especially in low income countries but they account for only 1.5% of the world GDP in 2007 (although they represent 20% of the population). As the inflation rate is higher in these economies, the difference of deflated GDPs (constant prices) is more substantial.

It is also important to note that agriculture contributes to the GDP in the highest way in low income countries (the rate is ca. 15-20 times higher compared to OECD countries) but the gross valued added of industry accounts for approximately the same rate of contribution (in relative terms) in all country groups. There is also empirical evidence that the more developed the countries are the more substantial gross value added of services is obtained.

As for the income side of GDP (gross capital formation and export), the differences among country groups are not significant compared to the production ap-

proach. Substantial growth rates can also be observed. The same holds true to FDI inflows, which can be interpreted as an effect of globalisation.

The growth rate of GDP per capita is a good proxy for economic growth, hence the analysis is based on this indicator. Data on PPS GDP per capita (and population) are available from IMF for 171[2] countries for the period from 1992 to 2008.

As there are not any other longer time series available in the database (or some data are missing) convergence is only measured by GDP per capita. Therefore, the factors having a significant impact on economic growth cannot be observed by this approach, only the speed and existence of convergence/divergence (which is also the purpose of this paper). At the same time, aggregated factors can be explained by classifications, dummy and constant variables, so the (average or aggregated) effects of omitted variables are also measured. Hence, the measurement of convergence is expanded for appropriate calculation.

As there are different definitions of convergence (e.g. converging to a rate or level of economy/economies, diminishing disparities), multiple methods have been applied. These indicators of convergence (especially the sigma and the beta) are basically based on theories of economic growth: on endogenous and exogenous models (Sorensen et al. 2005). However, as substantial differences exist among these theories, mostly non-theoretical (statistical) models were applied.

3. Data analysis

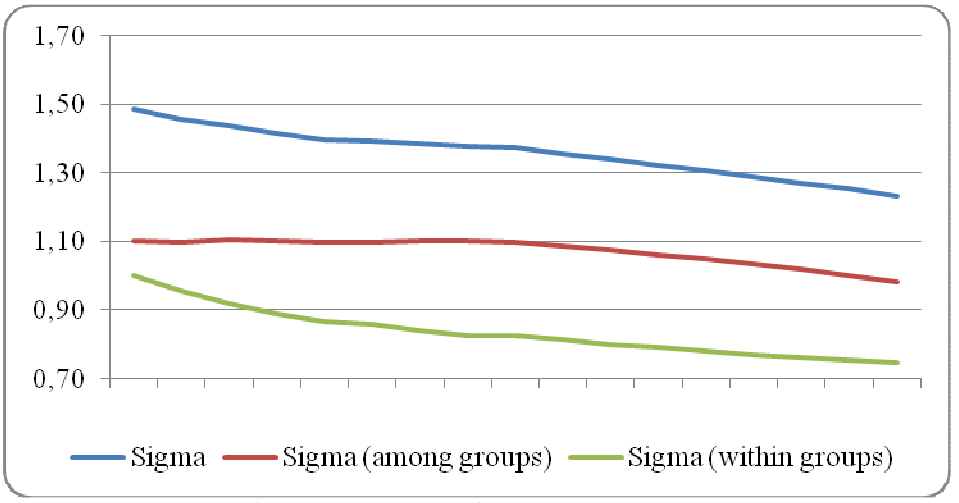
In the following five approaches of convergence are presented:

3.1. Sigma convergence

Chart 1 illustrates the standard deviation of GDP per capita of the world economies (in logarithmic form), which is called sigma (Sala-i Martin 1996b):

[2] Time series data of 11 countries were very short and were not taken into account: Afghanistan, Rep. of.; Bosnia and Herzegovina; Estonia; Georgia; Iraq; Liberia; Montenegro; Serbia; Slovak Republic; Timor-Leste, Dem. Rep. of; Zimbabwe.

Chart 1. Sigma convergence (1992-2008)



Source: own calculations based on data of IMF 2009

The indicator slightly differs from the classic one, as I wanted to modify the sigma in order to take certain factors into account. First, the log values are weighted by the number of population (as these differences are significant among world economies) and second, the sigma is divided into two groups: sigma among and within country groups (which are also weighted) in order to analyse the standard deviation in detail. For calculating the partial standard deviations three groups have been chosen: OECD (members of Organisation for Economic Co-operation and Development), LDC (members of least developed countries defined by the United Nations) and ROW (members of rest of the world countries – non-OECD and non-LDC)[3].

As you can see, *convergence* can be observed: the sigma shows a negative trend, i.e. inequalities were diminishing – approximately at a yearly rate of 1.4 percentage points – among world economies from 1992 to 2008. However, in 1992 the difference between sigma among and within country groups was not significant compared to 2008. This means that disparities among countries of the same group are diminishing quite fast, but disparities among country groups are diminishing at a much smaller rate.

Both tendencies can be accepted, as the convergence of world economies might be rather *conditional*: there are significant differences between countries and group of countries but the convergence might principally apply for countries of simi-

[3] As the time series are not long and there are not any significant changes between the base and the current year, the groups have been determined on the basis of the current year only.

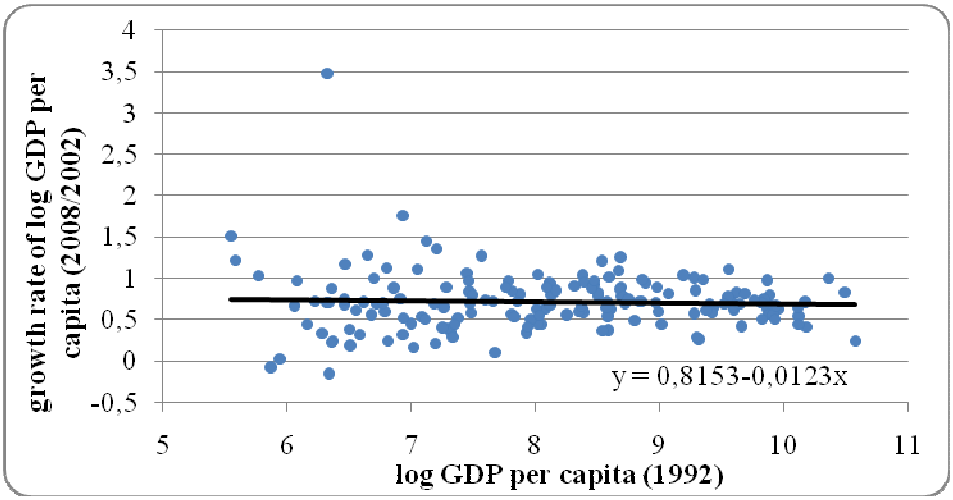
lar level, i.e. countries are converging to their own (group-) level of steady state. In such a way, there is also sign of *club convergence* as economies are not converging towards common, but towards separate steady sates, which may be the same for economies belonging to the same group.

On the basis of sigma I also calculated the number of expected years essential for significant *catching-up*, which is based on ARIMA models[4] (Maddala 2004) of the standard deviations. It can be assumed that natural inequalities will exist in the future as well, hence the comparison of deviations is of greater importance. On the basis of the models, the sigma within and among country groups will be equal in approximately 25-30 years and the total sigma is expected to be under 0.9 till then. Both tendencies are good signs of catching-up, hence strong convergence process is expected to take place, especially in the long-run. However, as past tendencies are actually extrapolated by the models (so long-run effects like trend shifts are not taken into account), the forecasts have to be accepted carefully.

3.2. Beta convergence

Chart 2 shows the log regression of GDP per capita and growth rate called beta (Sala-i Martin 1996a) among world economies:

Chart 2. Beta convergence (1992-2008)



Source: own calculations based on data of IMF 2009

[4] For the sigma among country groups an AR(1) (with constant and trend), for the (total) sigma an ARMA (p=1, q=1; with constant and trend) and for the sigma within country groups a double exponential smoothing model was applied. The (total) sigma and the sigma among country groups are stationary but the sigma within country groups is integrated of order one. No regression errors are detected in the residuals.

Where:

- y = growth rate of log GDP per capita
- x = log GDP per capita.

Convergence can be observed again as the rise of regression curve is negative, i.e. countries with lower GDP per capita tend to grow faster than the richer ones.

Though no autocorrelation and ARCH-effect are detected and the residuals are following normal distribution, the error term is heteroskedastic and one outlier – Equatorial Guinea – may also generate distortions (Ramanathan 2003). In addition, although several low income countries generated fast economic growth, the convergence is far from perfect as the high income countries were growing relative fast, too.

Therefore, the convergence might be rather *conditional* again, so other variables have to be applied for estimating the regression. As there are not any long times series of macroeconomic variables available in the IMF database (or values are missing), dummies are used for classifying the countries into groups mentioned previously.

The secondary calculations showed convergence again and no regression errors are detected:

$$y = 1.12 - 0.07 \cdot x + 0.25 \cdot dum1 + 0.17 \cdot dum2 \quad (1)$$

Where:

- y = growth rate of log GDP per capita
- x = log GDP per capita
- dum1 and dum2 = dummy variables for country groups.

All variables are significant but the outlier should be included for improving the results. The calculation assumes significant club convergence, an almost 7% catch-up rate among economies – exceeding substantially the growth rate confirmed by common empirical results and neoclassical models (2%). At the same time, the regression should be interpreted carefully as the adjusted R^2 is very low (below 5%). Still, there are signs of *club convergence* again.

On the basis of beta the *catching-up* process can be determined. It will take probably more than 100 years for the lowest income country to catch up with highest income country. However, there are very strong assumptions for both high and low income countries and the regression is biased, so the catching-up can only be interpreted as auxiliary indicator and other methods of convergence analysis should be applied.

3.3. Panel modelling

The analysis of inequalities is widely applied in the literature (mostly by panel data), and the modified sigma presented previously is a good proxy for inequalities in my opinion. Therefore, the sigma calculated within countries (which might be of the

greatest importance) have to be divided by the average GDP per capita of the groups and regressed with log GDP per capita:

$$V_{q_w,t} = c + \log(X_{ij,t}) + \varepsilon \tag{2}$$

$$V_{\sigma_{w,t}} = \frac{\sqrt{\sum_{j=1}^{n}\sum_{i=1}^{n_j} f_{ij,t} \cdot \left(\log X_{ij,t} - \log \overline{X_{j,t}}\right)^2}}{\log \overline{X_{j,t}}} \tag{3}$$

Where:

- c: constant variable
- ε: error term
- f: number of population
- x: GDP per capita
- i: country (1, 2… 171)
- j: group (LDC, OECD, ROW)

At the same time, the cross-sectional beta analysis is static, as only the base and the current period are compared. Therefore, panel data have to be applied in order to acquire new information. On the basis of different unit root tests (Levin, Lin, Chu; Breitung; Im, Pesaran, Shin; Fisher ADF and PP; Hadri) there is no presence of unit root in either of the variables.

I applied fixed and random effect models with different weights. The fixed effect model (fixed cross-sections and fixed period dummies) with cross-sectional weights explained the highest variation of inequalities (adjusted $R^2 = 98.1\%$):

$$\sigma_w = -0.27 + 0.04 \cdot x + a_i + b_t + \varepsilon_{it} \Rightarrow E(\varepsilon_i \varepsilon_i' \mid X_i^*) = \sigma^2 I_T \tag{4}$$

Where:

- σ_w: sigma within countries
- x: log GDP per capita
- a_i: cross-sectional fixed effects
- b_t: period fixed effects
- ε_{it}: error term
- i: country (1, 2… 171)
- t: time factor (1992, 1993, ... 2008)

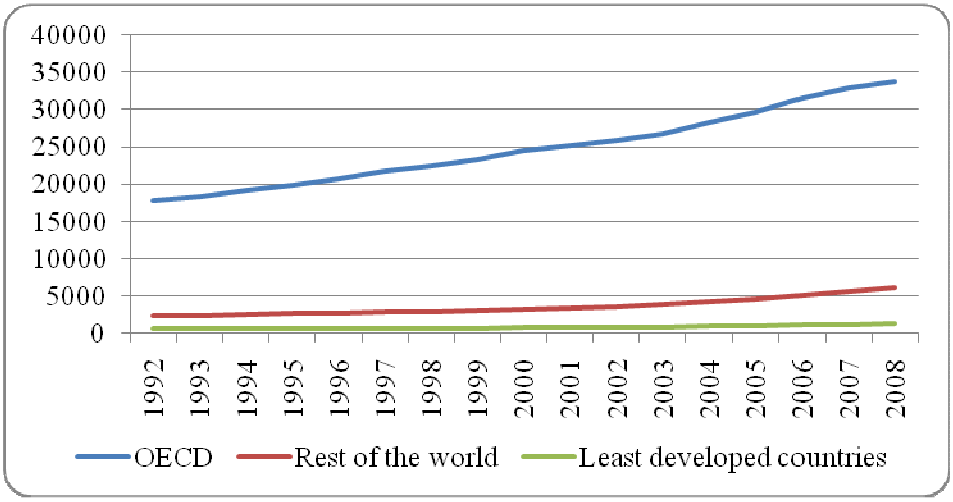
The fixed cross-sectional and periodic effects are tested by F statistics (Kőrösi et al. 1989) and are both significant. However, disparities among countries are much substantial in the whole time interval.

Although the GDP per capita is significant, it represents only 5.2% of the variation (adjusted R^2), i.e. the level of GDP per capita substantially influences the inequalities of economies but other (mostly cross-sectional) effects have a much significant role.

3.4. Cointegration

Chart 3 demonstrates the average log GDP per capita of the country groups:

Chart 3. Average GDP per capita (1992-2008)



Source: own calculations based on data of IMF 2009

Substantial differences can be seen between the countries, and the richer the countries are the faster they tend to grow.

However, some kind of link can be observed between LDC and ROW countries, which may be confirmed by testing the stochastic processes. By different type of unit root tests[5] (DF-GLS, KPSS, PP, ADF, ERS, NG-Perron) all three series are accepted as being integrated of order one, i.e. by differencing the series once stationarity can be achieved. But the linear combination of the series may be stationary, i.e. the series may be cointegrated (Maddala – Kim 1999).

After regressing each of the series by these variables the error terms (unit root, autocorrelation) showed sign of cointegration between the LDC and ROW series. To analyse the processes in the detail, error correction models were applied and cointegration was tested by the Johansen procedure. Cointegration can really be accepted between LDC and ROW series at a 10% significance level and the error correction

[5] Tests of higher power and higher significance levels (e.g. 10-20%) were primarily applied for robust estimation (Maddala – Kim 1999).

model underlines the diminishing disparities between countries. Therefore, long-run relationship between the average GDP per capita of least developed and rest the world countries exist, however, no relationship exist between OECD countries and ROW or LDC.

The expected catching-up process can also be extrapolated on the basis of ARIMA[6] models. There is no sign of convergence between LDC and ROW or LDC and OECD. But *convergence can be found between OECD and ROW*: for ROW countries it will take approximately 35-40 years to catch up with OECD economies on the basis of the models. Hence, inequalities are expected to diminish among high and middle income countries but there is little hope for the low income countries.

On Chart 3, we can see signs of divergence between ROW and LDC in the recent years, which also leads to convergence between ROW and OECD. So it can be assumed that cointegration cannot really be accepted, or more precisely: the cointegration between ROW and OECD will probably be accepted (but not between ROW or LDC) in the future. However, there is not any evidence for this approach but the high significance level of cointegration may also confirm this hypothesis.

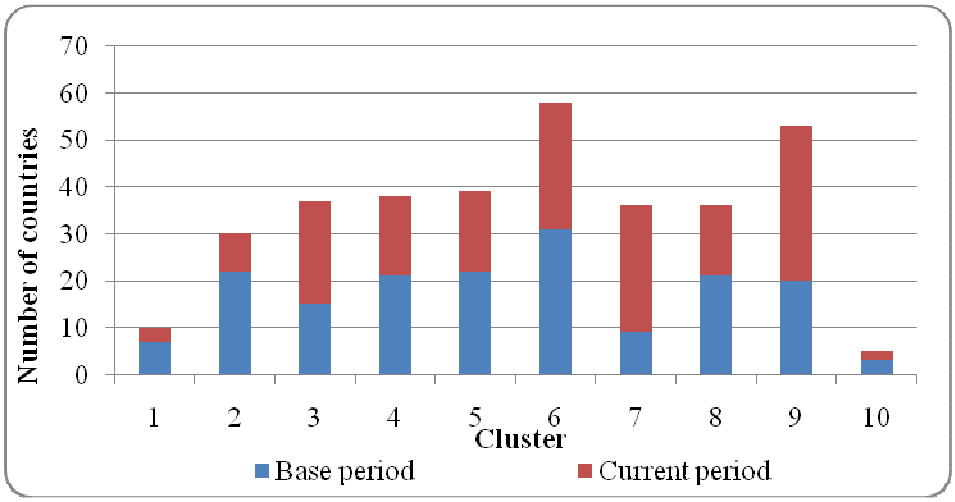
3.5. Omega approach

In the following, a different approach is applied: on the basis of clusters an indicator is presented to calculate convergence. As this type of analysis substantially differs from those presented previously, new information could be acquired by this approach. In such a way, we can take significant changes into account – which might be of great importance, as the other approaches presented previously are based on actual data.

At first, I clustered the GDP per capita of the countries in the base (2002) and in the current period (2008). Chart 4 demonstrates the results of clustering:

[6] For the OECD countries two MA variables (with constant and trend), for the LDC one MA variable (with constant and trend), and for the ROW countries two MA variables (with constant, trend and trend squared) have been applied. All three variables were also differenced once to achieve stationarity. No regression errors are detected in the residuals.

Chart 4. Omega approach (1992-2008)



Source: own calculations based on data of IMF 2009

As you can see, substantial differences can be observed between the base and the current period. Many countries originally clustered into groups of lower income, now probably joined higher income groups. However, we cannot be certain, as we do not know on the basis of this chart, what kind of process each country followed. Therefore, neither the existence, nor the rate of convergence or divergence can be measured by the chart or by the common indicators of descriptive statistics. Therefore, I created an indicator called omega to compare the base and the current period:

$$\omega = \frac{\sum_{j=1}^{n}\sum_{i=1}^{n_J} \alpha \cdot (K_{C\,ji} - K_{B\,ji}) \cdot \left[\left(\frac{f_{jiB} \cdot x_{jiB} + f_{jiC} \cdot x_{jiC}}{f_{jiB} + f_{jiC}}\right) - \left(\frac{f_B \cdot \bar{x}_B + f_C \cdot \bar{x}_C}{f_B + f_C}\right)\right]}{\left|\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\left(\frac{f_{jiB} \cdot x_{jiB} + f_{jiC} \cdot x_{jiC}}{f_{jiB} + f_{jiC}}\right) - \left(\frac{f_B \cdot \bar{x}_B + f_C \cdot \bar{x}_C}{f_B + f_C}\right)\right]\right|} \cdot \alpha \tag{5}$$

Restrictions:

- Optimal number of clusters
- $$\left(\frac{f_{jiB} \cdot x_{jiB} + f_{jiC} \cdot x_{jiC}}{f_{jiB} + f_{jiC}}\right) \neq \left(\frac{f_B \cdot \bar{x}_B + f_C \cdot \bar{x}_C}{f_B + f_C}\right) \tag{6}$$
- If: $(K_{C\,ji} - K_{B\,ji}) = 0$ (7)

Than:

$$\left| \sum_{j=1}^{n} \sum_{i=1}^{n} \left[\left(\frac{f_{jiB} \cdot x_{jiB} + f_{jiC} \cdot x_{jiC}}{f_{jiB} + f_{jiC}} \right) - \left(\frac{f_B \cdot \overline{x}_B + f_C \cdot \overline{x}_C}{f_B + f_C} \right) \right] \right| = 0 \quad (8)$$

Where:

- K: cluster
- C: current period (2008)
- B: base period (1992)
- f: number of population
- x: GDP per capita
- i: country (1, 2, … 171)
- j: group (LDC, OECD, ROW)
- α: other weights (no other weights have been applied in this analysis).

The indicator has been applied, as the less developed the countries are and the higher the rate of economic growth is (measured by the weighted deviation from the average), the more significant the convergence will be. The omega equals to zero, when the economy is clustered into the same group in the base and in the current period. That is essential in my opinion, as I wanted to take only significant changes into account – this idea is behind the indicator.

Table 2 shows the final results of calculations:

Table 2. Omega convergence (1992-2008)

Country group	**ω**	**V_ω** [7]
LDC	-0.98	1.27
ROW	-1.72	1.33
OECD	1.03	2.22
Total	-0.78	-

Source: own calculations based on data of IMF 2009

There is *convergence* among world economies, also within group of countries: while the convergence is *strong among ROW* countries (-1.72), among LDC and OECD countries only low convergence can be measured (-0.98 and 1.03 respectively). For OECD countries divergence is measured, as the (high income) economies either joined higher income clusters or remained in the same cluster for the current period.

Therefore, we can assume that countries with a GDP per capita significantly differing from the average tend to converge lower and leaving a country group may also be quite difficult.

[7] = Omega/total

4. Conclusions

The analysis of convergence is of great importance in economics, as the economic growth basically determines the welfare of countries. Therefore, it is important to see, whether a catching-up process among world economies exist or not.

This paper presented different types of analysis in order to measure the existence and rate of divergence or convergence among world economies. As the database covers only GDP per capita data, the factors behind the economic growth and convergence / divergence cannot be tested. However, the purpose of this paper is to determine the existence and rate of convergence / divergence, hence other variables are only needed to avoid distortions – but factors like constants, dummies ad clusters can be applied to avoid biased estimations.

Though some methods may be biased (especially the beta analysis), each model can be accepted. All of the analysis showed convergence among countries, which basically confirms the neoclassical approach, hence the main assumptions and conclusions of the growth model may apply for world economies. However, endogenous growth models cannot be rejected, either. As the convergence process is far from perfect and not only the level of GDP per capita determines the economic growth (confirmed by the beta and panel analysis), factors like research and development determined in the endogenous models probably influence long-run economic growth, too.

At the same time, the convergence might be rather conditional (which is also confirmed by neoclassical models). Therefore, the convergence process can be measured basically among countries, which are almost at the same income level, or more precisely: countries are converging to their own steady states. However, as for the low income countries, only very slow convergence can be measured. Therefore, club convergence can also be accepted – but absolute convergence can surely be rejected on the basis of these models. This means that middle income countries are especially converging towards each other (i.e. 'poorer' economies are catching up) but they are also converging towards high income economies.

The significant convergence of middle income countries can be interpreted by different ways. The OECD and LDC groups are primary economic categories but these are also influenced by other factors (e.g. political). It is also obvious that middle income economies follow other trends than LDC and OECD countries: they are more similar as they are actually determined by common, similar economic (e.g. GDP growth, inflation rate), political (e.g. democracy) and social factors (e.g. employment rate).

The variation among LDCs and OECD countries is higher as more differences exist among them – also within both groups. So it is very difficult for them to converge: lower income countries of LDCs cannot converge as they probably do not have significant economic driving forces, while higher income countries of OECD countries are hardly to be caught up as they have existing core driving forces

(mostly economic and political) significantly differing from the average and they will have / acquire such factors in the future, as well.

The catching-up process of economies was also forecasted in three ways. Long-run catching-up is expected to take place, at least 30 years are needed for significant convergence. But this primary applies for middle income countries, which mainly confirms the existence of club and conditional convergence.

References

Corrado, L. – Martin, R. – Weeks, M., 2004: *Identifying and interpreting regional convergence clusters across Europe*. Economic journal, University of St. Andrews, St. Andrews, Vol. 115, pp. 133-160.

Hobijn, B. – Franses, P. H., 2000: *Asymptotically perfect and relative convergence of productivity*. Journal of Applied Econometrics, Erasmus University of Rotterdam, Rotterdam. Vol. 15, pp. 59-81.

International Monetary Fund 2009: *World Economic Outlook Database*: http://www.imf.org/external/pubs/ft/weo/2009/01/weodata/index.aspx

Johnson, P. – Durlauf, S. N. – Temple, J. R. W. 2004: *Growth Econometrics*. Vassar College Department of Economics Working Paper Series, Vassar College, New York.

Kőrösi, G. – Mátyás, L. – Székely, I. 1989: *Gyakorlati ökonometria.* Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, Budapest.

Maddala, G. S. – Kim, I. 1999: *Unit roots, cointegration and structural change.* Cambridge University Press, Cambridge.

Maddala, G. S. 2004: *Bevezetés az ökonometriába*. Nemzeti Tankönyvkiadó Rt., Budapest.

Ramanathan, R. 2003: *Bevezetés az ökonometriába alkalmazásokkal*. Panem Könyvkiadó, Budapest.

Sala-I Martin, X. 1996a: *Regional cohesion: Evidence and theories of regional growth and convergence.* European Economic Review 40, pp. 1325-1352. Universitat Pompeu Fabra, Barcelona; Yale University, New Haven, Connecticut.

Sala-I Martin, X. 1996b: *The classical approach to convergence analysis.* The Economic Journal, pp. 1019-1036. Blackwell Publishers, Cambridge.

Sorensen, P.B. – Whitta, H. J. – Jacobsen 2005: *Introducing Advanced Macroeconomics: Growth and Business Cycles.* University of Copenhagen, Copenhagen.

World Bank 2009: *Key Development Data & Statistics, country profiles:* http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20535285~menuPK:1390200~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html

Testing Wagner's Law for Romania

Tudorel Andrei[1] *– Stelian Stancu*[2] *- Andreea Iluzia Iacob*[3] *– Claudiu Herteliu*[4] *– Bogdan Oancea*[5]

This study will test Wagner's law for Romania using both an aggregate measure of public expenditures and a disaggregated form (military expenditures). For the aggregate measure we will estimate the parameters for the models developed by Peacock-Wiseman, Gupta, Goffman, Musgrave and Mann. We will also verify the validity of the Wagner's law in the case of military expenditures for Romania.

Keywords: *Wagner's law, military expenditures, econometric models, cointegration.*

1. Introduction

Country's economic development has allowed a continuous growth of the of public expenditures volume. Economic analysis underlined a series of long-term relationship between economic evolution of a country and the volume of public expenditures. At the end of nineteenth century, Wagner noted that for the industrialized countries there is a faster growth of public expenditures related to the economic growth. A series of empirical studies based on data series at a country level highlighted this relationship also for the period after the Second World War. Among the most important classical studies that have highlighted this issue, we mention: (Gupta 1967), (Pryor 1968), (Goffman 1968), (Musgrave 1969), (Tarschys 1975) etc. New developments in econometrics, as the cointegration techniques (Granger and Engle), the analysis of Granger-type causality etc., have led to new

[1] Tudorel Andrei, PhD, professor, Bucharest Academy of Economic Studies, Department of Statistics and Econometrics, Romania
[2] Stelian Stancu, PhD, professor, Bucharest Academy of Economic Studies, Department of Economic Cybernetics, Romania
[3] Andreea Iluzia Iacob, PhD, professor, Bucharest Academy of Economic Studies, Department of Statistics and Econometrics, Romania
[4] Claudiu Herteliu, PhD, lecturer, Bucharest Academy of Economic Studies, Department of Statistics and Econometrics, Romania
[5] Bogdan Oancea, PhD, professor, Artifex University of Bucharest, Romania

Acknowledgement: This work was supported by CNCSIS –UEFISCSU, project number PNII – IDEI code 1793/2008, financing contract no. 862/2009 and project number PNII – IDEI code 1814/2008, financing contract no 763/2009.

studies to detect the long-term relationships between the public expenditures' dynamics and the economic sector. Among the most important studies, mention (Demirbas 1999), (Murthy 1993), (Oxley 1994), (Ahsan et al. 1996), (Halicioglu 2003), (Akitoby et al. 2006) etc.

Strategies used to test the validity of this law using new econometric methods to estimate parameters are based on regression models or on ECM models. The models are estimated using data series for various economic variables recorded for a single country or for several countries. For the second situation are used panel data.

We must mention that, for certain countries were obtained contradictory results. Moreover, in certain cases, this law was not respected. In (Ahsan et al. 1996) are presented a series of reviews on the factors determining the inconsistency of results and on the differences obtained by various authors: different time periods that define the data series used to estimate the parameters, the econometric techniques used -including here the model's specification, tests used and the influence of the omitted variables from the model etc.

Regardless the econometric model used to verify Wagner's law we must not forget that this one postulates the existence of a long-term relationship between public sector's size and the economic development of a country. Therefore, this dependence is not required to be verified for shorter periods of time.

2. Methodology used

In the economic theory are defined several models to verify the validity of the Wagner's law. To define these models, we consider the following macroeconomic variables: G - Total public expenditures, GDP - Gross Domestic Product, C - Private Consumption, P - Population.

We present a series of models used to verify Wagner's law. In the first category (Peacock-Wiseman and Pryor models) are included regression models based on absolute values of data series. In the second category (Gupta, Goffman, Musgrave and Mann models) are included regression models in which for one or more variables are used per capita values.

The (Peacock-Wiseman 1967) Model proposes a relationship in which the elasticity of public expenditures relative to GDP is constant. The model is defined as follows:

$$\ln G_t = \alpha + \beta \ln GDP_t + \varepsilon_{1t} \quad [1]$$

The (Pryor 1968) Model is based on the analysis of private consumption according to the Gross Domestic Product. In this case we consider that the elasticity of private consumption relative to GDP is constant for analyzed the period. The model is defined as follows:

$$\ln C_t = \alpha + \beta GDP_t + \varepsilon_{2t} \quad [2]$$

The (Gupta 1967) Model proposes a linear dependence between public expenditures per capita and GDP per capita $(GDP/P)_t$:

$$\ln(G/P)_t = \alpha + \beta(GDP/P)_t + \varepsilon_{3t} \quad [3]$$

The (Goffman 1968) Model uses a linear model to explain public expenditures, including GDP per capita $(GNP/P)_t$ as exogenous variable:

$$\ln G_t = \alpha + \beta(GDP/P)_t + \varepsilon_{4t} \quad [4]$$

The (Musgrave 1969) Model is based on the expression of public sector's size at the national economy aggregate (the share of government expenditures in GDP) depending on the evolution of a country's economic development (measured by GDP per capita):

$$\ln(G/GDP)_t = \alpha + \beta(GDP/P)_t + \varepsilon_{5t} \quad [5]$$

The (Mann 1980) Model proposes a linear dependence between the share of government expenditures in GDP and GDP:

$$\ln(G/GDP)_t = \alpha + \beta GDP_t + \varepsilon_{6t}. \quad [6]$$

In all these situations we must check if the two variables involved in the model are cointegrated, respectively whether the following conditions are verified:

1. there are two variables, X_t and Y_t, which are first order integrated (X_t and $Y_t \to I(1)$);
2. the two series are cointegrated if for these ones is defined the following regression model:

$$Y_t = aX_t + u_t, \quad [7]$$

where u_t it's a stationary series. We mention that it is vital that the residuals series must be zero order integrated; otherwise we have a spurious regression. If the series is not a stationary one, then between the model's variables there is no cointegration relationship and the regression model defines a spurious regression. The regression models mentioned above can be estimated. In this case we will verify the validity of Wagner's law in an aggregated form or by replacing government expenditures by their components, meaning that we will verify the validity of the law on components.

In order to verify the validity of Wagner's law we must test the stationarity of the data series included in the regression model. In this sense we can use the following tests: Dickey-Fuller test (Dickey-Fuller 1979, 1981), Augmented Dickey-Fuller test (Dickey-Fuller 1979, 1981), Philips-Perron test etc. In this situation we use the augmented Dickey-Fuller and the testing procedure proposed by (Bourbonnais-Terraza 2008). We consider X_t, a non-stationary series, which is first order integrated, $I(1)$. To test the presence of unit root test we use the following model:

$$\Delta X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 t + \sum_{i=1}^{p} \beta_i \Delta X_{t-i} + \varepsilon_t \qquad [8]$$

In the last relationship was noted by ΔX_t the first order difference of the X_t series. Depending on α_0 and α_2 parameter values we define particular models used for testing the presence of unit root for the series $(x_t)_{t=1,\ldots,T}$. If $\alpha_0 = 0$ and $\alpha_2 = 0$ we obtain the M_1 model without constant and trend. The M_2 model with constant is obtained from the general model if $\alpha_0 \neq 0$ and $\alpha_2 = 0$. In the [8] relationship we define the M_3 model with constant and trend. Testing the unit root presence in the data series is realized by an algorithm in cascade having the form: $M_3 \rightarrow M_2 \rightarrow M_1$. The natural parameter *p* value is determined using Akaike criterion. To test the presence of unit root - in which case the series is non stationary –we define the null hypothesis: $H_0 : \alpha_1 = \alpha_2 = 0$. The rejection of the null hypothesis recommends a data series with deterministic trend rather than one with stochastic trend.

3. The analysis of government expenditures in Romania's case

The validity of Wagner's law in Romania' case is verified using the above methodology. In this case are used data series for GDP and total budgetary expenditures for the period 1985-2000. Indicators are expressed in 1990 constant prices.

The parameters of the five regression models were estimated using OLS. As in all cases there was a significant first order autocorrelation of the residuals we used a first order moving average process, MA (1). Results are shown in Table 1.

Table 1. Peacock-Wiseman, Gupta, Goffman, Musgrave and Mann models for Romania's case

	α	β	t_α	t_β	R^2	LL	F
Peacock-Wiseman	−11.48 (2.04)	2.55 (0.31)	-5.62	8.17	0.90	11.36	52.61
Goffman	−3.7 (1.28)	2.61 (−0.79)	-2.88	6.95	0.88	10.23	44.41
Musgrave	−6.70 (1.21)	1.57 (0.35)	-5.51	4.42	0.75	10.70	17.8
Gupta	(−6.70 1.21)	2.57 (0.35)	-5.51	7.24	0.88	10.67	52.61
Mann	−11.47 (2.08)	1.55 (0.31)	-5.49	4.9	0.77	11.36	20.09

Observations: between parentheses are presented standard errors of the estimators.

Source: own creation

Analyzing the results from Table 1 we observe that Wagner's law is verified for all five models for aggregate budgetary expenditures.

To deepen the analysis of this law are considered two nonlinear models to explain the evolution of total budget expenditures. Thus, for aggregate public expenditures is defined the nonlinear model: $G_t = \omega GDP_t^{\beta}$. For its linearization we use the Box-Cox transformation. The model used to verify the validity of Wagner's law for public expenditures has the following form:

$$\frac{G_t^{\lambda} - 1}{\lambda} = \alpha + \beta \frac{GDP_t^{\lambda} - 1}{\lambda} \qquad [9]$$

where $\lambda \in R^*$. For $\lambda = 0$ we obtain the particular case which defines the Peacock-Wiseman model presented above. For the $\lambda = 0$ case the elasticity is constant: $e = \frac{\partial \ln ME}{\partial \ln GNP}$. In the opposite case, the elasticity is not constant for the entire period. This one is calculated using the following relation:

$$e_t = \beta \left(\frac{ME_t}{GDP_t} \right)^{-\lambda} \qquad [10]$$

To determine the best value of the λ parameter were used various values from the [-1, 1] interval. In Table 2 are presented the results for the $\lambda = 0$ and $\lambda = -0.8$ cases. The parameters are estimated in the situation of non correlated residuals. The value of Durbin-Watson statistics suggested the introduction of a first order autocorrelation.

Table 2. The parameters of total budgetary expenditures model

λ	α	β	t_α	t_β	LL	$-2\log\theta$	R^2	F	MA(1)
0	-11.59 (1.69)	2.57 (0.26)	-7.46	10.91	10.48	-	0.88	99.00	-
MA(1)	-11.48 (2.09)	2.59 (0.32)	-5.62	8.07	11.35	0.12	0.90	52.6	0.28^{*} $(1.08)^{1}$
-0.8	-7.82 (0.99)	7.27 (0.79)	7.9	9.14	72.37	-	0.86	84.89	-
MA(1)	-7.76 (1.30)	7.22 (1.04)	-5.99	6.94	73.1	0.27	0.88	44.07	0.25^{*} $(0.99)^{1}$

Observations: ***** Parameters are not significant for the 5% level of significance.

Source: own creation

Based on the above results we assert the following conclusions:

1. In both cases, β parameter is significantly different from zero. Moreover, the estimators are positive, indicating that the dependence between the two variables is positive;
2. Between the model for which $\text{cov}(\varepsilon_t, \varepsilon_s) = 0$ for $t \neq s$ and the model for which $\varepsilon_t \to MA(1)$ the differences are insignificant. To argue this conclusion was applied the likelihood ratio test, whose statistics, based on the relation $-2\log\theta$, follows a χ^2 distribution with one degree of freedom. For a 5% significance level resulted that: $\chi^2_{0.05} = 3.84 > -2\log\theta = 0.12$.

Finally, we state the following conclusions:

1. In Romania, Wagner law is valid for the aggregate budgetary expenditures;
2. The elasticity of total budgetary expenditures, which is defined by the following nonlinear function, is not constant over time:

$$\varepsilon_{G/GDP} = \frac{\partial \ln G_t}{\partial \ln GDP_t} = 7.22 \cdot (G_t / GDP_t)^{0.8} \qquad [11]$$

3. Throughout the entire period, the average elasticity is over unitary and is equal to 1.33.

4. The analysis of the Wagner law for military expenditures

Using the above mentioned methodology we will check if Wagner's law is verified for a component of public expenditures, respectively military expenditures (CM). The validity of this law is verified for Romania. First, we verify that the series used in the case of the above presented models are first-order integrated- see Table 3.

Table 3. The determination of the integration order for various data series

$\ln(GDP)$	$-0{,}869\ (p=1)$	$-4{,}350^{a}\ (p=1)$	-	I(1)
$\ln(CM)$	$-1{,}120\ (p=0)$	$-4.350^{a}\ (p=2)$	-	I(1)
$\ln(CM/GDP)$	$-2{,}720\ (p=0)$	$-3{,}690^{b}\ (p=1)$	-	I(1)

*Observation*s: Critical values of the integration test for 1%, 5% and 10% significance levels are, respectively, -4,67, - 3,73, - 3,31. a – significant for 1%. b - significant for 5%. c – significant for 10%.
Source: own creation

In Table 4 is presented the model used to represent the ΔCM stationary variable:

Table 4. The model for ΔCM representation

Romania	-0.1060 (−2.872)	-0.1060 (−2.872)	-1.0470 (−4.350)	0.0116 (3.320)	-4.350 (0.02)

Source: own creation

Using a Granger-type test we determine the type of causal relationship between GDP and CM variables. The obtained results from the application of this test are presented in Table 5.

Table 5. The analysis of causality

Romania	Lags	F-statistic	
L_GDP	2	3.04 (0.080)	$GDP \rightarrow CM$
L_CM	2	1.30 (0.307)	No causality

Source: own creation

The above results show that, in the case of Romania, there is a weak causal relationship between the level of GDP and the one of military expenditures. For Romania there is no causal relationship having the $CM \rightarrow GDP$ form.

5. Conclusions

Testing Wagner's law emphasizes some interesting aspects. In the case of all the models, Peacock-Wiseman, Gupta, Goffman, Musgrave and Mann, we observe that Wagner's law is verified for aggregate budgetary expenditures for Romania.

By applying Wagner Law for military expenditures, we also obtained a valid model. Moreover, in this case it resulted that there is a weak causal relationship between GDP and military expenditures. Similar calculus made for Bulgaria (Andrei 2010) show that this causal relationship is not verified for this country.

If we apply this law for education expenditures in Romania's case we obtain a regression model of Musgrave type, which is valid, but with negative slope. Moreover, by applying the Engle-Granger causality test it resulted that there is no causal relationship between GDP and education expenditures.

References

Ahsan, S. M. -Kwan, A. C. C. - Sahni, B. S. 1996: Cointegration and Wagner's Hypothesis: Time Series Evidence for Canada, *Applied Economics*, 28, 1055-1058. p.

Akitoby, B.- Clements, B.- Gupta, S.- Inchauster, G. 2006: Public Spending, Voracity, and Wagner's Law in Developing Countries, *European Journal of Political Economy,* 22, 908-924. p.

Andrei, T. - Bourbonnais, R. 2008: *Econometrie*, Editura Economica, Bucuresti.

Andrei, T. - Teodorescu, D. - Oancea, B. - Stancu, S. 2009: Some Comments about Wagner Law, *Economic Computation and Economic Cybernetics Studies and Research*, 44, 1, 45-62. p.

Andrei, T.- Iacob, A. I. 2003: Observations sur l'applicabilité de la loi Wagner, *XXXV èmes Journées de Statistique, 2-6 juin 2003, Lyon, France*, Société Française de Statistique.

Balls, A. 1999: Bombing Costs Present Allies with Hefty Bill, *Financial Times,* 9 avril.

Bardhan, P. 1997: Corruption and Development: A Review of Issues, *Journal of Economic Literature,* 35, 3, 1320-1346. p.

Barro, R. J. 1991: Economic Growth in a Cross-Section of Countries, *Quarterly Journal of Economics*, 106, 2, 407-443. p.

Bourbonnais, R. -Terraza, M. 2008: *Analyse des séries temporelles-applications à l'économie et à la gestion*, Dunod, Paris.

Chang, T. 2002: An Econometric Test of Wagner's Law for Six Countries**.** Based in Cointegration and Error**-**Correction Modelling Techniques, *Applied Economics*, 34, 9, 1157 – 1169. p.

Chletsos, M.- Kollias, Ch. 1997: Testing Wagner's Law Using Disaggregated Public Expenditure Data in the Case of Greece**:** 1958-93*, Applied Economics,* 29, 3, 371-377. p.

Demirbas, S. 1999: Co-integration Analysis-Causality Testing and Wagner's Law: The Case of Turkey, 1950-1990, *Discussion Papers in Economics series*, 3, Department of Economics, University of Leicester, UK. p.

Dickey, D. A. - Fuller, W. A. 1979: Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* 74, 366, 427-431. p.

Dickey, D. A. - Fuller W. A. 1981: Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica*, 49, 1057-1071. p.

Gemmell, N. 1993: Wagner's law and Musgrave's hypothesis. In: Gemmell, N. (Ed.): *The Growth of the Public Sector*. Edward Elgar, Cheltenham, 103–120. p.

Goffman, I. J. 1968: On the Empirical Testing of Wagner's Law: A Technical Note, *Public Finance,* 23, 3, 359–364. p.

Goffman, I. J. - Mahar, D. J. 1971: The Growth of Public Expenditures in Selected Developing Nations: Six Caribbean Countries, *Public Finance* 26, 1, 57–74. p.

Gupta S. P. 1967: Public Expenditure and Economic Growth: A Time Series Analysis, *Public Finance*, 22, 4, 423–461. p.

Halicioglu F. 2003: Testing Wagner's Law for Turkey, 1960-2000*, Review of Middle East Economics and Finance*, 1, 2, 129-140. o.

Hayo, B. 1994: No Further Evidence on Wagner's Law for Mexico. *Public Finance*, 49, 287–294. p.

Kolluri, B. R. - Panik, M. J. -Wahab, M. S. 2000: Government Expenditure and Economic Growth: Evidence from G7 Countries, *Applied Economics*, 32, 1059–1068. o.

Lin, C. 1995: More evidence on Wagner's Law for Mexico, *Public Finance*, 50, 2, 267– 277. o.

Mann, A. J. 1980: Wagner's Law: An Econometric Test for Mexico,. 1925-76, *National Tax Journal*, 33, 2, 189-201. o.

Murthy, N. R. V. 1993: Further Evidence of Wagner's Law for Mexico: An Application of Cointegration Analysis, *Public Finance*, 48, 1, 92-96. o.

Musgrave, R. A. 1969: *Fiscal Systems,* Yale University Press, New Haven and London.

Oxley, L. 1994: Cointegration, Causality and Wagner's Law: A Test for Britain 1870–1913, *Scottish Journal of Political Economy,* 41, 3, 286 – 298. o.

Peacock, A. T. - Wiseman, J. 1967: *The growth of public expenditure in the United Kingdom*. 2nd edition, George Allen and Unwin, London.

Peacock, A. T. – Scott, A. 2000: The Curious Attraction of Wagner's Law, *Public Choice,* 102, 1– 17. o.

Perotti, R. 1996: Redistribution and Non-Consumption Smoothing in an Open Economy, *Review of Economic Studies,* 63, 411– 433. o.

Pryor F. L. 1968: *Public Expenditure in Communist and Capitalist Nations*, George Allen and Unwind, London.

Ram, R. 1986: Causality between Income and Government Expenditure: A Broad International Perspective, *Public Finance*, 41, 3, 393-413. o.

Tarschys, D. 1975: The Growth of Public Expenditures: Nine Modes of Explanation, *Scandinavian Political Studies*, 10, 9-31. o.

New goods margin in international trade: empirical analysis for Visegrad countries

Miroslav Hloušek[1]

This paper deals with empirical analysis of international trade between Visegrad countries and EU-15 during past two decades. The goal of the paper is to find out if the growth in export is of intensive or extensive type. I follow methodology of Kehoe and Ruhl (2002) and use detailed trade statistics on the value of trade flows by commodity according to Standard International Trade Classification (SITC) codes. I find out that the goods that were traded the least in the benchmark year account for disproportionate share in trade after liberalization and reduction of trade barriers. The most significant increase was found for Poland and the Czech Republic. The set of goods which accounted for only ten percent of trade in 1993 accounts for nearly thirty percent of trade following the liberalization. Similar patterns were identified also for Hungary and Slovakia even it the analyzed period was shorter. The countries thus began to export goods that they had not been previously trading.

Keywords: international trade, trade barriers, liberalization, export growth, intensive and extensive margin

1. Introduction

This paper analyses international trade between Visegrad countries and EU-15. The goal of the paper is to find out if the growth in export is of intensive or extensive type. We follow methodology of Kehoe and Ruhl (2002) and use detailed trade statistics on the value of trade flows by commodity according to Standard International Trade Classification (SITC). The results show that the goods that were traded the least in the benchmark year account for disproportionate share in trade after liberalization and reduction of trade barriers. In other words the countries began to export goods that they had not been previously trading. This is growth on the extensive margin and should be reflected in models of international trade and also in decisions

[1] Miroslav Hloušek, Faculty of Economics and Administration, Masaryk University, Brno, Czech Republic

Acknowledgement: This work is supported by MŠMT project Research centers 1M0524. Many thanks to unknown referee for useful suggestion

of policymakers that usually focuses only on supporting of traditional export industries.

The rest of the paper is organized as follows. Section 2 briefly describes data used for analysis and section 3 focuses on measures of extensive margin growth. The results are presented and discussed in Section 4. Section 5 deals with sensitivity analysis of the measure of extensive margin and its influence on results. Final section concludes with prospects for further research.

2. Data

The data are obtained from OECD database. The measure is annual flow values of exports of particular country into EU-15. The data are disaggregated by commodities according to four-digit Standard International Trade Classification (SITC), revision 2. The data are quoted in thousands of U.S. dollars, however we are interested in relative quantities and thus the units are not important. The data sample for each country is determined by availability of data. The data sample starts in 1993 for Poland and for the Czech Republic and in 1997 for Hungary and Slovakia.[2] The last available period is 2006 for all countries. It must be mentioned that the process of trade liberalization was gradual and in some cases already started before period that we study. However, main reduction of trade barriers happened before accessions of Visegrad countries to European Union in 2004 which is covered in our analysis.

3. Measures of the extensive margin

For each country's exports into EU15, the SITC codes are ordered by their value of trade in the first three years of the sample.[3] Then the cumulative sets of the ordered codes are constructed. Every set includes one-tenth of total export. The first set starts with the smallest codes[4] and other codes are added to this set until the sum of their values reaches one-tenth of total export value. The next set is formed in similar way by summing the remaining smallest codes until the value of the set reaches one-tenth of total trade. This procedure produces ten sets of codes where each set represents one-tenth of total trade. The first set consists of the "least-traded" commodities – they have the smallest export value. Subsequent set contains less codes then the previous set as the (relative) trade value of the codes increases. Since the set comprises

[2] The data for Hungary are available from 1993 but methodological changes occured, especially between years 1996 and 1997, and therefore the analyzed period starts in 1997.
[3] Average value in three years is used for the sake of robustness of the ordering. Implications of the way of ordering (one year or three years average) are discussed in Section 5.
[4] The term 'smallest codes' means codes that account for the smallest value of total trade.

exactly one-tenth of total export, some codes are split into two sets. Therefore the number of codes in the set need not be integer number.

Given this partitioning of the SITC codes, two measures of export growth are considered. First measure corresponds to change in each set's share of trade over the sample period, second measure focuses on the time evolution of the least-traded set of codes with the aim to capture timing of the export growth of these goods.

The first measure is constructed by calculating the share of total exports for each of the ten sets of codes in the last year of the sample period. The interpretation of this measure is as follows. If the growth in trade is driven only by proportional increase in the value of goods already traded, each set of codes would retain its one-tenth share in trade. On the other hand, if the trade liberalization leads only to trade of goods that were previously untraded, the first set of codes would gain trade share, while share of other set would decline. The first case is intensive margin, the second case is extensive margin in trade growth.

The second measure uses the same partition of SITC codes but looks only on the share of least-traded goods in total export. This share is calculated for each year of the sample period. If the lowering of trade barriers leads to trade of goods not previously traded, there should be an increase in the share of trade accounted for by this set of goods. This measure should show the timing of any change it the trade of new goods. If an increase in the share of exports coincides with the implementation of trade reforms, we can think of it as an evidence of the link between trade liberalization and growth in the extensive margin.

4. Results

Overall results indicate significant export growth on the extensive margin for all Visegrad countries. Table 1 shows the end of sample export shares of the least-traded goods. The most significant increase in the extensive margin is observed in Poland and the Czech Republic. Least-traded goods comprise 28 % and 27 % of the total export share in 2006, respectively. The shares of least-traded goods in Hungary and Slovakia account for 23 and 19 percent, respectively. These lower values can be caused by shorter sample period, but even so the increase is quite considerable.

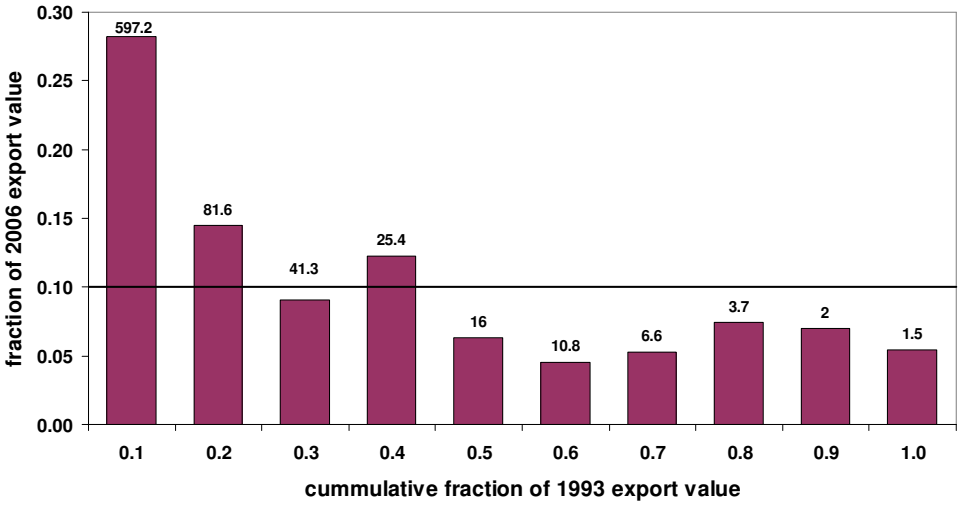
Table 1. Share of export value in 2006: least-traded goods in benchmark year

Country	Benchmark year	Share of least-traded goods
Poland	1993	0.28
Czech Republic	1993	0.27
Hungary	1997	0.23
Slovakia	1997	0.19

Source: own creation

Let us look at the trade share in individual countries one by one. The highest extensive margin growth was recorded for Poland. Graph 1 shows the decomposition of trade into individual sets of goods for this country. The set of least-traded goods that account for 28 % trade share in 2006 includes 597 SITC codes. Given the detailed structure of the measure, we can find which codes experienced the most significant changes. Two codes "7132: Internal combustion piston engines for propelling vehicles" (in the first decile) and "7611: Television receivers, colour" (in the second decile) increased their share on total export value by about 4 %. There is also more than proportional increase in the fourth decile where code "7849: Other parts & accessories of motor vehicles" increased its share by 5 percents.

Graph 1. Composition of Exports: Poland into EU 15



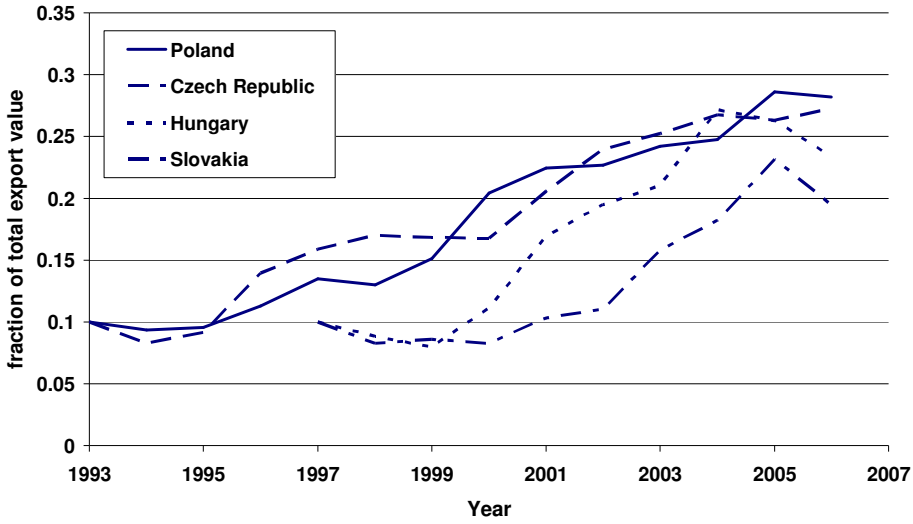
Source: own creation

On the other hand, bunch of categories related to textile products show about one percentage decrease of their share and they are quoted in the fifth, sixth and seventh decile. One category "3222: Other coal, not pulverized, not agglomerated" that can be found in the last set – the most traded goods – decreased its share by 6 percent. It is positive news that Poland turn away from exporting articles of small value added to more processed ones.

It is also interesting to look at the timing of the trade change. Graph 2 plots the trade share of the least-traded goods over the sample period for all countries. There is small decrease of the share at the beginning of every time series. This is caused by the method sorting of the codes (average of first three years) and not much of atten-

tion should be paid to this.[5] In case of Poland (solid line), the increase in the trade share is gradual, with short stagnation in 1997 followed by sharp increase between 1999 and 2000. Are there any events of trade liberalization that are connected to these changes? Lot of custom duties between EU and Poland was abolished in years 1994, 1997 and 2001, but no sharp increase is observed in the time series following these years. Even the opposite is the case – the time series is flatter after these years. Thus, the coincidence of trade barriers reduction with increase of the share of least traded goods is not completely confirmed. However, if we treat the period as a whole, we can conclude that the trade growth on extensive margin has certainly some sources in trade liberalization.

Graph 2. Time evolution of least traded goods: Visegrad countries



Source: own creation

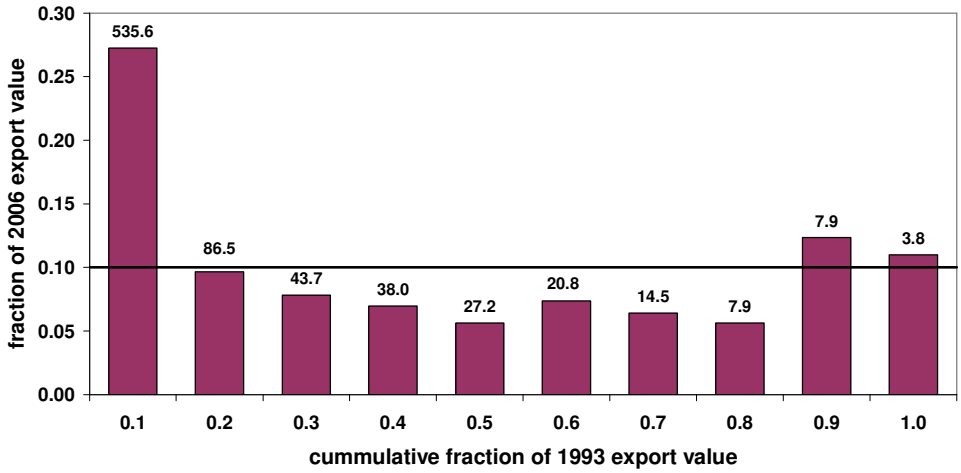
Graph 3 shows trade decomposition in the Czech Republic. The set of least-traded goods includes 535 codes and comprise 27 % of export share in 2006. There are not many single codes (among least-traded goods) that would achieve large share over the analyzed period.[6] However, it is worth to note that goods in two last deciles (most-traded goods) account for more than proportional increase in their sets. No wonder that these sets include codes "7810: Passenger motor cars for transport of

[5] Sensitivity analysis in Section 5 deals with this issue.

[6] The largest increase – by 4 % – was recorded for code „7523: Complete digital central processing units".

pass. & goods“ and „7849: Other parts & accessories of motor vehicles“ – traditional branches of Czech industry.[7]

Graph 3. Composition of Exports: Czech Republic into EU 15

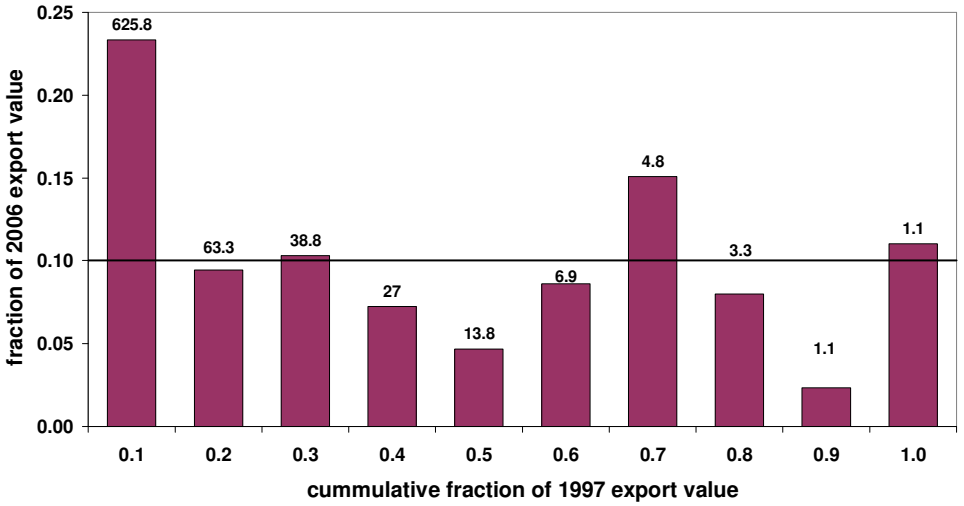


Source: own creation

For time evolution of least-traded goods in the Czech Republic we look again on Graph 2. Compared to Poland it is only slightly different. Regarding trade liberalization, the free trade agreement between EU and Czechoslovakia was arranged in 1992.[8] EU states abolished the customs and quantity limits unilaterally as the agreement came into force. Since we have data from 1993, the effect of this liberalization is partly distorted. We can see two waves (accruals) in the share of least-traded goods: one from 1995 to 1998 and the second from 2000 to 2004. But it is hard to find some concrete causes.

[7] These two codes increased their share by 5 % and 7 %, respectively.
[8] After splitting of Czechoslovakia it was automatically related to the Czech Republic.

Graph 4. Composition of Exports: Hungary into EU 15



Source: own creation

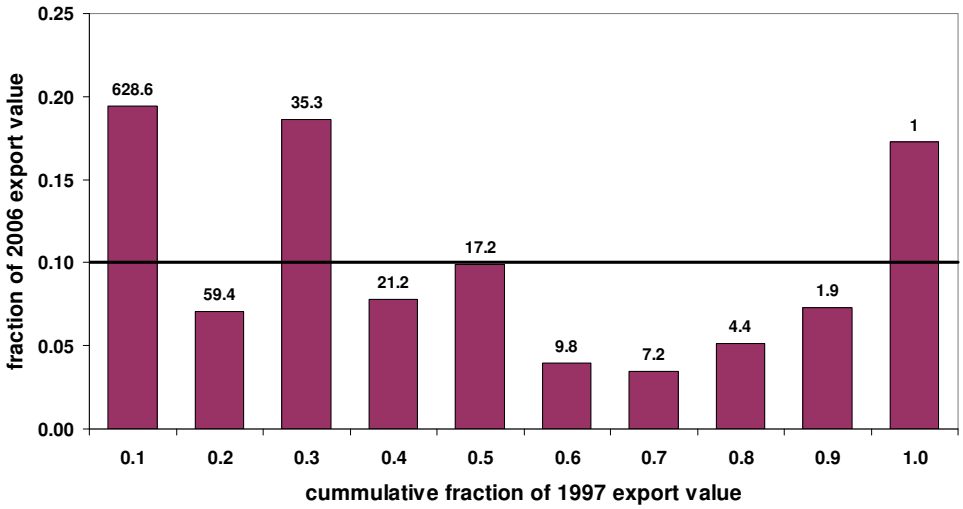
Trade in other two Visegrad countries – Hungary and Slovakia – were analyzed for shorter time period. In spite of this fact, the extensive margin is quite considerable. Trade decomposition for Hungary is shown in Graph 4. The set of least traded goods comprise 23 % of total export value. Among 564 codes from this group the only code "4236: Sunflower seed oil" achieved significant increase (by 6 %). Quite significant change happened also in the seventh decile. Its share is 15.1 % and it includes codes "8749: Parts and accessories for machines, appliances, instruments and apparatus n.e.s." and „6289: Other articles of rubber, n.e.s." that both contributed to increase of the share by 3 %. On the other hand, very significant decrease is observed in the ninth decile. The code "6931: Stranded wire, cables, cordages and the like" decreased by 8 % and is major source of this change.

For time evolution of least traded goods look again at Graph 2. As the analyzed period for these two countries is quite short, it is even more difficult to match reduction of trade barriers to changes in the share of least traded goods. Main steps to trade liberalization have been already done before 1997. Nevertheless, in case of Hungary, the group of least traded goods experienced steady increase from 1999 up to 2004 where it reached the share of 27 % comparable with Poland and the Czech Republic. Then the share little bit decreased to value 23 % as reported above.[9] The overall increase is more rapid and comes earlier than in case of Slovakia.

[9] There was another methodological change in Hungarian data in 2004; therefore the decrease can be ascribed to this fact. For further details about methodological changes, see:

The composition of trade in Slovakia is shown in Graph 5. It seems that Slovakia is country with growth in trade on both extensive and intensive margins. The set of least traded goods has 19 % share on total exports. The most important code is "7842: Bodies for the motor vehicles of 722/781/782/783" that recorded increase by 3 %. There is also significant increase in trade of goods included in the third decile. Share of this set is 18.6 % which is almost the same as of the first set. The code „7611: Television receivers, colour" with 12 % increase derived this change. And finally, the last decile (the most traded goods) has 17.3 % share on total trade. This increase (by 7 %) was caused by single code „7810: Passenger motor cars, for transport of pass. & goods". Intensive margin is therefore also significant source of export growth for Slovakia. From the detailed analysis of codes we see that car manufacturing industry had positive effect for related industry that became also well established. However, the extensive margin growth is still considerable which will be clearer in the next section that deals with sensitivity analysis. The time evolution of least traded goods exhibits flat pattern with gradual increase from 2002. It could be caused by some government export incentives than by reducing of trade barriers because the process of trade liberalization was nearly completed.

Graph 5. Composition of Exports: Slovakia into EU 15



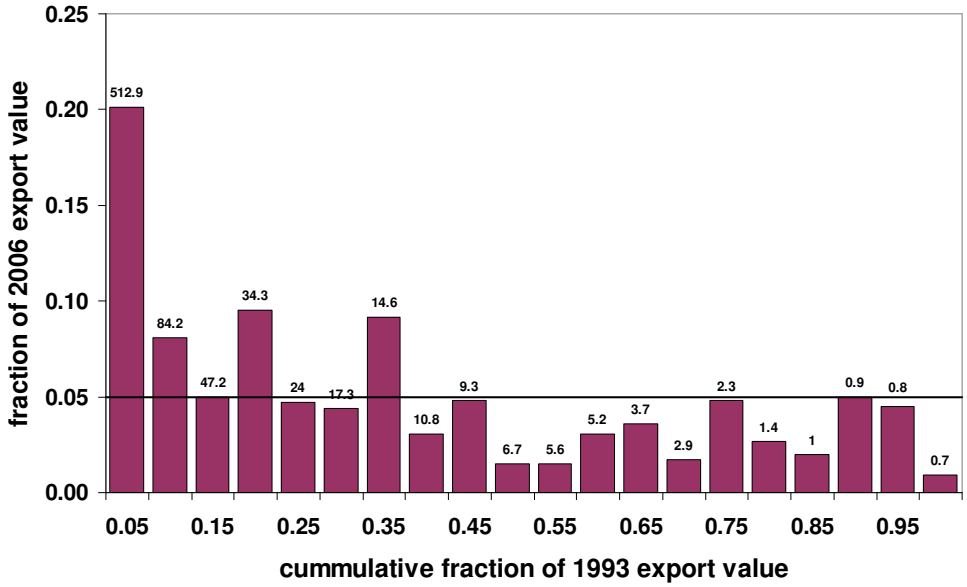
Source: own creation

http://portal.ksh.hu/pls/ksh/ksh_web.meta.objektum?p_lang=EN&p_menu_id=110&p_ot_id=100&p_obj_id=QKT&p_session_id=50231930

5. Sensitivity analysis

This section focuses on sensitivity analysis of our empirical measures. First, we check how the results depend on the choice of cutoff level. As an illustration, composition of trade for Poland using 5% cutoff is shown in Graph 6. This procedure produces twenty sets of good. Increase of least traded goods is apparent and corresponds to the previous pictures. Table 2 reports the extensive margin growth rates for 5%, 10% and 20% cutoffs for all Visegrad countries. Each column shows percentage growth rate of the least traded goods between benchmark year and year 2006. All countries exhibit the same pattern. If we consider smaller cutoff, the least-traded goods grow more. It again support the idea, that goods with very small trade shares drive the extensive margin growth and our measure that uses 10% cutoff can even underestimate size of the extensive margin growth. The measure calculated with 20% cutoff shows smaller growth and it is roughly one half of the measure using 10% cutoff. Even if larger cutoff makes the set of least-traded goods quite big, the increase is still significant.

Graph 6. Composition of Exports: Poland into EU 15 (20% cutoff)



Source: own creation

Table 2. Results under different cutoff values

		Cutoff		
Country	**Benchmark year**	**5%**	**10%**	**20%**
Poland	1993	302.1	182.1	113.4
Czech Republic	1993	267.4	172.4	84.5
Hungary	1997	256.3	133.3	63.8
Slovakia	1997	151.2	94.5	32.7

Source: own creation

Second sensitivity check deals with ordering of the goods at the beginning of the sample period. Primary analysis made in the paper used the average of export value in first three years. The goods were then ordered according to this measure and the first year was chosen to calculate the deciles. Alternative way is to order the goods according to their value in the first year. The least-traded goods thus contain little different set of codes (usually more of them). The results of this procedure are reported in Table 3.

Table 3. Share of export value in 2006: least-traded goods in benchmark year

Country	**Benchmark year**	**Share of least-traded goods**
Poland	1993	0.34
Czech Republic	1993	0.32
Hungary	1997	0.24
Slovakia	1997	0.36

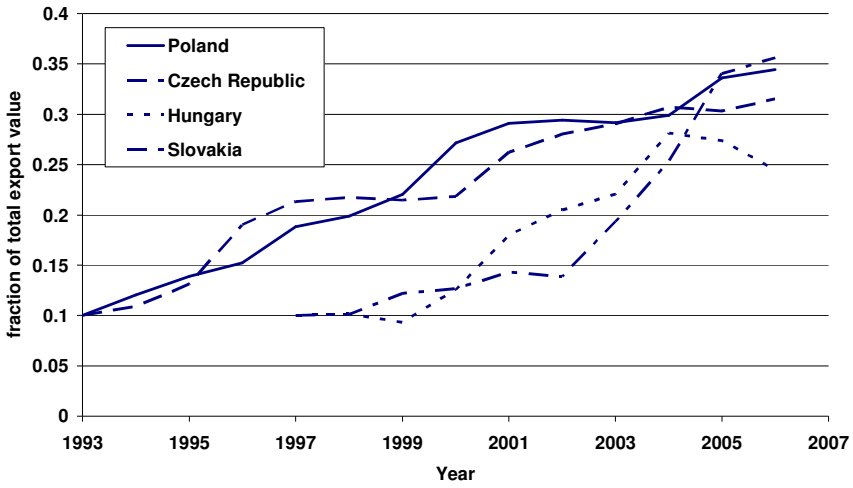
Source: own creation

The extensive margin growth is higher for all countries and thus our analysis understated the importance of this growth.[10] The largest difference is for Slovakia. Contrary to previous measure, the least-traded goods now comprise 36 % of total export value. It is even more than in Poland and the Czech Republic. Detailed analysis reveals that code „7611: Television receivers, color“ is responsible for this change.[11] Time evolution of least-traded goods for Visegrad states using alternative ordering is shown in Graph 4.

[10] Compare to Table 1.

[11] This code was included in the third decile when previous ordering was used (according to the average of first three years).

Graph 7. Time evolution of least traded goods: Visegrad countries



Source: own creation

What is different? We see that the share is growing already from the beginning – there is no initial decrease in first years that was present in Graph 2. Time series of Hungary is almost unchanged. Main trends for other countries are very similar to previous measure except of Slovakia. Here, the share of least-traded goods exhibits rapid growth from 2002 that finally overtook the share in other countries, as discussed above.

To summarize it, using smaller cutoff implies more significant extensive margin growth. Similarly, ordering of the goods according to the first year export value (instead of the average of three years) increases the importance of new goods margin in international trade.

6. Conclusion

This paper analyzed international trade between Visegrad countries and EU15. Results show that there is clear evidence of extensive margin in trade growth. Countries are exporting goods that they had not been previously trading. It is sometimes hard to find if trade liberalization was the main cause of this change because reduction of trade barriers was gradual. This issue deserves more attention. Foreign direct investment can be also important determinant of observed changes and will be subject of further study.

The implications for economic theory are clearcut. International trade models should focus on modeling of the extensive margin. The modified Ricardian model

used in Kehoe and Ruhl (2002) is one of the examples. Calibration of this model using data on intra-industry trade is also topic for further research.

References

Disaggregated Trade data. U.S.-China and China-U.S. Available at: http://www.econ.umn.edu/~tkehoe/classes/8401-07.html, [cit. 2009-05-12].

ITCS International Trade by Commodity Database – SITC, Revision 2. Available at: http://www.sourceoecd.org

Kehoe, T. J. 2002: *An Evaluation of the Performance of Applied General Equilibrium Models of the Impact of NAFTA*, Federal Reserve Bank of Minneapolis. Available at: http://www.econ.umn.edu/~tkehoe/papers/NAFTAevaluation.pdf

Kehoe, T. J. - Ruhl, K.J. 2002: *How Important is the New Goods Margin in International Trade?*Available at: http://www.econ.umn.edu/~tkehoe/papers/Kim.pdf

Regional Trading Agreements. Available at: http://intl.econ.cuhk.edu.hk/rta/index.php?did=14, [cit. 2009-05-12].

Intra-Industry Trade and Revealed Comparative Advantage: Empirical Analysis of New Members State`s Economic Competitiveness (EU-8) on a Single Market between 2000 and 2007

Edward Molendowski[1] *– Wojciech Polan*[2]

This article analyzes changes in the intensity of intra-industry trade between the new Member States (EU-8) and the EU-15 in the period of 2000-2007. The authors claimed that the results should help to assess adaptation processes which had taken place in these countries before and after the accession to the EU. By analysis of revealed comparative advantage and horizontal and vertical intra-industry trade further research examines the real adjustment processes in terms of EU-8 state economic competitiveness, which occurred during the pre- and post-accession period in order to prepare them to maximize the benefits of the Single European Market. Shares of intra-industry trade, and their dynamics was calculated on the basis of the Grubel and Lloyd`s intra-industry trade (IIT) index. Analysis of the competitiveness of goods from EU-8 in a trade with the EU-15 was based on the Balassa`s revealed comparative advantage index, RCA. Three trade types (one-way, two-way trade in similar products, two-way trade in vertically differentiated products) indictors were calculated using the Freudenberg and Fontagné methodology.

Keywords: intra-industry trade comparative advantage, EU enlargement, state economic competitiveness

1. Introduction

Intra-EU trade is based mainly on intra-industry specialization – which basically rely on factor substitutability rather than on factor complementarity. Moreover, intra-industry trade is a dominant form of exchange on the Single European Market (SEM) and the new Member States of the EU-8 group continue to show relatively low level of GDP per capita and the technological gap in relation to the EU-15 countries. Therefore, it seems vital to observe the adaptation processes in the EU-8 countries to the SEM in the area of intra-industry division of labor intensification.

[1] Edward Molendowski, PhD Hab., Cracow University of Economics, Faculty of Economics and International Relations Chair of International Economic Relations (Cracow)
[2] Wojciech Polan, PhD Student, Cracow University of Economics, Faculty of Economics and International Relations Chair of International Economic Relations (Cracow)

That is particularly important because of proved relation between the intensity of the share that a country or a group of countries take in the international division of labor – measured by the level of intra-industry trade and the efficiency of production factors allocation (Pluciński 2001; Zielińska-Głębocka 1996).

In this respect it seems important to examine to what extent trade flows between the new EU Member States (EU-8) and the EU-15 countries were the result of the development of intra-industry specialization. The share of intra-industry trade (IIT) and revealed comparative advantage (RCA) indicators reflect changes in the level of the competitiveness of transition economies, as well as the status of countries` adaptation to the global market in terms of intra-industry division of labor (Misala 2007). This article analyzes changes of trade structure between the EU-8 and the EU-15 in the period of 2000-2007 on the basis of those indicators to assess real adjustment processes in terms of EU-8 state economic competitiveness, which occurred during the pre- and post-accession period in order to prepare them to maximize the benefits of the SEM.

2. Methodology

The analysis based on the IIT and RCA indicators was conducted at 6-digit CN codes level[3] (Cieślik 2003) for total trade and by 5 groups reflecting the intensity of using production factors (1. material-intensive, 2. labor-intensive, 3. capital-intensive, 4. technology-intensive goods, easy to imitate, 5. technology-intensive goods, difficult to imitate)[4] (Misala 1992; Wysokińska 2001). The analysis of EU-8 and EU-15 trade specialization by three trade types (one-way, two-way trade in similar products - HIIT, two-way trade in vertically differentiated products - VIIT) was based on indicators calculated at 8-digit CN codes level due to requirement of product unit value calculation[5] (Śledziewska-Kołodziejska 1998; Michałek-Śledziewska-Kołodziejska 2000).

Shares of intra-industry trade, and their dynamics was calculated on the basis of the Grubel and Lloyd`s intra-industry trade (IIT) index. It shows the share of

[3] This corresponds to the theoretical concept of the industry, and meets the condition that only the highest level of data disaggregation allows the comparison of the same products and is able to reflect the actual degree of overlap between exports and imports, i.e. the real level of intensity of intra-industry trade. Since the primal objective of this study is an analysis of indicators` changes in time, the level of aggregation does not play the most significant role.

[4] This analysis was based on a study of the structure of international trade with emphasis on the intensity of using production factors proposed by: J. Misala and Z. Wysokińska.

[5] The researchers of horizontal and vertical intra-trade underline that it should be analyzed at the level of the product as the best guarantee of avoiding problems associated with empirical aggregation of sectoral data is the sufficiently detailed data.

intra-industry exchange in the total trade taking place within the industry i of the country j, i.e.:

$$IIT_{ij} = \frac{(X_{ij} + M_{ij}) - \left|X_{ij} - M_{ij}\right|}{(X_{ij} + M_{ij})}.$$

where Xij and Mij are respectively the value of exports and imports for the industry i in the trade with the country j.

This indicator is a relative measure receiving values from the interval [0, 1]. If IITij = l, then it is assumed that all trade is of intra-industry characteristics, i.e., Xij = Mij. If, however, IITij = 0, exports and imports do not overlap each other within the industry i, which means that intra-industry trade do not occur, i.e. either Xij = 0 or Mij = 0 (Cieslik-Śledziewska 2003; Czarny 2002; Misala 1985; Misala-Pluciński 2000; Molendowski 2006; Molendowski 2007).

The analysis of the competitiveness of goods from EU-8 in a trade with the EU-15 was based on the Balassa`s revealed comparative advantage index, RCA. It allows to simultaneously take into account the position of the goods from the analyzed countries and the presence of competitors from other countries in a given market. This study was to show for which commodity groups the EU-8 countries have an advantage in exports to the EU-15.

RCA index is defined here as the advantage of the export share of the good n from the county j in the world exports of this good on the EU-15 market above the share in total exports of the country j above the total world exports to the EU-15. The following formula was used for the calculation[6] (Balassa 1965):

$$RCA_n^j = \frac{\dfrac{X_n^{\ j}}{X_n^{extraUE-15}}}{\dfrac{X^{\ j}}{X^{extraUE-15}}}$$

where: X - exports to the EU-15 market; n - commodity group measured at the CN-6 level, j - the country; extra EU-15 - EU-15 external trade.

This index has the following interpretation: when the index takes values higher than 1, the country has comparative advantage in exports of products belonging to the tested group against foreign countries. If the value is less than 1,

[6] This formula is based on the RCA B. Balassa index which originally has the following form:
RCA = (Xij : Xi) : (Xj : X), where:
Xij — value of exports of the commodity group i from country j;
Xj — walue of total exports from country j;
Xi — value of world exports of the commodity group i;
X — value of total world exports.

then there is the opposite situation – the country does not posses comparative advantage (Jagiełło 2003; Marczewski 2003; Mroczek-Rubaszek 2003).

Three trade types (one-way, two-way trade in similar products, two-way trade in vertically differentiated products) indictors were calculated using the Freudenberg and Fontagné methodology (Fontagné-Freudenberg 1997).

Flows in each particular CN-8 industry (product) were identified as two-way trade if their minimal value represented at least 10% of maximum flows, or met the condition:

$$\frac{Min(X_{kk'it}, M_{kk'it})}{Max(X_{kk'it}, M_{kk'it})} > 10\%$$

where: X – eksports; M – imports; k – country; k' – partner country; i – product; t – year.

Flows were qualified as two-way trade in similar products (horizontal diversification) if there was a fulfillment of two conditions: the one relating to two-way trade and, at the same time, an additional condition relating to similar products. According to the latter condition, products being part of trade flows are considered as similar if the unit value of imports and exports change less than 15%. So, the products are treated as quality substitutes (vertical differentiation) if they have similar prices and they meet the following condition:

$$\frac{1}{1{,}15} \leq \frac{UV^{X}_{kk'it}}{UV^{M}_{kk'it}} \leq 1{,}15$$

where UV– is a product unit value.

3. Changes in the intensity of intra-industry trade

Multilateral intra-industry trade indexes were calculated for total trade between EU-8 and EU-15 countries and for the mutual trade of EU-8 countries in 2000-2007 (see: Table 1).

The analysis of IIT indicators confirms that during the analyzed period for most of EU-8 countries these indices have increased significantly both in mutual trade and trade with EU-15 countries.

The biggest growth took place in Latvia (where the share of intra-industry trade in 2007 was 69,4% higher than in 2000). In Estonia and Poland IIT indexes were higher by about one-forth (respectively 25,2% and 23,9%) and by one-sixth in Slovakia and Hungary (20,6% and 15,3%). The slowest rate of IIT increase in the analyzed period was in Slovenia, the Czech Republic and Lithuania. **Therefore, it might be clearly stated that between 2000 and 2007 there was a significant**

increase in the share of intra-industry trade of all EU-8 counties with EU-15. Moreover, calculated IIT indicators show that regardless of the upward trend in the intensity of intra-industry trade, still the biggest part of exchange between EU-8 and EU-15 continues is of inter-industry characteristics.

Comparative analysis was conducted for IIT indicators in EU-8 mutual trade. Results for the period 2000-2007 shows that the highest growth occurred in Lithuania, Slovenia and Latvia (by 69,4%, 59,3% and 51,4%). In turn, the Slovakian share of intra-industry trade in 2007 was 46,8% higher than in 2000. Growth in the other countries of the EU-8 group (in Czech Republic, Estonia, Poland and Hungary) was in the range from 28,9% to 11,6%. To sum up, **in 2000-2007 there was a significant increase in intra-industry trade between EU-8 countries**.

It is also important that growth rates of IIT indicators for 2000-2007 in EU-8 mutual trade (an increase of 29,6%) proved to be higher than the dynamics of these indicators in exports and imports of the EU-8 with EU-15 (an increase of 13,0%). The average increase of IIT indicators in recent years before the accession (2000-2003) was 10,5% in EU-8 mutual trade and 4,6% in the EU-8 trade with the EU-15, in the first years after accession (2004-2007) respectively 8,2% and 0,3% . This indicates that after the accession of the EU-8 countries to the EU growth rates of IIT indicators in trade between EU-8 and EU-15 had a greater slowdown than in the EU-8 mutual trade.

Table 1. Grubel and Lloyd`s intra-industry trade (IIT) indexes of EU-8 countries in trade with EU-15 and in mutual EU-8 trade in 2000-2007 (%)[a]

	Country	2000	2001	2002	2003	2004	2005	2006	2007	Dynamics		
										2007/ 2000	2003/ 2000	2007/ 2004
UE-8 and UE-15	**Czech Rep.**	46,3	47,1	47,5	47,1	50,1	48,3	48,7	48,1	103,8	101,7	96,0
	Estonia	22,5	24,9	26,2	27,7	30,3	30,3	26,7	28,2	125,2	122,9	93,1
	Lithuania	16,3	11,8	13,9	18,7	16,8	15,9	17,4	16,4	100,8	114,9	97,9
	Latvia	8,5	9,6	9,3	9,8	11,2	15,0	13,9	14,4	169,4	114,9	128,1
	Poland	33,1	33,2	35,2	36,8	38,7	38,9	40,0	41,0	123,9	111,3	106,0
	Słovakia	27,0	29,5	29,6	30,4	33,2	31,6	30,1	32,6	120,6	112,6	98,3
	Slovenia	35,8	36,7	36,2	37,2	36,8	36,7	38,5	38,4	107,3	103,9	104,5
	Hungary	35,9	36,4	35,6	34,0	39,2	39,3	40,2	41,4	115,3	94,9	105,4
UE-8 mutual trade	**Czech Rep.**	39,1	40,3	41,5	43,2	45,9	49,1	50,4	50,4	129,0	110,6	109,7
	Estonia	35,8	35,2	32,9	31,5	35,0	39,3	40,1	40,0	111,6	88,0	114,4
	Lithuania	24,3	25,0	30,6	32,0	34,7	38,5	40,4	41,1	169,4	131,5	118,7
	Latvia	29,4	28,8	28,4	29,0	34,4	45,2	44,5	44,5	151,4	98,9	129,2
	Poland	38,2	40,4	41,8	41,1	42,7	44,3	44,1	44,4	116,3	107,5	103,9
	Słovakia	30,9	35,1	37,2	38,2	40,8	41,8	44,3	45,3	146,8	123,9	111,1
	Slovenia	17,6	15,4	15,9	18,3	22,4	26,5	28,4	28,0	159,3	104,2	125,1
	Hungary	35,5	34,9	37,8	36,2	43,2	43,4	42,4	44,0	123,8	101,9	101,7

[a] Indicators calculated at 6-digit CN codes level

Source: Comext: Intra- and extra- EU trade Data, Eurostat, 2008. Own calculations.

Source: Own calculations based on Comext: Intra- and extra- EU trade Data, Eurostat, 2008.

Thus, we come to the conclusion that - compared with four years before the accession - in the analyzed period after accession (2004-2007) there were more favorable structural changes in the EU-8 mutual trade than in trade of these countries with the EU-15. This comes up mainly from the fact that before the accession the scope of trade liberalization – being a result of regional free trade agreements – between EU-8 countries was narrow in comparison with that between each of these countries and the European Union - which in turn was the result of the implementation of the provisions of Association Agreements. Removal of barriers to EU-8 mutual trade after their accession to the EU, therefore, created a foothold for the rapid increase in the level of intra-industry trade.

The share of intra-industry trade between EU-8 and EU-15 countries and the EU-8 mutual trade for almost all eight analyzed countries is still relatively low. IIT indicators exceeded 50% only for Czech Republic (in 2004 – 50,1% for trade with the EU-15 and in the years 2006 and 2007 – 50,4% in trade with other EU-8 countries).

While analyzing IIT indicators in both EU-8 trade with EU-15 and EU-8 mutual trade - apart from the leader in both groups, the Czech Republic - several groups of countries might be distinguished. In the EU-8 trade with the EU-15:

- Poland, Slovenia and Hungary had the highest share of intra-industry trade (between 33,1% - 35,9% in 2000 and 38,4% - 41,3% in 2007);
- IIT indicators for Slovakia and Estonia were at the average level (between 22,5% - 27,0% in 2000 and 28,2% - 32,6% in 2007);
- the lowest IIT indicators were in Lithuania and Latvia (shares of intra-industry trade between 8,5% - 16,3% in 2000 and 14,4% - 16,4% in 2007).

Further analysis of indicators for the EU-8 mutual trade put an emphasis of the following:

- average values of IIT indicators i.e. between 24,3% - 38,2% in 2000 and 40,0% - 45,3% in 2007 were in Estonia, Lithuania, Poland, Slovakia and Hungary;
- much lower values were in Slovenia – IIT indicators: 17,6% in 2000 and 28,0% in 2007.

The lowest IIT rates in Slovenia in the EU-8 mutual trade and in Lithuania, Latvia and Estonia in trade between EU-8 and EU-15 might put an emphasis on the fact that these countries have not yet adequately developed their intra-industry links. This may result from traditional orientation of Slovenian companies to EU-15 markets. On the other hand, for companies from Baltic states their experience in trade with other EU-8 countries (with similar level of development) should play an important role in shaping their subsequent relationship with the EU-15.

4. Intra-industry trade of selected CN product groups

In order to complement the analysis of IIT indicators for each country the distribution of these indicators for selected product groups was studied. To this end, IIT indicators were calculated for each CN section (Table 2, 3, 4 and 5).

One of the main conclusions derived from this analysis was the identification of section groups with the highest or lowest IIT values. In trade between EU-8 and EU-15, the most often occurring sections (from 6 to 8 countries) with higher IIT indicators than the IIT for total trade of each country were:

a) in the year 2000:

- plastics and rubber, articles thereof (section VII) – in all 8 countries,
- footwear, headgear, etc. (section XII) – in 7 countries,
- machinery and mechanical appliances, electrical equipment (section XVI) – in 6 countries,

b) in the year 2004:

- plastics and rubber, articles thereof (section VII) – in all 8 countries,
- base metals and articles of base metal (section XV) – in all 8 countries,
- footwear, headgear, etc. (section XII) – in 7 countries,
- transport equipment (section XVII) – in 7 countries,
- miscellaneous manufactured articles (section XX) – in 7 countries,

c) in the year 2007:

- plastics and rubber, articles thereof (section VII) – in all 8 countries,
- base metals and articles of base metal (section XV) – in 7 countries,
- miscellaneous manufactured articles (section XX) – in 7 countries.

In EU-8 mutual trade, the most often occurring sections (from 6 to 8 countries) with higher IIT indicators than the IIT for total trade of each country were:

a) in the year 2000:

- wood and articles of wood (section IX) – in 7 countries,
- products of chemical industry (section VI) – in 6 countries,
- plastics and rubber, articles thereof (section VII) – in 6 countries,
- footwear, headgear, etc. (section XII) – in 6 countries,
- machinery and mechanical appliances, electrical equipment (section XVI) – in 6 countries,
- miscellaneous manufactured articles (section XX) – in 6 countries,

b) in the year 2004:

- plastics and rubber, articles thereof (section VII) – in 7 countries,
- footwear, headgear, etc. (section XII) – in 7 countries,
- pulp of wood, paper, cardboard, etc. (section X) in 6 countries,
- transport equipment (section XVII) – in 6 countries,
- miscellaneous manufactured articles (section XX) – in 6 countries,

c) in the year 2007:

- plastics and rubber, articles thereof (section VII) – in 7 countries,
- transport equipment (section XVII) – in 7 countries,
- foodstuffs, etc. (section IV) – in 6 countries,
- footwear, headgear, etc. (section XII) – in 6 countries,
- miscellaneous manufactured articles (section XX) – in 6 countries.

The abovementioned listing confirms that 3 sections: VII (plastics and rubber, articles thereof), XII (footwear, headgear, etc.) and XX (miscellaneous manufactured articles) remained stable in the group of sections with the highest IIT values in the exchange between EU-8 and EU-15 countries in 2000, 2004 and 2007 and in EU-8 mutual trade in 2004 and 2007.

Table 2. Grubel and Lloyd`s intra-industry trade (IIT) indexes of EU-8 countries in trade with EU-15 and in mututal EU-8 trade in the year 2000 (at 6-digit CN codes level, in %)

CN Sections[a]		Countries[b]								Number of countries with the highest IIT in each section[c]
		CZ	**ES**	**LT**	**LV**	**PL**	**SK**	**SI**	**HU**	
UE-8 with UE-15	IV.	22,1	8,3	6,1	**10,9**	18,3	17,4	18,9	29,0	1
	VI.	17,3	9,5	2,4	3,6	12,4	7,0	19,0	22,5	0
	VII.	**51,5**	**29,6**	**17,2**	**16,3**	**35,0**	**38,9**	**43,2**	**42,3**	8
	VIII.	29,1	20,5	**21,6**	**9,4**	**39,6**	**30,9**	27,9	24,3	4
	IX.	32,9	9,3	7,1	2,2	29,8	19,2	33,4	26,2	0
	X.	**48,7**	14,6	9,8	7,7	28,5	17,5	**37,6**	24,6	2
	XI.	40,6	**23,8**	**19,1**	**13,8**	20,0	19,0	**53,5**	31,5	4
	XII.	**56,3**	**34,8**	**53,8**	**24,7**	**50,6**	21,3	**59,3**	**37,5**	7
	XIII	42,5	**23,2**	10,9	6,4	**37,5**	20,3	25,7	33,7	2
	XV.	**47,1**	**26,6**	14,9	8,4	**41,3**	25,7	**39,4**	**41,9**	5
	XVI.	**54,1**	**26,0**	12,2	**9,6**	**34,2**	**38,1**	35,3	**38,1**	6
	XVII.	**50,1**	16,6	**16,7**	6,2	**52,9**	23,2	**41,9**	**36,1**	5
	XVIII.	**48,1**	**33,2**	**33,7**	7,3	**34,3**	23,8	34,6	**48,8**	5
	XIX.	38,9	0,0	1,5	5,9	7,2	**52,5**	0,0	0,2	1
	XX.	45,9	**32,9**	16,1	**25,7**	**36,2**	**40,0**	34,7	**44,5**	5
	Total	***46,3***	***22,5***	***16,2***	***8,5***	***33,1***	***27,0***	***35,8***	***35,8***	
Number of sections with higher than total average IIT value		7	8	6	7	9	5	6	7	
UE-8 mutual trade	IV.	**52,9**	26,7	**30,4**	27,2	26,4	**48,4**	**19,8**	21,2	4
	VI.	32,6	**39,0**	**30,4**	**62,5**	**41,7**	**40,0**	5,8	**39,3**	6
	VII.	**52,1**	**62,2**	**49,7**	26,8	**41,5**	29,6	**29,3**	**37,9**	6
	VIII.	28,0	25,7	17,8	25,9	24,5	**43,2**	6,1	21,9	1
	IX.	**46,4**	**41,2**	**45,6**	**36,9**	**50,5**	**38,1**	**18,2**	31,6	7
	X.	38,8	27,7	**25,3**	**31,0**	**47,6**	**38,5**	16,8	**49,6**	5
	XI.	34,5	31,4	**30,5**	20,7	34,4	29,6	**20,9**	**43,0**	3
	XII.	**48,4**	**70,6**	**27,9**	15,4	32,4	**53,3**	**55,3**	**57,9**	6
	XIII	35,6	15,3	**43,5**	17,2	**38,3**	**39,5**	11,1	**43,1**	4
	XV.	**39,9**	26,1	**36,1**	22,5	36,9	21,4	15,2	**37,4**	3
	XVI.	**46,9**	**37,1**	**34,1**	26,3	**48,3**	**37,5**	**19,3**	28,8	6
	XVII.	33,7	**38,2**	**32,5**	13,3	**51,0**	**36,3**	**53,0**	29,7	5
	XVIII.	33,6	10,3	**37,2**	17,5	32,5	**38,4**	8,1	**38,0**	3
	XIX.	28,1	0,9	1,5	19,8	14,1	**38,6**	0,0	0,0	1
	XX.	**41,6**	**53,2**	**41,5**	21,4	31,3	**50,0**	**21,7**	**48,1**	6
	Total	***39,0***	***35,8***	***24,2***	***29,3***	***38,2***	***30,8***	***17,5***	***35,5***	
Number of sections with higher than total average IIT value		7	7	13	3	7	12	8	9	

[a] Without sections, where due to their nature it is difficult to apply intra-industry trade. It considers: AGRICULTURAL RAW MATERIALS (sections I-III), MINERAL PRODUCTS (section V), NATURAL OR CULTURED PEARLS, PRECIOUS OR SEMI–PRECIOUS STONES, PRECIOUS METALS, METALS CLAD WITH PRECIOUS METAL, AND ARTICLES THEREOF; IMITATION JEWELLERY; COIN (XIV) and WORKS OF ART, COLLECTORS' PIECES AND ANTIQUE (XXI).

Section marking:
IV. PREPARED FOODSTUFFS; BEVERAGES, SPIRITS AND VINEGAR; TOBACCO AND MANUFACTURED TOBACCO SUBSTITUTES
VI. PRODUCTS OF THE CHEMICAL OR ALLIED INDUSTRIES
VII. PLASTICS AND ARTICLES THEREOF; RUBBER AND ARTICLES THEREOF
VIII. RAW HIDES AND SKINS, LEATHER, FURSKINS AND ARTICLES THEREOF; SADDLERY AND HARNESS; TRAVEL GOODS, HANDBAGS AND SIMILAR CONTAINERS; ARTICLES OF ANIMAL GUT (OTHER THAN SILKWORM GUT)
IX. WOOD AND ARTICLES OF WOOD; WOOD CHARCOAL; CORK AND ARTICLES OF CORK; MANUFACTURES OF STRAW, OF ESPARTO OR OF OTHER PLAITING MATERIALS; BASKETWARE AND WICKERWORK
X. PULP OF WOOD OR OF OTHER FIBROUS CELLULOSIC MATERIAL; RECOVERED (WASTE AND SCRAP) PAPER OR PAPERBOARD; PAPER AND PAPERBOARD AND ARTICLES THEREOF
XI. TEXTILES AND TEXTILE ARTICLES
XII. FOOTWEAR, HEADGEAR, UMBRELLAS, SUN UMBRELLAS, WALKING STICKS, SEAT–STICKS, WHIPS, RIDING–CROPS AND PARTS THEREOF; PREPARED FEATHERS AND ARTICLES MADE THEREWITH; ARTIFICIAL FLOWERS; ARTICLES OF HUMAN HAIR
XIII. ARTICLES OF STONE, PLASTER, CEMENT, ASBESTOS, MICA OR SIMILAR MATERIALS; CERAMIC PRODUCTS; GLASS AND GLASSWARE
XV. BASE METALS AND ARTICLES OF BASE METAL
XVI. MACHINERY AND MECHANICAL APPLIANCES; ELECTRICAL EQUIPMENT; PARTS THEREOF; SOUND RECORDERS AND REPRODUCERS, TELEVISION IMAGE AND SOUND RECORDERS AND REPRODUCERS, AND PARTS AND ACCESSORIES OF SUCH ARTICLES
XVII. VEHICLES, AIRCRAFT, VESSELS AND ASSOCIATED TRANSPORT EQUIPMENT\\
XVIII. OPTICAL, PHOTOGRAPHIC, CINEMATOGRAPHIC, MEASURING, CHECKING, PRECISION, MEDICAL OR SURGICAL INSTRUMENTS AND APPARATUS; CLOCKS AND WATCHES; MUSICAL INSTRUMENTS; PARTS AND ACCESSORIES THEREOF
XIX. ARMS AND AMMUNITION; PARTS AND ACCESSORIES THEREOF
XX. MISCELLANEOUS MANUFACTURED ARTICLES

[b] CZ-Czech Republic, ES-Estonia, LT-Lithuania, LV-Latvia, PL-Poland, SK-Slovakia, SI-Slovenia, HU-Hungary

[c] Incorporates sections in which the IIT index was higher than the average index for each country (bold).

Source: Own calculations based on Comext: Intra- and extra- EU trade Data, Eurostat, 2008.

To sum up, an important trend characteristic for the analyzed period is the unchanged or smaller number of sections with the highest IIT values in the period after the accession rather than in the period before the accession. Moreover, the highest IIT values both in pre- and post-accession period occurred in less than half of the analyzed sections in EU-8 mutual trade and external trade between EU-8 and EU-15.

In the next stage of the analysis of the distribution of IIT indicators by CN sections the concentration in the sections with the highest values of the IIT indicators was studied (relevant data are summarized in the last rows of Table 2, 3, 4 and 5). It seems interesting that the conclusions from the analysis of EU-8 mutual trade differ from those derived from the analysis of trade between EU-8 and EU-15.

In trade between EU-8 and EU-15 in 2000-2003 concentration of trade with the highest values of the IIT indicators in the lower number of sections CN occurred in Poland and the Czech Republic. In four countries (Estonia, Slovakia, Latvia and Lithuania) the number of these sections has not changed, in case of two countries (Hungary and Slovenia) trade with the highest values of IIT spread on more sections. In the post-accession period (2004-2007) the concentration in fewer sections occurred in three countries (again, this was Poland, but also Latvia and Slovenia). Trade with the highest values of IIT spread on lower number of sections in the three countries (Czech Republic, Lithuania, Slovakia) and in two countries (Estonia and Hungary) the number of sections with the highest rates stood still.

Table 3. Grubel and Lloyd`s intra-industry trade (IIT) indexes of EU-8 countries in trade with EU-15 and in mututal EU-8 trade in the year 2003 (at 6-digit CN codes level, in %)

CN Sections[a]		Countries[b]								Number of countries with the highest IIT in each section[c]
		CZ	**ES**	**LT**	**LV**	**PL**	**SK**	**SI**	**HU**	
UE-8 with UE-15	IV.	30,49	10,95	8,13	**9,93**	27,82	18,01	15,55	24,55	1
	VI.	18,86	13,24	3,41	3,78	17,50	8,43	23,73	28,46	0
	VII.	**49,64**	**37,61**	**20,11**	**20,38**	**41,91**	**33,95**	**41,36**	**46,05**	8
	VIII.	29,57	**27,70**	**22,70**	5,75	**37,91**	23,05	32,38	24,70	3
	IX.	33,50	10,26	9,39	2,14	26,42	22,80	34,33	**35,44**	1
	X.	**53,10**	20,76	5,07	**9,81**	31,34	18,93	**43,30**	22,31	3
	XI.	42,22	24,00	17,31	**17,04**	21,41	22,68	**49,92**	27,01	2
	XII.	**63,92**	**38,14**	**42,56**	8,99	**47,56**	16,46	**57,96**	**35,42**	6
	XIII	45,42	**28,20**	11,43	7,25	**39,64**	24,38	27,50	**34,79**	3
	XV.	**48,44**	**35,54**	**19,48**	**13,19**	**45,33**	**36,31**	**39,97**	**39,72**	8
	XVI.	**51,65**	**37,64**	14,74	**10,05**	**37,70**	**32,52**	**40,12**	30,68	6
	XVII.	**53,62**	20,76	**39,10**	5,43	**55,17**	**36,96**	**39,10**	**53,88**	6
	XVIII.	**52,87**	**40,76**	**31,17**	9,55	**50,63**	27,48	36,72	**42,18**	4
	XIX.	31,34	0,28	1,44	1,13	6,55	**31,24**	0,00	**41,34**	2
	XX.	44,83	**29,52**	17,15	**30,15**	31,72	20,73	36,66	**43,99**	3
	Total	***47,12***	***27,70***	***18,71***	***9,76***	***36,85***	***30,44***	***37,19***	***34,04***	
Number of sections with higher than total average IIT value		7	8	6	7	8	5	7	9	
UE-8 mutual trade	IV.	**55,81**	29,80	**43,02**	**31,95**	34,89	**53,36**	15,39	29,64	4
	VI.	**44,34**	**38,04**	**36,04**	**60,27**	39,37	**41,02**	4,94	35,46	5
	VII.	**54,83**	**47,30**	**41,66**	**29,08**	**46,98**	**44,20**	**39,32**	**43,27**	8
	VIII.	26,33	22,52	**40,88**	**30,62**	**46,66**	**46,56**	9,83	**52,95**	5
	IX.	41,02	**40,18**	**33,18**	28,66	36,72	**40,05**	**20,83**	35,63	4
	X.	43,18	**33,00**	27,83	**30,25**	**53,03**	**44,23**	**20,34**	**39,66**	6
	XI.	34,85	30,59	**36,53**	25,22	37,16	**43,07**	**18,58**	**47,85**	4
	XII.	33,91	**58,17**	**43,50**	24,79	39,84	**58,20**	**38,78**	30,33	4
	XIII	38,53	12,25	**38,06**	15,67	**47,25**	**45,69**	16,60	**48,34**	4
	XV.	**46,77**	**32,33**	**32,70**	**29,97**	37,29	25,31	**20,96**	**45,88**	6
	XVI.	**46,63**	**39,08**	**38,70**	**34,51**	**50,29**	37,96	**20,12**	33,82	6
	XVII.	30,54	**45,57**	**63,41**	17,09	35,36	38,10	**36,87**	32,85	3
	XVIII.	35,67	19,48	**41,24**	23,82	30,97	**38,78**	14,15	33,72	2
	XIX.	21,93	13,70	23,46	8,40	7,06	22,66	0,00	**44,35**	1
	XX.	**52,70**	**47,03**	**61,86**	**29,71**	35,61	**49,55**	**26,82**	**46,09**	7
	Total	***43,20***	***31,54***	***31,95***	***29,04***	***41,06***	***38,24***	***18,30***	***36,21***	
Number of sections with higher than total average IIT value		6	9	13	8	5	11	9	8	

a;b,c Marking of section and countries - as in Table 2.

Source: Own calculations based on Comext: Intra- and extra- EU trade Data, Eurostat, 2008.

Table 4. Grubel and Lloyd`s intra-industry trade (IIT) indexes of EU-8 countries in trade with EU-15 and in mututal EU-8 trade in the year 2004 (at 6-digit CN codes level, in %)

CN Sections[a]		Countries[b]								Number of countries with the highest IIT in each section[c]
		CZ	**ES**	**LT**	**LV**	**PL**	**SK**	**SI**	**HU**	
UE-8 with UE-15	IV.	30,5	25,0	11,0	**17,5**	31,4	21,4	15,7	26,0	1
	VI.	20,0	12,8	3,2	3,0	19,4	8,3	25,0	30,9	0
	VII.	**52,0**	**35,5**	**20,7**	**21,4**	**41,9**	**35,2**	**37,6**	**47,2**	8
	VIII.	36,3	**38,7**	**28,0**	8,8	**39,1**	28,9	34,1	31,6	3
	IX.	36,6	12,4	12,1	2,6	27,9	20,7	**42,0**	36,5	1
	X.	**57,4**	25,6	8,3	10,4	33,2	18,9	**47,0**	29,3	2
	XI.	**54,5**	**34,7**	**22,1**	**24,7**	29,0	26,0	**39,5**	30,8	5
	XII.	**67,6**	**52,1**	**33,5**	**13,8**	**42,4**	25,2	**44,6**	**49,4**	7
	XIII	45,6	23,8	13,9	**13,6**	**42,5**	28,9	27,0	31,6	2
	XV.	**50,4**	**35,6**	**28,6**	**12,4**	**46,9**	**41,2**	**42,1**	**42,1**	8
	XVI.	**51,5**	**38,4**	15,9	9,9	37,2	**35,1**	**38,6**	37,5	4
	XVII.	**63,3**	21,6	**23,6**	**11,3**	**59,9**	**39,1**	**41,7**	**56,6**	7
	XVIII.	49,6	**38,3**	**38,8**	7,2	**43,6**	23,9	31,1	**43,7**	4
	XIX.	48,7	4,5	2,3	3,1	3,3	27,1	13,7	**52,8**	1
	XX.	**52,6**	**34,1**	**19,1**	**23,8**	29,5	**36,7**	**42,0**	**50,4**	7
	Total	***50,0***	***30,3***	***16,7***	***11,2***	***38,6***	***33,1***	***36,7***	***39,2***	
Number of sections with higher than total average IIT value		8	8	8	8	7	5	9	7	
UE-8 mutual trade	IV.	**58,4**	**37,1**	**49,5**	**38,6**	37,5	**45,5**	11,4	39,4	5
	VI.	**47,4**	31,5	34,4	**60,2**	40,7	32,6	10,8	37,6	2
	VII.	**53,6**	**51,0**	**42,8**	30,5	**49,3**	**44,7**	**35,8**	**49,8**	7
	VIII.	33,7	29,5	**54,9**	**38,3**	**44,7**	37,9	6,8	36,6	3
	IX.	44,9	**42,6**	**37,5**	29,1	32,4	37,4	16,1	37,2	2
	X.	**47,3**	**43,5**	34,1	**39,7**	**51,4**	**41,3**	**22,6**	37,7	6
	XI.	42,1	30,7	**42,7**	28,2	37,8	**48,3**	19,2	**47,5**	3
	XII.	**63,1**	**65,6**	**54,4**	**38,9**	38,9	**62,5**	**44,0**	**46,3**	7
	XIII	41,4	18,6	**39,2**	16,8	**50,6**	**45,9**	13,8	**51,5**	4
	XV.	44,4	**43,5**	34,0	28,9	41,7	30,6	**25,3**	**45,1**	3
	XVI.	**46,3**	**35,0**	**48,5**	32,5	**47,3**	39,9	19,4	41,9	4
	XVII.	39,6	**53,8**	**39,1**	18,0	**53,8**	**46,3**	**51,1**	**52,3**	6
	XVIII.	40,7	27,2	**38,9**	20,3	23,5	27,4	9,1	28,5	1
	XIX.	37,3	0,2	**35,0**	15,5	22,5	18,1	0,3	36,3	1
	XX.	**53,9**	**54,7**	**63,5**	27,6	36,2	**46,8**	**28,9**	**46,8**	6
	Total	***45,9***	***34,9***	***34,6***	***34,4***	***42,7***	***40,7***	***22,3***	***43,2***	
Number of sections with higher than total average IIT value		7	9	12	5	6	8	6	7	

a;b,c Marking of section and countries - as in Table 2.

Source: Own calculations based on Comext: Intra- and extra- EU trade Data, Eurostat, 2008.

Table 5. Grubel and Lloyd`s intra-industry trade (IIT) indexes of EU-8 countries in trade with EU-15 and in mututal EU-8 trade in the year 2007 (at 6-digit CN codes level, in %)

CN Sections[a]		Countries[b]								Number of countries with the highest IIT in each section[c]
		CZ	**ES**	**LT**	**LV**	**PL**	**SK**	**SI**	**HU**	
UE-8 with UE-15	IV.	36,1	16,5	**18,1**	**16,9**	**42,4**	25,1	17,2	31,4	3
	VI.	21,8	14,6	3,9	5,2	27,2	9,2	30,4	33,5	0
	VII.	**48,8**	**28,9**	**19,3**	**22,6**	**46,4**	**44,1**	**46,5**	**44,9**	8
	VIII.	**56,2**	**38,3**	**30,8**	8,0	39,8	18,9	25,1	27,4	3
	IX.	38,6	20,4	12,3	5,9	29,9	**38,9**	33,0	**42,2**	2
	X.	**58,0**	20,5	**18,3**	13,6	32,4	17,7	**46,4**	35,4	3
	XI.	**53,2**	**43,9**	**33,3**	**36,6**	38,5	**36,1**	28,2	38,3	5
	XII.	**61,2**	**45,4**	14,4	7,0	33,8	**34,5**	**56,8**	**44,0**	5
	XIII	**56,7**	23,0	**17,5**	11,3	**44,9**	**41,5**	37,3	36,4	4
	XV.	**50,1**	**38,3**	**33,3**	**26,8**	**47,8**	**42,2**	**41,4**	39,0	7
	XVI.	**60,7**	**33,3**	16,2	11,2	**43,1**	**38,7**	36,7	**43,0**	5
	XVII.	43,3	18,8	7,6	7,9	**54,0**	28,3	**47,6**	**50,7**	3
	XVIII.	**50,8**	**35,2**	**36,5**	6,2	35,6	12,5	36,6	**48,6**	4
	XIX.	43,2	7,0	**16,7**	0,0	0,5	0,0	4,3	22,5	1
	XX.	**51,6**	**35,3**	**26,4**	**31,8**	27,3	**47,9**	**38,7**	**48,9**	7
	Total	***48,0***	***28,2***	***16,4***	***14,3***	***41,0***	***32,6***	***38,4***	***41,3***	
Number of sections with higher than total average IIT value		10	8	10	5	6	8	6	7	
UE-8 mutual trade	IV.	**62,8**	**43,3**	**60,6**	**62,9**	41,0	**46,1**	9,9	**45,0**	6
	VI.	47,6	31,0	35,5	**57,4**	43,5	34,9	14,3	37,2	1
	VII.	**56,8**	**53,0**	37,0	**48,6**	**50,0**	**52,1**	**37,4**	**44,2**	7
	VIII.	**53,8**	**61,4**	**48,1**	**52,9**	41,8	44,6	**52,3**	43,0	5
	IX.	**51,5**	**50,0**	**40,0**	**46,4**	38,4	43,8	**32,2**	32,4	5
	X.	**51,8**	28,5	33,3	**45,9**	**52,3**	40,8	9,8	39,6	3
	XI.	**53,2**	34,5	**49,0**	42,6	**51,9**	**56,3**	**35,2**	39,1	5
	XII.	**74,1**	**55,0**	**67,3**	42,4	**68,7**	**58,2**	**45,0**	26,0	6
	XIII	46,5	28,0	39,9	24,7	**47,4**	**47,2**	18,1	39,2	2
	XV.	**50,3**	**41,7**	**43,9**	31,3	**46,6**	40,8	**39,0**	**44,5**	6
	XVI.	43,1	**58,0**	**48,6**	42,8	39,3	34,3	25,9	**45,7**	3
	XVII.	44,4	**57,6**	**47,6**	**58,5**	**60,6**	**55,0**	**40,9**	**57,9**	7
	XVIII.	42,4	**47,3**	**52,1**	**45,6**	**47,8**	41,4	14,5	**44,9**	5
	XIX.	20,6	0,0	**41,4**	0,0	0,1	38,3	0,0	15,1	1
	XX.	**59,1**	**55,0**	**68,3**	44,0	35,2	**53,1**	**36,9**	**50,4**	6
	Total	***50,3***	***39,9***	***41,1***	***44,4***	***44,4***	***45,3***	***27,9***	***43,9***	
Number of sections with higher than total average IIT value		9	10	11	8	8	7	8	7	

a;b,c Marking of section and countries - as in Table 2.

Source: Own calculations based on Comext: Intra- and extra- EU trade Data, Eurostat, 2008.

By contrast, analysis of the EU-8 mutual trade between 2000-2003 indicated strong tendency to trade concentration with the highest IIT indexes in an increasingly lower number of CN sections. It occurred in four countries (Czech Republic, Poland, Slovakia, Hungary). Number of sections did not change only in Lithuania, and in three countries (Slovenia, Latvia and Estonia) trade with the highest IIT indexes spread out in more sections. However, this trend reversed in the years after accession. Then, in five countries (Czech Republic, Estonia, Latvia, Poland and Slovenia) trade with the highest IIT values spread out in more sections, than in pre-accession period. In Hungary, the situation remained unchanged, and only in Lithuania and Slovakia concentration occurred in the lower amount of sections.

Throughout the analyzed period (2000-2007) both in EU-8 mutual trade and in trade between EU-8 and EU-15 there was a reinforcement in the tendency to increase the intensity of intra-industry trade in an increasing number of CN sections. Concentration of trade with the highest values of IIT indexes in increasingly lower number of CN sections occurred only in Latvia and Poland - in trade with the EU-15 countries and in Lithuania, Slovakia and Hungary - in a mutual exchange within the EU-8 group.

5. Intra-industry trade by 5 groups of the intensity of using production factors

In order to analyze the relationship between the intensity of intra-industry trade flows in pre- and post-accession period and the transformation in the external trade structure between EU-8 countries and EU-15 the analysis of IIT indicators by 5 groups of the intensity of using production factors was conducted. Appropriate calculations are summarized in Table 6.

The comparison of IIT indicators in total trade between EU-8 and EU-15 with the indicators for each of 5 groups of the intensity of using production factors, as well as with the structure of external trade by these groups allows to indicate that in the period between 2000-2007 there were significant transformations in the level of intra-industry trade.

In 2000, material-intensive products (group 1) in most of the analyzed countries were characterized by the lowest IIT indicators (with the exception of the Lithuania and Poland, where the lowest indicators at that time were recorded in group 3 and 4). At the same time, these goods were not playing important role in EU-8 – EU-15 trade flows. In exports, only for Lithuania and Latvia its share significantly exceeded the total export share for whole EU-8 group (10%) and accounted for 35% and 50% accordingly. Total import share of this commodity group for EU-8 was in 2000 the lowest (only 8%) in comparison to other analyzed groups.

Table 6. EU-8 external trade with EU-15

Country	Groups[a]	IIT (%)			Change[b]	RCA			EX_i/EX (%)			IM_i/IM (%)		
		2000	2004	2007		2000	2004	2007	2000	2004	2007	2000	2004	2007
Czech Republic	**1**	28,5	25,4	**25,0**	***87,7***	0,3	0,2	**0,2**	7,4	5,6	**8,5**	7,1	6,5	**10,7**
	2	47,4	55,2	**57,0**	120,1	1,5	1,2	**1,3**	30,4	24,5	**26,8**	23,3	22,5	**23,6**
	3	45,7	51,2	**43,6**	95,3	2,4	2,0	**2,1**	25,8	23,3	**31,7**	20,2	20,2	**25,2**
	4	36,6	34,7	**35,2**	***96,3***	0,5	0,8	**0,6**	8,3	14,9	**6,3**	17,3	17,7	**15,9**
	5	54,2	58,6	**58,7**	108,3	1,1	1,4	**1,9**	28,1	31,1	**26,1**	32,1	33,1	**24,6**
	Total	***46,3***	***50,1***	***48,1***	*103,8*						-			
Estonia	**1**	7,1	12,8	**21,7**	303,2	0,8	0,7	**0,8**	18,6	18,8	**26,9**	10,1	9,5	**13,1**
	2	29,1	34,2	**36,1**	124,2	1,5	1,8	**1,7**	31,3	35,7	**35,9**	24,8	24,1	**21,6**
	3	13,5	21,3	**17,4**	128,5	0,2	0,5	**0,6**	1,9	5,9	**9,3**	15,7	21,8	**34,6**
	4	24,3	36,8	**10,6**	43,8	2,1	1,1	**0,2**	38,4	21,9	**2,3**	15,5	16,9	**9,0**
	5	25,7	32,6	**34,4**	134,0	0,4	0,6	**1,3**	9,8	13,9	**17,3**	33,8	27,7	**21,7**
	Total	***22,5***	***30,3***	***28,2***	*125,2*						-			
Hungary	**1**	17,7	27,7	**22,1**	124,7	0,3	0,3	**0,3**	6,1	6,4	**10,3**	4,2	6,3	**9,9**
	2	37,1	41,4	**43,3**	116,7	0,8	0,6	**0,6**	17,0	12,7	**13,7**	23,8	20,1	**20,2**
	3	35,9	48,3	**45,1**	125,4	1,3	1,0	**1,3**	14,0	11,2	**19,2**	16,1	17,6	**24,7**
	4	39,5	33,6	**37,7**	**95,4**	1,5	1,4	**0,9**	28,1	26,7	**10,2**	20,7	17,5	**14,0**
	5	35,4	39,6	**38,8**	109,6	1,2	1,3	**2,4**	29,9	27,3	**32,7**	35,2	38,5	**31,3**
	Total	***35,9***	***39,2***	***41,4***	*115,3*						-			
Lithuania	**1**	14,4	7,6	**11,2**	**77,8**	1,4	1,6	**1,1**	35,3	39,3	**40,2**	12,6	9,1	**10,1**
	2	19,4	22,8	**29,2**	**150,3**	2,3	1,9	**1,6**	45,8	38,4	**35,1**	29,8	24,1	**19,7**
	3	6,0	12,1	**8,8**	147,1	0,2	0,3	**0,4**	2,4	3,7	**6,8**	20,0	22,6	**32,8**
	4	10,9	13,0	**10,5**	96,2	0,2	0,2	**0,9**	3,2	4,5	**9,7**	17,6	19,4	**18,6**
	5	17,7	23,6	**18,1**	102,6	0,5	0,6	**0,5**	11,9	13,9	**7,3**	19,9	24,9	**18,8**
	Total	***16,3***	***16,8***	***16,4***	*100,8*						-			
Latvia	**1**	2,8	3,8	**7,8**	**280,6**	2,1	1,7	**1,4**	50,4	43,8	**49,7**	11,9	11,1	**11,4**
	2	15,1	21,1	**27,1**	179,0	1,9	1,8	**1,2**	38,3	37,0	**25,4**	30,2	26,8	**24,8**
	3	4,2	7,4	**8,9**	209,1	0,6	1,1	**0,8**	6,5	12,9	**12,5**	18,0	20,2	**32,0**
	4	4,2	4,0	**9,0**	**217,7**	0,1	0,1	**0,3**	1,3	1,0	**3,3**	17,4	17,4	**9,3**
	5	10,9	12,4	**10,1**	92,7	0,1	0,2	**0,4**	3,4	5,2	**4,8**	22,6	24,5	**22,5**
	Total	***8,5***	***11,2***	***14,4***	*169,4*						-			
Poland	**1**	21,9	23,8	**33,8**	**154,1**	0,6	0,6	**0,5**	14,0	14,6	**18,1**	8,3	8,2	**13,8**
	2	34,5	39,7	**43,0**	124,9	1,8	1,5	**1,4**	36,4	30,0	**29,2**	25,9	22,4	**22,1**
	3	43,5	51,7	**47,9**	110,3	2,0	2,0	**1,7**	21,1	23,7	**26,3**	20,1	21,9	**25,7**
	4	18,6	20,3	**28,4**	**152,9**	0,4	0,3	**0,4**	7,0	6,5	**4,5**	19,7	18,9	**16,8**
	5	36,2	41,8	**42,5**	117,2	0,9	1,2	**1,6**	21,5	25,3	**21,8**	25,9	28,5	**21,6**
	Total	***33,1***	***38,7***	***41,0***	*123,9*						-			
Słovenia	**1**	10,0	11,7	**21,1**	**210,0**	0,1	0,1	**0,3**	3,3	3,7	**9,0**	13,8	13,6	**19,3**
	2	44,8	42,9	**43,1**	**96,3**	1,8	1,5	**1,2**	36,3	30,5	**25,2**	25,2	22,7	**21,2**
	3	37,5	41,6	**44,5**	118,9	2,7	2,6	**2,7**	29,0	30,8	**41,3**	24,0	24,5	**27,7**
	4	24,9	24,9	**36,0**	**144,6**	0,2	0,3	**0,5**	4,5	5,3	**5,2**	13,9	15,7	**13,0**
	5	36,9	40,6	**36,5**	98,9	1,1	1,4	**1,4**	26,8	29,7	**19,4**	23,1	23,5	**18,7**
	Total	***35,8***	***36,8***	***38,4***	***107,3***						-			
Slovakia	**1**	17,5	24,4	**31,3**	179,0	0,3	0,3	**0,2**	7,3	7,6	**8,5**	7,2	6,4	**9,4**
	2	28,5	35,7	**44,4**	155,7	1,3	1,2	**1,0**	26,6	24,2	**21,6**	22,7	21,0	**20,7**
	3	21,3	34,6	**27,9**	131,3	3,6	3,0	**3,1**	38,9	35,1	**46,9**	24,1	25,8	**27,6**
	4	25,1	22,4	**23,9**	**95,3**	0,4	0,5	**0,6**	7,2	10,1	**6,2**	14,5	14,4	**13,1**
	5	36,5	36,5	**34,8**	95,1	0,8	1,0	**1,1**	19,6	22,8	**14,9**	31,5	32,4	**29,1**
	Total	***27,0***	***33,2***	***32,6***	*120,6*						-			
UE-8 with UE-15[c]	**1**	19,3	21,4	**26,9**	**139,4**	0,4	0,4	**0,4**	10,2	10,0	**13,8**	7,7	7,7	**11,9**
	2	37,3	42,9	**45,8**	122,8	1,4	1,2	**1,2**	29,1	24,4	**25,0**	24,5	21,7	**21,2**
	3	38,1	46,3	**41,0**	107,6	2,0	1,8	**1,9**	21,3	21,2	**28,6**	19,6	20,9	**25,6**
	4	30,1	28,6	**30,5**	**101,3**	0,7	0,7	**0,6**	13,4	13,9	**6,3**	18,2	17,3	**14,8**
	5	40,0	44,9	**43,7**	109,3	1,0	1,2	**1,7**	24,7	26,6	**23,2**	29,2	30,6	**23,4**
	Total	***35,5***	***40,0***	***40,1***	*113,0*		-				-			

a Group marking: 1. material-intensive, 2. labor-intensive, 3. capital-intensive, 4. technology-intensive goods, easy to imitate, 5. technology-intensive goods, difficult to imitate.
b indicators where dynamics between 2004-2007 was higher than between 2000-2004 is marked with bolded font
c IIT indicators for trade between EU-8 and EU-15 were calculated as a weighted average where the weight is a share of the country turnover in total turnover of the whole group.
Source: Own calculation based on Comext: Intra- and extra- EU trade Data, Eurostat, 2008. and Z. Wysokińska, Dynamiczne współzależności wymiany handlowej krajów Europy Środkowej i Wschodniej w świetle teorii integracji i wymiany międzynarodowej, Wydawnictwo Uniwersytetu Łódzkiego, Łódź 1995.

Labor-intensive goods from group 2 had the highest IIT values (from 19% to 47%) in the half of analyzed countries (Estonia, Latvia, Lithuania, Slovenia) - average IIT of EU-8 for this product group was relatively high (37%). At the same time they played a significant role in the external trade of all 8 countries (first place in the case of exports from 4 countries - Czech Republic, Lithuania, Poland, Slovenia and import to 4 countries - Lithuania, Latvia, Poland, Slovenia) in trade with the EU-15 (export and import shares of EU-8 group of countries accounted for 29% and 24%).

Capital-intensive goods (group 3) were characterized by the highest IIT only in Poland (43%), and in most countries (except Poland and Czech Republic) this indicator leveled on lower position than for the total trade. Total exports and imports shares of EU-8 group for this product group accounted for 21% and 20% - were thus on average level.

In group 4 (technologically intensive goods, easy to imitate) in most countries (except Czech Republic and Hungary; in case of Hungary this group was characterized by the highest IIT value - 39%) these indicators had much lower values (between 4% and 25%) than for the total trade of EU-8 group (30%). These products played a relatively minor role in trade between EU-8 and EU-15 (shares: 13% in exports and 18% in imports).

It is worth to emphasize that the goods technologically intensive, difficult to imitate (group 5) only in Hungary were characterized by lower values of IIT than for the total trade of EU-8 group. These commodity group was distinguished by the highest IIT (40%) for total EU-8 trade with EU-15. At the same time they played a dominant role in imports of most of analyzed countries from EU-15 (with the exception of Slovenia and Lithuania – 3rd place and Latvia 2nd place).

To sum up, **in 2000 in most of EU-8 countries the relatively highest IIT indexes characterized their trade with the EU-15 in labor-intensive goods (group 2) and products technologically intensive, difficult to imitate (group 5). The lowest IIT values were indicated in group 1 (material-intensive goods). Export flows of EU-8 countries to the EU-15 were dominated by low-processed products, imports – by products with modern technology and in these commodity groups there were indicated the most favorable conditions for the growth of intra-industry trade**.

In the analyzed period until 2007 the share of intra-industry trade in all 8 countries increased significantly (the highest growth rates were in Latvia: 169% and Estonia: 129%). However, the dynamics of these indicators in 2000-2007 in different groups of the intensity of using production factors was differentiated. Comparing these indicators for various groups of commodities with these for total trade in individual countries, they grew much faster in Czech Republic in groups 2 and 5, Estonia in groups 1, 2, 5, Hungary and Slovakia in groups 1, 2, 3 , Lithuania in groups 2, 3, 5, Latvia in groups 1, 2, 3, 4, Poland in groups 1, 2, 4, and Slovenia in groups 1, 3, 4. It is also worth to mention that in the analyzed group of EU-8 countries indicators for certain groups decreased: in the Czech Republic (group 1, 3, 4), Estonia and Hungary (group 4), Lithuania (group 1 and 4), Latvia (group 5), Slovenia (group 2 and 5), Slovakia (group 4 and 5). Poland was the only country where all the indicators for the period 2000-2007 increased.

As a result, in the year 2007 goods from group 2 were still characterized by the highest IIT indicators in case of half of the analyzed countries. These goods were also important in exports (25%) and imports (21%) for whole EU-8 group. Moreover, in Lithuania and Slovenia IIT indicators for labor-intensive goods increased faster after the accession (their dynamics in 2004-2007 was higher than in 2000-2004).

Material-intensive goods (group 1) for all EU-8 countries were characterized in 2007 by lower IIT values (indicators between 7% - 33%) than IIT indicators for total trade in each country. The average intra-industry trade share for EU-8 group was also the lowest for this group of products (27%).

In group 3 (capital-intensive goods) IIT indicators in most countries were lower than for total trade (with the exception of Hungary, Poland and Slovenia - the first place). These goods, however, played in 2007 the biggest role in external trade of EU-8 with EU-15 (the highest shares in imports: 26%, and in exports: 29%).

Technology-intensive goods, easy to imitate (group 4) were not so significant in EU-8 trade with EU-15 (very low share of exports: 6% - the last place in all countries except Lithuania, and relatively low share of imports: 15%). Simultaneously, in all EU-8 countries IIT indicators in this group were lower than for total trade.

Technology intensive goods, difficult to imitate from group 5 were characterized in 2007 the second highest (right after group 2) average IIT indicator (44%). What is more, in most countries IIT indicators for this group was higher than for total trade (except Hungary, Latvia and Slovenia) and in the case of Czech Republic the highest at all. These products played an important role in EU-8 external trade with EU-15 (both exports and imports shares valued 23%).

It should be noted that between 2000-2007 there was definite increase of the importance of group 2 (labor-intensive goods) in intra-industry trade between EU-8

and EU-15 – IIT indicators in 2000 were in the range of 15% to 47%, in 2007 between 27% and 57%. Quite big significance of technology-intensive goods, difficult to imitate (group 5) was kept during this period. At the same time, despite an increase of IIT indicators (in 2000 in the range of 2% - 28%, in 2007 between 7% and 33%) material-intensive goods from group 1 did not change its position as the least significant in intra-industry trade between EU-8 and EU - 15. On the other hand, in this period the greatest improvement in terms of share in EU-8 imports and exports to and from EU-15 were recorded in groups 1 and 3 (material- and capital-intensive goods).

The observed changes seem to support the thesis that the improvement in intra-industry trade of EU-8 with EU-15 in 2000-2007 was largely the result of changes in the structure their external trade. In most of these countries the pace of change of IIT indicators in different groups of the intensity of using production factors differed significantly from the pace of changes of these indicators for total trade. Moreover, in the case of group 1 (material-intensive goods) and group 4 (technology-intensive goods, easy to imitate) IIT indicators in most of EU-8 countries increased faster after accession than in pre-accession period (their dynamics in 2004-2007 were higher than in 2000-2004). Thus we came to the conclusion that for these two product groups – material-intensive and technology-intensive goods, easy to imitate – changes in conditions for trade after the accession to the Single European Market had the most beneficial effect on their growing importance in intra-industry division of labor.

6. The analysis of RCA and IIT indicators between EU-8 and EU-15

Basic causes of low intensity of EU-8 intra-industry trade with EU-15 come from the asymmetry of demand factors (differences in GDP per capita) and supply factors (technology gap), which still exists between these two groups of countries. This is confirmed by the analysis of trade flows between EU-8 and EU-15 by RCA indicators (presence of comparative advantage indicates the competitiveness of the analyzed economy in each product group and is regarded as a determinant of intra-industry flows) and export and import shares of the various commodity groups of the intensity of using production factors.

On the basis of calculation results compiled in table 6 for the whole group of EU-8 countries in pre- and post-accession period the following trends were indicated:

- despite increasing IIT indicators and growing imports and exports shares to and from EU-15 material-intensive goods (group 1) in 2007 were still characterized by inter-industry specialization (IIT <50%). In addition, EU-8 countries did not reveal comparative advantage and the RCA index for this group remained at a low level (0.4) throughout the analyzed period;

- external trade flows in labor-intensive goods (group 2) were of inter-industry characteristics. The weighted average IIT indicator for this group in trade between EU-8 and EU-15 was, however, in 2007 the highest (46%) among all five analyzed commodity groups and showed an upward trend despite the declining (but still higher than 1) RCA indicator. EU-8 countries maintained revealed comparative advantage in trade in these commodities with EU-15, but exports and imports shares in 2000-2007 decreased;
- in case of capital-intensive goods (group 3) revealed comparative advantage of EU-8 countries remained at a relatively high level (despite a slight decline during the period RCA indicator hovered around the level of 1.95) and this commodity group increased also imports and exports shares, which together resulted in increased participation in intra-industry trade;
- the analysis of EU-8 external trade with EU-15 in goods from group 4 (technologically intensive products, easy to imitate) did not indicate an improvement in their competitiveness. What is more, exchange of these products still was of inter-industry characteristics (weighted average IIT indicator in 2007 was 30%). The lack of comparative advantage was deepening together with export and import shares;
- trade in technologically intensive, difficult to imitate goods (group 5) was also characterized by inter-industry specialization (IIT average indicator in 2007 was 43%) with growing in IIT values. EU-8 countries maintained and strengthened their revealed comparative advantage in trade with EU-15, but import and export shares were decreasing.

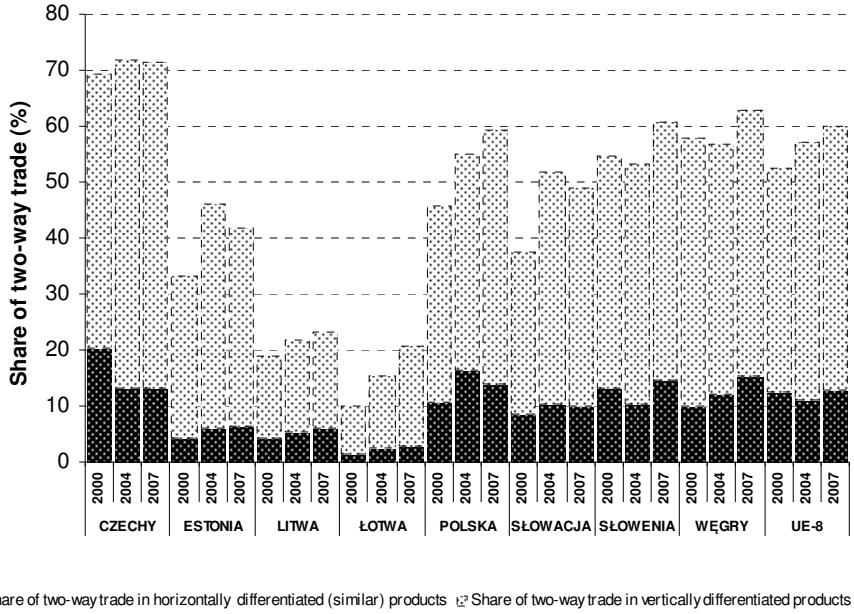
Reasons of low intensity of EU-8 intra-industry trade with EU-15 during the analyzed period were confirmed by the results of empirical analysis. As indicated by the analysis of growth of export and import shares in each group, new member states still did not bridge the technology gap - they increased only in group 1 (material-intensive goods) and 3 (capital-intensive goods) which are not characterized by the highest level of technological advancement. Adaptation processes to the SEM rules did not cause significant changes in trends of group 4 (technologically intensive goods, easy to imitate). Nevertheless, **some positive trends also shown up, which mean improving the EU-8 states` economic competitiveness and the substitutability of their economies inside the EU – intra-industry nature of the exchange enhanced by the upward trend of IIT values in all groups, and revealed comparative advantages in groups 2, 3 and 5**.

7. The analysis of EU-8 and EU-15 trade specialization by three trade types

On the basis of the results of calculations made for EU-8 trade with EU-15 in 2000, 2004 and 2007 (figure 1) it can be stated as follows:

- two-way trade included in 2007 already 60,0% of EU-8 trade with EU-15, with a remarkable increase even before the accession (52.5% in 2000 and 57.3% in 2004). It was however differentiated across countries – the largest in the Czech Republic, Slovenia, Poland and Hungary (59% - 71%), slightly smaller in Slovakia and Estonia (49% and 41%), and the smallest in Latvia and Lithuania (20% and 23%). Analyzing the dynamics of two-way trade shares for each country, it should be noted that in case of Latvia, Lithuania and Poland the improvement of indicators occurred both before and after accession. In turn, for the Czech Republic, Estonia and Slovakia a characteristic was the increase of indicators prior to accession and decrease in post-accession period. In other countries, so in Hungary and Slovenia, these indicators decreased before and increased after the accession;
- EU-8 intra-industry trade with EU-15, both before and after accession, was dominated by vertically differentiated good – so these flows were shaped mainly by the exchange of products that differ in quality;
- in post-accession period there was a significant increase in exchange of similar products (horizontally differentiated) - in 2007 it exceeded 21% of two-way trade between EU-8 and EU-15.

Figure 1. Two-way trade between EU-8 and EU-15 in the years 2007, 2004 and 2000 (%)



Source: Own calculations

EU-8 two-way trade with EU-15 in 2000-2007 was dominated by products of different quality, what proves the vertical specialization. However, after the accession there was increase in two-way trade in horizontally differentiated products (close quality substitutes). This may provide the emergence of the consumer preferences convergence process within the SEM and that increased effect of creation of intra-industry trade between EU-8 and EU-15 becomes apparent.

8. Conclusions

This paper was meant to analyze changes in the intensity of intra-industry trade between the new Member States (EU-8) and the EU-15 in the period of 2000-2007. The analysis focused on pre- and post-accession period indicated the following:

- significant increase of EU-8 intra-industry trade flows in almost all countries with EU-15. Levels of IIT indicators also presented that despite the upward trend in intra-industry trade still the main form of exchange between EU-8 countries and EU-15 was of inter-industry characteristic;

- meaningful increase of intra-industry trade flows in EU-8 mutual trade. Compared to the period before the accession, after that (between 2004-2007) there were more positive structural changes in the intra-EU-8 trade than in trade between EU-8 and EU-15. It seemed that this was due to the fact that before the accession the scope of trade liberalization resulting from the regional free trade agreements between EU-8 countries was lower than between EU-8 and EU-15. The removal of barriers to mutual trade of the EU-8 countries in 2004 had therefore a major impact on the formation of intra-industry flows among them;
- unchanged or smaller number of CN sections with the highest IIT values after the accession rather than in pre-accession period. Moreover, the highest IIT values both in pre- and post-accession period occurred in less than half of the analyzed CN sections in EU-8 mutual trade and external trade between EU-8 and EU-15;
- the tendency to increase the intensity of intra-industry trade within an increasing number of CN sections was significantly reinforced, both in the intra-EU-8 trade and in trade between EU-8 and EU-15. This process was accompanied by increased IIT indicators in the greater part of CN sections, especially that were characterized by the lowest IIT indicators in 2000;
- both in EU-8 external trade with EU-15, as well as in trade between the EU-8 countries there were sections of various countries with decreased IIT indicators in 2000-2007;
- the analysis of dynamics of intra-industry indicators within particular CN sections showed that within EU-8 mutual trade and trade between EU-8 and EU-15 its values in most sections before the accession (2000-2003) was higher than that in post-accession period (2004 -2007). Moreover, more often IIT dynamics after the accession was higher in EU-8 mutual trade than in trade between EU-8 and EU-15;

Further analysis examined the real adjustment processes in terms of EU-8 state economic competitiveness, which occurred during the pre- and post-accession period in order to prepare them to maximize the benefits of the Single European Market. The study confirmed that in the years 2000-2007:

- improvement of IIT indicators in trade between EU-8 and EU-15 was largely the result of transformations in the structure of commodity trade between these countries. In most EU-8 countries the pace of changes of IIT indicators in each groups of the intensity of using production factors differed significantly from that for total trade. In the case of group 1 (material-intensive goods) and group 4 (technology-intensive goods, easy to imitate) IIT indicators in most of EU-8 countries increased faster after accession than in pre-accession period (their dynamics in 2004-2007 were higher than in 2000-2004), thus, for these two product groups – material-intensive and technology-intensive goods, easy to imitate – changes in conditions for trade after the accession to the Single European Market had the most beneficial effect on their growing importance in intra-industry division of labor;

- upward trend in the intensity of intra-industry trade in most commodity sections and aggregates should be regarded as a positive direction of changes in the process of real adjustment of producers from the EU-8 to the requirements of the EU-15 market;
- trade in vertically differentiated products is still the major element of two-way trade flows between EU-8 and EU-15, simultaneously there was an increase in the share of two-way trade in horizontally differentiated products (close quality substitutes), which might provide the emergence of the consumer preferences convergence process within the Single Market and that increased effect of creation of intra-industry trade between EU-8 and EU-15 becomes apparent;
- the analysis of IIT and RCA indicators in trade between EU-8 and EU-15 showed positive trends, which in long term may result into increased competitiveness of the EU-8 countries within the EU. This, in turn, should facilitate better use of all positive effects resulting from the advantages of the single market.

References

Balassa, B. 1965: *Trade Liberalisation and Revealed Comparative Adventage*. The Manchester School, Vol. 3, 99-123. p.

Cieślik, A. 2003: *Handel wewnątrzgałęziowy Polski z Unią Europejską. Stan obecny i perspektywy*. In Michałek, J. J. - Siwiński, W. - Socha, M. (eds): *Od liberalizacji do integracji Polski z Unią Europejską. Mechanizmy i skutki gospodarcze*. PWN, Warszawa, 65. p.

Cieślik, A. – Śledziewska, K. 2003: *Krańcowy handel wewnątrzgałęziowy Polski z Unią Europejską: analiza procesów dostosowawczych w okresie przedakcesyjnym*. „Ekonomista", nr. 6.

Czarny, E. 2002: *Teoria i praktyka handlu wewnątrzgałęziowego*. SGH, Warszawa.

Fontagné, L. – Freudenberg, M. 1997: *Intra-Industry Trade: Methodological Issues Reconsidered*. CEPII, document de travail n° 97-01. p.

Jagiełło, M. 2003: *Wskaźniki międzynarodowej konkurencyjności gospodarki*. „Studia i Materiały". nr. 80, IKC HZ, Warszawa.

Marczewski, K. 2003: *Kierunki zmian w handlu zagranicznym Polski po przystąpieniu do Unii Europejskiej*. „Ekonomista", nr. 2.

Michałek, J. - Śledziewska-Kołodziejska, K. 2000: *Analiza i perspektywy handlu międzygałęziowego i wewnątrzgałęziowego oraz zmiany strukturalne handlu Polski z Unia Europejską.* In *Korzyści i koszty członkostwa Polski w Unii Europejskiej*. IKC HZ, Warszawa, 168-173. p.

Misala, J. 1985: *Wymiana wewnątrzgałęziowa w polskim handlu zagranicznym*. „Ekonomista", nr 3.

Misala, J. 1992: *Analiza porównawcza rozwoju gospodarczego Hiszpanii i Portugalii*. IKC, Warszawa, 93. p.

Misala, J. 2007: *Wymiana międzynarodowa i gospodarka światowa – teoria i mechanizmy funkcjonowania*. SGH w Warszawie, Warszawa, 287-309. p.

Misala, J. – Pluciński, E. M. 2000: *Handel wewnątrzgałęziowy. Teoria i praktyka*, Elipsa, Warszawa.

Molendowski, E. 2006: *Handel wewnątrzgałęziowy krajów CEFTA z państwami Unii Europejskiej*. „Gospodarka Narodowa", nr 5-6.

Molendowski, E. 2007: *Liberalizacja wymiany handlowej krajów Europy Środkowowschodniej w okresie transformacji ze szczególnym uwzględnieniem doświadczeń krajów. CEFTA*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.

Mroczek, W. – Rubaszek, M. 2003: *Determinanty polskiego handlu zagranicznego*. „Materiały i Studia", Zeszyt nr 161, NBP, Warszawa.

Pluciński, E. 2001: *Makroekonomia gospodarki otwartej. Wybrane zagadnienia teoretyczno-empiryczne z perspektywy integracji polskiej gospodarki z rynkiem Unii Europejskiej*. Elipsa, Warszawa 2001, 261-275. p.

Śledziewska-Kołodziejska, K. 1998: *Kształtowanie się handlu wewnątrzgałęziowego Polski, Czech i Węgier z Unią Europejską* [in:] *Rynki i regulacja. Polska u schyłku lat dziewięćdziesiątych*, Wydział Nauk Ekonomicznych Uniwersytetu Warszawskiego, Warszawa, 60. p.

Wysokińska, Z. 2001: *Konkurencyjność w międzynarodowym i globalnym handlu technologiami*. Wydawnictwo Naukowe PWN, 131. p.

Zielińska-Głębocka, A. 1996: *Handel krajów uprzemysłowionych w świetle teorii handlu międzynarodowego*. Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk, 102-104. p.

Perfect planning or the lack of effective lobbying? The analysis of participants of European gas market liberalization

Sarolta Somosi[1]

The EU has great medium term goals according to the common energy market. Its aim is to create a competition based productive, competitive gas market with respect to environmental aspects as well. Our aim is to examine with all of our restrictions, what (may) happen on this special field of energy market, closer on the gas market, by all the changes made within the framework of EU directed liberalization.

More precisely, the aim of this study is to find out whether it is possible to handle the EU as a unity and to create one energy policy for the 27 Member States? Using more than a dozen of indicators chosen mainly from OECD studies and national statistics of countries we would like to classify EU Member States into some possibly homogeneous groups. We try to find out, whether it is possible to make some homogeneous clusters, or there may be some Member States left, as outsiders, which do not produce the expected conditions for example in state control, public ownership, entry regulation, and so on? We try to answer on the question why the common energy policy is so sensitive topic, and why a common energy policy cannot really be created without its full perception by Member States? By the help of this research we would like to get closer to answering our basic question, which refers to that, whether it is possible to extract a real EU level competition by liberalization in a market that has so different characteristics among member countries?

Keywords: gas market liberalization, Member States, cluster analysis

1. Background of the research – European gas market liberalisation

If we look at the history of energy use, energy efficiency and its developments we can notice that the progress is intensifying and happening more and more rapidly by time passing. So what future will bring? Will we be burning fossil fuels, or will we use solar energy gathered high up in the atmosphere and beamed to us in the form of microwave radiation? The only reliable answer is that no one knows what will happen. Until someone will be able to answer these questions the participants, producers, suppliers and end users must adapt and adjust themselves to the actual

[1] Sarolta Somosi, assistant professor, University of Szeged, Faculty of Economics and Business Administration, Institute of Finances and International Economic Relations (Szeged)

circumstances. This is the reason why the EU has "only" medium term real goals considering its common energy market. It's several times published and emphasized aim to create a competition based productive, competitive gas market with respect to environmental aspects as well.

1.1. Economic reasons

Gas market liberalisation and the competition should solve all the problems of public supply in theory. Widely emphasized and accepted arguments are the *widening possibility of choices*, and the *improved quality services on lower prices* by the help of competition of suppliers. On the contrary the public property stands, which is often used as synonym of corruption and worsening quality. In reality we have to collect and add some more information and economic reasons in order to be able to decide which way is better to choose.

Among the upper mentioned economic reasons the first must be the reason of *competitiveness*. This criterion can be found both at micro – company or sectoral level – and of course macro – Member States and EU-wide – levels as well. This is not the place where we would like to decide which is more important. We must add that according to the theory of industrial organisations energy supply as a public service provided by a sector that is influenced by a natural monopoly seems to be inconsistent with the conditions of competitiveness. As a solution for this situation the EU does not expect Member States to "sell out" public properties, but to improve the market structure, at least by a separation.

In many Member States, where regulators decided to reorganise the originally monopolistic structure of gas sector, private investors have "great expectations" to gain higher profitability. The higher are these expectations are the bigger should be the *innovation constraint*. This pressure appears not only from the investors side, but from the fact that public services – be it energy services, telecommunication or water supply – are usually in bad conditions (Scheiring – Boda 2008).

Among economic reasons the following could be that these structures and services should or expected to be *run efficiently and in a competitive spirit*. A widely accepted solution is that a network could be more effective if competition is being created or let on at least some levels of the network, if not in the whole structure. The question is now changing, since we have to decide somehow that those structures seem to function "better" where market based competition characteristics appear or those where still natural monopolies dominate (Bakács 2003)?

1.2. EU policy reasons

It is easy to understand that with economic reasons mentioned in the previous paragraph the European Union must face and adapt itself to the present challenges and circumstances like the problems in the security of supply, higher gas prices, climate change and try to exploit as much from the market by the help of

liberalization and so competition as it can. Europe must go back to its roots and original aims to have a common answer on these problems which affect almost every Member States. Sustainability, security and competitiveness are not the ones which could be reached by single Member States themselves. Along with some other not less important goals the creation of a common energy strategy was also a central element of the European Coal and Steel Community (1952) and also of the European Atomic Energy Community (1957) (COM(2007) 1). Although conditions have changed a lot in the last more than 50 years, common answer, namely the creation of a common European gas market is more necessary and important than ever.

Unfortunately or not, the European Union seems to be a bit poor considering the reach of these acceptable goals. Although the beneficial effects of an integrated and common answer are all known and stressed regularly by the Commission and by scholars of energy policy *"Today's Union is still without what could be called Common Energy Policy."* (Pointvogl 2009, pp. 1) The European government has still tools only mainly in the sub- or co-areas of energy like environment, research, infrastructure and lately competition and Single Market. There can be several reasons lying behind the processes. The supranational level regulation's fate is sealed by the fact that until the 70-ies public services were not being regulated at Community level. Moreover, the EU has the right to regulate this field only by directives. The problem of using directives in regulating such a sensitive and strategically important sector of Member States is serious. Directives are said to be less effective in reaching goals compared to regulations, but unfortunately they are the most applicable tools in fields where Member States want to maintain some licenses and possibility of intervention.

Although they have adopted several directives in the field of common energy market, it does not seem to be enough to reach the EU's mid-term goals. The Council of the European Union has accepted the first two directives for the electricity (96/92/EC) and the gas market (98/30/EC) quite late. Later on because of ineffectiveness in decreasing monopolies' dominancy and unifying fragmented market two other directives replaced them in 2003 of which the more important from our point of view is the 2003/55/EC. Although development has started and there are positive effects of the liberalisation process, still not every advantages are being exploited in the market (COM(2009) 115). In the meantime Services Directive (2006/123/EC) became adopted in 2006, but it does not contain the complete regulation of services of general interest. The third package of measures must have been adopted because of the failure of the previous two attempts. The third package – with a report called *"An energy policy for Europe"* (COM(2007) 1) as a starting point – adopted by the Commission plans to ensure that all European citizens can take advantage of the numerous benefits provided by a truly competitive energy market. Consumer choice, fairer prices, cleaner energy and security of supply are at the centre of this third legislative package, adopted by the Commission on 19

September 2007. We have to add not only by the way, that the directive plans to decrease market concentration and the cutting up of giant vertically and both horizontally integrated energy providers in order to create symmetric and harmonized energy market. These steps seem to be necessary because of the above mentioned report published in 2007, which reflected highly concentrated market structure concerned state or regional levels, by which incumbent actors with the lack of effective liberalisation and competition are still able to control import/production, access of networks and maybe prices as well. We have to add that the success of these directives is highly influenced by the attitude of Member States and interested companies of the energy sector. The present state of legislation in the Member States will come up later.

1.3. Problems

Going ahead introducing the reasons of necessary gas market liberalisation we have to mention shortly the problems, of which solution seems to be only the creation of a really competitive and competition based market structure.

We all know that liberalisation does not mean deregulation but rather re-regulation in the sense that new regulations are needed in order to create the expected structure and competition. This is true in a greater extent for the market of public services. The most important reason for the liberalisation is that the state seems to be inefficient in providing these services. But we cannot forget that these markets seem to be so special that they need special handling as well. By the help of new regulation market failures like monopolies or abuse of dominant position can be handled and new coming actors can face better conditions than ever in order to be enough competitive compared to their incumbent competitors. Reduction of so called *market entry barriers* is the main task of decision makers, authorities, and of actors in order to let new comers entering the market. These barriers of market entry can be divided into three main groups. The group of *natural barriers* (1), covers mostly physical market access problems, and so delivery costs; the group of *artificial/administrative barriers* (2) of market entry means mostly active trade policy and some structural barriers that can come from the specific industry we are talking about. As we all know that the distances – because of decreasing transport costs – became less and less, and the EU, the WTO and so almost every countries of the world are somehow committed promoting liberal trade policies and towards the reduction of trade barriers, the first two groups of market entry barriers have less and less importance in reaching different markets. The rest of the conditions that may restrict trade can be mentioned as *strategic market entry barriers* (3). The importance of these latest mentioned barriers has increased along by the decreased importance of the first two groups (Török 2003). Their common characteristic is that they can be somehow related to strategic behaviour of actors of a market and so usually "only" national competition authorities and/or the Commission of the

European Union (DG Competition) deals with them. The problem that lies beneath this phenomena is that whether national governments on a sensitive field like the EU's common gas market accept their loose of control with resignation or they use their toolbar of competition regulation in order to maintain their influence. If they feel like it is necessary to hold this licence in their hand that will necessarily mean stronger resistance to every EU-level attempt that tries to liberalise the market and will appear as a legal barrier later.

Second type of problems in our opinion can be considered more important because in some aspects it is somewhat independent in short and medium-terms from every attempts and plans made by governments or by the EU itself. That is why *security of supply* of energy[2] – by acceptable reasons – has become one of the most popular keywords of political and economic debates. It is getting considered as the Holy Grail of common energy policy (Weisser 2007, Jamasb – Pollitt 2008). The Commission's Green Paper (COM(2000) 769) about European strategy for the security of energy supply drew attention to the worrying and growing level of dependence – it may reach 70 percentage by 2020 – on gas imports from sources outside the EU. The task of organising security of supply cannot be entrusted to just one player of the market. Member States are obliged to define the roles and responsibilities of all the players on the market with regard to security of supply (2004/67/EC). As it can be seen from the footnote definition security of supply has different sides and so different objectives to reach. One European-level *measure that may improve security of supply* would be the resolution of the above mentioned problem, namely the removal of barriers of supply-side competition and so creation of closer cross-border cooperation and interoperability of gas systems. From an infrastructural side further trade and investments within the internal gas market as well as interlinking of networks play a fundamental role in the flexibility of supplies. With these measures we could improve gas trade among Member States. Another measure that needs common answer but does not remain within the borders is the reduction of dependency from a single supplier. In this research further on we plan to map only present possibilities and not the future ones.

Both two above mentioned problems are so serious, complex and interconnected, that require common answer from the part of the EU. A problem or difficulty coming from this fact is only that whether Member States are willing to renounce their licenses in favour of a common policy. Without such a cooperative behaviour common energy policy seems to be impossible to create!

[2] According to the Communication from the Commission COM (1999) 571 security of gas supply in short-term includes the ability to maintain continuity of gas supply despite exceptional demand and difficult supply conditions including possible disruptions of gas supply whether of a technical, economic or political nature. In longer-term security of gas supply is the ability to ensure that future gas demand can be rated by a combination of indigenous and imported gas supplies. This requires adequate investments in production, transmission infrastructure and supply diversity and clearly has a geopolitical dimension.

2. Research – the analysis of participants of European gas market liberalisation

Our research is based on an assumption – made by previous experiences and by the reactions and answers given on the upper mentioned problems – that there are huge differences among EU Member States considering for example their gas market structure, the level of state intervention, the security of supply, the sources of energy and some other important factors. So our research is based on a somehow theoretical identification of the conditions in which there can be differences existing, and so defining different groups, parties among Member States.

The reason of our choice is based on the basic differences between the gas and the electricity sector. Although they are highly bounded and related to each other, we have to distinguish them. Meanwhile electricity is a produced good and every Member State and even households could be able to generate it by traditional tools, or using the "greenest", recently supported resources, natural gas is a type of energy that can be found in only one half of the Member States but far not in an enough amount neither related to their present nor to their future demand.

Our other reason for the object of our research is that what is presented in the study of Brekke et al (2008). They examined the net welfare effects of the liberalisation process both in the electricity and in the gas sector and found that the positive effects of liberalising domestic gas markets are much smaller compared to the other sector.

2.1. Basic assumptions

The idea that Member States are not in the same position to give the same answers and follow the same methods is quite old. It can be originated from the Tindemans-report (1976), but nowadays it is more actual than ever. We could hear it in connection with the Constitution of Europe, but also related to the Economic and Monetary Union, connected to general economic policy and the economic crises, EU-level institutional organization or in general in debates over the future of the European Union (Csaba 2006, Marján 2007).

This made us think over this phenomenon in connection with one special and sensitive market, the gas market. There are several reasons lying beneath the reaction of countries. We will now focus on those which can be basis of an objective measuring and try to exclude those which are coming from the sensitiveness of the sector.

There are three studies or researches which lead us on our way to define our questions, create the factors and summarize our results. First tool was the so called

Railway Liberalisation-index of IBM from 2004[3]. As a reason of the study, we have to mention that similarly to the gas sector, one of the main objectives of the European Union is the integration of national rail markets into a harmonized domestic European rail transport market. In view of this goal, the coexistence of de facto closed and open national rail markets was and is unacceptable. The main is to analyze how open the rail sector is to new competitors. For this reason, when analyzing the markets, the LIB Index[4] concentrated on the point of view of external railway undertakings (IBM 2007).

The second instrument dealing with market differences in a specific sector is closer to our topic. The Energy Policy Index (EPI) created by Röller et al. (2007) became published by Bruegel in the *Energy: Choices for Europe*. They have investigated the status quo of Europe's energy policy – both gas and electricity(!) markets –, in terms of the three energy policy objectives: *competitiveness, security of supply* and *environmental sustainability*. In order to get a comparable measure across Member States, they have constructed the above mentioned indicator which gives a position of each country in relation to each of the objectives. On the basis of the Energy Policy Index, they made a cluster analysis, by which they could group countries according to the three objectives. From this analysis they also noticed that national starting-points vary considerably.

The third analysis made by Pointvogl (2009) was based on a research focusing on the driving forces of the integration and on the differences revealed among member states of European energy policies, with the title *"Perceptions, realities, concession"*.

Not exactly assumptions, but many ideas and data came from the *Ten years of product market reform in OECD countries – insights from a revised PMR indicator* OECD Working Papers written by Wölfl et al. in 2009 and some previous ones like the work of Conway - Nicoletti (2006). We could obtain many ideas and influences creating the factors based on their approach.

2.2. Objectives

By the help of the above mentioned studies we tried to focus exactly on the gas market of the EU Member States. We could create three groups where the

[3] In the second edition of the 'Railway Liberalization Index,' the basis of the analysis was extended to include the new EU member states, apart from Cyprus and Malta, and of course Romania and Bulgaria at that time.

[4] The LIB Index comprised two sub-indices. The first, the LEX Index, analyzed to what extent legislation allows market access (law in the books). It therefore contains the organizational structures of the incumbent, the regulation of market access and the competencies of the regulatory authority. The second sub-index, known as the ACCESS Index, measures numerous market access barriers (law in action) encountered in practice as well as the percentage of accessible domestic market. A separate index, the COM Index, compares the competitive situation in the countries analyzed (IBM 2007).

answers of a Member State for different challenges may be interesting and determining in our opinion. This research is intended to be an exploratory analysis, focusing on those dimensions and factors that determine member states' approaches towards integration, and their sometimes contrasting view.

As a first group we planned to *examine the conditions and the circumstances that determine security of supply (SoS)* (1). In our opinion these are indicators that can be changed only in longer term, and so we can consider them constant in a short-time period. *Examination of different vertical levels and market structure of the gas sector* (2) is the second group that contains some information about production/import, transmission/distribution and providing service towards consumers. As we think, these conditions can be changed in medium term. Finally the *examination of government involvement* (3) contains important facts about the actual position of a market. These conditions can be mentioned as legal market entry barriers which are almost related only to government decisions and interests.

2.3. Sources of data

We agree with Pointvogl (2009) and know that the gap between the trinity of the EU's policy goals – competitiveness, security of supply and environmental sustainability – and the strategic behaviour of involved players – governments, companies, authorities – risk the success of a deep and updated analysis of gas market in itself. This is further more complicated in a period when member states are in their integration process.

Our work is theoretical as we have already mentioned, because we are convinced that it could function in practice if all the relevant information were published and publicly accessible at the same time. Unfortunately we have met many difficulties in collecting relevant data. We missed many updated information related to present state of different markets. We also met the problem of higher level data protection in connection with such a strategic sector and activity of countries or companies. Unfortunately competition authorities are also missing some relevant information referring to market shares that would be useful for them in investigating horizontal and vertical cartels in the sector.

Significant amount of the information we could collect from OECD Working Paper 36 (Wölfl et al. 2009), and from a DG TREN Staff Working Document (SEC(2009) 287) and information that refers to constant data like consumption and dependency we could find in Energy Pocket Book (CEC 2009). Unfortunately there were still important data missing, so we used Internal Market Fact Sheets of the Member States, Network Country factsheets of International Energy Regulation, the statistical basis of International Energy Agency and of Statistisches Bundesamt Deutschland if needed.

3. Methodology

As we have collected and – in case of missing ones – calculated all data we considered being important for a brief analysis of the gas sector, we had to decide the method of the comparison and of the grouping of Member States to reveal the differences and the similarities among them. In order to be able to perform the chosen method, cluster analysis we were forced to create the factors that constitute the basis of the comparison.

3.1. The variables

For the performance of a successful cluster analysis we had to collect those data, which tell us the most accessible information we could ever get. We have collected several, here not represented information considering the gas markets of the EU, but we have faced the problem of missing information so many times, that we had to reduce the number of factors used in the analysis. Finally – as Table 1 shows it – we have chosen those where the information were the most reliable and the most comprehensive. Having seen the remained variables, we created three groups of them; one refers to the basic conditions a country can face, other deals only with the structure of the market, and the third group of variables will give us information about the government involvement and the legal barriers of market entry. We have to add that we agree with Jamasb - Pollitt (2008) and Pointvogl (2009) who stress that variables though being exogenous are not absolutely independent from each other and from influences.

3.1.1. Conditions

This first group of variables cover mostly our first objective, namely to reveal and stress the differences in the basic circumstances if they exist at all. We consider these conditions important, because there are some which cannot be basis of a change in short terms, but these are almost the only ones determining security of supply[5].

So our first variable among this group is *Total energy consumption per capita* (it is measured in kg or equivalent per person). Our second, but maybe more important indicator is *Gas Import dependency*. It is given in percentage and getting calculated as Net Imports is being divided by the sum of Bunkers and Gross Inland Consumption. The third content of the first group of variables is *Diversity*. In our

[5] Of course we have to add that for example in Hungary compared to the previous three years a significant decrease could be measured in gas consumption when gas price increased in the middle of 2006. That shows a slight consciousness from the part of the population, and the earlier useless(?) wasting of energy, if prices are not forcing them to savings (Kaderják 2009).

opinion this is also an important indicator, since it shows the number of possible sources covering import necessity. A country's one-sided energy defencelessness could hold many problems in the supply side, as we have seen it two or three times now in connection with Europe's high dependency from Russian sources. As Jamasb and Pollitt stress it, *"The most commonly suggested remedy for increasing security of supply is diversification, to reduce overdependence on individual countries or regions, and on particular types of supply."* (Jamasb – Pollitt 2008, pp. 4585). In a way we could make further difference in the exact meaning of the word diversity. According to the cited sentence it may contain also various types of energy sources, so maybe Weisser's word "optionality" – the degree of viable alternative options – represents our intents better (Weisser 2007).

Our present research does not contain it, but in the future we think that measuring and comparing *Storage capacity* can be also useful in getting a clear picture of the market conditions.

3.1.2. Market structure

Within the second group of variables we have collected indicators which give us information about both the market share of the largest company in the gas production/import, the gas transmission and the gas supply industry.

With the first indicator – *Market structure* – we could get a quite broad view about the whole market, but we have collected further information as a control for each sub-market. The variable of *Number of companies with over 5% share of production/import capacity*, the *Share of 3 biggest companies (by available gas) (%)* and the *Share of 3 largest wholesalers in wholesale market (%)* give us information about the structure of gas import and production. The last three indicators – *Number of independent suppliers, Companies with market share over 5%* and *Market share of 3 largest companies in whole retail market (%)* – show the structure of gas retail market. Unfortunately most of the necessary estimations were made within these data.

3.1.3. Barriers to competition

As we have mentioned it above this group of indicators hold information mainly about the existence of legal barriers and about the level of government involvement in the different levels of the sector.

Scope of public enterprise sector contains information on the fact whether national, state or provincial government controls at least one firm in gas sector (gas production/import or gas transmission or gas distribution or gas supply). *Government involvement in network sector* reveals the percentage of shares owned

by government in the largest firm in the gas production/import, or gas transmission or gas distribution or gas supply sector. The *existence of price controls* gives us information about the practice of regulated tariffs in case of industrial users, small commercial users or households. We consider this variable to be important since price control represents one of the most common instruments for the regulation of public services (Dorigoni – Portatadino 2009). *Legal Barriers to entry* discovers those national, state or provincial laws or other regulations that may restrict the number of competitors allowed to operate a business in at least some markets in gas sector. *Antitrust exemptions for public enterprises or state-mandated actions* contains the information about existing rule or principle providing for exclusion or exemption from liability under the general competition law for conduct that is required or authorized by other government authority (in addition to exclusions that might apply to complete sectors). *Entry regulation in Gas industry* is a composite index of relevant information. It contains answer on the question how are the terms and conditions of third party access (TPA) to the gas transmission grid determined (1), what percentage of the retail market is open to consumer choice (2), and about whether national, state or provincial laws or other regulations restrict the number of competitors allowed to operate a business in at least some markets in the sector: gas production/import (3). The information about the *proportion of market open to competition* also belongs to this group of questions, but refers to the entire sector. *Vertical integration in gas industry* is also a composite index, that contains answers for three questions referring to the question of the degree of vertical separation between gas production/import and the other segments of the industry (1), the degree of vertical separation between gas supply and the other segments of the industry (2) and on the question whether gas distribution is vertically separate from gas supply (3).

Table 1. Summarizing list containing the three groups of variables used in the research

Variables of Conditions	***Variables of Market structure***	***Variables of Barriers to Competition***
Total energy consumption per capita	Market structure	Scope of public enterprise sector
Gas Import dependency	Number of companies with over 5% share of production/import capacity	Government involvement in network sector
Diversity/optionality	Share of 3 biggest companies (by available gas) (%)	Existence of price controls
	Share of 3 largest wholesalers in wholesale market (%)	Legal Barriers to entry
	Number of independent suppliers	Antitrust exemptions for public enterprises or state-mandated actions
	Companies with market share over 5%	Entry regulation in Gas industry
	Market share of 3 largest companies in whole retail market (%)	Proportion of market open to competition
		Vertical integration in gas industry

Source: own creation

3.2. Cluster analysis

In order to find out whether countries can be grouped according to their conditions and characteristics, we perform a cluster analysis by the help of the huge amount of standardised data. It is important to stress that the groupings which came out of the analysis are rather indicative and, of course, could be subject of further qualitative assessment of the data gathered from several sources.

Cluster analysis is an exploratory data analysis technique which attempts to identify natural groupings (clusters). We performed a hierarchical cluster analysis using the data of 23 Member States for creating natural groupings. Only 23 countries' data is involved. Because of the missing updated information we had to use data from 2006 until 2008. Almost every statistics from this period missed Romania and Bulgaria as new Member States. The explanation of the other two "missing" countries is more trivial. Because of climatic conditions and geographical background Cyprus and Malta simply do not use gas as energy (EUROSTAT 2009).

4. Results of the research

In order to find out whether countries can be grouped according to their conditions and characteristics, we perform a cluster analysis by the help of the huge amount of standardised data. It is important to stress that the groupings which came out of the analysis are rather indicative and, of course, could be subject of further qualitative assessment of the data gathered from several sources.

Cluster analysis is an exploratory data analysis technique which attempts to identify natural groupings (clusters). We performed a hierarchical cluster analysis using the data of 23 Member States for creating natural groupings. Only 23 countries' data is involved. Because of the missing updated information we had to use data from 2006 until 2008. Almost every statistics from this period missed Romania and Bulgaria as new Member States. The explanation of the other two "missing" countries is more trivial. Because of climatic conditions and geographical background Cyprus and Malta simply do not use gas as energy (EUROSTAT 2009).

4.1. The five clusters

On the basis of the collected data and relevant information and using hierarchical cluster analysis with Ward methods, we could classify Member States into five groups. The identification of the number of clusters was based on the coefficients of Agglomeration Schedule.

4.1.1. A "unique group"

We start our examination with the end of the results. We have chosen a 5 clustered outcome because we knew that there may be some potential states that can be unique in some senses. By this way the first group contains only Latvia.

It is easy to understand the result if we see the characteristics of this country. Small population with high energy consumption, which is combined with high import dependency (108.8%) which is covered from Russian sources only. It has a small concentrated market with a vertically integrated state owned company, which means high level of government involvement. As a summary in our point of view we could say that the Latvian gas sector is not considered to be mature in the sense of liberalisation, competition and development.

4.1.2. Group of states that need further development

The characteristics of *"Group 2"* that consists of Czech Republic, Estonia, Finland, Hungary, Luxembourg and Sweden are quite low energy consumption, with an exception of Luxembourg with its highest 10137.8 kgoe per person total energy

consumption. With the rate of 82.2% gas import dependency proved to be the lowest in Hungary. Although the average level of government involvement seems to be lower than in other groups, compared to other clusters' members, the overall market structure is not "developed" enough in the sense of level of competition. So concentrated market structure, medium level government involvement and vertical integration are combined with weak diversity position (except in Luxembourg).

4.1.3. Group of "country-mix"

For us, this third group is quite surprising since the members vary, and being so in general no one would put them together in one "party" in a debate over common energy policy. *"Group 3"* contains: Austria, Belgium, France, Greece, Ireland, Lithuania, Poland, Portugal, Slovakia, Slovenia and Spain.

As we have mentioned the characteristics of this group are varying somehow. For example in case of Total energy consumption we can find the lowest in Poland (2429 kgoe per person) and the highest in Belgium (5891.7 kgoe per person). Gas import dependency is the lowest again in Poland (around 71.9%) and the highest in Spain (101.3%). Only Ireland and Lithuania are the two countries of this group who can cover their natural gas demand from one source according to our data. The lowest level of government involvement we can find in Spain, meanwhile Ireland maintains the highest level of it that is also combined with the highest vertical integration in gas industry among the members of the group. Compared to the previous group they seem to be in better position because of their market structure.

4.1.4. Group of "good basics"

The fourth group consists of only Denmark and the Netherlands. Their medium level energy consumption is combined with low dependency, since they belong to that minority within the EU who have significant domestic resources considering natural gas. From the number of players being active in each level of the sector we can see that competition already exists, although government involvement is not the lowest compared to other clusters. Their market structure seems to be competitive, liberalised and so developed. From some aspects United Kingdom could also belong to this group, since it represents quite the same characteristics.

4.1.5. Group of "developed" nations

The last but not least group became the triad of Germany, Italy and the United Kingdom. They perform medium level energy consumption, around 90 percentage energy dependency (except the UK with its 11.8%), already competitive market

structure, and liberalised market with the highest number of independent companies in the supply side, and with higher government involvement only in Italy.

Although we can find differences among them, there are some common interests lying behind in case of large countries, or countries with relatively low – compared to other groups – foreign dependency. They might not derive significant benefits from giving a European dimension to their external policy (Röller et al 2007). Albeit within these countries almost everybody can mention at least one so called national champion to which companies' high level of national interests and emotion belongs, these are the countries that represent good performance if we measure development through liberalisation process and competitive state of the sector.

4.2. Positions of Member States related to each other

Multidimensional scaling gives the geometrical representation of our objects in a lower dimension but with the maintenance of the order and of original distances. Considering the goodness of fit S-stress will give the basis of the success and the possibility of interpretation of the outputs[6].

One dimensional scaling technique will provide the possibility of the creation of a rank of development if S-stress less then .2 and the dimension is definable. The value of S-stress of the examination is 0.12227, that can be considered good, so the model of the reduced number of dimensions seems to hold every relevant information. According to correlation between and the indicators providing the basis of MDS, we have to define to dimension. From Table 2, we can see the significant correlations between mds1 and those variables that seemed to be significant from the point of our research.

[6] Results of the MDS are considered to be good if S-stress shows a value lower than 0.1. If the value is between 0.1 and 0.2 the results could be acceptable, meanwhile in case of a value higher than 0.2 the output is not able to define.

Table 2. Correlations among those variables that seemed to be significant from the point of the research

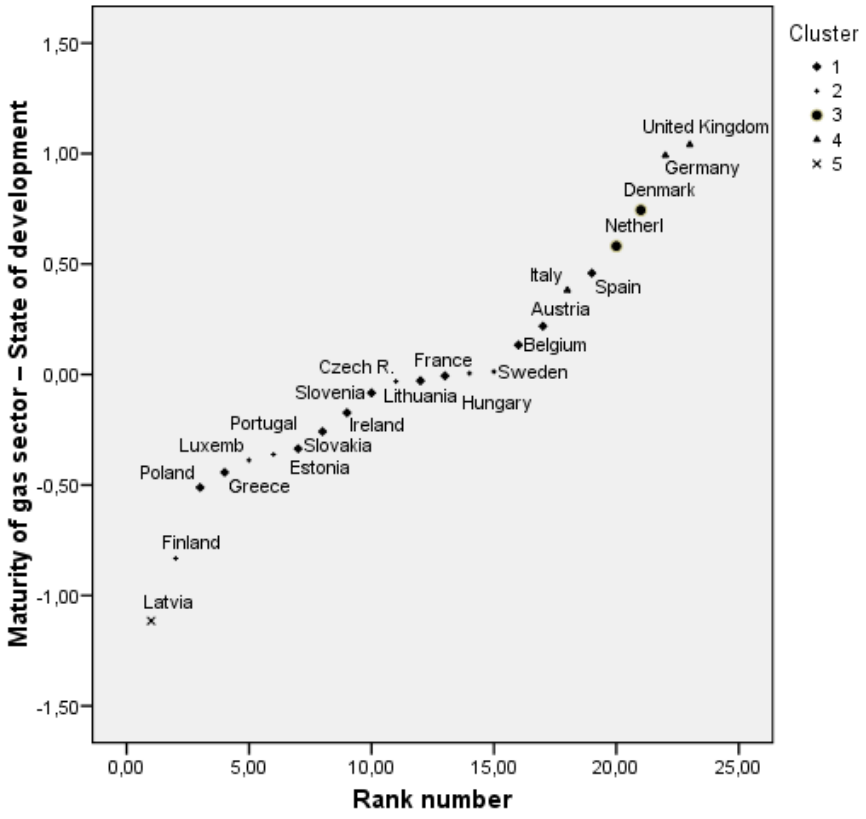
		mds1
Total energy Consumption (per capita)	Pearson Correlation Sig. (2-tailed)	.623(**) .001
Gas Import dependency (in %)	Pearson Correlation Sig. (2-tailed)	-.552(**) .006
Diversity: number of possible sources covering import necessity	Pearson Correlation Sig. (2-tailed)	.548(**) .007
Market Structure	Pearson Correlation Sig. (2-tailed)	-.822(**) .000
Number of companies with over 5% share of production/import capacity in 2007	Pearson Correlation Sig. (2-tailed)	.852(**) .000
Share of 3 biggest companies (by available gas) (%) in 2007	Pearson Correlation Sig. (2-tailed)	-.675(**) .000
Market share of 3 largest companies in whole retail market (%)	Pearson Correlation Sig. (2-tailed)	-.496(*) .016
Government involvement in network sector	Pearson Correlation Sig. (2-tailed)	-.418(*) .047
Proportion of market open to competition (2007)	Pearson Correlation Sig. (2-tailed)	.581(**) .004
Vertical integration in gas industry	Pearson Correlation Sig. (2-tailed)	-.535(**) .008

Source: own creation

**Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed)

As we can see bigger part of the variables determine the mds1. Analysing these correlations, we can point out that the dimension mds1 could really be considered as an indicator of development, by which we can determine the order of Member States' by the state of their gas markets. We must handle these results carefully since the country represented with twice bigger MDS coordinates does not mean two times more developed country on the scale of gas market maturity. So for getting the exact positions every Member States shall get a rank number (Figure 1). If two countries' MDS coordinates seem to represent the same value, we have to apply the mean rank.

Figure 1. Ranking of Member States according to the maturity of their gas sector



Source: own creation

We think that state of development is somehow related to the three stages of evolution of gas markets detailed by Austvik (2009). According to our expectations the ranking is led by the United Kingdom, who is often mentioned as one – beside the non-EU member USA – of the good examples considering energy market liberalisation. Those countries follow it where *"The monopolistic position of the transmission companies becomes less predominant, and market transactions are more diversified. Producers and customers have to a degree more purchasers to choose from (gas-to-gas competition). At the same time, however, companies start integrating horizontally (with competing firms) and vertically (with firms further up or down in the gas chain) by mergers and acquisitions, which may contribute to higher concentration around the large champions, if allowed by competition authorities.*" (Austvik 2009, pp. 92.). So countries which are quite mature in the

sense that there is some competition with alternative sources, routes of transport, and where competition within the market appears to lead the ranking of our research.

The order is being closed by Latvia and those countries where there is a need for investments both upstream and downstream to improve competition at markets. In case of countries from the second half (like Czech Republic, Estonia, Finland, Greece, Latvia, Poland, Slovakia, etc.) with low domestic resources and few possible sources covering their high import dependency, concentrated market structure, medium level government involvement and vertical integration are combined with weak diversity position. They might see the creation of domestic competition as a danger to their security of supply (Röller et al. 2007). Some small central European and Baltic countries could find themselves in a weaker position when negotiating with foreign upstream suppliers, and might be more eager to face such negotiations under a protective European umbrella.

The groups represented with different shapes in Figure 1 could provide the basis of understanding the differences among the cooperative behaviour in energy policies and reactions on common energy policy of the Member States. By the help of this indicative result we can focus on special characteristics and demands of countries with different energy conditions. We have to add that we consider the result only indicative that could be subject of further examination, since many things and conditions on this market changed during the last 2-3 years and are changing nowadays as well. The study must be updated regularly in order to get precise basis for the creation of an effective and workable policy.

5. Conclusions

Our aim, as we have mentioned it previously, was to reveal the differences among Member States if they exist at all. Our second objective was to find a possible classification to see the possible perceptions, common perspectives and negotiating positions of countries that are close to each other not geographically but in their gas market characteristics. We found this important because the debate on a common European energy policy cannot simply ignore the current situation of each Member State.

In fact we agree with Röller et al. (2007) that countries might find conflicts in pursuing all three energy policy objectives – *competitiveness, security of supply* and *environmental sustainability* – at the same time and at least in the short term, might be confronted with a number of trade-offs in favour of each, on each fields of energy sector. Although our examination concentrated only to the first objective, we can see from the results that policies designed to increase efficiency, secure supply and protect the environment might not necessarily be complementary, driven and fulfilled together at the same time. Increasing the strength of one might require relaxing the pursuit of other objectives.

As a result, we have found that there are significant differences among the state of development of Member States' gas sectors, considering their basic circumstances and the strength of market players – independent companies or governments – as well. At this point we agree with the statement of Pointvogl (2009) who found that the perception of Member States drives the integration of European energy policies. We think that this perception is highly determined by their state of development considering each sub-sectors of the energy industry. Our intention was to reveal these differences in especially the gas sector.

The conflict of deciding about the order of meeting the objectives is further more increased by the liberalisation "push" from the part of the EU. In many Member States this requirement faces old industrial model with high amount and vide range of market entry barriers, and with high level of dependency towards foreign countries. We join to the opinion of Vissi (2006) and Kaderják (2009) who say that the conditions of competition of general interest are unfortunately missing both at wholesale and retail level of the market in several Central and Eastern European countries or telling the truth in countries with low-level of diversity who are represented in almost every groups. Until the number of the input possibilities – as we have mentioned the level of diversity or optionality – will not increase, real competition and so competitive market is impossible to be obtained by any kind of force.

Our last but not least question is that does it make sense and is it possible liberalising on a market with such different conditions? The diversity of the input and the size and dynamism of wholesale market decides whether there is enough choice and prices competition possibility on a market. We risk asking that whether it makes sense liberalising until in many Member States dependency is so high and diversity of the input are so scarce?

References

Austvik, O. G. 2009: EU Natural Gas Market Liberalization and Long-term Security-of-supply and –Demand. In Gunnar Ferman (ed): *The Political Economy of Energy in Europe: Forces of Integration and Fragmentation.* Berliner Wissenshafts Verlag GmbH, October, pp. 85-118.

Bakács, A. 2003: *Versenyképességi koncepciók.* A Gazdasági és Közlekedési Minisztérium által finanszírozott, *„A hazai közlekedési hálózatok hatékonysága, versenyképessége növelésének lehetőségei a nemzetközi tapasztalatok alapján"* című kutatási projekt keretében készült tanulmány. MTA Világgazdasági Kutatóintézet, Budapest. http://www.vki.hu/~tfleisch/~haver/szakirodalom/haver-BAKACS-final-031109.pdf Letöltve: 2009. július 28.

Brekke, K. A. – Golombek, R. – Kittelsen, S. A. C. 2008: *Is electricity more important than natural gas? Partial liberalization of the Western European energy markets.* Memorandum No 1/2008, Department of Economics, University of Oslo, Oslo.

CEC 1999: *Security of EU gas supply.* Communication from the Commission of the European Communities COM (1999) 571 (final), Brussels.

CEC 2000: *Towards a European strategy for the security of energy supply.* Commission Green Paper of 29 November 2000 COM(2000) 769 (final), Brussels.

CEC 2007: *An energy policy for Europe.* Communication from the Commission to the European Council and the European Parliament of 10 January 2007, COM(2007) 1 (final), Brussels.

CEC 2009: *EU Energy in Figures 2007/2008*. Commission of the European Communities, Directorate-General for Energy and Transport (DG TREN), Brussels.

CEC 2009: *Report on progress in creating the internal gas and electricity market.* Communication from the Commission to the European Council and the European Parliament of 11 March 2009, COM(2009) 115 (final), Brussels.

Conway, P. – Nicoletti, G. 2006: *Product Market Regulation in non-manufacturing sectors in OECD countries: measurement and highlights.* OECD Economics Department Working Paper No.530., OECD, Paris.

Council Directive 2004/67/EC of 26 April 2004 concerning measures to safeguard security of natural gas supply, Brussels.

Csaba, L. 2006: *A fölemelkedő Európa.* Akadémiai Kiadó, Budapest.

Dorigoni, S. – Portatadino, S. 2009: Natural gas distribution in Italy: When competition does not help the market. *Elsevier, Utilities Policy* 17, pp. 245-257., doi:10.1016/j.Jup.2009.03.001

IBM 2007: *Rail Liberalisation Index 2007. Market opening: comparison of the rail markets of the Member States of the European Union, Switzerland and Norway.* IBM Global Business Services, Brussels.www.db.de/liberalisation-index

Jamasb, T. – Pollitt, M. 2008: Security of supply and regulation of energy networks, Energy Policy, Vol 36, 4584-4589 p.

Kaderják, P. 2009: *Energiafüggőség Közép és Délkelet-Európában.* Regional Centre for Energy Policy Research (REKK) Presentation hold February 19th 2009 in Szeged.

Marján, A. 2007: *Európa sorsa. Az öreg hölgy és a bika.* HVG Kiadó Zrt., Budapest.

Pointvogl, A. 2009: Perceptions, realities, concession – What is driving the integration of European energy policies? *Elsevier, Energy Policy*, August, doi:10.1016/j.enpol.2009.08.035

Röller, L. H. – Delgado, J. – Friederiszik, H. W. 2007: *Energy: Choices for Europe.* Bruegel Blueprint Series, Bruegel.

Scheiring, G. – Boda, Zs. 2008: *Gazdálkodj okosan! A privatizáció és a közszolgáltatások politikája.* Új Mandátum Könyvkiadó, Budapest.

SEC 2009: *Report on Progress in Creating the Internal Gas and Electricity Market.* Technical Annex to the Communication from the Commission to the Council and the European Parliament {COM(2009)115} Directorate-General for Energy and Transport (DG TREN) Staff Working Document No.287, Brussels.

Török, Á. 2003: A piacra lépési korlátok átalakulása és a világgazdaság kibontakozása. Kísérlet a kapcsolatteremtésre a piaci szerkezetek elmélete és a nemzetközi kereskedelem története között. *Közgazdasági Szemle*, L. évf., március, 195–208. o.

Vissi, F. 2006: *Verseny – energiapiacok.* Készült a Vezetékes Energiaellátási Tevékenységet vizsgáló Szakértői Bizottság részére, Budapest. http://misc.meh.hu/letoltheto/energiajelentes_melleklet_vissi.pdf Letöltve: 2009. augusztus 25.

Weisser, H. 2007: The security of gas supply – a critical issue for Europe? *Elsevier, Energy Policy* 35, December, 1-5, doi: 10.1016/j.enpol.2005.10.002

Wölfl, A. – Wanner, I. – Kozluk, T – Nicoletti, G. 2009: *Ten Years of Product Market Reform in OECD Countries – Insights of a Revised PMR Indicator.* Economics Department Working Papers No.695., OECD, Paris.

The impact of global economic crisis on the operations of Montenegrin economy

Mersad Mujevi[1]-Mirsad Nuković[1]

It is more than obvious that the world has become a global village. No economy is fully independent. The most influential countries, such as the United States of America, Canada, China, India, Russia, Japan, Great Britain etc., depend on global trends and oscillations in the world market.

There is an everyday dispute among world leaders, experts, managers of financial corporations, on how deep this crisis is going to impact the world economy. According to the opinion of the majority, it is biggest since the Great Depression from the early 30s of the last century. Financial crisis led to a serious crisis and recession, hence it resulted in decceleration of economic activities and exhaustion of one model of economy. From many countries, including ours, we can almost everyday hear the news on dismissal of workers, reduced production, limitation of spending etc., which represent basic elements and indicators of recession.

Since the crisis emerged, the situation has significantly changed, both in the country and abroad. The money has become more expensive, there are much less foreign investments, the real sector is gradually entering the recession in the majority of developed countries, which makes export more difficult for domestic producers. In addition to that, the investors are now, and they will be even more in the forthcoming period, much more cautious when it comes to investing their funds.

What is the response to recession?

Financial crisis and global recession in the world are just a chance for the development of enterpreneurship. Montenegro, dependent on import and being a small economy, just like many countries with similar features, can ensure mitigation of the crisis consequences by stimulating and developing the enterpreneurship area. Opening of small and medium enterprises would allow the creation of products intended to fulfill the needs of the population and, at the same time, it would reduce dependence on the import and ensure economic development (Adizes 2009). In the environment of pessimistic reports on the current situation and future perspectives of the world economy, the enterpreneurship and enterpreneurs appear as a response to the requirements of recession, growth and development. The enterpreneurs use changes as a trump card and they use them in the right way and at the right place in order to achieve their business breakthroughs. Upon that, technological improvements and innovations, although significant, are not the only basis for

[1] University of Novi Pazar, Faculty of Informatics and Information Technology

the ovrall technological progress. They can also include organizational improvements, innovations of marketing pr administrative procedures, which lead to business success, that is, to the changes in the environment in the form of creative destruction. All these changes are carried out in the purpose of placing the high quality goods and services at acceptable prices on the market.

The job opportunities are no longer in the public or local sector, but they are rather in the private sector. The private sector is much more flexible and that makes it more resistant to recessions. That is the case of a Japanese steel processing plant, which faced a reduced demand of its products from the basic assortment, and virtually over night it was reorganized in order to preserve its employees and its profit. The plant identified the demand on the market, reoriented its production, and now one half of the plant produces season salad, which it places on the local market. As you can see, it was a diametrical restructuring.

That kind of flexibility is not achieved over night. It is obvious that the complete program had been designed as an alternative solution long before the recession occured. In that regard, one must bear in mind that the crisis could not be ''survived'' without a strategical and fundamentally new way of thinking and approach. For each product or service, program, project - there has to be an alternative. An optimal alternative always has to be analyzed in the environment of the recession as the one which will ensure efficacy and effectiveness when the recession is over, and the normal economic situation takes its place.

Also, the employees in the government and in the local authorities may have their opportunities in launching their own businesses. The current crisis and possible threats that the employees could be dismissed from their jobs, in the conditions when the chances for getting a new job are scarce, put us all in a temptation and check the character and strength of a personality. Those who are positive, proactive, and who believe in the power of the entrepreneurial spirit and their own ability, think of turning that loss of a job into a new challenge and new opportunity, launching their own businesses and finding alternative jobs. An attitude that can often be heard – that the global crisis is in fact a new opportunity – describes just those who are thinking of starting their business and test themselves in the entrepreneurship. Upon that, one must bear in mind the following:

- *You must plan!*
- *Invest in yourselves now!*
- *Maintain and take care of the long-term investments*
- *Act like an entrepreneur and an innovator*
- *Spend less*
- *Reexamine your liaisons, contacts and*
- *Don't panic!*
- *Create a demand*
- *Separate your company from your personal or family expenditures*

Running your own business is much more complex than working for somebody else. If you lose your business or fail, you can lose much more than just a job. At the same time, your own business provides you with much more freedom, opportunities and possibilities. That is why we hope that these few pieces of advice will serve as a brief and a simple tool, intended to warn you and to help you.

1. The impact of global economic crisis on the operations of Montenegrin economy

The negative impact made by global financial and economic crisis onto the Montenegrin economy at the end of previous year was continued in this year as well, and it is reflected in the slowing down the overall economic activities in the real sector, in solvency downfall of enterprises and banks, a high level of mutual liabilities and claims, limited credit support to the business, decreasing recovery of debts and settlement of their own obligations, hampered marketing of finished products etc.[2]

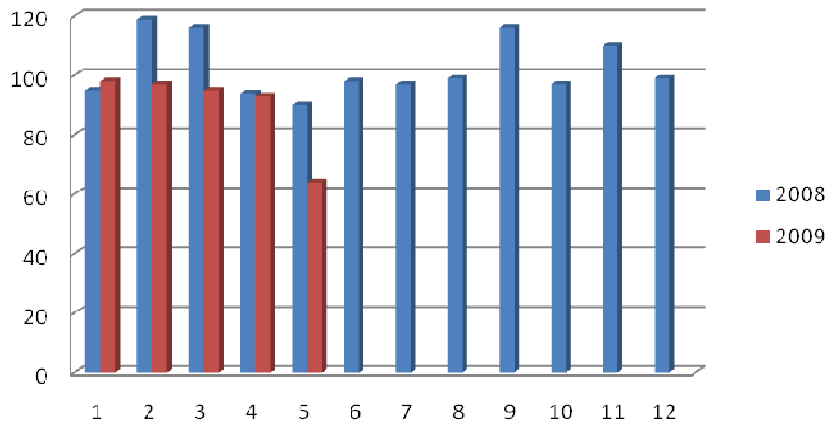
The decrease of production has been recorded in the manufacturing industry, forestry, building industry, and reduced volume of services in transport and tourism. However, the overall economic developments in Montenegro will depend in the following period on the development of global events and measures undertaken with a purpose of giving a support to the enterpreneurship, new investments and development plans as well as projects in specific sectors (Budgetary Low of Montenegro 2009).

Industrial production of Montenegro, in the period January-May 2009, compared to the same period last year, is 16.5% lower. In May of this year, compared to the same month of the previous year, the production was 25.3% lower, compared to April it was 15.7% lower, while compared to the monthly average from the previous year it was 35.9% lower.

In the sector of energy production, which participates with 23.3% in total industrial production, recorded a growth of 41.9%. The other sectors, during the observed period, record a decrease of production, namely in the mining and quarrying sector for 42.4% and in processing industry for 34.7%. Only six areas of industrial production, with their share of 18.8% in the total industrial production, recorded a growth and in the specific sub-areas of: coal production for 31.2%, manufacture of food products and beverages for 7.8%, manufacture of leather and leather products for 165.6%, papermaking, publishing and printing 10.1%, then prodction of chemical products and fibers for 6.4% and manufacture of machinery and equipment for 5.8%.

[2] www.pkcg.org

Figure 1. Industrial production indexes



Source: MONSTAT

Primary agricultural production in Montenegro records a growth during recent years which is, above all, a result of a overall support to the development of private sector through placement of financial resources allocated by Agrobudget and realization of the international projects in this branch.

The intense agrarian reforms were continued during 2009. In the implementation phase are: Strategy for food production and development of rural areas, National Program for food production and rural development, as well as Strategy for fishing industry. The Agrobudget was adopted for 2009 amounting €19,75 million, and it is 34,0% higher compared to 2008. The basic support is realized through dairy sector, a direct support to cattle breeding and crop farming, as well as through support measures aimed at infrastructure, investment projects, services in agriculture. There have been launched the activities with the European Commission on elaboration of Program for rural development and the negotiations were completed with the World bank related to the MIDAS project (project for instititutional development and strengthening the agriculture of Montenegro with a total value of € 14 million).

The agriculture of Montenegro together in conjunction with agroindustry has a share of about 15% in the national GDP, and in the employment that share is about 20%. Agroindustry has a share of 9.3% in the structure of industrial production of Montenegro, of which the production of food and beverages has 9,1%, while manufacture of tobacco products has 0,2%. In the first five months of this year, the manufacture of food products was 7.8% higher, and the manufacture of tobacco products recorded a 10.8% growth. The main problems for the manufacturers in agroindustry are: the lack of financial means for intensifying the production and

export, high insolvency and lack of possibilities of outstanding debts recovery, then the spending limits and the problem which is becoming more and more distinctive and it concerns the placement of products especially in surrounding countries, a modest marketing approach of manufacturers at the international market, design, packing, inexistance of an appropriate trade mark in export, the large number of small exporters, incompetitiveness of products, and due to numerous business barriers the placement of these products was hampered at the markets of signatory countries CEFTA 2006.

Production of *forest* assortments from state forests of Montenegro for the period January-May of this year records a downfall of 18% compared to the same period in previous year. Total available mass of timber cutting in 2009 is planned for just over 420 thousand m^3 , and during the five-months period a total of 45.932 m^3 of timber mass were cut. Wood processing and wood products for the same period was lower by 49%. There is an apparent drop in demand for these products in the market, both in export and domestic market, which led to reducing of production, creating stocks of finished products, reducing the employment of work force, deterioration of the solvency, hampered debts recovery and settlement of due obligations, etc. Reducing the workload of the construction industry, particularly in the housing, has got a negative impact on the timber industry through a decrease in demand for sawn timber, doors and windows, furniture production and other products of wood.

Value of *construction work* in the first quarter of 2009 amounted to 38.64 million EUR, which is 18.6% lower compared to the same quarter of the previous year, while the physical volume of measured effective hours of work is lower for 17.14%. The value of new contracts on dwelling buildings is lower for 52.1%, and on other buildings for 58.3%.

Almost all forms of *traffic* in the first five months of this year saw a drop of services in relation to the comparative period last year. The increase in this period was recorded only by road transport, where the number of passengers increased 16.3%, and the number of tone kilometers in freight traffic increased 5.3%, as well as mobile telephony services which recorded a 14.5% growth.

Number of passengers in railway transport is 22.9% lower, and the passenger kilometers for 15.4%. Rail transport of goods recorded a decrease of 53.1% and the passed tone kilometers of 49.7%.

Although a number of road traffic passengers increased, total number of passenger kilometers recorded a decrease of 13.2%, which implies that those were short trips. When it comes to the road freight transport, where an increase in the number of tone kilometers was recorded, the amount of freight is less than for 12.4%.

The volume of services in maritime transport is also decreasing, so the transport of goods in tones is 10.1% lower, and the number of tone miles is 27.5% lower, while the harbor traffic is 22.4% lower.

Passenger traffic by air transport was 21.9% lower in first five months, while the amount of transported goods in tones is 54.6% lower.

According to the preliminary Montstat data, turnover in *retail trade* in market rates was in first four months of this year 0.6% lower compared to the same period of previous year, and in April this year it was 3.55% lower compared to the April last year.

Activities in the field of trade are aimed at creating conditions for the development of competition in the market, improvement of the quality of offered goods, diminishing the business barriers, improvement and modernizing the trade sector as a basis for improvement of conditions for the inflow of foreign capital, further suppression of the gray economy, improvement in the area of protection of consumers' rights, improvement of the intellectual property rights, improvement of the quality infrastructure as well as completing the initiated transformation processes of commercial companies.

Tourism is a strategic branch which had a most dynamic development during recent years. There is no doubt that this branch was the first to found itself under the attack of economic crisis, because since the end of last year its impact has reflected on the overall results of operations.

For the five months of 2009, the realized outputs of the tourist trade demonstrate a decrease of the total turnover in tourism, that is: decrease in number of tourists for 8.97%, and total number of overnight stays for 13.8%. It is quite clear that normalization of tourist flows and traffic growth could be expected only when the stabilization of global economic processes is felt. Thence it is necessary, on all levels, to seek the solutions at least to improve the current situation. These solutions include both the Government's measures, and the measures of local governments and their public enterprises significant for establishing a quality touristic offer, hoteliers and all service providers. The success of the tourist season this year will, above all, depend on the extent of adjustment of all tourist trade subjects to the new situation both in foreign and in domestic market and attracting as many tourists to Montenegro.

The total *visible trade with foreign countries* in five months of 2009, was 722.04 million EUR, which is 37.8% lower compared to the same period of last year, and it is an indicator of the strong impact that the global economic and financial crisis have in this area. Total import amounted 621.10 million EUR, and it is 36.8% lower, while the total export was only 100.94 million EUR and it is 43.3% lower than the comparative period. Deficit in trade with foreign countries was realized in amount of 520.15 million EUR, which was, due to the significant decrease of value of imported goods 35.42% lower compared to the deficit realized during the same period of last year. Coverage of imports by exports was 16.25%.

Complementary economies, vicinity of markets and approximately same level of competitiveness represent a solid basis for improvement of mutual cooperation with neighboring countries and the region. Due to the impact of crisis on the

economic operations, the volume of this exchange was also reduced, because, compared to the same period of the previous year, the export was lower in all CEFTA signatory countries except in Former Yugoslav Republic of Macedonia, while the import was higher only from Albania. Total turnover with CEFTA countries in five months amounts 367.65 million EUR, of which the export is 39.65 million EUR and the import is 328.3 million EUR. The nontariff barriers contributed to this which were recorded since the Agreement entered into force, and among these barriers the most important ones were double certification of products, then sampling of each individual delivery and each packing , which makes the export more expensive and often puts into question the realization of signed contracts.

In the EU market, the Montenegrin economy places 52.22% of the total exported goods. The import from EU countries represents 31.57% of total Montenegrin import.

Inflation rate, measured by the growth of consumer prices, in May 2009 compared to December 2008 was 1.6%, while in comparison to the same month of the previous year there has been a growth of 4.8%. Consumer prices in May, compared to April were 0.1% higher. Growth rate of consumer prices during the period January-May 2009, compared to the same period of the previous year, was 5.2%.

Observed by groups of products and services according to the purpose of consumption, in May the highest growth was in the prices of culture and recreation area for 0.6%, also the communication area for 0.6%, then food and soft drinks for 0.2% and spirits and tobacco products for 0.1%.

The trend of increase in the number of *employees* continues. By the end of May 2009, the number of 174.218 employees was recorded, which is 5.4% higher than the previous year's situation. In the files of the Employment Bureau there were 27.785 unemployed persons, which is 6,8% less than in the same period of the last year. The unemployment rate was 10.8%.

Average *earnings* in Montenegro in May 2009 were 651 EUR, while the average earnings without taxes and contributions were 468 EUR. Considering the fact that the consumer prices in May this year compared to April was 0.1% higher, result is the 0.5% growth of real earnings in Montenegro in May, compared to April this year. An average earning without taxes and contributions in May, according to sectors of industry, was from 156 EUR in the fishing industry to 930 EUR in the sector of financial brokerage.

Global economic crisis led to reducing of all power units in the banking sector. Total *bank* assets amounted at the end of April 3.142,2 million EUR, and it was 5.1% lower than it was at the end of 2008, while at the annual level it was 2.9% lower.

Total deposits amounted 1.722,8 million EUR at the end of April, and they were 21.7% lower than in the April of previous year, and they were 13.4% lower compared to December 2008. The surveys conducted by the Central Bank, show that

the main problem is not withdrawal of deposits due to some psychological reasons, but the problems concerning repayment of credits, that is the usage of deposits for settlement of credit obligations towards the banks, which means that the money has not left the banks, but was used for the repayment of the obligations to those banks. In total deposits, the deposits of natural persons still have a dominant share and at the end of April they were amounting 762.4 million EUR or 44.3%, then the deposits of non-financial institutions 32.3%, etc.

The value of credits was 2.664.8 million EUR at the end of April 2009, which, compared to the end of 2008, represents a decrease of 4,7%, while the credits at the annual level had a 3% growth. Ratio credits/deposits has become worse and it amounted 1,55 (in March this year it amounted 1,52, in December 2008 - 1,41, and a year before it was 1,18). In the structure of approved credits, 95,5% are the credits approved to the economy and population.

In April, the active weighted average nominal interest rate (APPNS) was amounting 8,67% and compared to March it records a downfall of 0.09 percentage points, while the weighted average effective interest rate of active (APPES) was amounting 9,36% and it records a slight growth compared to the previous month.

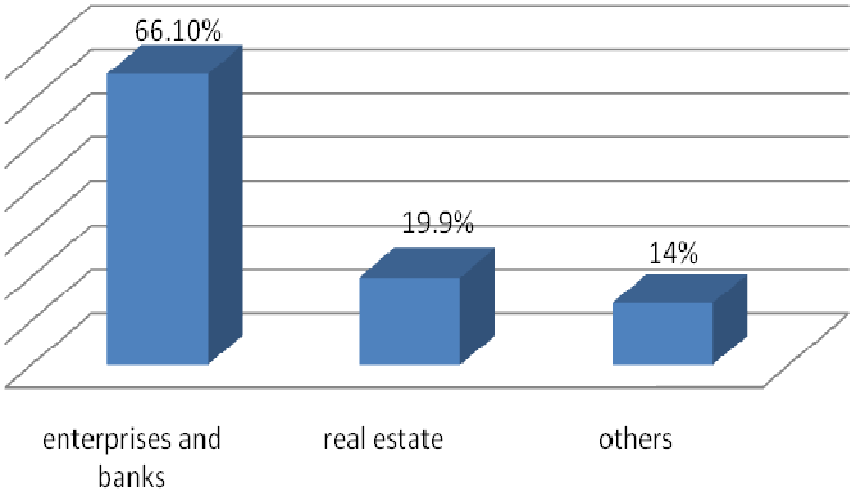
In the second quarter of this year there have been certain positive developments compared to the situation of March 31 of the same year. State of the monetary market is gradually improving, negative trends are slowing down and it is expected that this state will continue in the future. The primary expectation for the forthcoming period is the growth of deposits, due to sale of the part of the Electric Enterprise, then the recapitalization of banks and credit lines from abroad (KFW, FEP, FES, EIB and others).

- The Council of the Central Bank of Montenegro has adopted a Decision on the use of obligatory reserves with the Central bank for the period longer than one day, with a purpose of mitigating the negative effects of the global financial crisis and protection of banking sector in Montenegro (CG Ekonomist).
- There has been introduced a uniform rate for obligatory reserves which is 11% on all deposits, with no exemptions. Effect of this measure is a reduction of obligatory reserves for 32 million, as well as 32 million in treasury bills, which have been introduced as a novelty within the Decision, and it relaxed the banking sector for over 60 million EUR.
- The Central Bank of Montenegro will maintain its intense control over establishing a governance of the credit risks and solvency risks, especially regarding the system banks.

Last year there was achieved a record amount of *net foreign direct investments,* and they were 567,6 million EUR. Despite of the global financial crisis, net inflow of foreign direct investment recorded a slight growth in the first five months of 2009, compared to the same period of the previous year. According to the preliminary data in the possession of the Central Bank of Montenegro, net inflow of

foreign direct investments (inflow minus outflow) in the period January-May 2009 amounted 269 million EUR, which is 5% higher than in the same period of the previous year. The increase of net inflow was realized mostly due to the sale of shares of four joint investment funds in the Electric Enterprise of Montenegro to the Italian company A2A for about 122,7 million EUR. Sale of shares to the Italian company is a confirmation that, when it comes to good companies and projects, despite of crisis, it is not difficult to find investors and it is a good signal for the participation of foreign investors in the recapitalization process of the electro energetic company.[3]

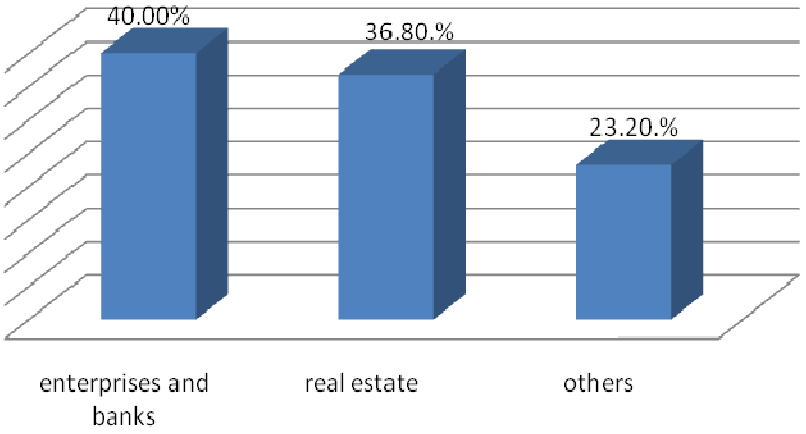
Figure 2. Structure of inflow of Foreign Direct Investments I-V 2009



Source: MONSTAT

[3] www.worldbank.org/montenegro

Figure 3. Structure of inflow of Foreign Direct Investments I-V 2008



Source: MONSTAT

Total inflow of foreign direct investments in Montenegro, in the period January-May 2009, amounted 308 million EUR. In the structure of this inflow 265,4 million EUR related to the influx based on investment in the form of equity capital, while the sum of 42,7 million EUR was realized in the form of intercompany debt. Regarding the structure of investment in the form of equity capital, more than 66% was related to investments into the local banks and companies (204 million EUR), which is 31,3% higher compared to the same period of the previous year. The inflow achieved by selling of the real estate has significantly fallen down – 56,8%. The investments Into real estate during the observed period amounted 61,2 million EUR.

Total outflow based on the Foreign Direct Investments amounted 39 million EUR, which is 71,7% lower than in the same period of 2008.

The level of Foreign Direct Investments should be significantly increased until the end of current year, provided that the projects of capital infrastructure building and recapitalization of banks are realized.

National debt of Montenegro on March 31, 2009 was 1.016,2 million EUR or 28,7% of the estimated GDP for 2009 (3.538,0 million EUR), of which the internal debt is 465,8 million EUR or 13,2% GDP, while the external debt is 550,4 million EUR or 15,6% GDP. (Economic policy for 2009 projected the participation of external debt of 19,1%) (Monstat, CGCG, Mininstry of finance 2009).

In the first quarter of the year there has been a drop of *budgetary revenues* compared to the plan for 13,8%, and compared to the same period of the previous year for 14,8%. The greatest drop of budgetary revenues was recorded in the VAT revenues on import and customs – between 30 and 40%, due to reduced volume of import of goods as a direct consequence of diminished economic activities which

was caused by the economic crisis impact. In contrast to it, VAT revenues realized through the internal transactions still show a growth tendency. After three years of positive balance, in the first quarter of this year, there has been a deficit in the budget of about 8 million EUR.
It has been assessed that *the Gross Domestic Product* in the first quarter of this year was achieved really at the level of the first quarter of the previous year (Directorate for Development of Small and Medium Sized Enterprise of Montenegro 2007).

Table 1. Major indicators of Montenegro, 2009

CONTENTS	**Estimation for 2009**	**Realization Months I-V 2009**
GDP in current prices, millions of €	3.715,0	-
Increase of real GDP, %	5,0	-
Inflation rate, %	4,5	1,6
Employment growth %	2,1	5,4
Unemployment rate, %	10,3	10,45
Goods and services exchange deficit, millions of €	1.151,6	520,15
Share of deficit in the GDP %	31,0	-
External debt, millions of €	710,0	550.4
External debt, in GDP %	19,1	15,6
Net Foreign Direct Investments, millions of €	460,0	269,0
Net Foreign Direct Investments, % in GDP	12,4	-

Source: MONSTAT

2. Government's measures

The set of measures for relieving the consequences of economic crisis was adopted by the Government's Budget for 2009. Since the beginning of the year, the following fiscal policy instruments have been applied:

Reduction of the current, unproductive expenditure through:

- reduction of current outgoings, that is, lower participation of public spending in GDP,
- freezing of 10% of expenditures for materials and services in the first half of this year.

The adopted saving plan envisages realization of savings in amount of 66.8 million EUR or 14,1% of total annual budget.

Support to economy and population:

- Diminishing the income tax rate of the citizens from 15.0% to 12%,
- Diminishing the health insurance contributions rate from 12% to 10.5%,
- Diminishing tax rate for retirement insurance from 21% to 20.5%,
- the obligation to pay contributions for compulsory social insurance for incomes based on severance compensations was abolished,
- motorway fees were abolished,
- the fees for usage of construction sites were abolished,
- the price of electric energy was reduced for small and medium enterprises by 10%, and subsidies for the most jeopardized categories of population were introduced,
- premature repayment of internal debt based on the restitution and repayment of old foreign currency savings,
- broadcasting taxes were abolished,
- The conditions for ensuring the long-term credit lines were created, with an aim to support small and medium enterprises,
- Strengthening investment in infrastructure by increasing expenditures in capital budget.

Besides fiscal policy measures, socio-economic support measures are being realized through project ''A job for you'', having a value of 18.5 million EUR for 2009.

A support measure aimed at ensuring additional solvency of small and medium businesses as well as stimulation of employment is realized through borrowing from abroad and providing credits for small and medium enterprises (ŠTIBLER 2009):

- Representatives of Government and European Investment Bank (EIB) signed on 3 July a Contract of guarantee of 91 million EUR for Montenegrin business banks intended for financing projects of small and medium enterprises. The largest sum of money will be lent to the Montenegrin Commercial Bank – 33 million EUR and to NLB Montenegrobank – 20 million EUR. Atlasmont Bank will receive nine million EUR, Podgorička Bank will receive seven million, Opportunity Bank

six, Hypo Alpe Adria Bank five, Hipotekarna Bank four million EUR, Commercial Bank three million and First Financial Bank will receive two million EUR. Grace period is four years with time of payment 12 years. For infrastructure projects grace period is five years with time of payment 15 years. EIB reserved the right not to approve the individual credit if it assesses that the interest rate or the time of payment are out of proportion compared to interest rate and time of payment under which the bank obtained that credit.

- However, EIB is already present by crediting the Montenegrin Railway Enterprise, projects concerning waste water disposal, as well as possible investments into regional landfills.
- The issuing of Government guarantees worth 167,5 million EUR was envisaged with an aim to support large systems for sustaining the economy, such as: Željezara (Steel Processing Plant) Nikšić (25 million EUR), Aluminum Plant Podgorica (135 million EUR), and the rest goes to Montenegroairlines and Pobjeda;
- The Erste Bank provided 35 million EUR for support to the budget;
- There was also envisaged the issuing of guarantees for credits by the European Bank for Reconstruction and Development (EBRD) for the project Construction of regional water supply on the Montenegrin coast, with the amount of three million EUR and the European Investment bank (EIB) for railway project – phase II, amounting seven million EUR;

The guarantees will be issued also to the German Bank for Reconstruction and Development (KfW) for the projects Thermal Power Station Pljevlja I and Substation Ribarevina, amounting 15 million EUR, then, this bank will also obtain the guarantees for the project Hydroelectric Power Station Piva, amounting 9,5 million EUR.

- The state will issue the guarantees to KfW bank also for the support to the banking system in total value of 50 million EUR, and the beneficiaries are Montenegrin business banks, then to the Abu Dhabi Fund for development for the project Regional water supply worth 15 million EUR and to NLB Montenegrobank for five million EUR with the Bauxite Mines as a beneficiary.
- In the course of this year, the Government signed a Contract on guarantee with EBRD for the third portion of credit for the project Urgent rehabilitation of railway infrastructure, which is four million EUR worth.
- The Government, up to the present, was issuing the guarantees for the projects which were aimed at the improvement of infrastructure and which represent a basis for a sustainable economic development.[4]

[4] www.pkcg.org

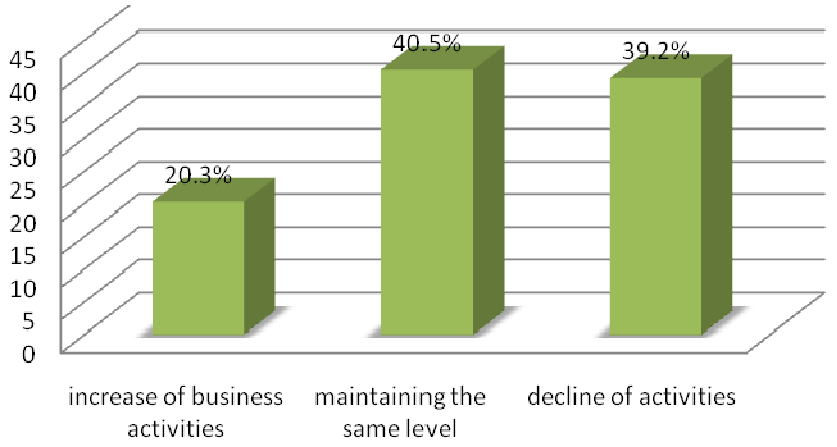
3. Overview of situation based on conducted surveys (Vukčević 2009)

Chamber of Economy of Montenegro continuously monitors operations of companies through its association committees, then by visiting the companies and in direct contacts with economic operators and by the already prepared questionnaires.

The survey conducted by the Chamber and direct contact with economic operators, included over 320 economic enterprises from all economic activities. Based on immediate discussions on problems and limitations encountered in the course of business operations and based on the presented answers to the questionnaire, some main limitations to the economic operations were reported along with the recognition of business expectations for the current year.

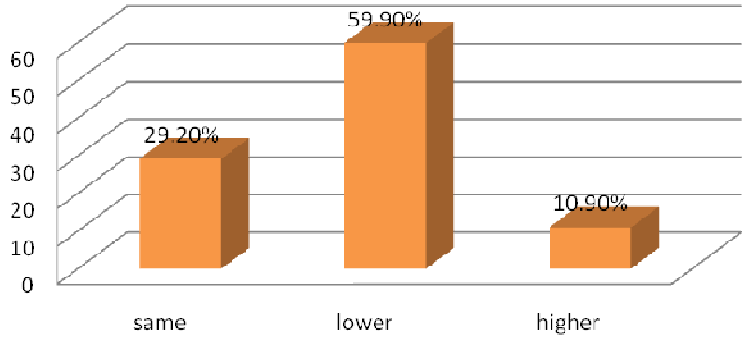
1. By processing the data obtained from the questionnaires, it can be concluded that the business expectations of the interviewed economic operators are still encouraging until the end of the year, since almost half of the interviewees (40.5%) expect keeping their business activities at the previous year's level. About 20.3% of the interviewees expect the business activities to grow, and 39,2% already face a downfall of the production volume and expect that trend to continue until the end of the year.

Figure 4. Business expectations by the end of 2009



Source: MONSTAT

Figure 5. Volume of production/turnover in the first quarter of the year compared to the same period of the previous year



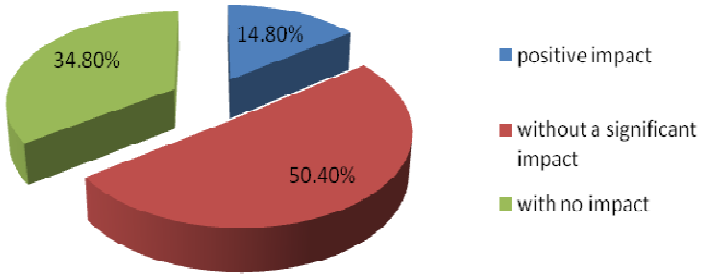
Source: MONSTAT

Reduction of turnover volume in the first quarter of this year compared to the same period of the previous year was recorded by more than half (59.9%) of the interviewed, the same volume has 29,2%, and higher volume was recorded by only 10.9% of the economic operators.

Out of the enterprises with declining volume of production (turnover), the most of them, about 48.9%, have the turnover reduced from 1-30%. In the zone 31-60%, reduced turnover is recorded by 31.8% of the interviewed, while only 6.8% have a turnover reduced by over 60%. Amongst the enterprises which have an increased volume of turnover, the majority of the interviewed (87.5%) has got an increase up to 30%, while 12.5% of them have got higher turnover from 30-60%, but no enterprise succeeded to increase their turnover over 61%.

2. Over half of the interviewed consider that the Government's measures, aimed at suppression of the negative effects caused by the crisis, did not have some significant impact to their business operations, 34.8% consider that they had no impact whatsoever, while only 14.8% of the interviewed consider that the measures taken by the Government had a positive impact on their operations.

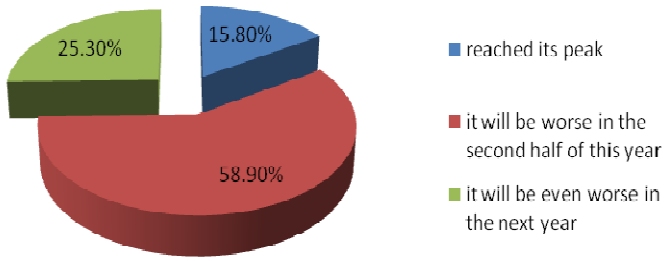
Figure 6. Assessment of the impact made by the Government's measures to the operations of economic enterprises



Source: MONSTAT

3. When it comes to the assessment of the intensity degree of the crisis which overtook the Montenegrin economy, only 15.8% of the interviewed consider that the crisis reached its peak, while the rest of the interviewed expect more complex conditions for performance of their operations in the forthcoming period, that is, the crisis yet has to reflect onto the Montenegrin economy.

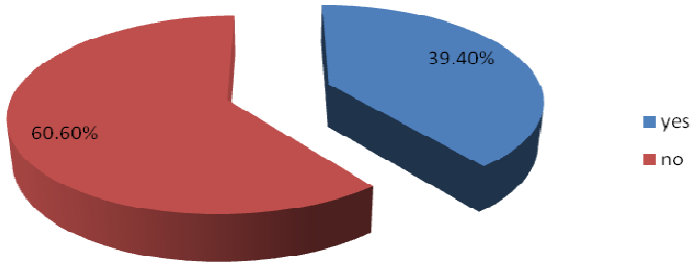
Figure 7. The intensity of crisis



Source: MONSTAT

4. The larger number of the interviewed (60.6%) did not ensured, or, better to say, they did not contracted their operations until the end of business year, while 39.4% of them already have signed contracts for this year.

Figure 8. Contracted business arrangements for 2009



Source: MONSTAT

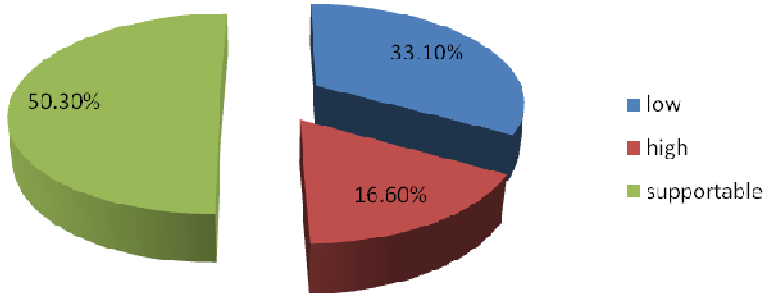
Compared to the survey from March this year, the situation remains almost unchanged

5. The contracts were denounced due to: hampered placement, inability to pay obligations, hampered collection of debts, insolvency of business partners, high cost of raw materials, falling prices of finished products, reduced spending capacities, due to untimely tender for award of concessions, inability of investors to close the financial construction, lack of financial resources, situation in business banks, impact of the global crisis on the foreign partners, etc.

6. The largest negative impact on business has a difficult collection of debts (21%) and lack of credit resources (19%), then duties (16.5%), costs and hampered placement of products (16%), while the smallest impact have the hampered purchasing conditions (11.5%).

7. Regarding credit debts of economic enterprises, it is, according to their assessment, found supportable by the half of the interviewed, it is low by 33%, while 16.6% is highly credit indebted. There were no significant changes of the situation compared to three previous months (the previous survey).

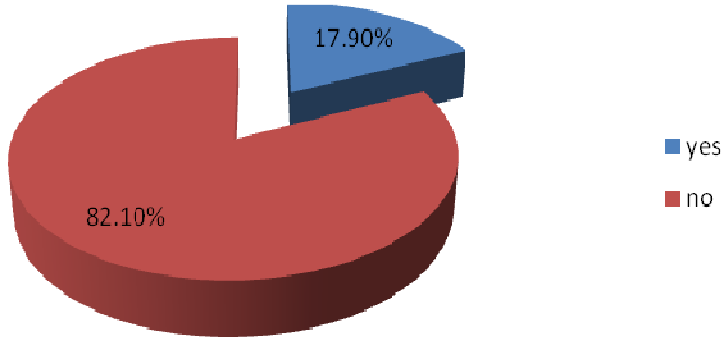
Figure 9. Assessment of credit indebtedness of economic enterprises



Source: MONSTAT

Over 82% of the interviewed did not submit a request for rescheduling of credit.

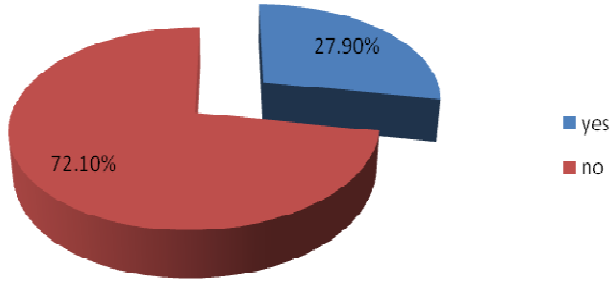
Figure 10. Request for rescheduling of credit



Source: MONSTAT

Most of the total submitted requests for rescheduling of credit (72.1%) were not approved.

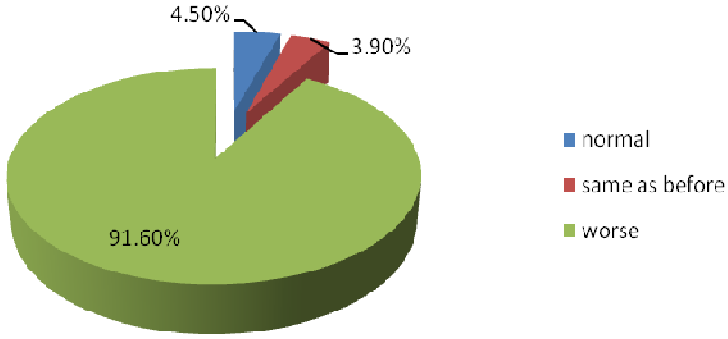
Figure 11. Approved requests for rescheduling of credit



Source: MONSTAT

8. About 91.6% of the interviewed consider that the conditions for obtaining credit during the past period of this year were more adverse, while the rest consider them same as before (3.9%), or normal (4.5%). Compared to the survey from March this year these assessment are even more negative, since at that time 75% of the interviewed considered those conditions more adverse.

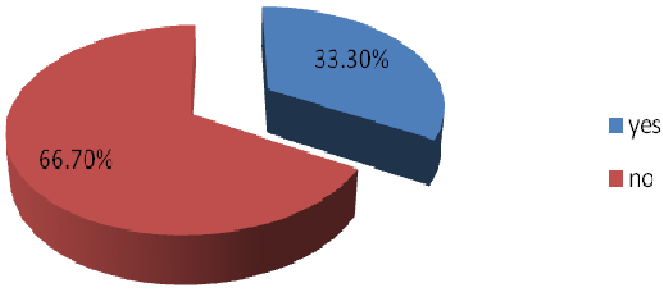
Figure 12. Assessment of the possibilities for obtaining credit in 2009



Source: MONSTAT

About two thirds of the interviewed did not submit their requests for obtaining credit during this year.

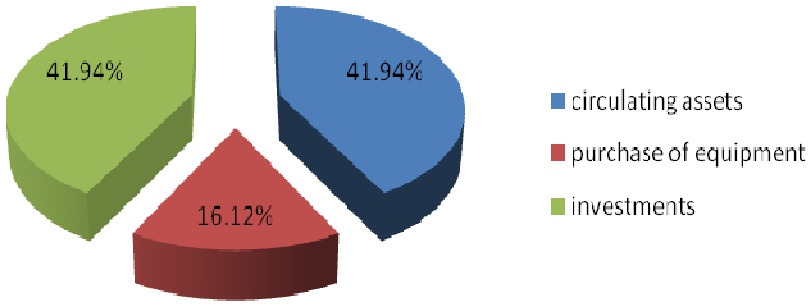
Figure 13. Submitted requests for obtaining credit in 2009



Source: MONSTAT

8.1 Of one third of the interviewed which submitted requests for credit, the largest number sought credit for circulating assets (41.9%) and for investments in fixed assets and expansion of business (41.9%), while only 16.1% for purchase of equipment.

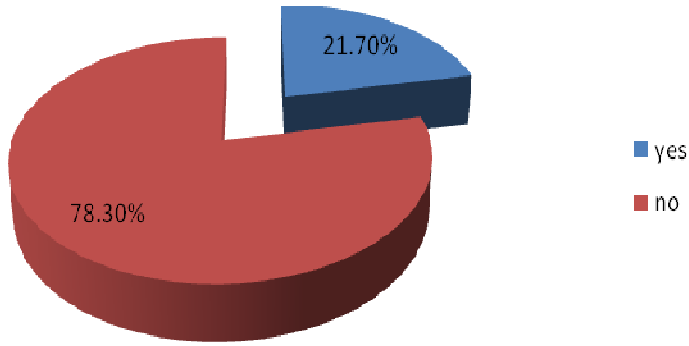
Figure 14. Credit purpose



Source: MONSTAT

8.2 For the most of the interviewed (78.3%), the credits have not yet been approved, but the majority (71.7%) expects their approval.

Figure 15. Approved credits

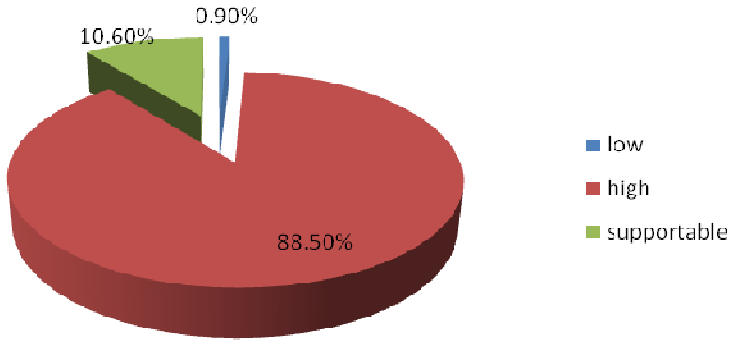


Source: MONSTAT

8.3 About 2.3% do not expect that their requested credits would be approved, and as the main reasons they mention: the new situation significantly changes the position of the company and affects its solvency, insufficient and irregular execution of certified realization (building construction), poor financial balance from the previous year, which results in poor solvency of companies (transport) and more strict bank procedures for giving new credits.

9. Interest rates for credits are mostly high in the opinion of 88.5% of the interviewed, while for 10.6% of them they are supportable, and an insignificant number (0.9%) consider them to be low.

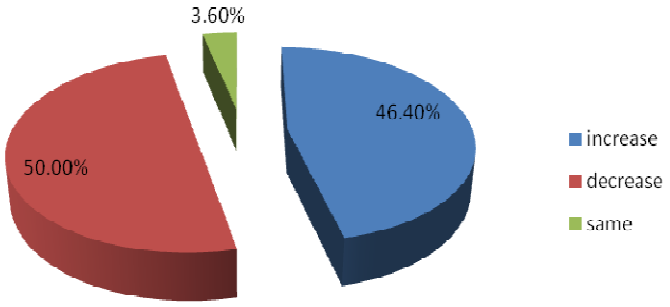
Figure 16. Assessment of interest rates and other banking services in 2009



Source: MONSTAT

10. Regarding the number of employees in the companies and plans until the end of the year, almost half of the interviewed (46.4%) plans to increase the number of employees and mostly within the range 1-30%, while the other half of the interviewed plans to reduce the number of employees within the range 1-30%, so that it is expected that the total number of employees will remain the same.

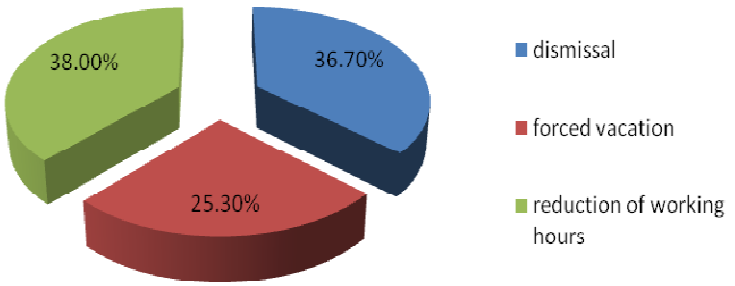
Figure 17. Trends in number of employees until the end of 2009



Source: MONSTAT

10.1 The interviewed with redundant workers have dominant attitudes that this issue can be solved by: shortening the working hours (38%), dismissal of employees (36.7%) or forced vacation (25.3%).

Figure 18. Settlement of the redundant employees' problem

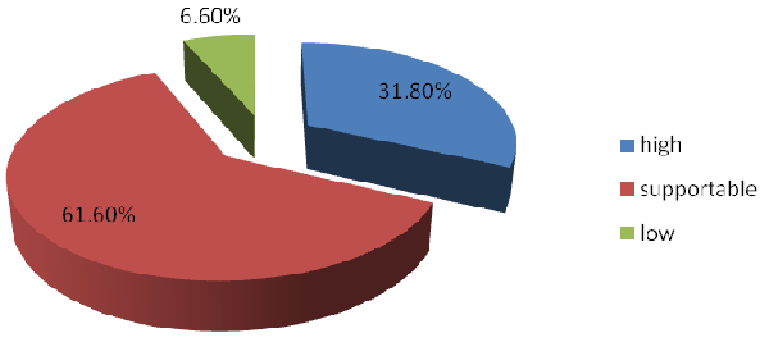


Source: MONSTAT

10.2 The largest number of the interviewed (87.3%) regularly pays their obligations towards the employees, while the others are late for some 2-3 months.

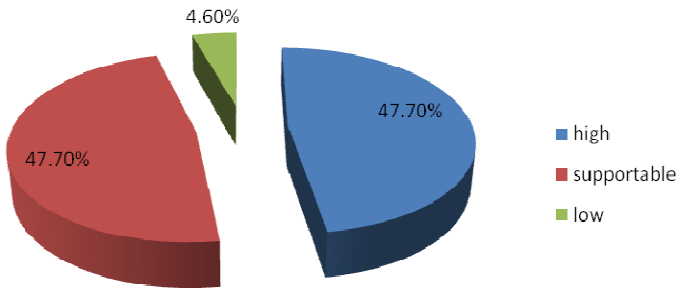
10.3 Regarding the assessment of liabilities and claims by the end of the year, the largest number of employees (61.6%) considers that they have a supportable level of obligations, while the obligations are high for 31.8% of the interviewed.

Figure 19. Assessment of the level of obligations by the end of 2009



Source: MONSTAT

Figure 20. Assessment of the level of claims by the end of 2009

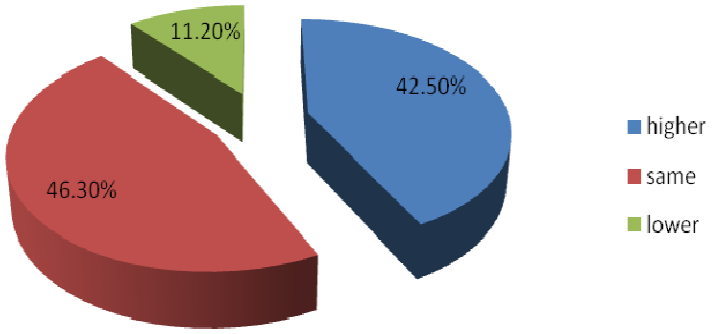


Source: MONSTAT

The claims, however, are equally high and supportable in the opinion of 47.7% of the interviewed, and a very small number of about 4.6% has a low level of obligations and claims.

11.1 Level of obligations compared to the same period of the previous year remained the same at 46.3% of the interviewed, while 42.5% have greater obligations, mostly within the range of 1-30%.

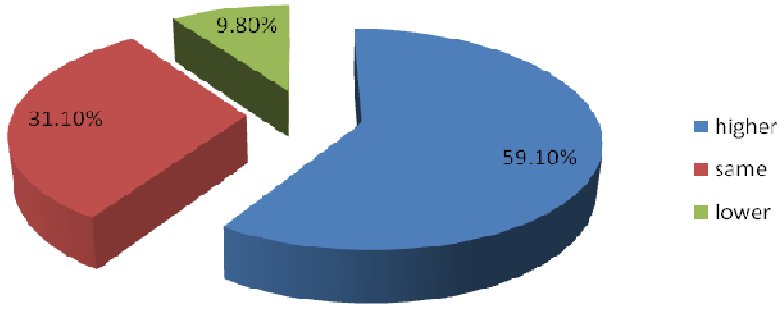
Figure 21. Level of obligations in the 1st quarter 2009/1st quarter 2008



Source: MONSTAT

The claims were higher for 59.1% of the interviewed and mostly within the range 1-30%, while 31.1% have the same level of claims as they had at the end of the first quarter of the previous year, and only 9.8% have a lower level of claims.

Figure 22. Level of claims for the period 2009/2008

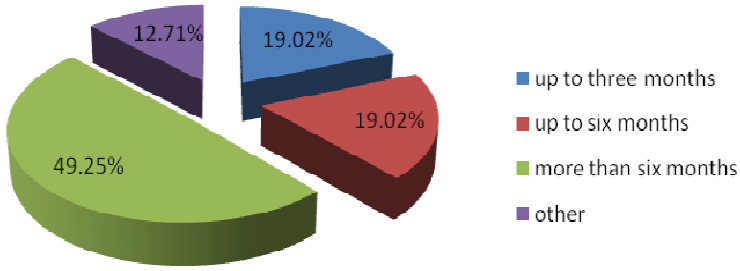


Source: MONSTAT

11.2 Obligations/claims ratio is equal for 25.8% of the interviewed, while 32.5% have claims double compared to obligations, and the rest of the interviewed have got claims which are three, four or five times higher than the obligations, while in ten cases there are obligations double compared to claims.

12. About 66% of the interviewed launched a lawsuit for payment of claims, and the court proceedings generally last for more than six months (49.2%), while considerably lower number of the interviewed had court proceedings with duration of up to three (19%) or up to six months (19%).

Figure 23. Duration of the court proceedings



Source: MONSTAT

13. Other adverse effects expected by the interviewed until the end of the year are: difficult collection of claims from the commercial sector; increase in the number of blocked accounts; increase in the number of dismissed workers and reduction of salaries; poor tourist season; lack of subsidies for agriculture and difficult repurchase of agricultural products as well as their placement due to uncertain debt collection; lack of working capital to pay associates; delays in granting concessions, which prevents the contracting business; weak occupation of the capacities and labor due to reduction of ordering and a lack of raw materials; decline in prices of finished products (timber); decline of investment activities in Montenegro and in the region which have negative effects on the placement of domestic equipment; general insolvency; inability to pay off credits; reduced financial power of the population; increase of interest rates by the business banks; reduced turnover in all forms of transport as well as the reduction of profit; unfair competition and insufficient ability to work in appropriate industries.

14. Government measures which, in the opinion of economic operators, facilitated the operations in the first quarter of this year are:

- postponing obligations for taxes and contributions for salaries,
- postponed payment of VAT on import, although that Government measure is hardly applicable in practice, since it requires special guarantees from banks,
- reducing obligatory reserves by the Central Bank of Montenegro,

- Subsidies to the prices of energy, and reduction of price of petroleum products,
- Faster issuing of building permits,
- Grant Schemes of the Directorate for development of small and medium enterprises,
- obtaining guarantees from international financial institutions for credits for SMEs by force of the Government,
- Abolition of fees for city construction land,
- Abolition of weight tax charged by custom duty.

15. Measures of local administrative bodies which, in the opinion of business operators, facilitated their operations in the first quarter of 2009:

- Reduction of utility fees,
- Abolition of fees for city construction land.

16. Measures taken by economic entities with a purpose of rationalizing their operations:

- Savings in the primary and auxiliary materials, reducing waste products,
- Rational management of stocks,
- Reduction of operational costs on each level,
- Increase of productiveness,
- Introduction of new services,
- More intensive contacts with the producers aimed at launching some joint activities for sales improvement, promotional prices, longer repayment period,
- Maintaining the same level or reduction of prices of the finished products or services,

- Increase of activities in the area of marketing,
- Reduction of working hours,
- Reduction of the number of employees and suspension of recruiting new employees,
- Maximum involvement of the managements in putting into function all business capacities, on obtaining deals which are at the level of simple reproduction,
- Reduction of earnings,
- Continued investment in projects that have already started or are in their final stages,
- Careful selection when entering into new projects,
- Maintaining construction operations as well as the quality of performed works,
- Selling products on deferred payment only to verified buyers,
- Attempt of placement of the products on foreign market,
- Extension of transactions volume which does not require large investments,
- Savings at the depreciation at the cost of fixed assets,
- Strict control of debts collection,
- Selection of buyers and ensuring of claims,
- Adoption of standards.

3. Conclusions and suggestions

Based on the indicators of the general state of the economy in the first five months of this year (Montstat and the Central Bank of Montenegro) and on the basis of the survey continuously conducted by the Chamber of Commerce, direct contacts and

conclusions by associations committees, decline in economic activity since the beginning of the year is evident. The similar situation is expected by the end of the year, so it is necessary to reexamine the projected macroeconomic indicators and to readjust them to economic reality. Measures adopted at the end of last year and at the beginning of this year, according to assessment of the majority of economic operators (over 85%) were without any significant impact, or they had no positive impact to operations of those entities. Therefore, they need to be improved and adjusted to the new economic reality, because negative consequences of the financial crisis become more and more evident. Insolvency of economic operators is increasing, and it is demonstrated by the results of the survey conducted during the second quarter, compared to the results of the first quarter survey. More and more companies have high outstanding liabilities (about 32%), and that fact is confirmed by data provided by the Central Bank of Montenegro regarding blocking of the accounts and increased level of due outstanding liabilities.

Credit support

- Government's assistance programs for sound companies should be directed towards ensuring necessary bank guarantees for obtaining new credits, regression of interest rates and rescheduling of credits.
- Increasing the amount of funds for subsidies within the budget's capacities, and these subsidies are intended for interest rates of the credits for sustaining the solvency and financing of circulating assets, as well as for the import of raw materials and semi manufactures which do not have a substitute in domestic production, for the two following years. Extension of time limit for approval of these credits is necessary considering that the operating conditions have got worse, as well as that short term credits are not sufficient for financing of permanent circulating assets.
- With a purpose of generating domestic demand, there should be introduced credits with subsidied interest rate for consumer credits which are intended for purchasing certain permanent consumption goods.
- Continue to support the SMEs through Directorate for Development of SMEs and the Development Fund, which base their production programs and business plans on the final production, manufacturing and placement of higher-value products, increase of the production volume and the quality of export goods.
- To establish a Bank for Development which would, by achieving profits from selling the state property and the funds borrowed from the international capital market with the state guarantees and Budget resources, finance the strategic development projects as well as the projects regarding sectors and companies which would in short term ensure some positive economic effects. The advantage is to be given to the programs which ensure savings in energy and raw materials, automation of processes, introduction of standards, quality control, extension of the market

(export), strengthening of domestic manufacturers which will result in increase of the level of their competitiveness.

Improvement of the financial discipline

- Tightening of measures in payment operations in the part concerning respecting the payment deadlines and providing instruments for payment guarantees (checks, bills of exchange, bank guarantees). The introduction of new methods or financial instruments (order of acceptance) should be encouraged, or, for example, extrajudicial collection, through specialized agencies, factoring and forfeiting.
- Due to big problems encountered by Montenegrin economy, for the duration of the crisis, as a short-term measure there should be introduced a multilateral compensation with the public sector, and by that reduce the volume of internal debts.
- It is necessary to introduce an obligation for users of budgetary funds to regularly pay their dues to the economy, for conducted works and services from the contracts awarded through public procurement procedures, in accordance with the allocated funds and contracted dynamics of payment for no longer than 60 days.
- Impose an obligation to the Commercial Court to, when registering the company, it must make prior examination of data with the Tax Administration and the Central Bank of Montenegro regarding the existence of eventual due liabilities of the future owners of economic entities. That would prevent abuses, that is the extinction of existing economic entities and registration of new ones by the same persons and that would prevent avoiding of payment of due liabilities.
- In purpose of strengthening the financial discipline, the efficacy of judicial system has to be improved in terms of court deadlines for recovery of claims. In the court proceedings for recovery of claims, due to elimination of delays of process to the detriment of creditors, the obligation of payment based on the court decision should be introduced, after which it would be possible to start the complaint mechanisms. It would be necessary to increase the level of default interests so that the debts could not be paid, and by that reduce the duration of the court proceedings.
- It should be continued with a consistent carrying out of the measures aimed at the suppression of gray economy and in the fields of transport and production. Synchronized and efficient measures, increased technical and staff capacities with a coordinated activities of customs, inspection and other institutions should be fully engaged in the control and prevention of illegal import, elimination of illegal forms of business operations and prevention of unfair competition on the market.

References

ADIZES, I. 2009: *How to Manage in Times of Crisis.* Novi Sad, 82-83.

Budgetary Low of Montenegro 2009

Directorate for Development of Small and Medium Sized Enterprise of Montenegro 2007, Development Strategy for Small and Medium Sitze Enterprise 2007 – 2010, Podgorica, 91-120.

Governing Laws in Montenegro, Official Gazette of Montenegro, The Third Amended Edition Volume I and II, Podgorica 2008;

Municipal Regulations, Official Gazette of Montenegro, Chronological Register of Legal Acts and Other Acts, Podgorica 2008;

Public Procurement Directorate of Montenegro / Ministry of Finance of Montenegro, Reports on Public Procurement for 2007 and 2008

ŠTIBLER, F. 2009: Impact of Global Crisis on Montenegro and Western Balkans, Podgorica 2009, 151-171.

VUKČEVIĆ, M. 2009: EU Internal Market and Basic Freedoms, Podgorica, April 2009, 144-145.

Other used resources:

www.worldbank.org/montenegro
www.pkcg.org
www.poslovniforum.hr
www.djn.gov.me

Russian SMEs in different types of settlement: what is the influence of global crisis?

Yulia Guleeva[1]

The aim of the paper is to analyze the changes in the structure and strategies of small early entrepreneurs under crisis.

The research structure is based on the Global Entrepreneurship Monitor (GEM) definition of entrepreneurship and its application to the analysis of entrepreneurship potential in Russia. Novelty of this work would include the differentiation of some important issues of early entrepreneurship by settlement types into five groups (rural areas, small towns, medium-sized towns, big cities and megapolises).

The socio-psychological characteristics of Russian potential and early-stage entrepreneurs have the most significant but rather ambiguous impact on total business activity. On the one hand, the fear of failure is very strong: the maximum level is expected among potential entrepreneurs in rural areas. On the other hand, the vast majority of potential and early-stage entrepreneurs considers to have knowledge and skills to open a new business.

The main indicators of starting business are social networks, perceived capabilities and a principal source of income. Under crisis the impact of perceived capabilities and social networks has increased in rural areas and cities.

The results of the research can be applied in various governmental and regional assistance programs in Russia where the type of settlement would be considered as one of the main differentiating factor.

Keywords: Global Entrepreneurship Monitor (GEM), early-stage entrepreneurship, settlement aspect, social networks, perceived capabilities

1. **Introduction**

When it comes to the all-too-human problem of recessions and depressions, economists need to abandon the neat but wrong solution of assuming that everyone is rational and markets work perfectly." These words of Nobel prize laureate Paul Krugman are more than actual in Efficiency-driven economies. That is why we try

[1] *State University - Higher School of Economics, Russia*

to figure out the indicators important for businessmen in irrational and imperfect economic world.

The main aim is to examine whether the global crisis has a different effect on the personal and socio-demographic determinants of early entrepreneurial activity in different types of settlements.

Recently the problem of support to small business has been placed in the forefront because of the recession and unemployment rate growth. Nevertheless, most of recent analyses fall short of complete adequate suggestions for the efficient government support to Russian small business.

Early-stage entrepreneurs and nascent entrepreneurs in Russian settlements as the most weak and unstable part of businessmen are the main objectives of the study.

The logic of the study is the following. The first part of the paper concerns the differences of early entrepreneurial structure in rural areas, towns and cities. Then, I would show what factors are important to start a business. And finally, it would be revealed what are significant differences of the determinants of decision to become an entrepreneur in different settlements before 2009 and in crisis in 2009.

2. Global economic crisis in Russia and conditions of entrepreneurial start: Settlement aspect

The problem of crisis shift from one country to the other is not new. Thus, after Latin America crisis in 1994 and Asian crisis in 1997-19989 the problem was actively discussed in economic literature. And already 10 years ago the economists have showed that at the age of free capital movement even slight faults of economic policy can cause incommensurably strong reaction all national markets and lead to serious economic shocks (Krugman 1999). And the most serious problems arise in the countries that have their own internal economic problems.

The financial recession in Russia affects the most seriously the entrepreneurial start conditions of Russian SMEs. The abilities to attract formal investors to finance a new business fall. The growth of financial barriers increases already high dependence on informal investments and "love" capital. Such credit sources are less risky but at the same time they are less efficient (Murzacheva 2008). Moreover, the resource of "love" capital – savings of households – has also reduced because of the inflation, the growth of unemployment rates and decrease of salaries. Particularly, the results of official statistical surveys indicate the growth of consumer prices (for more than 2% in January, 2009 and for 8,1% from January to

September of 2009) [2], the growth of part-time employment and the fall of industrial production (more than 10% in 2008).

Under such conditions the development of entrepreneurial potential in Russia becomes the most important economic and social goal. On the one hand, the growth of entrepreneurship leads to the economic growth of the country. On the other hand, business activity is the mechanism of adaptation for unemployed and individuals who have lost high salary incomes.

At the same time, the crisis dictates rigid constrains to government support. So, the strongly differentiated approach is necessary. And that is why the first step for new support programs involves the analyzing of the influence of various factors on the entrepreneurial choice in different settlements.

Moreover, GEM methodology allows to separate spatially and dynamically comparable entrepreneurial groups for the purposes of the study (Obraztsova 2007).

As it has been showing repeatedly in economic literature, the position of small entrepreneurship is weaker and more unstable at the earlier stages of business activities (Arenius-Ehrstedt 2008). It concerns financial conditions, the amounts of capital and so on. Nevertheless, the support to nascent entrepreneurs as the main addressee of all government programs cannot become the most efficient strategy if the country is not homogenous by the tendencies of social and economic growth itself. Especially the efficiency of such programs of support falls if the irregularity is set in the development of entrepreneurship and the design of framework business conditions. And this is the case of Russia (Gabelko 2008, 2009). The recent example of failure in the field of entrepreneurship support is the program of favorable privatization terms for small business in Russia in 2008.[3]

In order to design a really efficient program of a government support for small business it is necessary to conduct a comparative analysis. The study should analyze the tendencies and factors that are significant for early-stage and nascent entrepreneurs. Also it would take into account socio-demographic and cultural peculiarities of real economic conditions in different parts of Russian Federation.

Meanwhile, settlement aspect is more actual to the analysis than regional one at the context of small entrepreneurship development. For example, the comparison of Moscow or Saint-Petersburg with their suburbs and small towns not far from the city is inappropriate because of different socio-economic conditions and living standards. On the contrary, the application of spatial economics shows that community of socio-economic environment in the settlements of the same type forms the homogenous entrepreneurial groups. GEM methodology of adult population surveys allows to prove this hypothesis empirically.

[2] http://www.gks.ru/bgd/free/b04_03/Main.htm

[3] Federal statute "About the peculiarities of alienation of real assets in the regional ownership …" from 22 July, 2008.

The results of the analysis based on GEM data that has been carried out to find the determinants of the entrepreneurial strategies in Germany (especially after reunion of Western and Eastern parts) has shown up the community of the factors influencing the entrepreneurial choice on different types of settlement (Sternberg-Wagner 2002).

Moreover, Arenius and DeClercq (2005) argue that individuals differ in terms of their perception of opportunities because of the differences between the networks they are embedded in. The theoretical foundations of this study are network theory and human capital theory. In this paper we propose that different *types of settlement* have differently structured networks. We make the distinction between rural and urban areas and big agglomerates.

We are proposing that rural areas have networks which are characterized by strong relationships among a limited number of people. It is more likely that the current residents of rural areas have been living there for a significant amount of time. Therefore, networks in rural areas have a high level of network cohesion.

Big agglomerates or, to a lesser extent, urban areas are proposed to have networks with loose ties among their residents. These areas are characterized by a higher number of potential contacts and by a higher likelihood that new contacts move into the area.

In short, agglomerate areas will be more likely to have extensive networks of 'loose' contacts among their residents and thus their networks characterized by a lower level of network cohesion.

Arenius & Minniti (2005) found that, across all countries and across genders, perceptual variables and, in particular, the perception that individuals have of their own entrepreneurial abilities is very important. Their results suggest that those who perceive themselves as possessing the necessary skills are more than 6 times more likely to be nascent entrepreneurs than those who do not believe to have the necessary skills. Knowing other entrepreneurs and perceiving entrepreneurial opportunities also had a significant positive effect. Fear of failure on the contrary decreased the probability of entrepreneurial activity.

3. Data sources and GEM methrology

The data of Russian entrepreneurial potential surveys based on GEM methodology and questionnaire have allowed to carry out the analysis of the tendencies connecting early-stage entrepreneurial dynamics and behavioral choice that has arisen in the period of 2006-2009, just before the start of global crisis (2006-2008) and under Russian crisis (2009).

First of all, I would point out some basic principles of GEM methodology[4].

[4] http://www.gemconsortium.org/

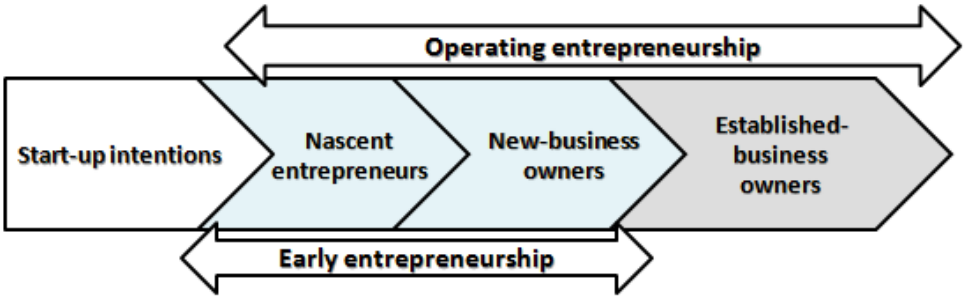
It divides entrepreneurs in three groups on the base when the business got the income for the first time:

- nascent entrepreneurs (starting business, income less than 3 months)
- new/baby business (income from 3 to 42 months)
- established business (functioning more than 42 months).

Nascent entrepreneurs and baby businessmen form the group of early-stage entrepreneurs. Start-up intentions are defined as the individuals who want to launch a business and have some skills and knowledge for it (Reynolds-Bosma 2005).

The figure 1 shows the main steps of business growth in accordance with GEM Methodology.

Figure 1. The stages of entrepreneurship development, GEM methodology



Source: own creation

The data in my work is divided into five groups by settlement type:

- rural areas,
- small towns (population under 100 thousand),
- medium-sized towns (from 100 to 500 thousand people),
- big cities
- megapolises.

Then, during the research, I have managed to regroup the data into three clusters (rural areas, towns and cities).

Novelty of this work would include the differentiation of some important issues of early entrepreneurship by settlement types. GEM Russia APS data file for 2006-2009 was divided into five groups: rural areas, small towns (population under 100 thousand), medium-sized towns from 100 to 500 thousand people, big cities and megapolises.

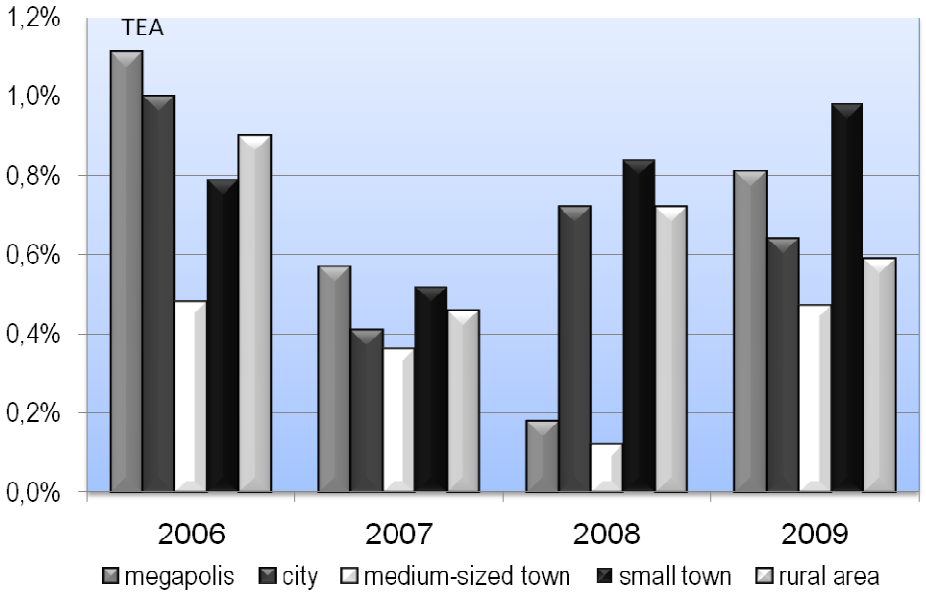
4. Comparative analysis of nascent entrepreneurial activity dynamics under global crisis conditions: Settlement aspect

At the first stage of this research early-stage entrepreneurs have been divided into five clusters by settlement type. Standard demographic classification has been used: rural areas (in accordance with the glossary), and small towns (from 10 to 100 thousand inhabitants), and medium-sized towns (from 100 to 500 thousand residents), and cities (from 500 to 750 thousand inhabitants), and megapolises (more than 750 thousand residents).

The structural and dynamic analysis of entrepreneurial activity at early stages has revealed certain settlement features of business development under crisis conditions in 2006-2009. The entrepreneurial reaction on toughening financial and economical conditions can be analyzed in the context of entrepreneurial activity itself as well as in the context of various qualitative indicators.

The following graph (figure 2) shows the dynamic analysis of nascent entrepreneurial activity through four recent years in different types of settlement.

Figure 2. Early-stage entrepreneurial activity: dynamic analysis



Source: own creation

Early entrepreneurial activity in cities, megapolises and rural areas has decreased from 2006 to 2009. Otherwise, the nascent entrepreneurial activity in small towns has increased. The rates are significantly different at 5% level in 2006-2009. Also, it is necessary to mention that the decrease of the indicator has started

in 2007 – earlier than the official Russian statistics has shown the production decline and recession in the economy.[5]

All in all, the heterogeneity of dynamics of qualitative indicators testifies to the effect that settlement peculiarities exist for those who have chosen small business as the behavioral strategy in the period of crisis.

5. Qualitative indicators of early-stage economic activity in different types of settlements: Dynamic analysis

In terms of the entrepreneurship theory (Chepurenko 2007) the crucial indicators of entrepreneurial potential quality that determine the individual choice to start a business are very closely associated with individual judgment of three main factors:

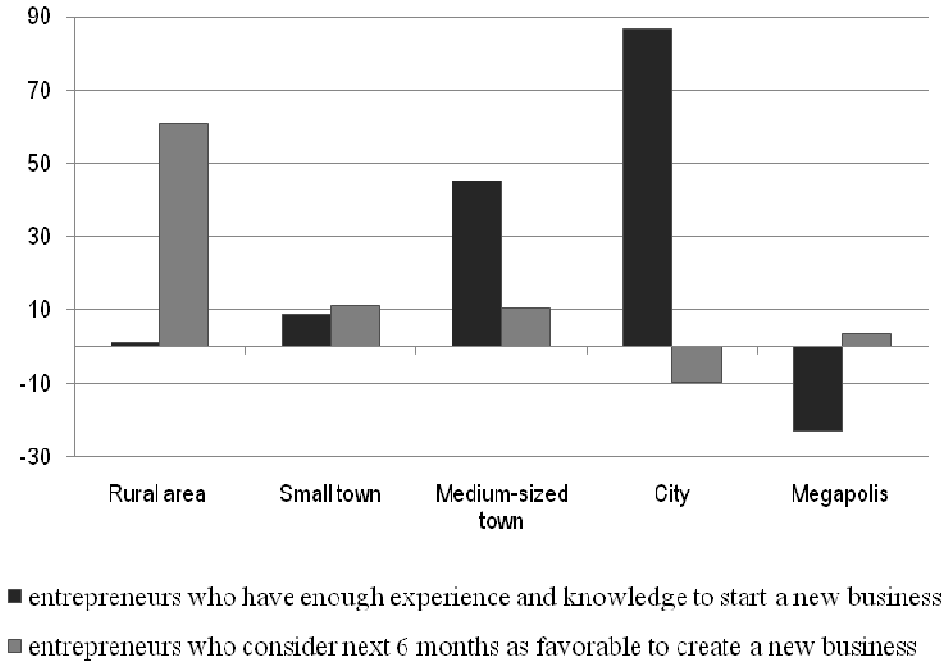
- Framework conditions (i.e. the quality of economic environment) – so-called indicator of perceptual abilities;
- Individual knowledge and experience (i.e. the quality of human capital) – so-called indicator of perceived capabilities;
- Economic risk rate of business activity (on the basis of self-appraisal).

Firstly, the individual assessment of factors specified above has been estimated for nascent entrepreneurs by each settlement type. The group of nascent entrepreneurs has been chosen for the qualitative analysis as the most representative from the cluster of operating entrepreneurship. Furthermore, it is the most unstable and unprotected businessmen.

Secondly, the dynamics of the constructed indicators has been calculated on the basis of growth rates of minimum value (5% significance). The results of the analysis are presented at the figures 3-4.

[5] www.gks.ru

Figure 3. Qualitative indicators of early-stage economic activity: dynamic analysis of future conditions and perceived capabilities, % to all nascent entrepreneurs

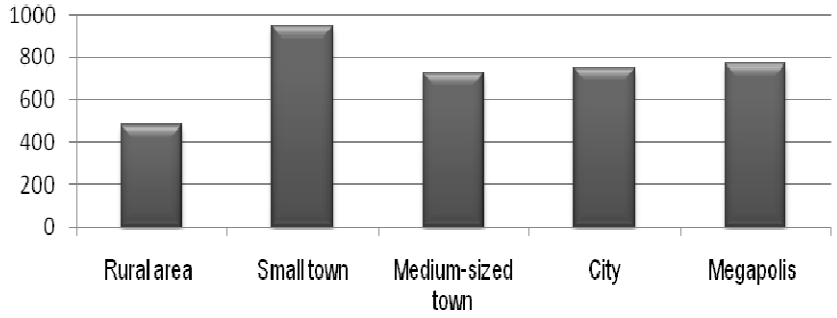


Source: own creation

Qualitative indicators of nascent entrepreneurs, analyzed in the 2006-2008 period show the high level of confidence in the proper skills and knowledge (perceived capabilities) in all settlements, excluding megapolises.

All, except cities see good opportunities to start a new business in next 5 months. On the contrary, the level of fear and lack of knowledge among early entrepreneurs is very high..

Figure 4. Qualitative indicators of early-stage economic activity: dynamic analysis of the rate of "fear", % to all nascent entrepreneurs



Source: own creation

The impact of socio-psychological characteristics of Russian early-stage entrepreneurs on total business activity is rather ambiguous. On the one hand, the fear of failure is very strong: the maximum level is expected among potential entrepreneurs in rural areas. On the other hand, the vast majority of potential and early-stage entrepreneurs consider to have knowledge and skills to open a new business. The rate of "fear" is higher in small towns and rural areas.

Thus, the results of analysis are very controversial. So, we need to check the significance of each factor to the entrepreneurial choice. For this purpose the ordinal regression for three types of settlement is constructed.

6. SME as a behavioral strategy in settlements before a global crisis and under the crisis conditions

First of all, nonparametric statistics has been selected to analyze behavioral strategies in different groups of settlement. The method which has been chosen for the comparative factor analysis is stipulated by the structure of basic data. Secondly, the independent variable (IND) has been formulated in the form of attributive indicator. It is measured in the ordinal scale and possesses three values depending on the extent of individual involvement in the business activities (according to GEM methodology of selecting involvement stages, as shown in figure 1).

So, the following thresholds have been determined for the independent variable:

1. who does not relate to entrepreneurship (IND = 0),
2. who has start-up intentions (IND = 1),
3. who is involved in operating entrepreneurship (IND = 2).

All tested independent variables can be divided into three groups: economical, social and behavioral characteristics. And all these factors are measured at the ordinal scale.

The logic of the study how various factors affect the individual entrepreneurial choice was the following:

- The estimate how significant the differences between quantitative and qualitative indicators of entrepreneurial activity (criterion of rank sums);
- Regrouping of settlements in order to aggregate clusters (enlarge group occupancy) by consolidating insignificantly different types of settlement;
- Contingency table analysis (on the basis of Pearson contingency coefficients);
- Regression analysis

Now consider the successive steps of factor analysis of the settlement involvement in the business activity. The results have been obtained with the use of the analytical software SPSS.

The first step of this research analysis has allowed regrouping settlements into three clusters which are homogeneous by qualitative and quantitative indicators of early entrepreneurial activity:

1. Rural areas
2. Towns (up to 500 thousand inhabitants)
3. Cities (more than 500 inhabitants, including megapolises)

This aggregated classification reflects significant differences in the conditions of entrepreneurial start and development. The selected grouping is homogenous by the human capital of early entrepreneurs and by economic environment but at the same time the groups are significantly different between themselves.

At the second step of the factor analysis the cluster formation has been done to form a primary data file for each aggregated group and for each of four analyzing years. Then the impact of 44 indicators on the predictor has been tested in each primary data file (on the basis of Pearson contingency coefficients).

At this stage of analysis it is obvious that the differences in the evaluation of the future for an enterprise affect the choice to start a business or not. This tendency is typical to cities and megapolises, rural areas and small towns on 2006. The income sources are also important for a decision to become an entrepreneur.

The level of education is surprisingly insignificant; a weak impact is only tested in small and medium-sized towns. In rural areas a certain influence can be rendered by the stereotype the entrepreneurship is a desirable career. Despite the controversial dynamics of "fear" for nascent entrepreneurs, the impact of the percentage of those who is afraid to start a new business has not been tested.

The evaluation of Pearson contingency coefficients has determined the factors that are significant for making a decision to start a business. And this result has allowed passing to regression factor analysis.

Thus, the most stable impact on the entrepreneurial involvement has been demonstrated by the indicators of social networks and perceived capabilities irrespective of the settlement type and the year.

These factors have turned out to be significant in all constructed regressions. The interpretation of the variables that has been used in the regression is presented at the table 1.

Table 1. The factors of behavioral choice: interpretation of the variables

Mesure	Description
Indicator of social networks	Percentage yes on item: You know someone personally who started a business in the past 2 years?
Indicator of perceived capabilities	Percentage of 18-64 population (individuals involved in any stage of entrepreneurial activity excluded) who believe to have the required skills and knowledge to start a business
Salary as a main source of income	Percentage of 18-64 population who consider a salary as a principal source of her/his income
Profit as a main source of income	Percentage of 18-64 population who consider a profit as a principal source of her/his income
Another principal source of subsistence	Percentage of 18-64 population who consider another sources of income as the main ones

Source: own creation

The ordinal regression (binding function – logit) has been used as all independent variables detected at the stage of contingency analysis are measured at ordinal scale. The dependent variable that is also the same for all regressions is ordinal. As it has been pointed out earlier, the regressant divides the respondents into three thresholds: the population not involved in the entrepreneurship, those who have start-up intentions (latent entrepreneurs) and operating entrepreneurs. Table 2 represents the summary of all model estimations

Table 2. The consolidated information about regression estimations: Pseudo R-Square

Type of settlement	**Pseudo R-Square (Nagelkerke)**			
	2006	**2007**	**2008**	**2009**
Rural area	0,498	0,565	0,694	0,678
Town	0,574	0,611	0,572	0,555
City	0,533	0,571	0,508	0,825

Source: own creation

As the summary information of tables 2-3 shows, all factor models are significant, the level of total dispersion explanation by the business involvement feature varies (by years and by settlement 50-70%). So, model assessment is quite good. The significance chi-square is less than 0,0001.

Table 3. The consolidated information about regression estimations: chi-square significance

Type of settlement	**Model fitting information (Chi-square sig.)**			
	2006	**2007**	**2008**	**2009**
Rural area	3,00E-26	4,30E-24	1,50E-32	4,41E-34
Town	8,00E-55	1,60E-61	2,30E-49	3,65E-01
City	1,00E-56	2,90E-50	2,00E-48	6,64E-61

Source: own creation

Additionally, the observed values differ insignificantly from expected frequencies that have been estimated on the basis of the regression models. Moreover, the significance of Pearson chi-square test decreases from 2006 to 2009 (see table 4). That means a very high approximation degree.

Table 4. The changes in structure and significance: 2006-2009

Type of settlement	changes in factor structure through 2006-2009 (+\–)	changes in significance of factors through 2006-2009(+\–)
Rural area	+	+
Town	–	+
City	–	+

Source: own creation

We would remind that estimators of regression parameter are thresholds for dependent variable and locations for factors. Analyzing locations can interpret the influence of selected factors and evaluate the rate of such influence assessing the probability that the regressor would attain the location level under certain conditions of the model.

For inhabitants of **rural areas**, in 2006 the most significant impact renders the indicators of perceived capabilities (the appraisal of knowledge and skills to launch a business) and of social networks (the question whether a respondent is acquainted with an entrepreneur who has started a business in the recent two years). Income indicators are insignificant for this group. In 2007 the indicators of perceived capabilities and social networks are less significant in rural areas. The profit as a main source of income has the highest influence. In 2008 the list of significant factors changes again. The impact of social networks is positive whereas the impact of perceived capabilities is negative this year. In 2009 the significance of these two factors is much stronger. Equally important that the impact of both social networks and perceived capabilities is positive. So, in crisis (2009) proper skills and knowledge can positively affect the decision to become an entrepreneur.

For **town** residents observed and expected frequencies differ insignificantly, as Pearson chi-square equals 0,025 in 2006 and falls to 0,00001 in 2009. Over a four-year period of analyzing settlements (from 2006 to 2009) the regression compiles the same four factors. The indicators of social networks and perceived capabilities are significant in towns as well as in rural areas. But the impact is not absolutely the same because these indicators affect negatively in towns (and in different ways in rural areas).

Besides, the regression models include three indicators of main income source – a salary, a profit and other sources (pensions, unemployment benefits, scholarships and so on).

All these variables affect positively on the entrepreneurial involvement in 2006-2008.

The composition of significant factors in **cities** over an analyzing period is rather similar to the combination of variables for the towns. The list of factors remains the same during the analyzing four-year period but the significance is changing. Thus, the indicators of social networks and perceived capabilities affect negatively on the entrepreneurial involvement in 2006-2007. Then, in 2008 the perceived capabilities become the positive and the most significant factor. And as "fortune favours the bold", certain group of people may become businessmen under crisis conditions. Among variables of income sources the profit as a main source of income influence positively in cities over an analyzing period. In 2006 this factor has the most significant impact.

Table 5 presents the main results of all regression models for 2006-2009 in three settlement types.

Table 5. SME as a behavioral strategy: parameter estimates

Year	Factors estimate/ Type of settlement	Indicator of social networks	Indicator of perceived capabilities	Salary as a main source of income	Profit as a main source of income	Another principal source of subsistence
2006	Rural area	0,97	-17,08	not sig.	not sig.	not sig.
	Town	-15,53	-18,14	0,64	3,58	0,49
	City	-15,65	-18,38	-0,14	2,48	-0,01
2007	Rural area	-1,9	-2,09	0,45	1,63	-0,45
	Town	-17,15	-18,92	0,62	2,09	0,11
	City	-0,43	-3,18	-0,34	1,2	-0,16
2008	Rural area	2,53	-4,99	not sig.	not sig.	not sig.
	Town	2,88	-21,9	0,31	2,53	0,48
	City	0,05	-24,63	0,29	2,36	-0,18
2009	Rural area	16,07	20,25	no data	no data	no data
	Town	-0,09	-0,3	no data	no data	no data
	City	1,29	20,88	no data	no data	no data

Source: own creation

On the whole, the regressions show that in pre-crisis period from 2006 to 2008 the positive impact have main sources of income and social networks in rural areas. In contrast, under crisis, in 2009, indicators of social networks and perceptual capabilities have become positive factors in rural areas and cities. So, these variables have changed their initial impact in the model. In towns in 2009 social networks and perceptual capabilities have stayed the negative factors whereas the income indicators have become insignificant.

7. Conclusions and implications

The study has established the generalities and differences in the development of Russian early-stage entrepreneurs before recession and under crisis conditions.

The analysis of the factors that have set conditions for observed differences have revealed the fact that under crisis conditions the decision to start or continue a business activity depends not only on economic but also on social and psychological conditions

The study that has been conducted has exposed a limited set of factors influencing an individual choice of behavioral strategy.

After the comparative analysis of early-stage entrepreneurial development in different types of settlement we have chosen and tested more than forty variables in order to prove whether there is a connection between these indicators and entrepreneurial involvement of Russian population. In all settlements the interconnection between entrepreneurial choice and indicators of social networks and perceived variables have been found. Additionally, principal sources of income can affect the business activity. Still, the influence of the factors in unstable over an analyzing period and differs in three settlement groups. In fact, from 2006 to 2008 main sources of income and the indicator of social networks in rural areas affect positively. On the contrary, under crisis the impact of perceived capabilities and social networks has increased in rural areas and cities.

So, the strengthening influence of the indicator of social networks has become the common feature of all settlement types. The future research would allow to figure out whether it is a national characteristic of Russian entrepreneurs or a common indicator for all businessmen. At the same time it is obvious that the informational, fiscal and other support and the development of social networks (including microfinance centers and business-incubators) are much more important than just additional resources for crediting small Russian business.

Under detailed consideration the received results can be implied as the base for development of complex, differentiated by settlement type approach government support for Russian SMEs, especially under crisis conditions. The results of the research can be also used for further elaboration of social and information policy in this field, organization and distribution of educational and re-training programs.

References

Acs, Z. – Desai, S. – Klapper, F. 2007: A comparison of GEM and the World Bank Group entrepreneurship data. *Third GEM research conference: entrepreneurship, economic development and public policy*.

Arenius, P. - DeClercq, D. 2005: A network-based approach to opportunity recognition. *Small Business Economics*, Vol. 24 No.3, pp.249-65.

Arenius, P. - Minniti, M.. 2005: Perceptual Variables and Nascent Entrepreneurship. *Small Business Economics Journal*. 24(3):233-247.

Arenius, P. – Ehrstedt, S. 2008: Comparative study of the entrepreneurial startup process in the GEM countries: Turning potential into entrepreneurial activity. *// IX International Academic Conference "Economic modernisation and Globalisation", SU HSE, Moscow, 3-5 April*.

Chepurenko, A.Yu. 2007. Sociology of Entrepreneurship. Publishing house SU HSE, Moscow (in Russian).

Gabelko, M.V. 2008: Evaluation of perspectives to entrepreneurship development by Russian population: optimists, pessimists and those who doubt – who are they? *Finances and Business*, N 3, Saint-Petrsburg (in Russian).

Gabelko, M.V. 2009: To be or not to be an entrepreneur: spatial comparisons of the entrepreneurial start estimation by Russian population. *Questions of Statistics*, N 5, Moscow (in Russian).

Jessen, R.J. 1985: *Statistical survey techniques*, F&S, Moscow.

Krugman, P. 1999. *The Return of Depression Economics*. Eksmo, Moscow.

Obraztsova, O.I. 2007: The Entrepreneurial potential of the Russian society: Russia in Global Entrepreneurship Monitor. *Questions of Statistics*, N 7, Moscow (in Russian).

Murzacheva, E. Strategies of Financial Resource Attraction by Nascent Entrepreneurs in Russian Federation Compared with World Wide Tendencies *// Publishing house of Tallinn University of Technology: special issue on Entrepreneurship and Small Business Research of The Journal "Research in Economics and Business"*. Tallinn, 2008

Reynolds, P. – Bosma, N. 2005: Global Entrepreneurship Monitor: Data Collection: Design and Implementation. *Small Business Economics* (2005) 24: 205–231

Sternberg, R. - Wagner, J. 2002: Personal and Regional Determinants of Entrepreneurial Activities: Empirical Evidence from the REM Germany. *Discussion Paper No. 624, IZA*, November.

The official site of GEM consortium http://www.gemconsortium.org/

The official site of Federal Agency of Russian Statistics www.gks.ru

Credit Risk Evaluation in the Small Business Activity: Individual Aspects at the Macroeconomic Level

Ekaterina Murzacheva[1]

The key area of the controversy among researchers is the potential impact of the small business activity on the national economic growth. The background of such a prerequisite lies in the Schumpeterian hypothesis, which marks out the role of an entrepreneur as an inevitable element for the technological innovation. In this way the growing Gross Domestic Product (GDP) is fungible with the progress and well-being improvement. However, last events showed that high growth rates and favourable competitiveness indicators were not the cure for the financial distress in the era of globalization.

Following this idea, one should take the role of the small business activity into consideration not only with respect to its potential growth input, but also as a possible «economic cushion» that mitigates the overall risk in the economy. Such a view on the problem is hugely exacerbated especially during the periods of economic downturn.

With this in mind the main objective of the paper is to identify basic risk factors perceived by small businesses, which impede the early entrepreneurial development as well as to draw a line between risk perception by entrepreneurs and economic progress. Additionally, the focus is intended to provide a rationale for a leading indicator construction, which is able to anticipate possible economic downturn. This issue is based on the assumption that early entrepreneurial activity is the most susceptible to the economic fluctuations structure.

The scope of the research is confined to the early business activity and the financing problem in particular. The opportunity to attract financial funds enables an entrepreneur to put the business into action, thus drawing the line between bright idea and the first material step towards its realization. Among various dimensions of small business support the financial issue is the most challenging and urgent component of every start-up.

To shed more light on the subject of the small business financing and its impact on the socioeconomic development of a country, one should take a structured approach:

Firstly, it is necessary to drill down into the background of the assumption that the small business activity fosters the GDP growth. Moreover it is important to bring on the surface the explanation of the relationship between the financing source and the level of entrepreneurial activity in the country.

Secondly, as far as the financial issue is concerned, it is crucial to identify risk factors which rein back or, conversely, put forward the actions of an entrepreneur. The counterpart

[1] Ekaterina Murzacheva, BA, researcher, State University-Higher School of Economics (Moscow)

of this process is the investor's outlook whether to sponsor small business or not, which also stems from the observed risk environment.

Thirdly, after economic risk factors are cleared out the analysis of the revealed risk factors – their likelihoods and consequences should be applied. Additionally, with the help of estimation and evaluation techniques the risks need to be ranged according to their significance on the national level, forming an indicator of the possible economic downturn.

Lastly, in order to achieve the stated goal and to condense the information into a single number it is essential to combine the revealed indicators into a complex index. The sense of this result is to absorb both internal characteristics (i.e. motivation, company profile) of small business units, given their perception of economic risks, and external environment with its overall risk concentration.

In this way, evolution of the idea to regard the small business activity as an analytical tool for the risk determination in the economy is developed in the research background section. Subsequently, two approaches are suggested in the paper. The theoretical foundation is presented in the form of structural models in order to provide an economic reasoning for the proposed views. The empirical evidence is outlined in the second part of the paper based on the statistical models' results.

Keywords: entrepreneurship; credit risk; informal capital; GEM

1. Research background overview

In a growing stream of academic literature on the topic of small business activity and financial resources attraction the concern about entrepreneurial contribution to the growth and competitiveness is evident. Keeping this statement in mind, two explanations of the relationship between entrepreneurial growth and overall economic development are set forth in order to cover the first point marked out in the previous section.

1.1 Entrepreneurial activity and its influence on economic performance

On the one hand the dissemination of small firms in the economy triggers off an increase in production rates and output growth. For this reason the nature of such a relationship should be understood. One of the arguments lies in the conceptual framework elaborated by Paul Reynolds who managed to establish «a comprehensive assessment of the role of entrepreneurship in national economic growth» (Reynolds 2005).

On the other hand there is one more channel for the impact of entrepreneurial activity on the socioeconomic development – creation of additional workplaces. This theory runs smoothly in practice unlike the previous one. It is generally accepted that the rise in the number of employed people stimulates the production growth as well as individual welfare improvement in terms of higher income rates. Admittedly, the ability of small business firms to generate jobs is not constant under different

circumstances. According to these findings it turned out that an increase in the labour force does not always lead to the growth of afirm (North 1995). On the contrary, entrepreneurs in urban areas are likely to minimize the number of additional workers to achieve higher growth due to the better productivity rates.

Both of considered theories are in line with the attempt to explain the priority role of the small business activity. As far as this study is concerned it is crucial to evaluate the primary impetus, which gives rise to a new firm birth. Moreover, the main explicit step towards a start-up organization is initial investments into the business process. Hence the point of the interest is a firm at the early stage of its development. Additionally, the most acute issue concerns the incentives of an entrepreneur to launch business activity and to attract financial sources for this purpose. In this way the latter theory refers to a mature firm which is capable to generate additional jobs over an extended time period. Such a firm has overcome most impediments successfully and is not able to reflect the changing environment in terms of dealing with economic risks. By contrast the former theory allows catching the essence of the starting point and seems to be a clear support for the chosen macro view on micro decisions at the individual level.

To set the stage for the risk identification in the small business financing it is crucial to clarify the influence of financial opportunities on the business expansion. The financial strategy of an entrepreneur is no more significant (from the theoretical side) unless it has some economic reasoning, which generates special expectations, behaviour models, informal rules, affecting the entrepreneurship environment as a whole. And conversely, specific features of small business in some area can leave an imprint on the business strategic view, including the choice of financial capital. In the last sense, the dominance, or the permanency, or a regular combination of different financial sources can turn out to be the most adequate and appropriate indicator of the economy's soundness.

For an entrepreneur as for a single economic unit two ways of financing are available: own funds and a debt. However, the notion of the debt financing implies a vast range of unique alternatives in distinction to a corporate organization. Informal capital, as one of those peculiar options, can be posed between self-financing and debt-financing and also can be regarded as a two-fold problem of both financing and investment decisions.

The side of formal financial support is thoroughly investigated in the literature and taken into consideration by policy makers. The focus lies in the special legal, tax and information environment which enforces lending technologies through the supply of different types of credits. Until recently it has been considered that the small business financing through banking facilities is the most convenient and effective source. Nevertheless, the study of Thorsten Beck reflected upon the contentious issue about the positive role of banking sector «in enhancing economic growth through more efficient resource allocation» (Beck et al. 2004).

Informal capital investigations are in tune with the drive towards the growing demand for alternative financial sources due to the lack of available loan products. The prevailing form of informal investments is venture capital which is the subject of an acute interest for both researchers and policy makers. Such an investment source fills in the financial gap during the growth of a new firm. Anyway, there are some peculiarities connected with the consideration of venture capital. First and foremost, classic venture capital implies the allocation of financial funds among young entrepreneurial firms with a high growth potential (Mason, 2002, p.430), in other words, among technologically innovative small businesses. Thus a large portion of entrepreneurial activity is out of the focus, although it is equally important for the overall economic development.

According to C.M. Mason and R.T. Harrison classification (2002) informal venture capital market comprises, firstly, private individuals (so-called business angels), then direct forms of investment by non-financial companies (known as corporate venture capital) and, lastly, institutional venture capital (investments from pension funds, banks and insurance companies). All of these types shape the aggregate supply of informal funds which is not distributed evenly across the economy and therefore is seen by policymakers as a key instrument for the economic growth and employment enhancement.

On the hypothesis of the relationship between the financial choice of entrepreneurs and the development of entrepreneurial activity in the country it is remarkably, that each financial source alludes to both internal motivation and external environment. As a result the way of small business financing (either formal or informal) reflects the level of risk perception by an entrepreneur. In the aftermath of understanding the background of the financial choice it is necessary to disclose the assigned risk which underpins the economic behaviour of an entrepreneur. By managing the supply of financial funds (through the encouragement of various institutions) and by changing the surrounding conditions (political, social context) one is able to achieve risk mitigation in the economy.

The necessity for the risk investigation and its further measurement techniques are captured by the risk-management theory, which began to evolve from micro approach and has reached the macro-level view nowadays. The next step concerns reasons and analytical tools for the risk consideration and is presented in the following section.

1.2 Risk management outlook

In recent years all the sectors of the economy in highly-developed countries have focused on the management of risk as the only way to make an organization successful in increasing its economic value. To handle risk appropriately at the firm's level is the task within one business unit. Such «micro» approach is not consistent with the aim of this study in dealing with the incentives for start-ups.

Keeping in mind, that the opportunity to attract financial funds enables an entrepreneur to put the business into action, such features as credit risk, business risk and strategic risk should be embedded into the analysis in order to tackle the issue of capital attraction for start-up funding.

Credit risk evaluation based on the calculation of minimum capital requirements for credit risk was elaborated by Basel Committee on Banking Supervision and presented in the document called «International Convergence of Capital Measurement and Capital Standards». The main objective of the standards was to ensure the stability of financial system and prevent excessive risk-taking by financial institutions. To translate the goal into reality it is assumed that ex ante methods should be applied to prevent possible financial distress, since ex post systems seem to be delayed.

The main point perceived by the Basel Committee while accepting the Revised Framework in June 2006 was to enhance available provisions by means of more risk-sensitive capital requirements elaboration. This initiative was triggered off by the growing demand for the risk integration into the decision-making process not only at the single firm's level but also according to the needs of supervisory authorities.

The document "provides a range of options for determining the capital requirements for credit risk to allow banks and supervisors to select approaches that are most appropriate for their operations and their financial market infrastructure" (Basel Committee on Banking Supervision 2006).

The same view can be applied to the entrepreneurial activity consideration. It is no matter whether risk-management aspects are used by a single business unit, the more important issue is the ultimate outcome. If it is possible to determine how much financial risk is absorbed by the small business activity in the economy, then it is an appropriate indicator of the behavioural expectations and intentions among economic units.

2. Conceptual framework

In compliance with the risk management theory business risk refers to «uncertainty about the demand for products, the price that can be charged for those products or the cost of producing and delivering products» (Crouhy 2005). In relation to the beginning of the small business activity an entrepreneur faces the possibility of lower earnings than he has initially anticipated all along of changing market conditions. The effect of business risk determines whether it is possible to maintain the consistent performance during the early stage of a small firm.

One of the ways to eliminate presumable losses resulting from the business risk acceptance is the financial choice of an entrepreneur. Furthermore, negative

consequences stemmed from the uncertain event of worsening market conditions encompass either the waste of own investments or the liability to formal/informal creditor. For each of the mentioned alternatives special risk metrics can be assigned according to the risk criteria of associated costs and benefits:

Self-financing – the lowest risk level, individual savings is a risk exposure (the lowest capital price);

Informal capital – medium risk level, the liability to an investor who is ready to accept a large portion of credit risk due to the specific business nature (medium capital price);

Formal capital – demonstrates the highest risk level. The liability to a creditor whose strategy implies limited credit risk acceptance because of the external regulation and internal policy (the highest capital price).

As far as the investor is concerned there are two major types of economic risks: strategic risk and credit risk. «Strategic risk refers to the risk of significant investments for which there is a high uncertainty about success and profitability» (Crouhy 2005). This uncertainty is clearly illustrated by the decision making of business angels (private investors) as opposed to corporate venture capital. Since this kind of creditors calls the shots of the business, their investments are incorporated into the whole venture as well. Moreover, if a small firm without the growth potential is assumed the outcome is much vaguer.

On the part of the corporate venture capital and banking institutions credit risk is on the alert. Generally, it involves a shift in the credit quality of a borrower: not necessarily a default but also any reduction in expected returns on investment. What is more, the crucial interest lies in the field of potential exposure, in other words in case of default a bank (or a venture capital organization) will lose either the whole position or only a part of it if some amount of the loan is recovered (Crouhy 2005).

Credit institutions are supposed to participate in the business process under contractual conditions, thus the probability of losses is substantial for a concrete asset. That is why venture capital investors are seeking for innovative firms and banks prefer large companies that maintain the consistent profitability - it is the only way to limit the adverse consequences of taking credit risk.

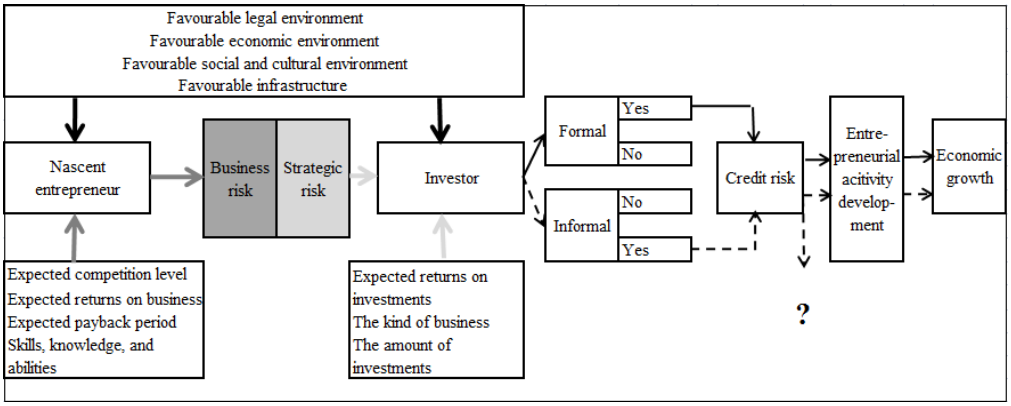
Following this idea it may be stated that at the stage of risk identification the strategic risk is more relevant with regard to private investors since they directly participate in the small business process, thus reducing possible adverse outcomes. Institutional and corporate venture capital as well as credit organizations tend to set bounds for the small business financing because of their inability to keep informed enough in order to assume more credit risk.

The snapshot of a conceptual model is presented in figure 1 below. An entrepreneur at the initial stage of the business development is influenced by both internal motives and external characteristics (Chepurenko 2007). The degree of these factors' perception determines the extent to the business risk acceptance. The outcome of this process is the financial choice of a nascent entrepreneur. At the

same time an investor encounters a risk of the unknown future returns, gained from the investments. As a result the model equilibrium is one of the two options: formal business financing or informal support.

If the formal credit is a dominant alternative, given the external environment and internal factors, then it speaks for the rational economic reasoning, based on the value maximisation by both sides (a borrower and a creditor) and leads to the economic growth encouragement. Although there is one exception: credit risk concentration. In case credit institutions assume excessive credit risk, neglecting due diligence procedures, that is a sign of the financial distress, as the whole mechanism has been damaged by wrong incentives.

Figure 1. Conceptual model: the interaction between nascent entrepreneurs and investors given the financial risk environment



Source: own creation

The choice in favour of informal funds characterises two events: either a favourable outcome or a negative effect, depending on the same issue – credit risk concentration. Moreover, high credit risk indicators can be regarded as inadmissible by formal credit institutions, thus passing them over to informal investors who are forced to take this excessive risk at their own account. If this portion of risk purports the ineffectiveness of the business implied, then it can turn out to be irrecoverable losses for the economy. And vice versa, if this excessive risk is accepted by informal investors because of the immaturity of the financial system then it is a single way to realise business opportunities.

Eventually, an entrepreneur will perceive the worsening conditions for the business differently, depending on the kind of the financial source attracted. In this way, it is necessary to set a range of specified parameters in order to estimate three kinds if financial risks in the economy mentioned above: business risk, strategic risk and credit risk.

According to Risk Management Standard elaborated by Federation of European Risk Management Associations the internal risk management process consists of three major blocks: risk assessment, risk treatment and risk monitoring (2003). To transmit the idea of the study down the line it is proposed to modify micro approach into macro one, preserving main principles and techniques.

The starting point is risk assessment, which comprises a series of stages following one after another. The first step is risk identification and subsequently risk description. The scope of risk encompasses events, which initiate the realization of adverse outcomes. The significance of the consequences for a person or an organization concerned is boiled down to risk tolerance. High level of tolerance indicates a sustained response of an entrepreneur to the changing environment, in other words his readiness to assume much risk until circumstances alter substantially. As it was mentioned before, private investors (as well as venture capital providers) prefer firms with high-growth potential. Whilst banks (and other direct financial institutions) more often reject in small business financing due to uncertainty and a danger of credit risk concentration.

Risk appetite implies the quantification of risk tolerance and includes three main parameters: possible losses, price of risk and risk control. The last item is embedded into the financial choice of an entrepreneur in case of business risk. As far as strategic and credit risk are concerned there are two possible explanations. Firstly, private informal investor is a direct participant in a small business process, so he is able to monitor the allocation of financial resources. In its turn banks establish comprehensive scoring systems in their business processes in order to track outstanding positions. Secondly, if the scope of risk is not confined to the risk appetite a creditor has an opportunity either to accept an excessive risk exposure or to refrain from the small business financing.

Table 1. Credit risk evaluation in compliance with Basel II IRB advanced approach

Basel II risk elements	**Adjusted indicator**	**Definition**
PD (Probability of default)	Probability of stopping business activity (ex post)	The share of entrepreneurs, who shut down business
	Probability of default among nascent entrepreneurs (ex ante)	The estimated share of entrepreneurs, who will possibly shut down business
LGD (loss given default)	The share of losses in the total sum invested into small business activity	The estimation of the expected returns on business by nascent entrepreneurs, which are less or equal to the 100%.
EAD (exposure at default)	The outstanding sum of money invested by any investor	The amount of money (as a part of the total sum of the start-up capital) attracted in the form of a loan

Source: own creation

Credit risk analysis stems from the requirements, assumptions and techniques in the framework of the advanced Internal Ratings-Based approach suggested by Basel Committee. The approach relies on the special metrics which provides a way for further capital calculations to cover bank's credit risk (Basel, 2006, p.52). In terms of this study a parallel shift of proposed metrics is implied, keeping all the requirements unchanged at the same time. Thus there is an opportunity to receive an intricate and interconnected system of indicators that captures all the credit risk parameters where both internal and external characteristics are embodied.

The correspondence between indicators calculated according to the Basel methodology and those that are appropriate for the current research is presented in the table 1 above.

3. Database description

The following study is a part of the international project Global Entrepreneurship Monitor (GEM), aimed at the investigation of entrepreneurial activity in the world. The Russian team's participation in the project made possible to carry out several researches concerning small business development in Russian Federation (2006-2008) and to analyze world-wide tendencies for the year 2006.

The opportunities of GEM data compared with the scope of official statistics allow to capture a deeper field of entrepreneurs' and their sponsors' internal

incentives. The example of Russia, as well as other countries, where the statistics is based on the United Nations Organization concept, enables researchers to investigate small and medium size enterprises only after 2 years have passed since their official registration.

The strength of GEM lies in the opportunity to categorize the group of potential entrepreneurs (the stage before the official registration) and nascent entrepreneurs (the stage just after registration and further for 3 months functioning).

The time period of the survey concerns the dataset for 2006 with indicators harmonized and standardized for comparisons in 42 countries – GEM project participants that year.

The object of the study is a group of nascent entrepreneurs marked out in accordance with the accepted classification of GEM methodology. These are people aged between 18 and 64, actively involved in managing a business which they already own solely or jointly, but who have not had any income from the business at all, or have only been receiving it for no more than 3 consecutive months.

According to the GEM methodology formal resources include capital in the form of banking credit or in the form of the financial governmental support. The GEM methodology also makes it possible to factor in various sources of financial resources, including the funds of business angels and love capital. The category of "love capital" (money from relatives, friends, neighbours, work colleagues) differs from the business-angels' investments (third parties who invested their own funds into detached business activity). Although both categories are combined into a group of informal investors: people who have personally provided funds for start-ups of others in the last 3 years.

The methodology used to check the stated hypotheses includes firstly a full descriptive analysis in order to reveal definite tendencies and to form expectations from the research and its possibilities. The regression analysis was applied for the model construction. Factor and variation analyses were used in dealing with the factors identification, influencing the financial behaviour of an entrepreneur.

4. Empirical evidence

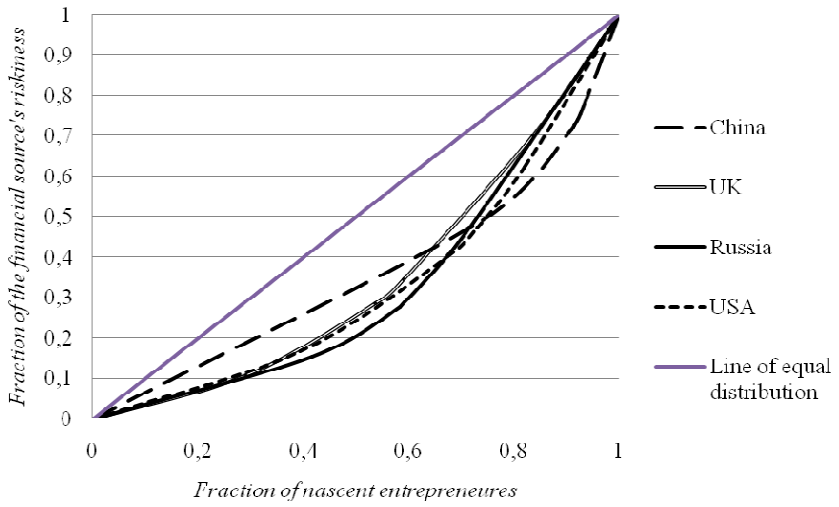
4.1 Business risk evaluation

As far as a small nascent firm is concerned, business risk is an inevitable element in the entrepreneurial performance. This kind of risk implies possible losses which can be regarded as a worse financial result compared to what has been initially expected. To mitigate the consequences of this risk acceptance an entrepreneur manages his financial choice by putting together costs of financial capital and expected losses in case of an adverse outcome. The results are a debt to a bank, a liability to an informal investor or a waste of the own funds.

In this way, taking into the account the entrepreneur's expectations about future changes in the economy the exact financial source will be demanded. External financial sources (which are available in the GEM database) were ranged according to their business risk reflection for the purpose of the further analysis (based on the special risk metrics, introduced earlier).

The graph below (fig. 2) shows the risk distribution among nascent entrepreneurs in several economies. This illustrative example covers 4 countries from different parts of the world with different economic development: China – the Eastern part of the World, Great Britain – a European representative, the USA – a separate continent, and Russia – a transitional economy. Remarkably, all of these countries represent different business risk performance. For Russia the concentration of the "safe" capital is obvious – more than a half of all nascent entrepreneurs attract "biased" capital for the business financing. As for the UK and the USA the same distribution is performed in regard to informal funds, and then an obvious distinction is observed: the demand for formal capital among English entrepreneurs is higher than among American ones.

Figure 2. The distribution of financial sources among nascent entrepreneurs: a differentiation of the riskiness



Source: own creation

The perception of the business risk by nascent entrepreneurs depends on the environment appraisal and internal motives (Sternberg 2005). To clear out the connection between the choice of a financial source (as an indirect characteristic of the business risk tolerance) and some factor variables a nonparametric test was used.

The following indicators were analyzed in the ordinal scale in order to ensure comparability and consistency of the results.

1. *Market conditions* – the estimation of the opportunities' availability in the market for a business launch.
2. *Skills and knowledge* – the internal parameters which drive the entrepreneurial behaviour, justifying his motivation.
3. *Payback amount* – the estimation of the business' future profitability, which characterises the potential of the chosen activity from the entrepreneur's point of view.
4. *Start-up capital* – the scope of the business: designates the welfare of an entrepreneur and his confidence in the success.

The ability to use the nonparametric Kendall-Smith test in order to assess the contingence between two variables in the ordinal scale is justified. The verification of the same distribution of all indicators was achieved: the right-sided asymmetry is confirmed at the 5% significance level.

The most crucial results are presented in the table 2 below (the results for other considered countries did not represent statistically significant values). The most crucial ones were achieved in 5 economies among 42 countries. Market conditions are essential for nascent entrepreneurs in Russia and the Netherlands. That is the degree of business risk taken is the reflection of the economic soundness. Entrepreneurs' skills and knowledge influence the financial choice in Ireland and the Netherlands. In all the considered countries the start-up sum is not significant in terms of the business risk assessment. As far as a payback amount is concerned it is of a substantial importance for nascent entrepreneurs from all these 5 countries.

Table 2. The contingence estimation between the financial choice of nascent entrepreneurs and various motivation factors, based on the Kendall-Smith nonparametric test (with a significance level in brackets)

Country	*Market conditions*	*Skills and knowledge*	*Start-up capital*	*Payback amount*
Iceland	0,129 (0,214)	0,063 (0,434)	0,138 (0,24)	0,889 (0,014)
Ireland	0,165 (0,204)	0,272 (0,02)	0,104 (0,439)	0,577 (0,014)
Netherlands	0,407 (0,066)	0,216 (0,29)	0,078 (0,684)	-*
Russia	0,426 (0,872)	0,039 (0,868)	0,115 (0,5)	-*
Great Britain	0,01 (0,872)	0,033 (0,592)	0,182 (0,01)	0,202 (0,152)

*- there are not enough observations for the analysis

Source: own creation

According to these scattered results it turned out that a set of the distinguished business risk parameters was a quite indefinite entrepreneurial characteristic for the

statistical use, which could not be aggregated to cover all economies. However, it is a useful element for the financial choice interpretation. For example, in Great Britain there are two extremes: formal credit business financing and informal support. Amazingly, but neither perceived market conditions nor entrepreneur's skills and knowledge explain this fact. Nevertheless, a payback amount influences the financial choice, so that, supposedly, more confident entrepreneurs are ready to take banking credit. The thing is that almost one quarter of nascent businesses are supported by the banks' products – the indicator of high risk tolerance of British entrepreneurs because of the low cost of the business risk.

The opposite situation is in Russia: more than a half of nascent entrepreneurs are going to attract informal funds under the influence of market conditions. Low risk tolerance is explained by the high cost of the business risk taken (immaturity of financial system, difficult access to the credit funds and others).

4.2 Strategic risk evaluation

The scope of the strategic risk refers to the downgrade of the returns on investments or to the lack of those returns, or to the total loss of the invested funds. This kind of risk is thoroughly mitigated by credit institutions with the help of the developed scoring systems and elaborated selective process. However, informal investors are not secured from the strategic risk consequences for two reasons. Firstly, because of the risk appetite ignorance and secondly, owing to the residual principle in the small business financing after the refusal of credit institutions.

In both cases informal investments can be regarded as an indicator of the economic situation in the country, although it is not clear if there is a rational reasoning for the strategic risk acceptance: either it is accepted without being identified or it is involuntarily taken.

In order to develop this view on the strategic risk, a widespread attraction of informal funds and, more precisely, the love capital support should be corroborated by the theoretical evidence. The key interest lies in the following: is the informal investment a complementary or additional source for small business' financing, and what are the consequences in both cases, what does it mean for the future development of the entrepreneurial activity and for the economy's evolvement as a whole?

To answer these questions a model was constructed. The parameters were specified up to the quadratic regression model, where all the coefficients are significant at the 5% level. The model views as follow:

$$Y=a+b\cdot X+c\cdot X2+\varepsilon$$

The independent variable is an indicator of the social and economic development (GDP per capita in real prices, USD) (World Bank 2006) and the dependent variable is a level of SME start-ups' informal funding (the proportion of

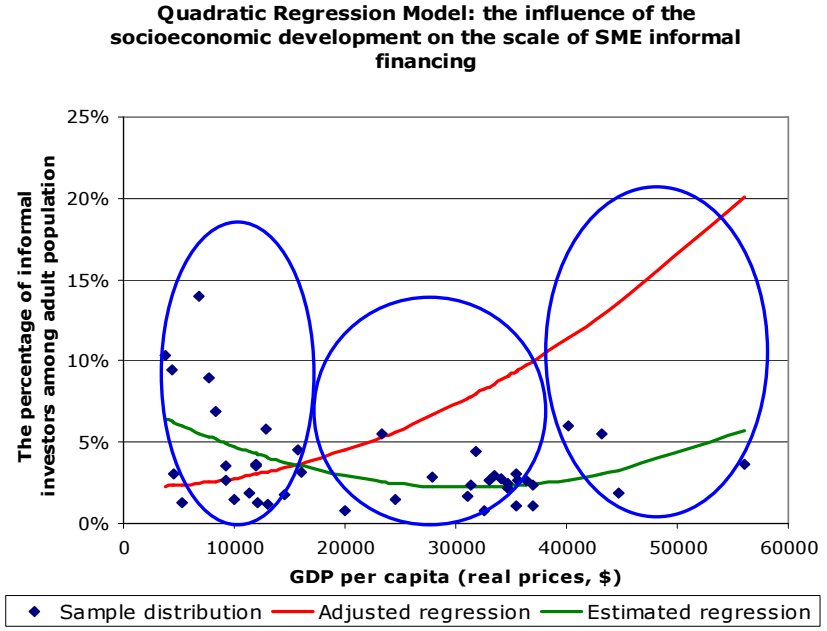
informal investors among the adult population). After the specific procedure (Weighted Least Squares method) an adjusted model was received with the maximum validity achieved (fig. 3).

Countries such as Slovenia, India, China, Brazil, Russia, Latvia, where the level of social and economic development can be said to be low or medium, have typically a large share of informal investments.

Germany, Finland, Czech Republic, Hungary, Denmark, Canada, Japan, Austria, Belgium, France, UK, Sweden on the other hand, have a very small percentage of informal investors, hardly 3% of employable adult population who invest enormous sums (providing the predominance of informal capital), though a level of the population's prosperity is sufficiently high.

And lastly, the USA, Norway, Iceland and Ireland formed the third group, where both socio-economic development and informal investments tend to be higher than average.

Figure 3. Dependence between social-economic development and the scale of informal investments for the 2006



Source: own creation

The constructed model represents the dependence between a level of socio-economic development and small business informal financing. It allows revealing the potential threat to the entrepreneurial activity by means of the excessive informal

support. In other words, identifying the place of a particular country in the parabola one can assess the role of informal capital in SME financing: whether it is a desirable way of funding or unnecessary source.

Therefore, the first group of countries demonstrates high levels of informal funding as the only possible way to promote business activities because there is no other possibility to set up an optimal infrastructure (Saemundsson 2003). At this stage the strategic risk is taken forcedly and reflects its inevitable role in the small business financing.

The second group of countries is in «balance»: a low fraction of informal investors is recognised due to the availability of alternative mechanisms for the distribution of the financial capital among entrepreneurs, which props the high level of prosperity. Possibly, informal investors are not aware of the strategic risk, thus their risk tolerance does not matter and the risk price is not efficient one.

As far as the third group is concerned the role of informal investments as well as the scope of strategic risk is not defined properly. The availability of spare financial resources and high expected returns on investments are able to encourage strategic risk acceptance and to enforce informal investments. At the same time, biased relations with borrower can trigger off the underestimation of risk factors and lead to the excessive risk taking. However, the strategic risk concentration is no more harmful for the economy until it is perceived by informal investors as their voluntarily deposits of available financial resources into the business activity. The occurrence of losses will correct the situation – inefficient investments will not be repeated.

The consideration of the business risk is reined back by their either biased or forced nature, which is difficult to fetch out from the available dataset. The strategic risk tolerance gives a chance to explain the economic development only for a part of the nations, keeping uncovered highly developed countries with a large portion of the informal support.

5. Credit risk evaluation: a leading indicator of the economic downfall construction

Unlike business and strategic risks credit risk arises after the funds have been invested, thus causing unexpected losses because of the downgrade of the creditor's quality. Both formal and informal investors encounter this kind of risk to a different degree irrespectively whether they are "morally biased" to provide funds or not. Over the last decade a lot of banking organisations tried to model credit risk – as a main source of unforeseen troubles.

Basel Committee on Banking Supervision elaborated a range of minimum requirements and a compulsory system of indicators for the credit risk evaluation in

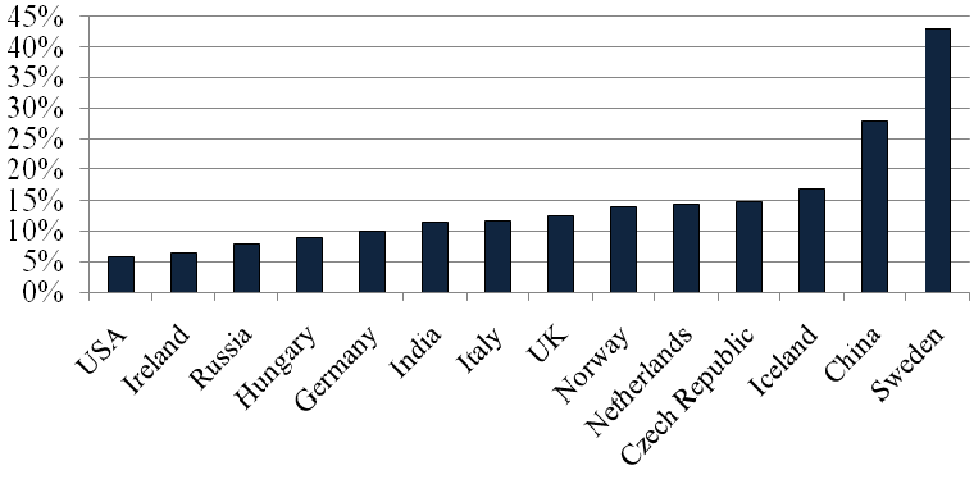
order to calculate the amount of the financial capital needed for a bank to cover unexpected losses.

To conduct a preliminary credit risk assessment not all the countries were investigated. The USA, Ireland, Iceland and Norway were included into the analysis because of their ambiguous position in the World scale (it was cleared out after the strategic risk estimation). Russia, India and China were taken into the account as the largest representatives of the first group, distinguished in the previous analysis. The UK, Sweden, Italy, Germany, Hungary, the Netherlands and Czech Republic were considered to be differing by their location and history European countries.

Firstly, the probability of default was estimated based on the sample number of defaulted entrepreneurs. Factors, influencing the intention to shut down business activity were distinguished with the help of discriminatory analysis (Engelmann 2006). Then these factors were used in the regression model construction in order to determine the probability of default among nascent entrepreneurs. For each investigated country a specific set of significant variables was selected (at the 5% level of significance according to the Fisher criterion).

Among such ex ante factors, which presumably have an impact on the intention to stop business activity, as age, gender, knowledge and skills, market conditions perceptions, skills and knowledge are prevailing in all the considered economies. Only in the UK age and gender turned out to be relevant whereas the rest of the indicators did not demonstrate a statistical significance.

Figure 4. The probability of default estimation among nascent entrepreneurs: low bound



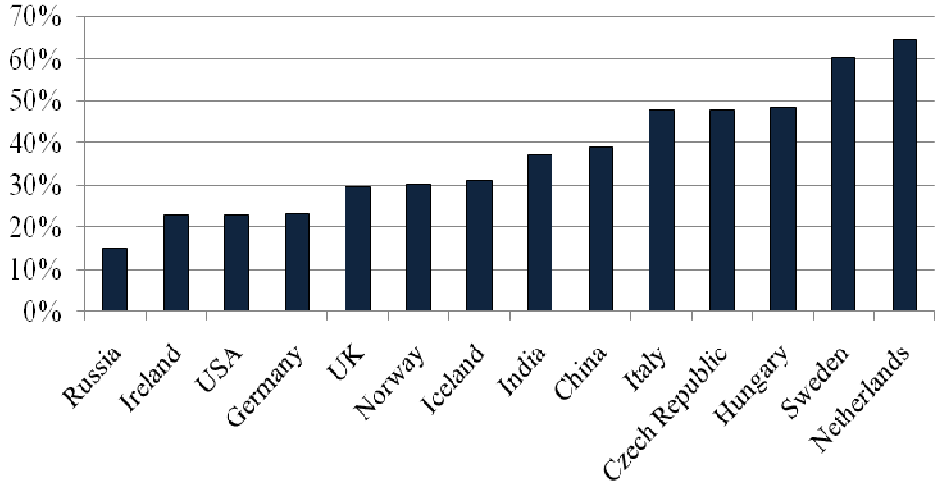
Source: own creation

After the selected factors were applied, the probability of default estimation among nascent entrepreneurs had been received. A range of countries by the estimated probability of default is presented in the fig. 4. Remarkably, the countries where the highest demand for the informal funds is observed (Sweden, China, and Iceland) demonstrate the highest instability of the business activity. Moreover, this result was achieved without taking into the consideration any indicators concerning the choice of financial sources. In this way, the survivability of small business depends not only on the internal perceptions of an entrepreneur but also on the external environment assessment.

The indicator of Loss Given Default was calculated based on the expected returns on the business activity by nascent entrepreneurs as a mathematical expectation of the business profitability. The results can be found in the fig. 5. The main tendency remains stable: countries with unstable business activity are characterised by the highest level of expected losses in the total amount of the borrowed capital. Moreover, these indicators are not interconnected, so the result can be regarded as being confident.

The case of Russia justifies the previous findings concerning an inevitable need for informal resources when the financial system is immature and does not ensure sufficient capital funds. This is the reason for the stable business activity and insignificant losses at the initial stage. Such countries as Czech Republic, Norway, and the Netherlands are at the edge of the stable entrepreneurial activity: a considerable share of expected losses is able to instigate further variability.

Figure 5. Loss given default estimation among nascent entrepreneurs: low bound



Source: own creation

As far as Exposure at Default is concerned, it was estimated as a share of the borrowed capital given the risk concentration (measured according to the Herfindahl-Hirshman index) in this sum. Among all the countries there is a significant (at the 5% significance level) relationship between the financial source chosen and the start-up sum needed to launch a business according to the nonparametric Spearman correlation criterion. Keeping this fact in mind the assumption about the riskiness of the concrete financial source was integrated into the analysis. The concentration of the sources bearing a high risk (that is banking credits and governmental support) in the economy identifies a risky nature of the borrowed capital. The more risky the invested capital is the more considerable the exposure at risk should be, and the more unstable environment is observed in the economy.

Summing up, an indicator of the exposure at risk is received by means of multiplying the total borrowed sum by the concentration coefficient (Gini ratio) of the risk in the economy (based on the financial structure).

In order to construct an ultimate indicator which absorbs all credit risk in the economy connected with the entrepreneurial activity numerically, the following formula should be applied:

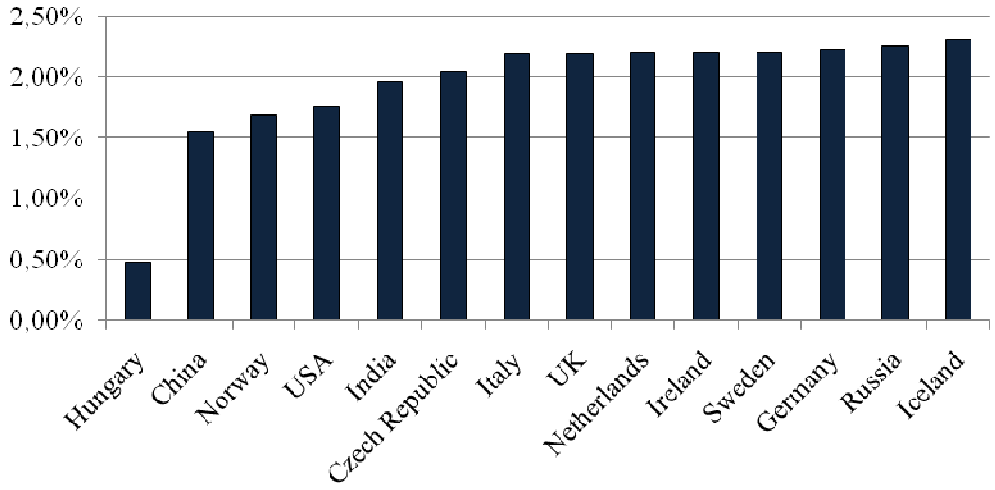
EL = PD*LGD*EAD EL – expected losses in the economy

PD – probability of default

EAD – exposure at default

Taking into the account the scale effect, the received indicator should be adjusted to the size of the country in terms of the GDP volume.

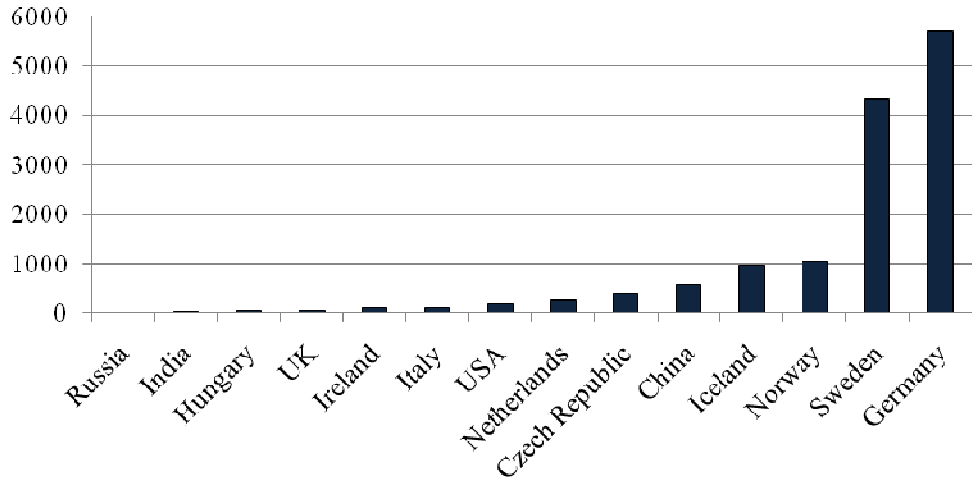
Figure 6. The estimation of the expected economic losses from the entrepreneurial activity in the total amount of GDP (PPP): upper bound



Source: own creation

Further calculations are based on the adjustment of expected losses to the single currency in order to relate the received amount to the total GDP rate, expressed in the USD (PPP). It is possible to get comparable values by finding the proportion between the GDP volume in the national currency and the GDP rate in the PPP (issued by the World Bank community) and applying these proportions to the amount of losses, received in the national currency for each country. Thus, the Global Default Index has been constructed. It shows the maximum share of the Gross Domestic Product needed to cover risks caused by the entrepreneurial activity in order to make the economy more sustainable to the systematic fluctuations. Fig. 6 throws light on the fact that in some European countries with a highly developed economies expected losses (or an entrepreneurial burden on the economy) are significantly higher than in some transition nations: India and China. As for the UK and the USA, due to their stable business financial structure (low PD and LGD indicators) the perception of the risky environment by nascent entrepreneurs softens the adverse consequences of the excessive risk taking. No more that 0.5% of the GDP is needed for Hungary to ensure a safety "cushion" and no more than 2.3% of its GDP is essential for Iceland to keep the economy afloat. Thus, a combination of financial sources determines whether the economy is able to stand against possible financial distress or not.

Figure 7. The estimation of the expected economic loss caused by the entrepreneurial activity per nascent entrepreneur (thousands, USD): low bound



Source: own creation

The other indicator provides more distinct economic sense for further interpretations in terms of numeric and dimension value (see fig. 7). Notably, estimations are provided given the percentage of nascent entrepreneurs among the adult population. A low bound of the confidence interval is taken with the significance level of 5%.

The presented indicator reveals potential losses among nascent entrepreneurs if they go bankrupt at the same point of time. Moreover, it can be interpreted as the mean level of losses among nascent entrepreneurs, which can be also regarded as a global indicator of default, because it comprises both default rates (riskiness of the financial source, internal entrepreneurial motivation) and the level of the economic development. Additionally, it is suitable for international comparisons and ranking. Furthermore, an economic interpretation is possible as well as the measurement of the problem's depth is available. In other words, the mean level of losses reflects the sensitivity of the economy to the losses occurred in the small business performance. Remembering the idea that small business units are the most mobile and vulnerable economic structures, greater sensitivity means quicker response to the subsequent deterioration of the economic situation.

In Russia the expected loss per nascent entrepreneur is not less than 54 thousands dollars to the accuracy of 95% and it is the minimum amount among the considered countries, obviously due to the low level of the early entrepreneurial activity. The expected loss in Germany is not less than 7 million dollars, in the USA – 228 thousands dollars (at the significance level of 5%). These figures disclose the risk concentrated in the economy where small business is functioning. The correspondence between the micro level analysis and the macro view on the problem can be drawn: as in the case of a bank that has to keep enough capital to cover risks arising from its credit affairs, the economy should guarantee that at least 2.25% of GDP can be released to support small business activity in Russia, Germany and so on.

In conclusion, if the economy is quite sensitive to small business fluctuations and nascent entrepreneurs respond to the environmental changes in a swift manner, it can be regarded as a sign of the possible deterioration of the economic performance, expressed as economic downturn conditions. Financial choice of nascent entrepreneurs is a kind of a hedging strategy in terms of the perceived unfavourable circumstances. Unavailability of formal financial sources can be viewed as a reflection of the risky environment and excessive risk concentration in informal funds, which do not require tough screening procedures when providing nascent entrepreneurs with financial capital.

Table 3. Summary of results: risk concentration vs. financial choice of nascent entrepreneurs

	Overall risk level in the economy			
Small business financial choice		***Low***	***Medium***	***High***
	Self-financing	Stable (UK, India)	Stable (Netherlands)	Economic downturn (Germany)
	Informal capital	Economic upturn (Russia, Hungary, Ireland)	Ambiguous (China, USA)	Economic downturn (Iceland, Sweden)
	Formal capital	Economic upturn (Italy)	Stable (Czech Republic)	Volatile situation (Norway)

Source: own creation

According to the research results all outcomes can be divided into 6 groups with a breakdown into 2 axes: financial source and the overall risk level in the economy (see table 3). Pursuant to the mean level of losses among nascent entrepreneurs it is possible to cluster the considered countries into 3 categories: low risk group, medium risk group and high risk group. The sense of the clusterisation is the following: the more capital is needed to cover the risk, caused by the entrepreneurial activity, the more risky environment is, and vice versa.

Remarkably, but a stable situation in terms of the entrepreneurial activity development along with anticipated economic stability is observed in India and the UK (low risk and self-financing). The Netherlands represent a benign picture of the small business favourable influence on the economic welfare (medium risk level and self-financing). Entrepreneurial activity gains momentum in Hungary, Ireland and Russia (informal capital and low risk performance). Italy demonstrates sound financial system whereas China depicts volatile situation (medium risk level along with the informal capital concentration).

These results are especially crucial given the interest to the ongoing crisis development. The "credit crunch" began in the USA in 2006, when the excessive risk concentration in the banking sector became apparent (Reinhart 2008). Meanwhile, the probability of default among nascent entrepreneurs was the lowest that year (see figure 7), as well as the loss given default rate. It is not surprisingly, as it was quite easy to get financial support to launch business activity, possibly, in order to cover increasing liabilities. However, the Global Default Index along with the mean level of losses reined back the optimistic expectations. The amount of capital that was equal to no more than 1.7% of the American GDP was needed to

offset the risk in the economy. An increase in the two last indicators' values, compared to the low level of PD and LGD rates, is explained by the magnified EAD factor. The growth of this indicator was connected either with a rebound of the borrowing rates among nascent entrepreneurs or with an increase in the capital required to launch business activity. More detailed analysis revealed that exactly the former hypothesis was not rejected: entrepreneurs became more cautious and refrained from excessive self-financing by leaving more savings for the future. That is the reason for the USA to be in the medium risk cluster, which gave grounds for the economic downturn forecasting.

6. Conclusions and implications

Thus, it was considered that by choosing a source of financing at the initial stage of development, entrepreneurial activity shaped the general economic environment. On the whole the fundamental idea of the study throws light on the fact that each small business entity is making decisions under the influence of both environmental changes and own motivation how to organise the process. After individual risk profiles were marked out they were aggregated into a single indicator which reflected risk concentration in the economy according to the actions of the primary risk bearers – entrepreneurs.

Therefore special techniques were adopted from the risk management practice, which is thoroughly developed in the corporate sector. In compliance with a standard methodology possible risks were identified and described as well as analytical tools were suggested for their assessment. The concrete metrics was specified for the indicators construction with a preliminary empirical validation.

To identify whether an economy is susceptible to the worsening outcomes or not it is useful to embark on the risk consideration caused by the entrepreneurial activity. These findings are crucial at the macro level when identifying gaps in the economic functioning and seeking for levers to manage the economic system.

Moreover, the ongoing research is a kind of support for credit organisations which are going to implement IRB advanced approach under the Basel II Framework. The assistance in the PD, EAD, and LGD estimation implies a range of internal factors and external characteristics selection which should be taken into consideration during the forecasts construction. In order to detect shortcomings in the modelling tools, GEM data opportunities come as an appendix to the internal parameters elaboration at the micro level within a single bank. The presented macro characteristics can be regarded as stressed estimators of the borrowers' default, depending on the overall surrounding conditions. Further specification based on the internal banking procedures will enable bank managers to assess credit risk more properly.

Small business is the most flexible and mobile economic structure, at the same time it is the most fragile and susceptible to the external fluctuations one. It is necessary to capture both the inner side of a small business activity and the outer impact on its perceptions in order to imprint the economic situation. It will give an opportunity to reveal certain movements in the process, thus controlling tendencies in the behavour of small business as a beacon for the economic pitfalls.

There are some restrictions and perspectives for the study. First of all, it is a sampled analysis only for a one time period (2006). Dealing with statistical estimators one should be accurate, moreover, the revealed interconnections might have changed through the time – so additional analysis for other time periods is anticipated to check the stability of tendencies. To use multiple regression methods a wider sample is needed to achieve the most appropriate results. The solution will be incarnated when international bases for 2007 and 2008 are available for their combining and expanding the number of observations. The study lacks the detailed environmental analysis of the chosen countries – their legal, political, economic and socio-cultural context. Such materials could enlarge the range of possible hypotheses; nevertheless the mentioned data should be harmonized with GEM data-set to receive correct and comparable results – that is the question of future research. And of course there are inevitable obstacles connected with the nature of data collection: longitudinal research implies the invariability of the questionnaire through the years even if the necessity in more correct and precise information is needed. To compare results among different countries the standardized methods are applied, although each nation is unique in its development and requires different research approaches.

References

Basel committee on banking supervision, 2005: *Studies on the Validation of Internal Rating System*. Bank for International Settlements, Switzerland.

Basel committee on banking supervision, 2006: *International convergence of capital measurement and capital standards*. Bank for International Settlements, Switzerland.

Beck, T. - Demirguc-Kunt, A. - Maksimovic, V. 2004: Bank competition and access to finance: international evidence. *Journal of Money, Credit and Banking,* 36, 3, 627-648. p.

Chepurenko, A. 2007: *Sociology of the entrepreneurship*. Publ. house "SU-HSE", Moscow.

Crouhy, M. - Galai, D. - Mark, R. 2005: *The essentials of risk management*. McGraw-Hill, New-York.

Engelmann, B. - Rauhmeier, R. 2006: *The Basel II Risk Parameters, Estimation, Validation, and Stress Testing*. Springer, Berlin.

Federation of European Risk Management Associations, 2003: *A risk management standard*. AIRMIC, London.

Mason, C.M. - Harrison, R.T. 2002: *The geography of venture capital investments in the UK*. Transactions of the Institute of British geographers, New Series, 27, 4, 427-451. p.

North, D. - Smallbone, D. 1995: *The employment generation potential of mature SMEs in different geographical environments*. Urban Studies, 32, 9, 1517-1534. p.

Reinhart, C.M. - Rogoff, K.S. 2008: Is the 2007 US Sub-Prime Financial Crisis So Different? An International Historical Comparison. *American Economic Review*, 98, 2, 339-345. p.

Reynolds, P. - Bosma, N. - Autio, E. 2005: Global Entrepreneurship Monitor: Data Collection Design and Implementation 1998-2003. *Journal of Small Business Economics*, No. 24.

Saemundsson, R. - Sigurdartottir, G. 2003: *Informal investors: whose money are they investing?* Third GEM research conference: entrepreneurship, economic development and public policy.

Sternberg, R. - Wennekers, S. 2005: Determinants and Effects of New Business Creation Using Global Entrepreneurship Monitor Data. *Journal of Small Business Economics*, No. 24., 193-203. p.

World Bank Data, 2006: *GDP per capita*. Available at: www.worldbank.org.ru

Direct investment and costs of production: empirical results of modelling costs structure applied to Russian industry in 2005–2009

Eugenia Nazrullaeva[1]

In the applied theory of exogenous or endogenous economic growth it is usually assumed that investment leads to capital accumulation with no influence on technological progress. On the other hand, at the firm level investment is aimed at developing innovative technologies (process innovations) which lower firms' production costs per unit of output. However, when a firm decides to make an investment it may consider other reasons besides improving production process. Firms may also invest in expanding the range of goods produced, i.e. invest in product innovations associated with the extensive growth. The question arises: do investments matter?

In our study we analyze a relationship between investment processes and the dynamics of production costs. We discuss the influence of different types of investment, including fixed investment, R&D investment, and foreign direct investment, on the structure of production costs (costs of raw materials, wages, etc.). The aim we pursued was to find out whether higher investments lower costs of production per unit of output controlling for the dynamics of relative prices. If our hypothesis of investment significance for production process and technological progress is not rejected, it means investments are efficient.

We base our study on the official statistics on seventeen key industries in Russian mining and manufacturing sectors, electricity, gas, and water supply sector. The data is taken from the Federal State Statistics Service (Rosstat), and it covers the period from 1st quarter of 2005 to 3rd quarter of 2009. Until 2005 Russian industrial statistics was based on the OKONKh industrial classification. Starting from 2005 the OKVED classification (harmonized with NACE European Classification of economic activities) has been introduced. Unfortunately, the continuity of official industrial statistics failed after 2004. Econometric modelling encounters serious limitations due to the small sample problem. Our previous empirical results suggest that there is the statistically significant relationship in several manufacturing industries. To verify our hypothesis of investment efficiency we assume that the long-run equilibrium found in 1995–2004 remains after 2005, as it takes time to develop and use innovative technologies which lower production costs. We test for cointegration between costs per ruble of output and investment, taking into account the possibility of structural breaks in the data which covers the beginning of the current crisis. Our results suggest that the effect of investment on costs varies across the industries, and the overall efficiency of investment can be seriously questioned. No investment efficiency found

[1] Eugenia Nazrullaeva, PhD student, State University – Higher School of Economics, Economics Department. Moscow, Russia

in mining sector of energy producing materials which accounts for nearly half of total investment in Russian industry.

Keywords: costs per unit of output, technological progress, direct investment, Russian industry, cointegration, structural break

JEL Classification: C22, E22, E23, O30

1. Introduction

In the applied theory of exogenous or endogenous economic growth it is usually assumed that investment leads to capital accumulation with no influence on technological progress. On the other hand, at the firm level investment is aimed at developing innovative technologies (process innovations) which lower firms' production costs per unit of output. However, investment behavior motives go far beyond securing technological progress. Firms may also invest in expanding the range of goods produced, i.e. invest in product innovations associated with the extensive growth instead of the intensive growth. So economic growth accompanied by the growth of private investment may lack the so-called investment nature. According to the Federal State Statistics Service (Rosstat) in the mid-2000s industry accounted for around 80 per cent of total investment in Russian economy. About 40 per cent was invested into fuel industry, particularly into oil industry[2].

So a question arises: do investments matter? For example, if we consider the Cobb – Douglas production function, widely used in theoretical and empirical analysis of economic growth and technological progress, output elasticities of factors used in production are assumed to be constant. If investments matter this no longer holds. The idea is as follows. Suppose the production function determined by capital (K) and labor (L):

$$Y = AK^{\alpha}L^{1-\alpha},$$

The assumption that investments have an effect on the production function may be formalized in the following way:

$$\dot{\alpha} = \gamma I,$$

where I - investment, and $0 \le \gamma \le 1$. Elasticities change proportionally to investments. This framework was introduced in Peretto & Seater (2007) as a model of endogenous growth.

The next question which arises is: what kind of investment processes may influence technological progress? Peretto & Setter (2007) consider cost-reducing R&D investment. We discuss the influence of different types of investment, including R&D

[2] «Investment in Russia», 2005. Source: Rosstat.

investment, fixed capital investment, and foreign direct investment on the structure of production costs.

The empirical analysis of the dynamics of the costs structure together with investment in fixed capital at the disaggregated economic activity level may reveal the nature of economic growth. In this paper a simple framework for this phenomenon is developed and tested. The main hypothesis of the paper can be stated as follows: higher investment leads to lower production costs per ruble of output as a result of technological progress. Its verification assumes modelling the costs structure as a function of investment.

This paper is probably one the first, based on Russian statistics, to look at the investment efficiency and technological progress from the point of view of the costs structure. There is literature that looks at the related but rather distinct theoretical and empirical questions concerning growth, technological progress, investment volatility, and costs of production.

Theory of exogenous economic growth considers the effect of capital accumulation and technological progress on growth. Endogenous growth theory, in turn, questions the nature of technological progress. Aghion et al. (2004) analyze the effects of productivity shocks on output growth in the presence of credit constraints. The transmission channel is the composition of investment: long-term and short-term investment. Short-term investment takes relatively little time and does not have a substantial effect on productivity. Long-term investment, in turn, has a direct effect on productivity. Short-term investment, according to Aghion et al., concerns maintenance of the existing equipment, while long-term investment is connected with investment in R&D and adopting new technologies of production. It can be inferred from the results of their research that under complete financial markets long-term investment and productivity are countercyclical and short-term investment are procyclical. In the presence of financial constraints long-term investment turns to become more procyclical and more sensitive to the exogenous productivity shocks.
Costs of production are analyzed in investment theory in terms of adjustment costs, i.e. costs of changing the capital stock. The nature of adjustment costs is rather close to the ideas proposed in this paper concerning the effect of investment on the costs structure dynamics. Higher short-term investment may bring higher costs per ruble of output because this type of investment is associated with maintenance costs and costs of new equipment installation. Long-term investment, in turn, is associated with developing new technologies of production. In other words, in long-term period higher investment is expected to stipulate technological progress.

Nadiri and Mamuneas (1994) analyze the effect of public sector investment on costs of production structure at the industry level. Public sector capital is being disaggregated at two components: infrastructure and R&D. Private sector costs are modeled as a function of relative factor prices, output produced, and public capital. The main hypothesis of the paper is that higher public investment in infrastructure

and R&D leads to lower costs per unit of production. The cost function is estimated for panel data for twelve two-digit US industries (by Standard Industrial Classification, SIC) during the period 1956–1986. The results obtained by the authors suggest that in case of investment in R&D public capital has a direct impact on the costs structure: it lowers the costs of production per unit of output and leads to higher productivity.

We base our study on the official statistics on seventeen aggregated and disaggregated economic activities in Russian mining and manufacturing sectors, electricity, gas, and water supply sector.

1. Mining and quarrying (C)
2. Mining and quarrying of energy producing materials (CA)
3. Mining and quarrying, except of energy producing materials, i.e. stone, sand, chemical and fertilizer minerals, etc. (CB)
4. Manufacturing (D)
5. Manufacture of food products, beverages and tobacco (DA)
6. Manufacture of textiles and textile products (DB)
7. Manufacture of leather and leather products (DC)
8. Manufacture of wood and wood products (DD)
9. Manufacture of pulp, paper and paper products; publishing and printing (DE)
10. Manufacture of chemicals, chemical products and man-made fibre (DG)
11. Manufacture of rubber and plastic products (DH)
12. Manufacture of other non-metallic mineral products (DI)
13. Manufacture of basic metals and fabricated metal products (DJ)
14. Manufacture of electrical and optical equipment (DL)
15. Manufacture of transport equipment (DM)
16. Manufacturing not elsewhere classified (DN)
17. Electricity, gas and water supply (E)

The data is taken from the Federal State Statistics Service (Rosstat), and it covers the period from 1st quarter of 2005 to 3rd quarter of 2009. Until 2005 Russian industrial statistics was based on the OKONKh industrial classification. Starting from 2005 the OKVED classification (harmonized with NACE European Classification of economic activities) has been introduced. Unfortunately, the continuity of official industrial statistics failed after 2004. We discuss a possible way of recalculating industrial time series in 1995-2004 retrospectively using tables of correspondence; however, the results do not seem to be very fruitful.

According to our hypothesis improvement in technologies as a result of higher investment implies lower costs per ruble of output. In order to measure the effect of investment on the costs structure cointegrating equations are estimated at the disaggregated economic activity level.

The rest of the paper is organized as follows. Section 2 describes the main problems offered by the data and its drawbacks. Section 3 outlines the model, the main results and their economic interpretation. Section 4 draws conclusions.

2. The data

2.1 The data: methodology

The data used is the official statistics borrowed from Rosstat (www.gks.ru). The statistics analyzed in the paper is the following:

- production costs, quarterly data, 2005–2009, in current purchaser prices[3];
- fixed capital investment, quarterly data, 2005–2009, in current prices;
- output produced (turnover from the supply of goods and services as a proxy), quarterly data 2005–2009, in current producer prices;
- R&D investment, quarterly data, 2005–2009, in current prices;
- producer price index (PPI), quarterly data, 2005–2009;
- investment deflators, quarterly data, 2005–2009;
- foreign direct investment (in USD), quarterly data, 2005–2009;
- RUR/USD exchange rate, average, quarterly data, 2005–2009.

According to the methodology of Rosstat investment in fixed capital is aimed at creation and reproduction of fixed capital stock, reconstruction and technical reequipment, acquiring machines, etc. So fixed capital investment is closely connected with a firm's level of technology.

Investment in R&D is a part of investment in non-financial assets not included in fixed capital investment. Unlike fixed capital investment this type of investment represents costs of developing and testing new prototypes of machinery and equipment, i.e. describes a firm's level of technological innovativeness.

Foreign direct investment (FDI) is defined as the accumulation of tangible and intangible assets, cash, capital lease, and credits from foreign co-owners of a firm. We suppose that motivation behind FDI should be more transparent than motivation behind domestic investment. The advantage of using FDI over domestic investment when modelling the costs structure is that we do not need to construct investment deflators. It was discussed in Bessonov & Voskoboynikov (2006) that official investment deflators published by Rosstat are biased and do not reveal the actual dynamics of the prices of fixed assets. Foreign direct investment does not need to be deflated because the dynamics of prices is eliminated by converting investment in rubles using quarterly average ruble exchange rate.

[3] i.e. prices which exist at the moment when goods and services are used in production process

Finally, costs of production constitute an item which includes all expenditures excluded from investment. Therefore, the differences between costs and investment in statistics are rather subtle and depend on how carefully accounting principles are abided by a firm. Costs of production include several components: material costs, wages, depreciation, and other costs. The dynamics of production costs is determined mostly by the dynamics of its main component: material costs. The composition of material costs includes:

- costs of acquiring raw materials, intermediates, and accessories for production process, including costs of acquiring import raw materials;
- transportation costs;
- fuel and energy consumption.

According to our hypothesis technological progress involving higher investment in development of less resource-consuming technologies should result in lower costs per ruble of output due to lower costs of acquiring raw materials used in production. Wages are excluded from the analysis of the costs structure. According to real business cycles theory positive technology shock leads to an increase in wages and labor supply (Romer, 2006). This fact contradicts our hypothesis of lower costs as a sign of technological progress.

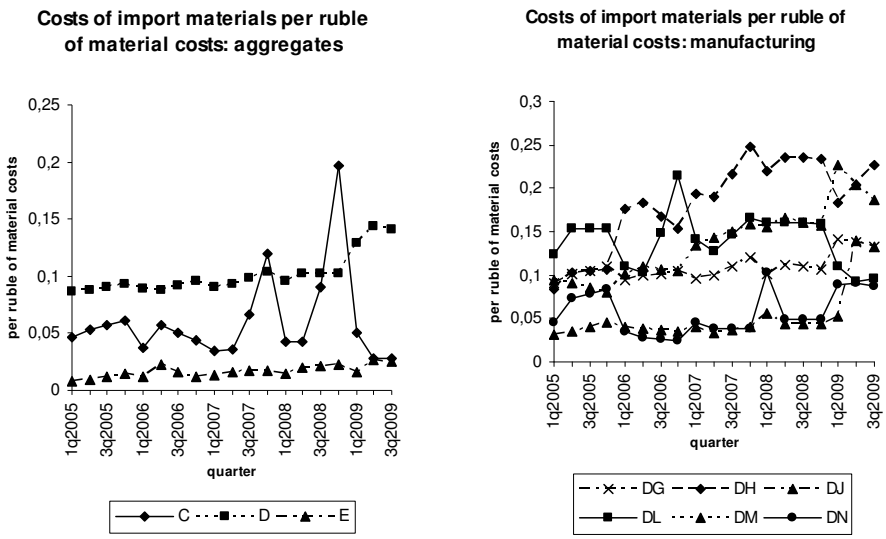
The main problem which arises when working with Russian industrial statistics is that until 2005 Russian industrial statistics was based on the OKONKh industrial classification, but starting from 2005 the OKVED classification (harmonized with NACE European Classification of economic activities) has been introduced. In the OKONKh classification an industry is an aggregate of firms according to their main economic activity. The object of classification is a firm. The drawback of such classification is that a firm has non-core activities that may constitute a large share of its profit. So in OKVED the object of classification is an economic activity. OKVED was officially introduced in 2003, transition period lasted for 2003–2004. So the continuity of official industrial statistics failed after 2004. Nevertheless, there is a possible way of recalculating industrial time series retrospectively for the period 1995–2004 using tables of correspondence published by Rosstat. These correspondence tables allow converting data from the OKONKh classification to the OKVED classification, and based on the principle that an institutional industry may be represented as a weighted aggregate of several economic activities. This approach produces estimates for the main production and financial indicators, though the quality of these estimates remains unknown and thus may be questioned. The weighting matrices used for recalculation are given exogenously. As a rule, the weights are obtained using the data of the transition period 2003–2004: the hypothesis of constant exogenous weights lies behind recalculation. We limited our analysis to the period 2005–2009 because our estimates obtained for the earlier period based on the approach discussed appear to be biased and thereby numerically incomparable, i.e. the approach itself requires further refinement.

2.2. The data: 2005-2009 dynamics and crisis

Because of methodological problems with official industrial statistics time series analyzed in this paper are rather short. Period from the 1st quarter of 2005 to the 3rd quarter of 2009 covers 19 observations. The impact of crisis hit Russian industry in autumn 2008 is supposed to be revealed through the dynamics of costs of production per ruble of output and the dynamics of real investments.

From Figure 1 it may be inferred that wages per ruble of output in mining and manufacturing industries witnessed substantial growth at the end of 2008. Unlike mining and manufacturing electricity, gas, and water supply sector faced a decrease of wages in the first half of 2009. From the figures it also can be seen that the level of wages per ruble of output is higher in manufacturing than in mining. The proportion between costs of labor in manufacturing and costs of labor in mining remains unchanged during 2006–2009 and accounts approximately 4 to 1, i.e. total costs of labor in manufacturing are four times higher.

Figure 1. Costs of production: wages and costs of acquiring import raw materials



Source: Rosstat

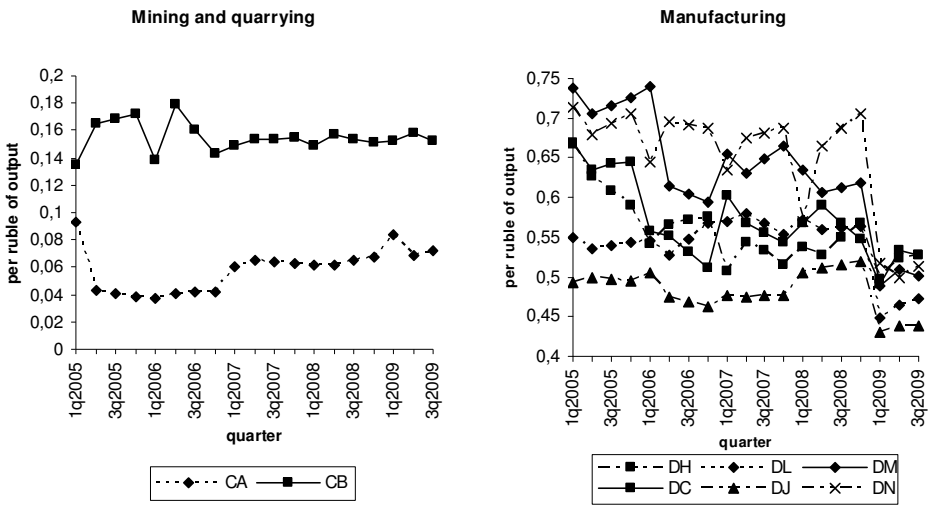
Another factor which should be taken into account in the analysis of the costs structure is the share of import materials used in production process. Costs of import raw materials per ruble of total costs of materials are higher in manufacturing as it can be seen from Figure 1. The autumn crisis of 2008 led to the fall in the percentage of import materials used for production in manufacturing industries. The

activities excluded are manufacture of transport equipment (DM) and manufacture of fabricated metal products (DJ).

One of the first conclusions is that manufacturing industries seem to face severe problems during 2008: real output decreased, import raw materials used in production are substituted for domestic ones.

The dynamics of costs of raw materials per ruble of output (fuel and energy consumption and transport costs eliminated) that we believe being an indicator of the level of technology is shown in Figure 2. Dynamics of costs per ruble of output in mining does not suggest any changes occurred after the crisis. Once again manufacturing industries appear to suffer more from the consequences of the crisis. It can be observed from Figure 2 that material costs per ruble of output in manufacturing plummeted after the 3rd quarter of 2008. So there is a clear structural break in mean. If we compare the effect of the 2008 crisis on material costs with the effect of the 1998 crisis, we see that the situation is quite different. As it was shown in Kantorovich & Nazrullaeva (2009) structural breaks in the dynamics of material costs per ruble of output occurred with a substantial lag at the beginning of the 2000-s. So the crisis of 1998 appears to have no immediate effect on the dynamics of costs, its influence was accumulated by real output and relative producer prices dynamics.

Figure 2. Costs of acquiring raw materials per ruble of output



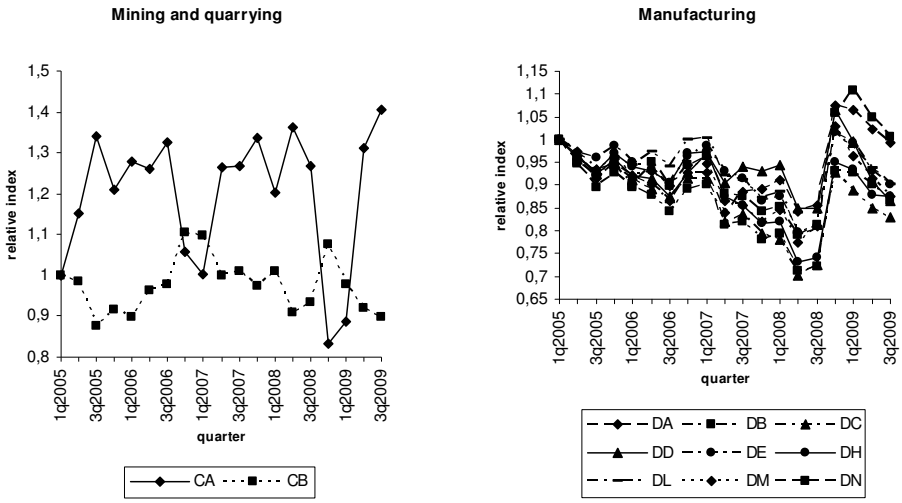
Source: Rosstat

Costs per ruble of output analyzed above are calculated in nominal terms, therefore, when modelling the relationship between investments and the costs structure this

indicator should be adjusted for the dynamics of relative prices, i.e. prices entering into costs relative to prices of output. No deflator of costs is being officially published by Rosstat. In our paper we suggest aggregate producer price index as a proxy for purchaser prices in which costs are measured. Then the proxy which accounts for the dynamics of relative prices in the dynamics of costs per ruble of output is relative PPI index: PPI of an economic activity relative to the aggregate PPI. In the previous version of Russian industrial classification the aggregate PPI was the PPI for industry as a whole. With the introduction of OKVED the aggregate PPI is the PPI which accounts for the dynamics of prices for three aggregate economic activities: mining (C), manufacturing (D), and electricity, gas, and water supply (E). This aggregate is calculated by Rosstat.

Figure 3 represents the dynamics of relative PPI in 2005–2009. As we can infer from Figure 3 there was a sharp reduction of relative prices in mining of energy producing materials at the end of 2008 followed after the fall in oil prices. Manufacturing appears to respond to the crisis faster than mining: a decline for manufacturing activities is witnessed in the 2nd quarter of 2008 followed by an increase after the 3rd quarter. At the end of 2008 we see that prices in manufacturing grew higher than average prices for industrial activities (C, D, E), i.e. relative PPI index exceeds 1. In mining of energy producing materials relative price index is high during the whole period 2005–2009. And the situation seems to remain unchanged if we compare it to the earlier data for 1995–2004. In mining the crisis of 2008 led to the same changes in relative prices as the crisis of 1998: a sharp decline followed by a gradual rise in relative PPI. However, most of institutional industries now aggregated into manufacturing economic activities showed another pattern of dynamics. For machinery, chemicals, food, and light industry relative PPI increased immediately during the crisis of 1998. So current crisis brings some new patterns which suggest that unstable situation in manufacturing industries is quite alarming. After the 3rd quarter of 2008 relative PPI in manufacturing exceeded its level before the crisis. During the year 2009 we observe a gradual decline in relative prices which appear to be a positive sign.

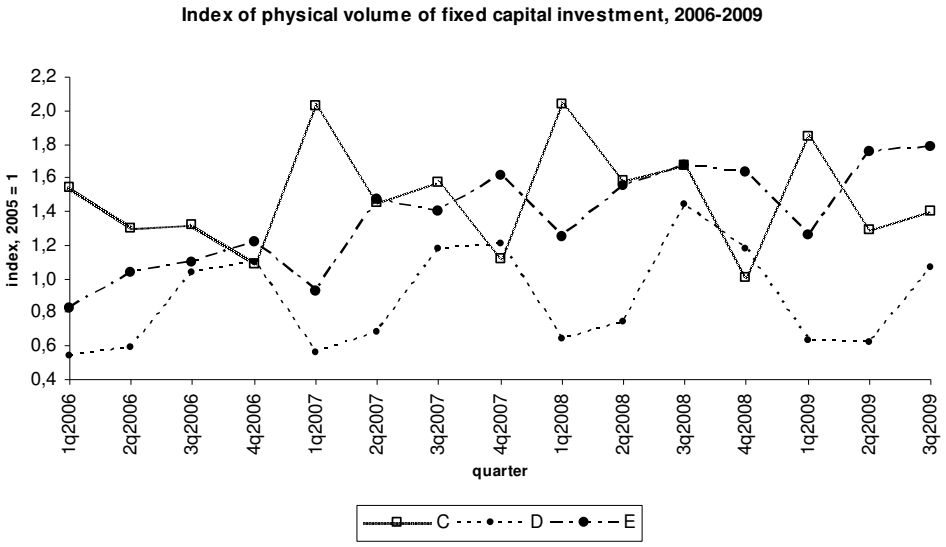
Figure 3. PPI by economic activities relative to aggregate PPI (C, D, E)



Source: Rosstat

Finally, the dynamics of investment and output should be analyzed. From Figure 4 we may see the dynamics of physical volume of production aggregated at the level of manufacturing and mining during 1995–2009 when Russian industry faced the two crises of 1998 and 2008 (the value in 1st quarter 1995 as 1). For the period 1995–2005 the indices of physical volume of production for economic activities are calculated retrospectively by Eduard Baranov and Vladimir Bessonov (SU – HSE) [Baranov, Bessonov (1999)]. The conclusion which may be inferred is that the negative consequences of the 2008 crisis for Russian industry are quite severe: the worst situation is in manufacturing where physical volume of output in the second half of 2008 plummeted more than in mid-1998. The effect of the crisis of 2008 on mining is roughly the same as in 1998, and the dynamics of physical volume of output in electricity, gas and water supply sector seems to be almost unaffected by the both crises.

Figure 4. Indices of physical volume: production and fixed capital investment.



Source: Rosstat.

The situation with investment is rather puzzling. During the period from the 2000-s Russian industry was characterized by stable growth of investment. Observing this growth the hypothesis that higher investment leads to lower costs per ruble of output was proposed. However, it is clear now that after the last year's crisis the situation has changed. Figure 4 suggests that in terms of physical volume of fixed capital investment the growth we witnessed does not exist anymore. The physical volume of investment seems to be rather stable during the last three years (the data officially published by Rosstat) if we neglect the seasonal component.

3. Empirical modelling of the costs structure

3.1 The model framework

Before we proceed to discuss modelling the costs structure the hypothesis we test should be formalized. The key assumption of our paper states that if higher investment leads to lower costs per ruble of output (relative prices dynamics eliminated) we may estimate the degree of technological progress in an industry.
The dynamics of costs per ruble of output is determined by the following components:

$$\frac{C_t}{Y_t} = A\frac{P_t^C}{P_t^Y},$$

where $\frac{C_t}{Y_t}$ is the share of costs in output in nominal terms, $\frac{P_t^C}{P_t^Y}$ is price index of costs relative to price index of output, and A stands for the real costs structure which is referred to as a technology component.

The initial approach applied for the data 1995–2004 was to verify the hypothesis by estimating autoregressive distributed lag models $ADL(p,q,r)$ at an industry level ($\forall m = \overline{1,M}$ industries):

$$\alpha_p(L)\,smc_t^m = \theta + \beta_q(L)\,inv_t^m + \gamma_r(L)\,rppi_t^m + \sum_{i=2}^{4}\delta_i d_{it} + \varepsilon_t^m, \text{ where}$$

- smc_t^m is the logarithm of costs of raw materials per ruble of output;
- inv_t^m is the logarithm of investment, deflated;
- $rppi_t^m$ is the logarithm of relative PPI index;
- d_{it}, $i = \overline{2,4}$ - seasonal dummy variables: 1 in quarter i, and 0 otherwise.

When the processes analyzed are integrated of order 1, $I(1)$, the ADL model allows for error correction model representation which is more informative from the point of view of economic interpretation in terms of short-run fluctuations and the long-run equilibrium. Therefore, first, formal tests for unit roots are needed to be applied; second, cointegration should be tested.

For the period 1995–2004 short-term fluctuations around long-term equilibrium described by cointergration were proved to be insignificant, except for chemical industry, i.e. investment influence the costs structure in the long-run only. Small sample does not allow adopting the same approach for the period 2005–2009. Previous results from Kantorovich & Nazrullaeva (2009) suggest that in this case the analysis may be limited to estimation of possible relationship between the costs structure and investments using cointegration approach. It should be mentioned, though, that cointegration phenomenon being approached in econometric literature as a long-term equilibrium requires longer time intervals. The applicability of cointeration concept may be doubted for the period which account for 4 years. Additional assumptions are required. We assume that the long-run equilibrium found 1995–2004 remains after 2005, as it takes time to develop and use innovative technologies which lower production costs, however the parameters of the equilibrium change.

Modelling empirical relationship between costs and investments controlling for relative prices requires the following steps.

1. Stationarity: unit root tests
   - Dickey – Fuller test (1979)
   - Zivot – Andrews test (1992) (”univariate”) for an endogenous structural break
2. Cointegration tests:
   - Two-stage Engle – Granger procedure (1987)
   - Gregory – Hansen test (1996) (”multivariate”) for an endogenous structural break in cointegrating equation
3. Statistical inference: asymptotic inference which, however, faces small sample problem (bootstrap inference was proved to be no better in this case).

3.2 Unit roots and structural change

The data first tested for a unit root using the augmented version of the Dickey – Fuller test (1979). The choice of lagged differences in the test is exogenous and aims at finding the trade-off between the power of the test and small sample problem. The results of the Dickey – Fuller test suggest that there is a unit root in $\left\{smc_t^m\right\}_{t=1}^T$, $\left\{inv_t^m\right\}_{t=1}^T$, and $\left\{rppi_t^m\right\}_{t=1}^T$ $\forall m = \overline{1, M}$, $M = 17$ economic activities.

However, as the period analyzed in the paper covers 1st quarter 2005 to 3rd quarter 2009 there may exist structural break in the series on $\left\{smc_t^m\right\}_{t=1}^T$, $\left\{inv_t^m\right\}_{t=1}^T$, and $\left\{rppi_t^m\right\}_{t=1}^T$ $\forall m = \overline{1, M}$ caused by exogenous shocks. In our case the crisis which hit Russian economy in the autumn of 2008 may be treated as such an exogenous shock. Therefore, formal procedure which allows testing for stationarity in case of a structural break is necessary: the Andrews – Zivot (Zivot – Andrews1992) test is applied. This test suggests three different types of structural breaks: level shift, shift in trend, and regime shift (both shift in level and in trend). Under the null hypothesis of stationarity with an endogenous structural break the Andrews – Zivot test has three specifications (see Table 2).

Table 2. The Andrews – Zivot (1992) test for an endogenous structural break: hypothesis specification

Alternative hypothesis $H_1: \alpha < 1$	$y_t = \mu + \alpha y_{t-1} + \beta t + \theta DU_t(\tau_B) + \gamma DT_t(\tau_B) + \varepsilon_t$	**Null hypotheses**
Level shift	(A) $y_t = \mu + \alpha y_{t-1} + \theta DU_t(\tau_B) + \varepsilon_t$	$H_0: \theta = \beta = 0$, $\alpha = 1$
Trend shift	(B) $y_t = \mu + \alpha y_{t-1} + \beta t + \gamma DT_t(\tau_B) + \varepsilon_t$	$H_0: \beta = \gamma = 0$, $\alpha = 1$
Regime shift	(C) $y_t = \mu + \alpha y_{t-1} + \beta t + \theta DU_t(\tau_B) + \gamma DT_t(\tau_B) + \varepsilon_t$	$H_0: \theta = \beta = \gamma = 0$, $\alpha = 1$

Source: own creation

Where $DU_t(\tau_B) = \begin{cases} 1, t > \tau_B; \\ 0, t \le \tau_B. \end{cases}$ and $DT_t(\tau_B) = \begin{cases} t - \tau_B, t > \tau_B; \\ 0, t \le \tau_B. \end{cases}$ are dummies for additive shift in level and shift in trend, τ_B – break moment.

The results of the Andrews – Zivot test for endogenous structural change with the specifications of the alternative hypothesis A, B, C are presented in Table 3 below.

Table 3. Endogenous structural break: results of the Andrews – Zivot test (1992), min t-statistics

Economic activity	Costs of materials per ruble of output (SMC) (break moment, year: quarter)			Relative PPI index (RPPI) (break moment, year: quarter)			Fixed capital investment, in 2004 prices (INV) (break moment, year: quarter)			R&D investment, in 2004 prices (INV) (break moment, year: quarter)			Foreign direct investment (break moment, year: quarter)		
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
1. C	-11.1*** 06:3	-7.53** 05:2	-10.1** 06:3	-3.50 08:2	-2.96 05:2	-5.11** 08:2	-4.29 06:1	-3.21 05:2	-4.56 08:3	-5.69** 08:3	-4.31* 05:2	-5.29** 08:3	-4.45 07:4	-3.72 05:2	-6.66** 07:2
2. CA	-11.3*** 06:3	-6.98** 05:2	-9.48** 06:3	-3.28 08:2	-2.95 05:2	-4.99* 08:2	-4.15 06:1	-3.10 05:2	-4.56 08:3	-6.00** 08:3	-5.12** 05:2	-5.30** 08:3	-4.44 07:4	-3.72 05:2	-6.58** 07:2

3. CB	-7.47*** 06:2	-6.03** 05:2	-6.99** 06:2	-2.91 05:4	-2.34 05:2	-3.81 06:2	-5.12 06:1	-3.94 05:2	-6.62** 05:3	-4.06 08:2	-2.63 05:2	-4.67 07:3	-5.80** 08:4	-4.25* 05:2	-6.28*** 08:1
4. D	-3.09 08:3	-2.41 05:2	-6.23** 07:3	-3.13 08:2	-2.95 05:2	-4.90* 08:2	-4.36 07:1	-4.03 05:2	-4.69 05:3	-6.59** 06:3	-5.77** 05:2	-6.58** 05:2	-4.80** 05:4	-4.53** 05:2	-5.67** 05:2
5. DA	-4.82** 08:1	-3.62 05:2	-8.50** 07:3	-5.55** 08:2	-2.48 05:2	-6.86** 08:2	-6.30** 08:3	-4.64** 05:2	-6.45** 08:3	-4.40 05:3	-3.56 05:2	-4.31 05:3	-8.31** 05:3	-6.26** 05:2	-8.85*** 06:4
6. DB	-4.58** 08:3	-1.48 05:2	-4.67 08:2	-2.60 06:4	-2.05 05:2	-4.84* 08:2	-6.04** 08:4	-4.22* 05:2	-6.31** 08:2				-6.42** 05:3	-3.91 05:2	-6.51*** 05:3
7. DC	-3.59 05:3	-3.26 05:2	-4.75 06:3	-3.06 06:4	-2.27 05:2	-4.75 08:2	-4.65* 07:3	-4.14 05:2	-5.01* 05:4						
8. DD	-3.99 08:4	-3.26 05:2	-4.39 08:3	-3.77 08:2	-3.47 05:2	-6.34** 08:2	-3.97 07:1	-3.49 05:2	-4.04 07:1						
9. DE	-3.91 08:3	-1.45 05:2	-3.18 07:3	-2.88 06:4	-2.51 05:2	-4.00 08:2	-4.14 06:1	-3.84 05:2	-3.92 08:1	-4.48 06:2	-3.30 04:2	-4.30 07:2	-4.64* 07:4	-4.28* 05:2	-11.5*** 07:4
10. DG	-2.96 07:3	-1.83 05:2	-3.27 06:4	-2.83 07:3	-2.32 05:2	-3.62 08:1	-4.54 08:3	-3.56 05:2	-4.97* 08:3	-5.16* 08:1	-5.06** 05:2	-5.19** 06:3			
11. DH	-4.20 08:3	-2.86 05:2	-3.31 07:3	-3.14 06:4	-2.23 05:2	-3.66 07:4	-3.72 05:4	-2.38 05:2	-5.07* 06:4	-2.11 06:2	-2.52 05:2	-4.70 07:3			
12. DI	-4.44 08:1	-3.40 05:2	-6.83** 07:3	-3.59 06:2	-1.83 05:2	-4.10 08:2	-4.52 07:1	-3.44 05:2	-4.35 08:1	-3.51 06:2	-2.94 05:2	-5.01* 07:4	-4.97** 08:3	-4.65** 05:2	-6.06*** 08:1
13. DJ	-3.78 08:3	-2.01 05:2	-3.43 07:3	-2.40 06:1	-2.08 05:2	-2.64 07:1	-3.98 07:1	-3.69 05:2	-4.59 08:1	-4.37 08:3	-3.22 05:2	-4.28 05:2	-5.10** 08:1	-4.33* 05:2	-5.86*** 08:1
14. DL	-7.46*** 08:3	-1.68 05:2	-6.36** 08:3	-3.13 06:4	-2.59 05:2	-3.46 07:4	-4.29 05:4	-3.80 05:2	-6.43** 08:2	-4.05 05:3	-3.74 05:2	-4.21 06:3			

15. DM	-2.73 08:3	-1.72 05:2	-2.93 06:3	-2.96 06:4	-2.47 05:2	-5.05** 08:2	-4.85** 07:1	-3.99 05:2	-5.18** 07:2	-5.56** 06:3	-4.78** 05:2	-5.34** 05:2	-4.94** 08:4	-3.83 05:2	-5.38** 08:3
16. DN	-6.00** 08:3	-2.68 05:2	-4.77 08:1	-4.35 08:2	-2.06 05:2	-3.88 08:2	-5.48** 08:3	-3.76 05:2	-4.99* 08:3	-4.66* 08:1	-4.36* 05:2	-5.91** 08:1	-4.65* 08:4	-4.31* 05:2	-5.66** 07:2
17. E	-2.75 05:2	-2.01 05:2	-3.41 05:2	-5.05* 08:2	-2.36 05:2	-4.34 07:4	-4.98** 07:1	-4.62** 05:2	-4.69 07:1	-4.05 05:3	-3.33 05:2	-4.17 08:1			

Notes

1. The null hypothesis of endogenous structural change is rejected ***, **, * at 1, 5 and 10 per cent significance levels respectively.
2. The Andrews – Zivot asymptotical critical values: 1 per cent significance level (A) –5.34, (B) –4.93, (C) –5.57; 5 per cent (A) –4.80, (B) –4.42, (C) –5.08; 10 per cent (A) –4.58, (B) –4.11, (C) –4.82.
3. Results obtained using Gauss 6.*0*.

Source: own creation

According to the results of the Andrews – Zivot test the moment of structural break varies across the economic activities analyzed and depends on the time interval for which the test is run. Nevertheless, we may infer from Table 3 that the formal test for an endogenous structural break for economic activities, especially in manufacturing, confirms that the break in costs and investments, as well as in relative PPI, occurred in the 3–4th quarters of 2008, i.e. when Russian industry finally faced the world economic crisis.

However, a univariate analysis for testing the hypothesis of relationship between costs and investments is not enough. As we see from the results of the Andews – Zivot test even if we assume that the series analyzed are stationary with structural breaks the moments of these breaks differ. So the empirical model we estimate should account for breaks. The approach which may solve this problem is using the results of the Dickey – Fuller test, i.e. assuming that the series are non-stationary (even if we know that structural breaks may lead to the biased results of the Dickey – Fuller test), and analyze cointegrating equations with a single structural break for the series. Instead of analyzing different moments of breaks we search for a single break for all the processes at once, i.e. the break which would describe the dynamics of costs, investments, and relative PPI in equilibrium.

3.3 Cointegration: costs of production and investments

Before discussing tests for cointegration it should be mentioned that the hypothesis we test in this paper implies the exogeneity of investment, i.e. technological progress assumes that the dynamics of investment determines the dynamics of the costs structure. The dynamics of investments (and relative PPI) is supposed to precede the costs dynamics.

In case of endogenous structural change standard tests for cointegration are replaced by the residual-based approach introduced in Gregory and Hansen (1996). The Gregory – Hansen test implies three possible specifications of cointegrating equations: level shift (C), level shift with trend (C/T) and regime shift (C/S). The specification of the Gregory – Hansen cointegrating equation is the following:

$$smc_t^m = \mu + \alpha inv_t^m + \phi rppi_t^m + \begin{bmatrix} \beta t + \theta DU_t(t_B) + \gamma DU_t(t_B) inv_t^m + \\ + \lambda DU_t(t_B) rppi_t^m \end{bmatrix} + \varepsilon_t, \quad (1)$$

where

- t is time trend;
- $DU_t(\tau_B) = \begin{cases} 1, t > \tau_B; \\ 0, t \le \tau_B. \end{cases}$ and $DT_t(\tau_B) = \begin{cases} t - \tau_B, t > \tau_B; \\ 0, t \le \tau_B. \end{cases}$ are dummies for

level shift (C) and shift in trend for (C/T) and (C/S), τ_B – break moment.

Table 4 below contains the estimates of cointegrating equations from the residual-based Engle – Granger and Gregory – Hansen two-step procedures. As it was already mentioned we analyze different types of investment processes which may influence the dynamics of the costs structure.

Table 4. The results of tests for cointegration (CI): Engle – Granger (1987), Gregory – Hansen (1996)

Economic activity	1st step: Parameter estimates (t-statistics in parethesis)					2nd step
	const	inv_t^m	$rppi_t^{im}$	t_B (year: quarter)	DU_t	CADF t-statistics (2nd step)
inv_t^m : **Fixed capital investment, deflated, in 2004 prices**						
3. CB	-6.32***	0.235***	-0.259	2006:3	-0.0883**	
	(-5.24)	(3.15)	(-1.27)		(-2.59)	-8.22***
11. DH	0.263	-0.0971*	-0.157		-0.113**	
	(0.386)	(-2.08)	(-0.854)	2009:1	(-2.33)	-5.41**
12. DI	0.0584	-0.101***	0.181**		-0.103***	
	(0.284)	(-7.75)	(2.30)	2008:4	(-8.19)	-5.07**
14. DL	-2.49***	0.0848**	0.0620		-0.149**	
	(-4.80)	(2.47)	(0.394)	2008:3	(-5.71)	-5.65**
16. DN	-3.51***	0.178**	-0.109		-0.164**	
	(-3.91)	(2.82)	(-0.483)	2009:1	(-2.50)	-5.12**
inv_t^m : **R&D investment, deflated, in 2004 prices**						
1. C	-1.36	-0.186**	-0.318		0.301***	
	(-1.68)	(-2.58)	(-1.13)	2006:4	(4.65)	-7.60***
2. CA	-1.54	-0.190*	-0.412		0.350***	
	(-1.52)	(-2.09)	(-1.44)	2006:4	(4.22)	-7.15***
11. DH	-0.545***	-0.0814***	-0.435**			
	(-3.87)	(-4.43)	(-2.52)	-	-	-3.99**
12. DI	-1.49	-0.00983*	0.171		-0.0799***	
	(-0.0799)	(-2.07)	(1.14)	2008:2	(-3.74)	-4.85*

Notes:

1. ***, **, * – OLS parameter estimates are significant at 1, 5, and 10 per cent level respectively. Asymptotical critical values for t-statistics with 15 degrees of freedom: (10 per cent) 1.75, (5 per cent) 2.13, (1 per cent) 2.95. Critical values for t-statistics with 16 degrees of freedom: (10 per cent) 1.75, (5 per cent) 2.12, (10 per cent) 2.92.

2. ***, **, * - the CADF test (2nd step of the Engle – Granger and Gregory – Hansen procedures) suggests there is no unit root in the residuals of cointegrating equation at 1, 5, and 10 per cent level respectively. The MacKinnon asymptotical critical values for the CADF test for T=19: (10 per cent) -3.28, (5 per cent) -3.68, (1 per cent) -4.54. The Gregory – Hansen asymptotical critical values: (10 per cent) -4.69, (5 per cent) -4.92, (1 per cent) -5.44.

4. Results obtained using Gauss 6.0.

Source: own creation

It can be seen from Table 4 that cointegration between the costs structure and fixed capital investment exists at the aggregated level in mining except of energy producing materials and for several manufacturing industries (manufacture of rubber and plastic products; non-metallic mineral products; electrical and optical equipment). For manufacture of non-metallic mineral products (DH) and manufacture of electrical and optical equipment (DI) we observe a statistically significant negative relationship between the costs structure and investment. Moreover, the additive structural change occurred in these industries at the end of 2008 and lowered costs of materials per ruble of output. As we already inferred from Figure 2 costs per ruble of output did decline at the end of 2008. Several explanations besides technological efficiency may be mentioned in this case. First of all, firms began to cut costs during the crisis. Secondly, relative producer prices also decreased meaning that prices of final goods grew faster than prices of costs. In order to avoid losses firms were forced to increase their prices of final goods as a response to higher prices of raw materials.

Table 4 shows that R&D investment has a significant influence on the dynamics of the costs structure in the same activities in mining and manufacturing. The elasticity of the costs structure to R&D investment (as our model is specified in logarithms) is negative when the influence of additive endogenous structural break is controlled for in cointegrating equations. Relative prices seem to have no influence on costs, except for manufacture of non-metallic products (DH). Therefore, despite the crisis there are technological improvements in production process of several economic activities, i.e. more R&D investments lead to lower unit production costs. However, positive results in terms of production and investment efficiency for only three economic activities do not appear to be very fruitful. The situation with investments in Russian industry after 2005 has undoubtedly changed.

It should also be mentioned that no significant results were obtained for foreign direct investments. So our hypothesis that this type of investment may have more transparent motivation than domestic investment fails.

4. Conclusion

In this paper the relationship between unit costs of materials and different types of investments was analyzed. Decreasing costs per ruble of output were revealed for several economic activities after the crisis of 2008. Comparing the dynamics of costs with the dynamics of investments was supposed to help explaining how the situation in Russian industry changed in the second half of the 2000-s. Previous research suggests that before 2005 investment growth led to technological progress and consequent output growth, i.e. production and investment efficiency was observed. However, after 2005 the situation in Russian industry began to change: we observe

the slowdown of investment growth and high costs of production per ruble of output. The crisis seems to redouble the problems Russian industry faces with almost no evident signs of recovery observed yet. Manufacturing appears to suffer most.

The analysis concerned modelling the costs structure at the disaggregated economic activity level for quarterly data 2005–2009 and its link to investments and relative prices (controlling for endogenous structural breaks). However, econometric modelling encounters serious limitations due to the small sample problem which is accounted for by the changes in official industrial statistics starting from 2005 with adoption of the OKVED industrial classification harmonized with NACE. The results suggest that if the effect of the crisis is eliminated there are a few economic activities which may be referred to as efficient: higher fixed capital investments and R&D investments lead to lower costs per ruble of output. However, for the majority of activities the hypothesis of investment and production efficiency is not verified. The effect of investment on costs appears to vary across the activities, and the overall efficiency of investment can be seriously questioned. The positive result, however, is that investment appears to have a partial effect on costs in mining sector: mining except of energy producing materials shows R&D investment efficiency. Our previous research suggests that during the transformation period in Russian economy and even at the beginning of the 2000-s no technological progress in mining existed. As it was mentioned mining in Russia accounts for more than a half of total investment, so the question whether these funds are spent efficiently is crucial for Russian industry.

The core result of this paper is that the relationship between investments and costs previously found in 1995–2004 appears to continue its existence meaning that there is a long-term equilibrium between costs and investments. Therefore, from a theoretical point of view investment is not simply a source of capital accumulation, it influences the level of technology in industry through production costs as endogenous growth theory predicts. From a practical point of view this result means that the worst fears of investments spent in the ways that fail to develop advanced technologies may be mitigated. Technological progress in 2005–2009 is still observed for several mining and manufacturing economic activities. Costs per ruble of output become lower with the rise in investment.

Future research in this field implies controlling the relation between investments, technological progress, and economic growth. A model framework that would allow capturing both the investment effect of lower costs as a technological component and the effect of higher output as a growth component should be developed and tested. Moreover, the analysis of technological progress and the level of innovativeness in industry requires working with the firm-level data together with the data at the economic activity level of aggregation.

References

Aghion, P. – Angeletos, G.-M. – Banerjee, A. – Manova, K. 2005: Volatility and Growth: Credit Constraints and Productivity Enhancing Investment. *NBER Working Paper Series*, 11349.

Anatolyev, S. 2003: Advanced Econometrics. NES Lecture Notes, New Economic School (NES).

Banerjee, A. – Lumsdaine, R.L. – Stock, J.H. 1992: Recursive and Sequential Test of the Unit Root and Trend Break Hypothesis: Theory and International Evidence. *Journal of Business and Economic Statistics*, 10, 3, 271–87.

Baranov, E. – Bessonov, V. 1999: Production intensity indices. *Promyshlennost v Rossii*, 3, 4–12. p.

Bessonov, V. A. – Voskoboynikov, I.B. 2006: Prakticheskij analiz. On Dynamics of Fixed Capital and Investments in the Russian Economy in Transition. *HSE Economic Journal*, 2, 193–228. p.

Cook, S. 2003: Spurious rejection by cointegration tests incorporating structural change in the cointegrating relationship. *Discussion paper series*, Swansea University, Department of Economics.

Dickey, D.A. – Fuller, W.A. 1979: Distribution of the Estimators for Autoregressive Time-Series with a Unit Root. *Journal of the American Statistical Assiciation*, 74, 366, 427–431. p.

Dickey, D.A. – Pantula, S.G. 1987: Determining the Order of Differencing in Autoregressive Processes. *Journal of Business and Economic Statistics*, 5, 4, 455–461. p.

DiNardo, J. – Johnston, J. 1997: Econometric methods. 4th ed., McGraw-Hill. New York.

Engle, R.F. – Granger, C.W.T. 1987: Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55, 2, 251–276. p.

Gregory, A.W. – Hansen, B.E. 1996: Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, 70, 1, 99–126. p.

Kantorovich, G. 2002: Times Series Analysis. Lecture Notes. *HSE Economic Journal*, 1-5.

Kantorovich, G. – Nazrullaeva, E. 2009: The Reduction of Costs per Unit of Output in Russian Industry: Direct Investment as a Reason. *HSE Economic Journal*, 13, 1, 59–79. p.

Maddala, G.S. – Kim, I. M. 1998: Cointegration, Unit Roots and Structural Change. Themes in Modern Econometrics. Cambridge University Press, Cambridge.

Leybourne, S. – Newbold, P. 2003: Spurious Rejections by Cointegration Tests Induced by Structural Breaks. *Applied Economics*, 35, 9, 1117–1121. p.

MacKinnon, J.G. 1991: Critical Values for Cointegration Tests. In Engle R.F. and Granger C.W.J. (eds): Long-run Economic Relationships: Readings in Cintegration, Oxford University Press, Oxford, 267-276. p.

Nadiri, I.M. – Mamuneas, T.P. 1994: The Effects of public infrastructure and R&D capital on the costs structure and performance of U.S. manufacturing industries. *The Review of Economics and Statistics*, 76, 1, 22–37. p.

Peretto, P.F. – Seater, J.J. 2007: Factor-Eliminating Technical Change. Economic Research Initiatives at Duke (ERID) *Working Paper*, 17.

Pollok, D.S.G. 1994: A Handbook of Time Series Analysis. Princeton University Press, Princeton.

Romer, D. 2006: Advanced Macroeconomics. 3rd ed. McGraw-Hill, Berkeley.

Van Dijk, F. 2007: Advance Time Series Econometrics. Lecture Notes, State University – Higher School of Economics, Helsinki.

Zivot, E. – Andrews, D.W.K. 1992: Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, 10, 3, 251–270. p.

Modelling of Factors Influencing Corporate Performance on the Bases of Domestic Empirical Studies

Erik Korom[1]

In this paper, I reviewed the development, from year 2001 to 2006, of the effects influencing the performance of businesses operating in the Hungarian economy. Studying the international special literature, I discovered a number of works focusing on this subject using a wide range of methodologies and tools. The novelty of my study is reinforced by both the subject group of businesses, and the sector, in which the same operate. Small and medium size enterprises account for a considerable economic weight in Hungary as well, just like in the European Union.

The analysis of the data shows correlation, where in practice, the parties concerned also assume and perceive – without sufficient theoretical background – that the intensifying competition makes the price-cost margins – and therefore profitability as well – decrease. Through the analyses performed on the database containing the data made available to me, and related to the subject period, I managed to support this unequivocally. My former studies already indicated the verifiability of the assumption related to the effect of labour productivity on profitability. There may be a positive relationship between labour productivity and profitability, meaning that an increase in labour productivity brings about increase in the profitability, while companies demonstrating a decreasing labour productivity also show decrease in their profitability.

Keywords: competition, performance, profitability, productivity

1. Introduction

It is a widely argued question whether competition and an intensified market concentration have a favourable effect on the businesses' profitability, productivity or efficiency. Many researchers found a positive relation between competition and the development of corporate performance. Both theoretical and empirical studies support, however, that on markets characterised with moderate technological progress and innovation, there are many players operating, and competition tends to work efficiently, with the margin converging to zero. In such cases, the price-cost margin can be an effective indicator of the intensity of competition. In this paper, I will ex-

[1]Erik Korom PhD, assistant professor, Department of Accountancy, College of Finance and Accountancy, Budapest Business School (Budapest)

amine, which factors have an effect on the development of corporate performance, based on a model built on the accounting figures of companies keeping double-entry books from 2001 to 2006.

2. Competition and its measurability

Researchers dealing with economics have been interested in the role of competition and its influence on economic efficiency for a long time. Most of them emphasize the 'beneficial' role of competition, although competition is not omnipotent, which is proved by the financial world crisis at the end of 2008, for example. Competition can be linked to certain markets (e.g. the market of services).

A market can be described on the basis of its structure, form, size and the number of players. These factors determine the possibilities of the players and the strength of the competition. The forms of the market can be judged by *the number* of sellers and buyers, whether it is easy or difficult for *new players* to enter the market and whether a player is able to influence the market price alone. The two extreme forms of market competition are perfect competition and pure monopoly. In one of the most well-known transitional market situations (Stackelberg describes nine possible market situations in his book. (Leontief 1936)), some companies knowing one another and being able to adapt to one another compete on the given market. This is called a multi-participant *oligopolistic* market.

A basic issue of the literature dealing with the relationship between competition and the performance of companies is defining the intensity of competition numerically. Competition is hard to measure directly. For the lack of exact index numbers, empirical studies attempt to demonstrate the intensity of competition based on an observable factor relating to competition indirectly. On the literature studied, I can declare that there are no universal index numbers that could give a really reliable picture of the extent of competition on a given market.

2.1. Index numbers of competition intensity

Of the indicators used for measuring competition, I can distinguish between *static* and *dynamic* index numbers. Static indexes provide information on the state of a given market at the time of the investigation, whereas dynamic indexes are applied to show the *changes* taking place on the market. The indicators based on the demography (establishment and closing down) and the market share changes of companies can provide further useful information on the intensity of competition.

In addition to and completing static and dynamic index numbers, several other approaches to assessing the extent and trends of competition can be found while studying the literature. Among others are index numbers demonstrating the existence of certain institutional conditions, the *regulation* level of the market or *freedom* of international trade and *innovation*.

I focus on the following indicators used for measuring market structure: a) concentration rate (CR), b) Hirschman–Herfindahl index (HHI) and c) Lerner index (L) or price–cost margin (PCM).

2.1.1. Concentration rate

The *concentration rate* shows what percentage the largest companies have of the total production or total turnover of a market. The concentration rate can be calculated on the 2, 4, 8, 16, etc. largest companies relating to the division and class concerned. The CR4 and CR8 indexes of the logistics division are in columns 3 and 4 of Table 1.

2.1.2. The Hirschman-Herfindahl index

The *Hirschman-Herfindahl index* is the index number of market concentration in the industrial sector, which is calculated by squaring and summing the market share of each firm. The maximal value of the index is 10 000 (100^2).[2] The HHI index is shown in the last column of Table 1.

[2] In compliance with the Merger Guide, we speak about moderate concentration if the index is between 1000 and 1800, and concentration if it is over 1800.
http://www.usdoj.gov/atr/public/guidelines/horiz_book/15.html

Table 1. The ten most highly concentrated classes in the service sector in Hungary in 2006

Industries	Company	CR4	CR8	MCR8	HHI
Air transport	67	93.51	97.05	3.54	5,785
Post and telecommunications	1,026	69.43	80.69	11.26	1,603
Water transport	79	64.14	77.52	13.38	1,182
Renting of machinery & equipment without operator	1,673	40.32	48.89	8.57	776
Recreational, cultural and sporting activities	8,269	41.45	44.25	2.80	752
Sewage and refuse disposal, sanitation and similar activities	749	32.76	40.62	7.86	422
Land transport; transport via pipelines	6,902	30.66	38.57	7.91	328
Sale, maintenance & repair of motor vehicles, etc.; retail sale of automotive fuel	10,994	21.45	32.24	10.79	169
Retail trade, except of motor veh. etc.; repair of personal & hh. goods	35,821	21.38	28.36	6.98	166
Other service activities	3,318	17.91	26.65	8.74	132

Source: own creation

There is a strong correlation between the concentration rates calculated with Hungarian corporate data on class level. The strongest relationship ($r = 0.981$) is between CR4 and CR8, but there is a close connection between CR4 and the HHI index ($r = 0.849$), as well as CR8 and the HHI index ($r = 0.777$). Correlations are significant at level $p<0.001$.

2.1.3. Lerner index

The *Lerner index* is an indicator determining the market power of the company as the difference between price and marginal cost. (Lerner 1934) The higher the Lerner index, the more market power the company has, as the company needs more and more power to be able to deviate from the price equal to the marginal cost.

2.1.4. International empirical studies on the above indicators

In empirical studies, both the concentration rate and the Lerner index are frequently applied (to measure competition). See for example: (Collins – Preston 1966, 1969; Saving 1970; Ornstein 1975; Rosenthal 1980; Domowitz et al. 1986, 1988; Amir – Lambson 2000; Nevo 2001), etc.

However, not all the publications we have studied led to an unambiguous result. Ornstein's investigations (1975), for instance, justified Collins and Preston's (1966, 1969) conclusions, stating that, in the industrial sector manufacturing of consumer goods,[3] the Lerner index shows statistically significant results with market concentration trends, but in sectors producing durable consumer goods, the results did not prove to be significant. Similarly, no significant relations could be detected in more capital-intensive divisions.

Domowitz (1986) and his co-authors came to the conclusion that, in the 1970s, there was a dramatic narrowing of the spread between the average price–cost margins of concentrated and non-concentrated divisions in the USA. The change in the spread between the margins of concentrated and non-concentrated sectors can be largely attributed to the different changes of adaptation and demand, which suggests that, in the explanation of price–cost margin fluctuation, the aggregate impact of demand is of greater importance than the effects of local demand (Domowitz et al. 1986).

Nevertheless, both theoretical and empirical research supports the hypothesis that, on markets characterized by moderate technological development and innovation, with lots of players, where competition functions efficiently, the margin approaches zero. In such cases, the price–cost margin can indicate the strength of competition well. A value close to zero indicates strong competition, whereas higher values point to the weakness of competition. The higher the index value, the weaker the competition. (I proved this thesis through utilizing the – independent – international research results of the 1990s.)

The vast majority of the empirical studies found on the effect of market share on profitability published in the last decades confirmed a positive correlation. The increase of productivity showed different extents taken as a function of the features of the samples, the methodology and the model applied, but according to the conclusions drawn from the studies, the increase in market share is equal to the rise in profitability (Szymanski et al. 1993).[4]

- Shepherd (1972) also found a positive and significant link between market share and profitability, but the relationship among companies in the service

[3] For more about 'nondurable goods', see for example: http://www.census.gov/epcd/naics02/def/NDEF424.HTM

[4] The authors examined 28 other studies on the relationship between market share and profitability, but they did not include the extent of the effect.

sector appeared to be stronger than among those in the production sphere. He pointed out the variability of the connection between market share and profitability, dividing the period under survey into parts. More profitable companies lost some of their market share later (Shepherd 1972).

- Nickell et al. (1992) believe that productivity growth is usually stronger in the case of companies with a large market share (Nickell et al. 1992). Nickell later (1996) adds that this growth effect is a cross-sectional result; therefore, it is not free from the problem that, in the long run, companies producing a relatively high rise in productivity usually grow faster and gain a larger market share. Thus, it cannot be regarded as an original competition effect (Nickell 1996).
- Hay and Liu (1997) say that the shifts of market shares are connected with the levels of efficiency, the price flexibility of demand and the company management, as well as with the number of companies on the market. If the competitor uses similar prices, the equation applied by the authors suggests that the market share of a company is in direct proportion to its relative costs; furthermore, the share of the players on the market decreases if the number of companies goes up. The shifts of market shares, or in other words, the reallocation effect, is less marked if the number of companies was originally high. It can also be concluded from the equation that, if the price flexibility increases or competitors' reaction regresses, efficient companies gain market share at the expense of the less efficient ones (Hay – Liu 1997).
- Halpern and Kőrösi (2001) examined the relationship between efficiency and market share in the Hungarian corporate sector over the period 1990–1998. In their study, in order to measure concentration, they used the reciprocal of the number of companies operating in the classes (Halpern - Kőrösi 2001). In the period I examined, the concentration rate derived from the number of companies can mainly be connected with the HHI index. If I make the calculations concerning groups and classes (less aggregate data), the relationships between the indicators slightly weaken.

2.2. The basic model of my calculations

Besides the concentration rate, the price–cost margin is one of the most frequently used index numbers for measuring the intensity of competition. In the model of economic theory describing a perfectly competing market, none of the players depart from the market price, which will be equal to the marginal cost in the long run. In a monopoly, it is only the demand that can limit the trend of prices; therefore, the less elastic the demand for a product, the bigger the market power and the price–cost margin of the company in a monopolistic position.

Collins and Preston (1966) explain the differences in the price–cost margins (PCM) of industries with the following model:

$$\mathbf{Y_{tu} = \alpha + \beta_1\, CR4_{tu} + \beta_2\, CR4_{tu}^2 + \beta_3\, IGEOD_{tu} + \beta_4\, K\_S\_R_{t0} + \beta_5\, G_{tu/t0}}$$

Y_{tu} = price–cost margin (PCM) in the last year of the examined period.
$CR4_{tu}$ = concentration rate calculated on the four largest companies of the class in the last year of the examined period.
$IGEOD_{tu}$ = index of geographical dispersion.
$K\_S\_R_{t0}$ = value of total assets output in the first year of the examined period.
$G_{tu/t0}$ = percentage change in the output of classes over the examined period.

The price–cost margin (PCM) was defined in two different ways. When calculating the numerator of PCM1, staff costs and material costs are deducted from the periodic output corrected with the variation in stocks; in the case of PCM2, besides material costs, material-type expenditures are also deducted. In both cases, the denominator is the periodic output corrected with the variation in stocks.

The index of geographical dispersion was calculated concerning the last year of the examined period as the sectoral sum of the absolute differences between the aggregated percentage of industry output on a regional level and the percentage of population in that region. The higher the geographical dispersion, the smaller the index.

In the value of total assets output ($K\_S\_R_{t0}$), the difference between the asset intensiveness of class output is taken into consideration. The quantity of assets (particularly the ratio of current assets and tangible assets, and intangible assets) can influence the trend of the price–cost margin differently.

3. The database examined

I worked with the data of 2001–2006, aggregated on class level, of companies having double-entry book-keeping. A total of *95,214* companies with double-entry book-keeping submitted legitimate corporate tax returns in each year of the subject period. This group of companies accounted for 30.08% of all companies having double-entry book-keeping in 2006, and employed 69.74% of all workers employed by all companies having double-entry book-keeping. In 2006, the books of these 95,214 companies accounted for 81.64% of the net sales revenues (which was 90.67% of the net export sales revenues!), and 84.30% of the pre-tax profit of all companies submitting tax returns with double-entry book-keeping.

I measured corporate performance using the average values, calculated in the class, of profitability indexes widely applied in foreign literature (ROE, ROA, ROS)[5] and the gross value added per capita6 index.

[5] **ROE** = EBIT / [(Equity $_t$ + Equity $_{t-1}$) / 2], **EBIT** = Company's usual profit + Interests to be paid and other interest-type expenditures – Other interests received and interest-type revenues; **ROA** = EBIT / [(Total assets $_t$ + Total assets $_{t-1}$) / 2]; **ROS** = EBIT / Net sales revenues.

In the empirical studies performed during my research, I surveyed the performance of companies in the service sector, on the bases of the figures included in the tax returns submitted by companies with double-entry book-keeping assigned to industries 50 to 93 (except Financial mediation) under the TEÁOR'03 system, as such data were collected by HCSO (Hungarian Central Statistical Office)[7]. Based on the data of 2006, 76.88% of all companies operated in these industries in Hungary producing 51.91% of all net sales revenues, employing 53.23% of all labour force. Additionally, this sector generated 47.29% of the total gross added value. *In this respect*, the service sector forms the largest sector of the Hungarian national economy in our days.

4. Empirical results of the investigation

In the course of the literature review, I explored the ways of demonstrating corporate competition. Based on the databases available, using the SPSS 15.0 for Windows program package, I carried out descriptive statistical calculations, comparing analyses of divisions and classes and comparing ratio analyses, as well as regression and variance analysis (ANOVA), in accordance with the mathematical methods used in the literature. Following the development of the basic model to be described, in addition to the compliance conditions of the regression analysis, the validity of the model in use was checked, too.

Similarly to the sample of Collins and Preston (1966), the values of the parameters received indicate a negative relation among concentration, the value of total assets output and the price–cost margin; moreover, the release growth of the classes shows a significant relation with the price–cost margin.

The parameters received on the basis of the model run on the 2001–2006 class data of Hungarian companies, if $R^2 = 22.8\%$,[8] are as follows (see Table 2).

[6] With the balance of taxes on products and subsidies on products, the **gross value added** index shows the company's contribution to GDP.
[7] The HCSO database does not include data for financial and off-shore enterprises.
[8] The unbiased estimator of the explained proportion of the statistical population.

Table 2. Testing the initial model with Hungarian data

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,536	,018		29,700	,000
	CR4t6	-,005	,001	-,659	-7,321	,000
	NNCR4t6	7,78E-006	,000	,117	1,347	,178
	IGEODtu	,003	,000	,318	15,100	,000
	K_S_Rt0	,010	,005	,038	2,183	,029
	Gtut0	2,38E-007	,000	,015	,863	,388

a. Dependent Variable: PCM106

Source: own creation

In this form, the model with explanatory power of about 20% does not meet the requirements of regression analysis perfectly. The concentration rate ($CR4_{t6}$) and its square indicate a rather strong correlation (r = 0.978) and, looking at Table 2, I find that the percentage change in the output of classes over the examined period ($G_{tu/t0)}$ are not significant.

During the further development of the model, considering the models of Collins and Preston (1969), Ornstein (1975) and Domowitz et al. (1986), by omitting the square of the concentration rate ($NNCR4_{t6}$) from the basic model and replacing the dependent variable with the price–cost margins of the specific business years, I obtained the following regression equation:

$$\mathbf{Y_{ti} = \alpha + \beta_1 CR4_{t0} + \beta_2 IGEOD_{tu} + \beta_3 K\_S\_R_{t0} + \beta_4 G_{t/t-1} + \beta_4 VTK\_M1_{ti} + \beta_6 Sz}$$

Y_{ti} = price–cost margin (PCM) in class *i* in year *t*.
$CR4_{t0}$ = concentration rate calculated on the four largest companies of the class in the last year of the examined period.
$IGEOD_{tu}$ = index of geographical dispersion.
$K\_S\_R_{t0}$ = the value of total assets output in the first year of the examined period.
$G_{t/t-1}$ = percentage change in output of classes over the examined period.
VTK_M1ti = labor productivity in class *i* in year *t*.
Sz = a dummy variable with a value of one when services sector and of zero otherwise.

The explanatory power of the model slightly fell (R^2 = 58.57%), but its reliability improved remarkably. The standard error of the estimation is rather low,

which indicates that the model can estimate quite well. The significance of the F-test also verifies the existence of the relation, and in the t-test, on the basis of the significance of the variables determining the slope, the explanatory variables really influence the value of the dependent variable (see Table 3).

Table 3. Summarizing table of the model supplemented with the productivity index

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,767[a]	,588	,587	,15248	,588	596,792	6	2508	,000

a. Predictors: (Constant), Szolgáltató, IGEODtu, Gtt1, K_S_Rt0, VTK_M1, CR4to

Source: own creation

According to the beta values, the price–cost margin is the most strongly influenced by geographical dispersion ($IGEOD_{tu}$) and the percentage change in class output ($G_{t/t-1}$). Similarly to the model of Collins and Preston (1966), concentration ($CR4_{t0}$) shows a negative relation to price–cost margin and to the value of total assets output ($K\_S\_R_{t0}$) as well.

In the course of regression analysis, the compliance of the model can only be declared after checking the fulfilment of numerous conditions: a) there are no correlation coefficients of high value among the explanatory variables and the 'tolerance' indexes are also favourable, so there is no multicollinearity either; b) residual variance seems constant, too, and the residuals have a normal distribution according to the histogram, which is confirmed by the Kolmogorov–Smirnov (K–S) one-sample test.[9]

Beyond graphical representation, I also verified the stability of the variance of remainder terms through calculations. The standard deviation of the residuals of ten percentage samples taken after sorting the independent variables in ascending order were tested using F-probe. As a result of the calculations performed, in can be concluded, with the level of significance being either five percent or *one percent*, that **t**he *condition of homoscedasticity* is also met.

Beyond the published verifications of the calculations carried out for individual industries, I also verified the operation of the model several times through dividing the full sample into two parts comprising the same number of elements. By applying the regression model to the first part of the sample, I provided estimations for the values of the dependent variable in the second part of the sample using theses

[9] The calculations were made by the SPSS 15.0 program package.

regression factors. The tests performed during the comparison of the standard deviations of residuals indicated no significant differences among the variance values, whereby the *validity* of the model is unquestionable.

The explanatory power of the model can be further increased if the labour productivity of classes is also involved in the independent variables (R^2 = 59.4%). The effect of the headcount (L_A) on the price/cost margin (PCM1) is clearly negative, while such headcount also slightly decreases the effect of labour productivity on the dependent variable (see Table 4).

Table 4. Parameters of the developed model

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,468	,013		35,665	,000
	CR4to	-,002	,000	-,200	-12,912	,000
	IGEODtu	,000	,000	,052	3,654	,000
	K_S_Rt0	-,018	,004	-,065	-4,772	,000
	Gtt1	,000	,000	,045	3,510	,000
	VTK_M1	3,86E-006	,000	,148	10,741	,000
	Szolgáltató	,326	,007	,665	47,706	,000
	L_A	-2,51E-006	,000	-,088	-6,568	,000

a. Dependent Variable: PCM1

Source: own creation

Proceeding with the examination of the effects of the individual factors in the corporate database related to the service sector, one may conclude that, in classes where the intensity of competition rose in the examined period (PCM fell), the profitability indexes and labour productivity are *significantly* lower than in classes where the intensity of competition lessened.[10]

Dividing the examined period into two parts, years 2001–2003 and 2004–2006, in classes showing *increasing competition*, the average profitability (ROA, ROE) is significantly lower than in less competing ones in the first half of the examined period. In classes with higher concentration (over 70%), *bigger values* of profitability can be detected than in less concentrated ones.

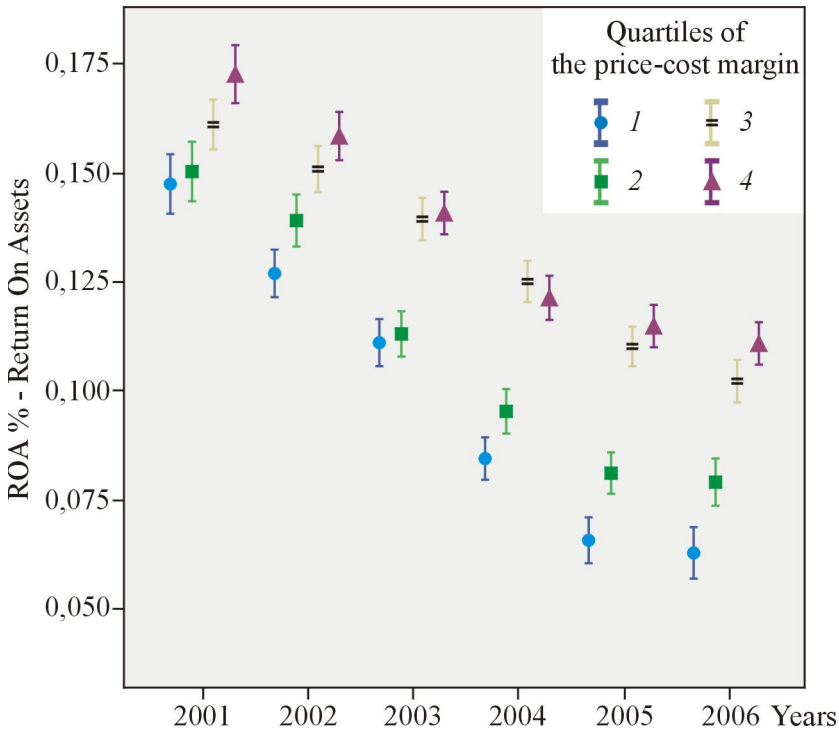
In the second half of the examined period (2004–2006), however, in classes with lower concentration and increasing competition, the average profitability (ROA, ROE, ROS) is higher, whereas in classes with higher concentration (over

[10] For lack of normality of the analysed variables, besides variance analysis and a t-test, I made the calculations with non-parametric tests, too.

70%) and decreasing competition, these values are higher. Consequently, in the whole period examined, in *highly concentrated* classes, the relation between price–cost margin and profitability indicators is *much stronger*.

There is a positive, significant relationship between labour productivity, profitability indicators and the price–cost margin of company data, too, within the examined period. Dividing the price–cost margin into quartiles, based on ROA and ROE indicators, the profitability of companies in the lower quartile is smaller than in the upper quartile (see Figure 1).

Figure 1. Profitability indicators in the quartiles of price–cost margin



Source: own creation

While the profitability indicators (ROA, ROE) decline in the quartiles of the price–cost margin, labour productivity rises. In higher price–cost margin (PCM) quartiles, the average value of labour productivity is higher. A *significant difference* can be seen between the values of individual quartiles. In the examined period, indicators increase in every quartile year by year, which, with some exceptions, indicate a significant difference even between individual years and quartiles.

Looking at the tendency of labour productivity and profitability, I can state that the profitability of companies able to enhance labour productivity in each year of the examined period also increased, whereas the profitability of companies with decreasing labour productivity decreased. A *significant difference* can be observed between return on sales and return on equity (ROS, ROE) as well.

I can point out increasing labour productivity among companies showing rising profitability (ROA, ROE, ROS); the labour productivity of companies showing decreasing profitability, however, decreases in a strictly monotone manner. Between the two categories of companies, a *significant difference* can be noticed in the trend of labour productivity.

A survey of the trends of corporate market shares clearly indicates that, sooner or later, the price of growth *has to compensated* by the decrease in profitability. While those loosing markets also have continuously decreasing profitability, such profitability started to decrease for companies gaining market share as well, but during the first half of the subject period only. While, for years 2001 to 2003, there is a significant difference between those showing decreasing and increasing trends, the difference is not *substantial*, as regards return on sales (ROS) and return on assets (ROA), for years 2004 to 2006.

Companies capable of increasing their corporate market share could also boost their labour productivity, while the indicators turned downwards for those loosing market share. While this decrease was not significant for the "losers" between 2001 and 2003, the tests showed *significant* difference for the winners. The results of the calculations also indicate that the average labour productivity indicator stands higher for companies gaining market share, than for those losing market share. At the same time, companies belonging to the group losing market share had higher labour productivity, on an average, at the beginning of the subject period, and for years 2004 to 2006, there is no significant difference between the initial years.

5. Conclusions

Companies increasing their market share were able to raise their labour productivity, whereas the indicators of 'market losers' began to fall. While this drop was not significant for the 'losers' in the years of 2001–2003, the tests showed a significant difference in the case of market gainers. I can also conclude from the results of the calculations that the average labour productivity of companies increasing their market share exceed the labour productivity of market share losers by the end of both the first and the second halves of the examined period. This rather refers to the fact that an *increase in market share leads* to the enhancement of productivity. (Market share, however, can be increased much less by enhancing productivity!)

This statement is also reinforced by the results obtained from the investigation into labour productivity and corporate market share. The market share of companies showing decreasing labour productivity in the examined period decreased, too, but companies able to enhance their productivity could not increase their market share significantly, either. The results of the separate analysis of the first and second parts of the examined period *provide further support to the above.*

The calculations in connection with profitability have similar results. The market share of companies with decreasing profitability fell significantly, while not even a slight increase in the market share of companies with growing profitability can be detected in the case of every indicator.

Looking at the above results, I can imagine that these will be somewhat surprising for company strategy makers. Nevertheless, these are only Hungarian data. I am aware of the need for international comparison, which I intend to execute in the following period.

References

Amir, R. – Lambson, V. E. 2000: On the Effects of Entry in Cournot Markets, *The Review of Economic Studies*, Vol. 67, No. 2. (Apr., 2000), pp. 235–254.

Collins, N. R. – Preston, L. E. 1966: Concentration and Price-Cost Margins in Food Manufacturing Industries, *The Journal of Industrial Economics*, Vol. 14, No. 3. (Jun., 1966), pp. 226–242.

Collins, N. R. – Preston, L. E. 1969: Price-Cost Margins and Industry Structure, *The Review of Economics and Statistics*, Vol. 51, No. 3. (Aug., 1969), pp. 271–286.

Domowitz, I. – Hubbard, R. G. – Petersen, B. C. 1986: Business Cycles and the Relationship between Concentration and Price-Cost Margins, *The RAND Journal of Economics*, Vol. 17, No. 1. (Spring, 1986), pp. 1–17.

Domowitz, I. – Hubbard, R. G. – Petersen, B. C. 1988: Market Structure and Cyclical Fluctuations in U.S. Manufacturing, *The Review of Economics and Statistics*, Vol. 70, No. 1. (Feb., 1988), pp. 55–66.

Halpern, L. – Kőrösi, G. 2001: Efficiency and Market Share in Hungarian Corporate Sector, *Economics of Transition*, Vol. 9, No. 3. (2001), pp. 559–592.

Hay, D. A. – Liu, G. S. 1997: The Efficiency of Firms: What Difference Does Competition Make?, *The Economic Journal*, Vol. 107, No. 442. (May, 1997), pp. 597–617.

Leontief, W. 1936: Stackelberg on Monopolistic Competition, *The Journal of Political Economy*, Vol. 44, No. 4. (Aug., 1936), pp. 554–559.

Lerner, A. P. 1934: The Concept of Monopoly and the Measurement of Monopoly Power, *The Review of Economic Studies*, Vol. 1, No. 3. (Jun., 1934), pp. 157–175.

Nevo, A. 2001: Measuring Market Power in the Ready-to-Eat Cereal Industry, *Econometrica*, Vol. 69, No. 2. (Mar., 2001), pp. 307–342.

Nickell, S. J. – Wadhwani, S. B. – Wall, M. 1992: Productivity Growth in U.K. Companies, 1975-1986, *European Economic Review*, Vol. 36. (June, 1992), pp. 1055–1085.

Nickell, S. J. 1996: Competition and Corporate Performance, *The Journal of Political Economy*, Vol. 104, No. 4. (Aug., 1996), pp. 724–746.

Ornstein, S. I. 1975: Empirical Uses of the Price-Cost Margin, *The Journal of Industrial Economics*, Vol. 24, No. 2. (Dec., 1975), pp. 105–117.

Rosenthal, R. W. 1980: A Model in which an Increase in the Number of Sellers Leads to a Higher Price, *Econometrica*, Vol. 48, No. 6. (Sep., 1980), pp. 1575–1579.

Saving, T. R. 1970: Concentration Ratios and the Degree of Monopoly, *International Economic Review*, Vol. 11, No. 1. (Feb., 1970), pp. 139–146.

Shepherd, W. G. 1972: The Elements of Market Structure, *The Review of Economics and Statistics*, Vol. 54, No. 1. (Feb., 1972), pp. 25–37.

Szymanski, D. M. – Bharadwaj, S. G. – Varadarajan, P. R. 1993: An Analysis of the Market Share-Profitability Relationship, *Journal of Marketing*, Vol. 57. (July, 1993), pp. 1–18.

Modeling some Entrepreneurship Factors

Annamaria Bianchi[1] *– Silvia Biffignandi*[2]

Entrepreneurship is increasingly recognized as a major factor of economic growth, productivity and competitive economy. Many countries are making efforts to support entrepreneurship and are interested in knowing how government policies and other factors can influence the amount and type of entrepreneurship. For this purpose they need to understand the determinants of and obstacles to entrepreneurship.

In spite of the large interest in entrepreneurship, due to the lack of internationally comparable data, the understanding of this phenomenon and its determinants remains still an open problem.

In 2006 OECD launched the Entrepreneurship Indicators Programme (EIP), which was joined by Eurostat in 2007. In Measuring Entrepreneurship: A Digest of Indicators (2008) a common set of concepts and definitions is presented. Furthermore, consistent data across different countries are published: even if these data do not represent the whole set of indicators which is needed for studying the entrepreneurship process, they represent a preliminary database of internationally comparable statistics.

Using this database (Structural and Demographic Business Statistics (SDBS)) and others (R&D database, Market Regulation database and Education at a Glance) we perform initial analysis of entrepreneurship across countries. Our interest is in understanding its determinants and in particular those related to education. Preliminary conclusions about the role of different educational level on entrepreneurship are obtained as a reference theoretical frame for more detailed analyses based on single country data.

Keywords: Entrepreneurship, education, regression analysis, multidimensional scaling

[1] A. Bianchi, Reseach Assistant, University of Bergamo, Faculty of Economics and Business Administration (Italy)
[2] S. Biffignandi, Full Professor, University of Bergamo, Faculty of Economics and Business Administration (Italy)

The paper has been developed within the Bergamo University ex 60% financial grant 2007 and 2008, responsible Silvia Biffignandi.

1. Introduction

The recognition that entrepreneurship and entrepreneurs are important drivers of economic growth, employment, innovation and productivity is generally accepted. However, the development of policies related to entrepreneurship has been limited for long time since the statistical base for entrepreneurship research was weak, especially in terms of international comparability. The increasing interest in how government policies may affect the amount and types of entrepreneurship has drawn more and more attention on this issue.

In order to address this information-gap, in 2006 OECD launched the Entrepreneurship Indicator Programme (EIP), which was joined by Eurostat in 2007. The purpose of this programme was to build internationally-comparable statistics on entrepreneurship[3].

The first challenge for the EIP was to give a definition of entrepreneurship in such a way to enable valid indicators to be introduced and collected across countries. Indeed, there was no widely-accepted definition of the word 'entrepreneurship': on the one hand, many definitions had an essentially theoretical basis with little concern for measurement; on the other hand, many papers bypassed the discussion of the definition of entrepreneurship and simply equate entrepreneurship to a specific empirical measure (readily available).

---

[3] The advantage of this project is to try to set up some harmonized definitions. Indeed, many sources and approaches on entrepreneurship are available, each one using its own concepts and definitions and/or with reference to selected countries.

For instance, the International Consortium on Entrepreneurship (ICE) recently released a report ranking selected countries on four performance indicators and 66 business environment indicators. Furthermore, this report provides a comprehensive overview of all available policy-relevant indicators relating to entrepreneurship, along with a quality assessment in order to enable policy-makers to evaluate the quality of policy analysis based on the given indicators.

The Kauffman Foundation releases an Index of Entrepreneurial Activity which measures the rate of business creation at the individual owner level. Presenting the percentage of the adult, non-business owner population that starts a business each month, the Kauffman Index captures all business owners, including those who own incorporated or unincorporated businesses, and those who are employers or non- employers. This index is published for years 1996 to 2008 for each state in the US (Fairlie 2009).

The *World Bank Group Entrepreneurship Survey* measures entrepreneurial activity around the world. The database includes cross-country, time-series data on the number of total and newly registered businesses for 84 countries (Klapper et al. 2008).

The OECD-Eurostat approach (Ahmad-Seymour 2008) has tried to combine both approaches, that is to give a theoretical definition paying attention to the measurement issue. In this context, the following definitions were established:

- *Entrepreneurs* are those persons (business owners) who seek to generate value through the creation or expansion of economic activity, by identifying and exploiting new products, processes or markets.
- *Entrepreneurial activity* is enterprising human action in pursuit of the generation of value through the creation or expansion of economic activity, by identifying or exploiting new products, processes or markets.
- *Entrepreneurship* is the phenomenon associated with entrepreneurial activity.

From its definition it is clear that entrepreneurship is a multi-faceted phenomenon both in its manifestations and in its impacts and that it cannot be measured by a single indicator but rather by a set of indicators. In this direction a first model was proposed in Ahmad and Hoffman (2008). This model identifies three separated but inter-connected stages, all of which are important in the formulation and assessment of policy measures: the first stage identifies various '*determinants*', which policy can affect and which in turn influence '*entrepreneurial performance'*; '*entrepreneurial performance*' reflects the target indicators that policy makers believe have an impact on some ultimate objectives ('*impacts*'), such as economic growth, job creation or poverty reduction.

Within this framework, the EIP has proposed a range of indicators that allow to understand and distinguish different types of entrepreneurial performance. The list of adopted indicators reflects relevance and measurability and is introduced in Section 2.1. As far as determinants are concerned, the EIP has not created specific databases, but other sources are available.

The present work falls within the framework presented above. We try to investigate the relationship between determinants and entrepreneurial performance. In particular, we focus attention on factors related to education and how they influence different types of entrepreneurship.

The rest of the paper is organized as follows. In Section 2 we introduce the set of indicators that we consider. Section 3 is devoted to the statistical analysis and Section 4 contains some preliminary conclusions.

2. Indicators

In this section we present the indicators that we use in the rest of the paper. In Section 2.1 we introduce indicators measuring the entrepreneurial performance of coun-

tries. In Sections 2.2 to 2.4 we present some indicators of the determinants, paying particular attention to those related to education.

2.1 Entrepreneurial Indicators

From the definition given above, it is clear that entrepreneurship is a phenomenon difficult to measure and that, given the diversity of its manifestations, no single indicator can ever measure it adequately. It is also clear that entrepreneurship has to do with something 'new': the creation of new markets and new value through new products and processes.

Before introducing the indicators that have been proposed by OECD-Eurostat, we highlight some aspects that help understand how they should be interpreted.

First of all, all the indicators are to be considered as proxies for the definition of entrepreneurship, each one describing a specific aspect of this phenomenon. They have to be interpreted as measures of entrepreneurship that have loose or strict interpretation of the word 'new'. For example, if one takes a liberal interpretation of 'new', all new businesses (enterprise birth) could be considered as creating new markets. On the other hand, it is clear that not all businesses are necessarily entrepreneurial.

Taking a stricter interpretation of the word 'new', one can consider that those firms that have had rapid growth (high-growth enterprise) are more likely to have demonstrated 'pure' entrepreneurship. In this case it is assumed that there was something significantly different about their product or process or market that led to significant growth. It is also assumed that firms do not need to be new to be entrepreneurial.

One could tighten further the definition of 'new' and require that high-growth firms also need to be young (gazelles).

In the light of the previous observations, we consider three indicators of entrepreneurship: those related to enterprise birth, high growth enterprises and gazelles. In the following we give exact definitions of these indicators. Other manifestations of performance are also considered in the EIP, such as firms death and survival rates, but we are not going to take them into consideration. The interested reader is referred to OECD/Eurostat (2007) and OECD/Eurostat (2008).

We start by defining an enterprise. *Enterprise* is the smallest combination of legal units that is an organisational unit producing goods or services, which benefits from a certain degree of autonomy in decision making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit. (Council Regulation (EEC) No 696/93 of 15 March 1993 on the statistical units for the observation and analysis of the production system in the Community). By *employer enterprise* is meant an enterprise with at least one employee.

An *employer enterprise birth* refers to the birth of an enterprise with at least one employee. This population consists of enterprise births with at least one employee in the birth year and of enterprises that existed before the year under consideration but were below the threshold of one employee.

The indicator *employer enterprise birth rate* is compiled as the number of births of employer enterprises as a percentage of the population of active enterprises with at least one employee.

High growth enterprises as measured by employment (or by turnover) are all enterprises with average annualised growth in employees (or in turnover) greater than 20% a year, over a three-year period, and with ten or more employees at the beginning of the observation period.

The formulas describing high growth enterprises are

$$\sqrt[3]{\frac{employees_{(xx)}}{employees_{(xx-3)}}} - 1 > 0.2 \quad \text{(as measured in employment)}$$

$$\sqrt[3]{\frac{turnover_{(xx)}}{turnover_{(xx-3)}}} - 1 > 0.2 \quad \text{(as measured in turnover),}$$

where $employees_{(xx)}$ denotes the number of employees in year xx and $employees_{(xx-3)}$ the number of employees in year $xx-3$. Similarly $turnover_{(xx)}$ and $turnover_{(xx-3)}$ denote the turnover in year xx and $xx-3$, respectively. In practice, average annualised growth of 20% over three years would be equal to 72.8% growth from $xx-3$ to year xx.

The *rate of high-growth enterprises* is computed as the number of high-growth enterprises as a percentage of the population of enterprises with ten or more employees.

Gazelles are the subset of high-growth enterprises which are up to five years old. In other words, they are enterprises up to five years old with average annualised growth (in employees or in turnover) greater than 20% per annum, over a three year period and with ten or more employees at the beginning of the observation period.

The *share of gazelles* is computed as the number of gazelles as a percentage of the population of enterprises with ten or more employees. For details on the different approaches to high growth firms and gazelles definition see Ahmad and Gonnard (2007).

The source of data about entrepreneurial performance is the OECD Database Structural and Demographic Business Statistics (SDBS), which has been developed within the EIP. In this database activities are classified according to the International Standard Industrial Classification of all Economic Activities (ISIC) Revision 3.

We considered data (and performed the analysis) on the Manufacturing sector (ISIC3 15-37).

2.2 Education Indicators

Indicators about education are taken from Education at a Glance 2007, an OECD publication that provides a comparable set of indicators on the performance of education systems in the OECD's member countries and in a number of partner economies.

In particular, we considered indicators about educational attainment (Education at a Glance, Table A1.1a): they are computed as percentage of the population aged 25-64 having obtained a certain degree (as highest level of education).

The classification is based on the International Standard Classification of Education (ISCED) 1997, which supplies the instruments for computing internationally comparable indicators.

In the following we define the indicators that we took into consideration:

- *Upper secondary education* (ISCED LEVEL 3): this level of education typically begins at the end of full-time compulsory education for those countries that have a system of compulsory education. The entrance age to this level is typically 15 or 16 years.
- *Tertiary education* (ISCED LEVEL 5): this level consists of tertiary programmes having an educational content more advanced than those offered at the lower level 3. Entry to these programmes normally requires the successful completion of ISCED level 3 or a similar qualification. This level corresponds to Bachelor's degree in English-speaking countries. Tertiary education is subdivided into:
  - *Tertiary type A* (ISCED LEVEL 5A), which is largely theoretically based and intended to provide sufficient qualifications for gaining entry into advanced research programmes and profession with high skills requirements.
  - *Tertiary type B* (ISCED LEVEL 5B), which is typically shorter than type A and focuses on occupationally specific skills geared for entry into the labour market.
- *At least upper secondary education:* this category comprises the population that has attained at least an upper secondary education.

More detailed information about these education levels are found in UNESCO (1997).

2.3 Research and Development (R&D) Indicators

A factor that we thought could be a determinant of entrepreneurship is the investment in new knowledge, since this is the basis of innovation and technological progress. And it is known that innovation and technological progress are the main sources of economic performance.

A major input that stimulate innovation is the investment in R&D. The indicator that we considered is built by considering business R&D by size classes of firms. Indeed, both small and large firms play an important role in countries' innovative performance, but their relative importance for business R&D varies. In OECD countries, the share of R&D performed by small and medium-sized enterprises (SMEs) (defined here as firms with fewer than 250 employees) is generally greater in small economies than in larger ones.

Small and medium-sized firms play an important role in innovation: they are a constant source of renewal of technology and of competitive pressures for large firms. However, SMEs face specific problems for innovating and for adopting new technologies (access to funds, markets and skilled labour).

On the other hand, the role of large firms should not be ignored: they play leading role in structuring markets, carrying out large-scale innovations and even in co-ordinating smaller firms. The complementary roles of small and large firms may vary across industries and across countries.

We argue that countries where a lot of SMEs perform R&D have a more active economy. Indeed it is quite obvious that large firms perform R&D. This is the reason why we considered the indicator '*SME share of business R&D*', computed as the number of SMEs performing R&D as a percentage of the business R&D.

The source of data about R&D is the OECD R&D Database.

2.4 Product Market Regulation (PMR) Indicators

PMR measures the degree to which policies promote or inhibit competition in areas of the product market where competition is viable.

The source for indicators of PMR is the OECD Market Regulation Database, where the indicators cover regulations in three areas: state control of business enterprises, legal and administrative barriers to entrepreneurship and barriers to international trade and investment (see Conway et al. 2005 and Nicoletti et al. 2000).

For the purpose of identifying determinants of entrepreneurship we considered indicators relative to *barriers to entrepreneurship*.

This indicator assumes values increasing with the degree of the restrictions imposed on market mechanism and it is constructed as the combination of detailed indicators that coincide with more specific features of the regulatory regimes. The detailed indicators can be classified into three main sub-domains: administrative burdens on start-ups, regulatory and administrative opacities (including the features

of the licences and permits system and the communication and simplification of rules and procedures) and barriers to competitions (including legal limitations on the number of competitors and exemptions to antitrust provisions for public enterprises).

3. Statistical analyses and results

In this section we present the analyses that have been performed. Section 3.1 contains the results of the regression analysis, which has been carried out using the SAS system, version 9.1. Section 3.2 is devoted to multidimensional scaling, performed by means of the program ALSCAL, developed by Forrest W. Young.

3.1 Regression Analysis

The hypotheses under study is that education plays a role as determinant of entrepreneurship. In order to validate this hypothesis we perform some regression analysis using as measures of entrepreneurship the variables employer enterprise birth rate and rate of high-growth enterprises. It would have also been interesting to consider the share of gazelles, but, due to insufficient number of observations, it has not been possible. Even for the other variables, the number of observations (countries) is not high; for this reason the models have to be kept quite simple. In the regression models, we consider the following independent variables: barriers to entrepreneurship, SME share of business R&D, tertiary, upper secondary and at least upper secondary educational attainments.

The results of the analysis are summarized in Table 1. Concerning the variable employer enterprise birth rate we found that the most significant education variable is the upper secondary education (p-value=0.0269) and that another significant factor is R&D through the variable SME share of business R&D squared (p-value=0.0591). The R^2 for this model is 0.6925.

As far as the variable rate of high-growth enterprises is concerned, we found a significant model containing the variables at least upper secondary education squared (p-value=0.1033) and the logarithm of barriers to entrepreneurship (p-value=0.0598). The R^2 for this model is 0.6032. The relations between dependent variables and explanatory variables are plotted in Figures 1 to 4.

The p-values for the model are quite high, but since the number of observations is very low we can perform tests at the 0.1 level of significance.

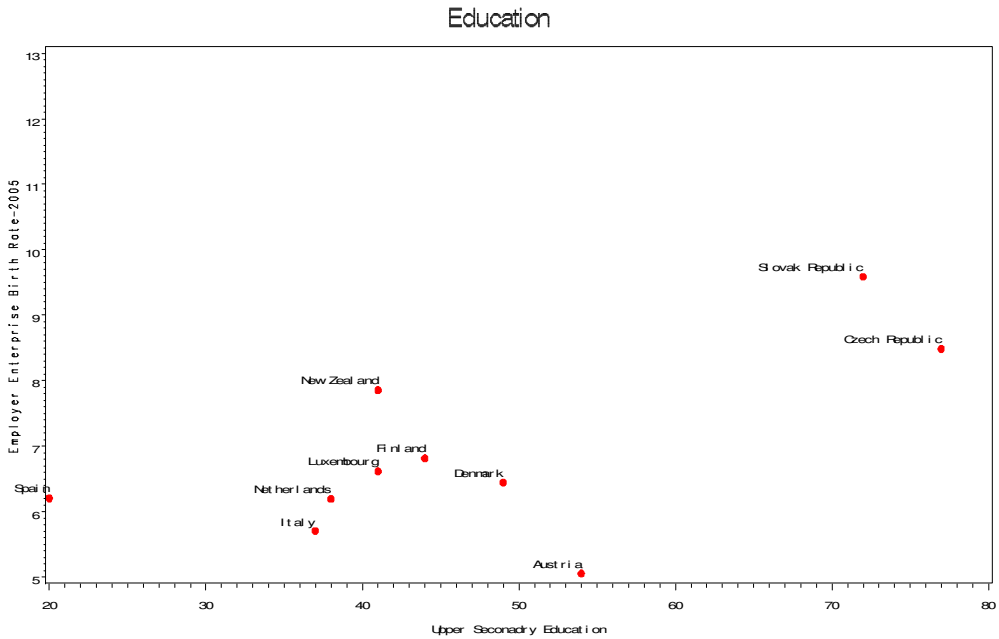
Notice that for each model we checked for multicollinearity using the Variance Inflation Factor (VIF) index and performed residual analysis (see Table 2).

From these analysis we can conclude that the birth of new enterprises is influenced by quite a low level of education (upper secondary) and by the percentage of SMEs which perform R&D. Indeed this variable denotes an active economy, which

promotes the birth of new enterprises. On the other hand, in order to have high-growth enterprises a higher level of education is needed. In fact we found that the most significant education variable is at least upper secondary education (which contains upper secondary, tertiary and higher education). This means that highly skilled workers are needed to have an enterprise quick growth, proving that education play a significant role in entrepreneurship competitiveness.

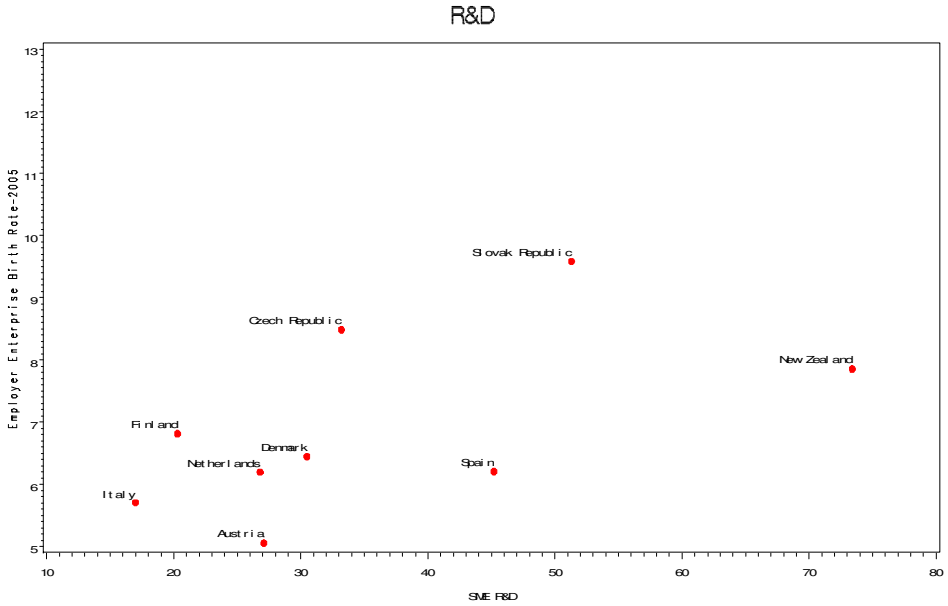
A second factor significant for high-growth enterprises is barriers to entrepreneurship. This fact can be interpreted as follows. Analysing how the variable barriers to entrepreneurship is defined, we see that this variable mainly interests enterprise births. Therefore, in countries where barriers to entrepreneurship are high it is more difficult to have an enterprise birth and this helps existing enterprises to grow faster.

Figure 1. Relation between upper secondary education and enterprise birth rate.



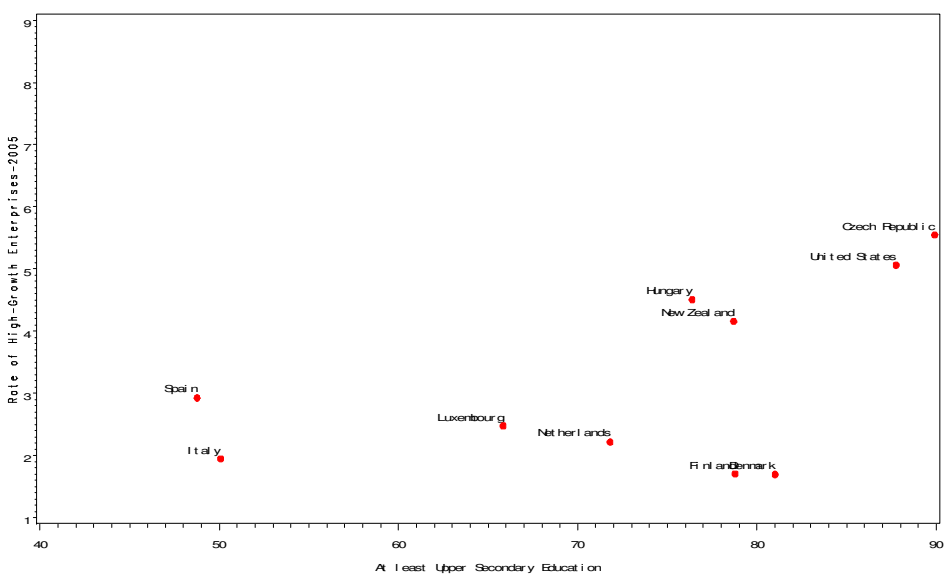
Source: own creation

Figure 2. Relation between SME share of business R&D and enterprise birth rate.



Source: own creation

Figure 3. Relation between at least upper secondary education and rate of high-growth enterprises.



Source: own creation

Table 1. Regression analysis.

Manufacturing (ISIC3 15-37)		
Model	**R^2**	**N. Observations**
Employer Enterprise Birth Rate= 3.618+0.05358 (Upper Secondary Education)+0.00047(SME R&D)^2 (p-value=0.0269) (p-value=0.0591)	0,6925	9
Rate of High-Growth Enterprises= 0.00034(At least Upper Secondary Education)^2+6.77ln(Barriers to Entrepreneurship) (p-value=0.1033) (p-value=0.0598)	0,6032	10

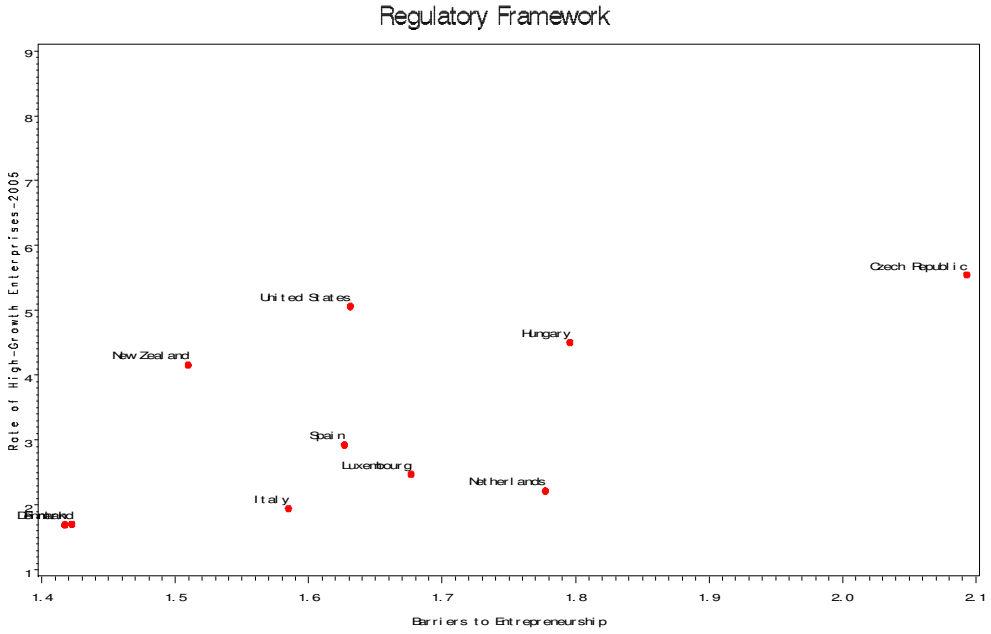
Source: own creation

Table 2. Residual analysis (Shapiro-Wilk test and Kolmogorov-Smirnov test for testing normality) and Variance Inflation Factor for multicollinearity.

Manufacturing (ISIC3 15-37)			
Dependent Variable	**S-W**	**K-S**	**VIF**
Employer Enterprise Birth Rate	0.1599	>0.1500	1.00079
Rate of High Growth Enterprises	0.6933	>0.1500	1.04161

Source: own creation

Figure 4. Relation between barriers to entrepreneurship and rate of high-growth enterprises.



Source: own creation

3.2 Multidimensional Scaling

To shed further light on the relationship among countries and variables we applied multidimensional scaling (MDS) techniques to produce a spatial representation of the data. This exploratory data tool allows to take advantage of the fact that for each observation we have many variables and moreover it is not affected by the small number of observations.

The obtained mapping shows the 'hidden structure' of the data and allows to understand the degree of relation between variables and countries.

Among the various types of MDS techniques we choose the classical multidimensional unfolding (CMDU). This type of MDS is suitable when the data form a rectangular matrix. CMDU represents both sets of objects (in our case countries and variables) as points in a single space, called joint Euclidean space.

The information in the data that is used to form the space is the degree of relation between rows and columns (countries and variables). The higher the value of a variable for a certain country, the nearer the country point will be to that variable point in the joint space.

Therefore country points will be located close to points for variables that take high values for that country, and far from points for variables that take low values for that country.

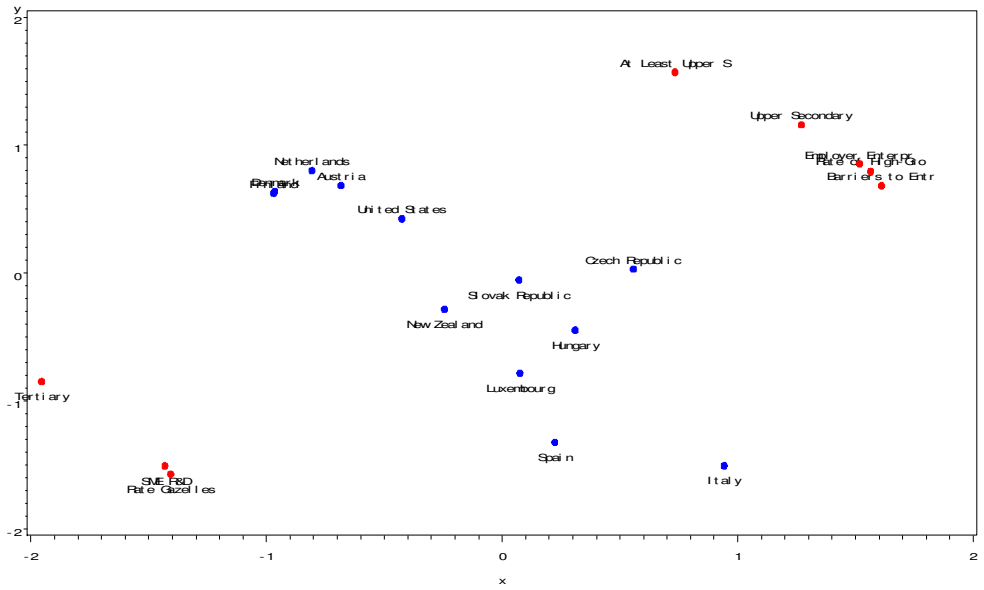
Conversely, each variable point will be located close to points for countries which assume high values of that variable and far from points for countries which assume low values.

Moreover, those countries which are near each other in the joint space have a similar behaviour with respect to the variables that are considered; those which are far from each other have different behaviour. For more details about MDS please refer to Borg-Groenen (1997), Kruskal-Wish (1981) and Schiffman et al. (1981).

The variables that we consider concern entrepreneurship indicators (employer enterprise birth rate, rate of high-growth enterprises and share of gazelles), education indicators (upper secondary, tertiary and at least upper secondary education), R&D indicators (SME share of the business R&D) and PMR indicator (barriers to entrepreneurship). Countries have been chosen in order not to have more than two missing values.

After standardization of the variables, we performed CMDU by means of the program ALSCAL. The result is displayed in Figure 5. In order to facilitate the interpretation of the map, disparities (which correspond to distances between points in the map) are shown in Table 3.

Figure 5. CMDU map.



Source: own creation

Looking at the picture and looking at the data in Table 3, we see that country points form some clusters, denoting similar behaviour: Denmark and Finland, Netherlands and Austria, Italy quite close to Spain.

Denmark and Finland are the countries closest to tertiary education, meaning that in these countries tertiary education is the highest among the countries represented. They are also not far from at least upper secondary education. Concerning the entrepreneurship indicators, they are closer to the variable share of gazelles than to employer enterprise birth rate and rate of high-growth enterprises. This fact clarifies which kind of entrepreneurship characterizes these countries: it is dynamic and involves highly skilled workers.

Netherlands and Austria are quite close to the variable at least upper secondary education, denoting that the education level of these countries is quite high but still with a good percentage of people having an upper secondary education. Indeed, with respect to Denmark and Finland they are farther from tertiary education and closer to upper secondary education.

As far as the entrepreneurship indicators are concerned, these countries are equidistant from the three indicators and quite far away from all of them, meaning that entrepreneurship is not very much differentiated and it does not present a high level.

Table 3. Disparities for CMDU

	EEBR	RHGE	SG	BE	R&D	TE	USE	ALUSE
Austria	2.208	2.251	2.368	2.294	2.315	1.990	2.011	1.670
Czech Republic	1.264	1.264	2.534	1.238	2.514	2.661	1.334	1.547
Denmark	2.492	2.534	2.252	2.576	2.194	1.784	2.295	1.937
Finland	2.498	2.540	2.237	2.581	2.179	1.769	2.303	1.948
Hungary	1.773	1.763	2.053	1.720	2.040	2.301	1.869	2.058
Italy	2.427	2.381	2.350	2.286	2.374	2.972	2.683	3.081
Luxembourg	2.180	2.168	1.679	2.121	1.672	2.032	2.279	2.441
Netherlands	2.325	2.372	2.446	2.421	2.389	2.007	2.108	1.722
New Zealand	2.096	2.105	1.735	2.091	1.705	1.801	2.091	2.093
Slovak Republic	1.706	1.716	2.119	1.705	2.091	2.177	1.704	1.751
Spain	2.530	2.505	1.650	2.436	1.666	2.230	2.692	2.935
United States	1.990	2.024	2.222	2.052	2.176	1.988	1.848	1.629

Note: EEBR=Employer Enterprise Birth Rate, RHGE=Rate of High Growth Enterprises, SG=Share of Gazelles, BE=Barriers to Entrepreneurship, R&D= SME R&D, TE=Tertiary Education, USE=Upper Secondary Education, ALUSE=At Least Upper Secondary Education

Source: own creation

Czech Republic is the closest to employer enterprise birth rate and rate of high-growth enterprises; this fact is probably related to its rapid economic growth. As far as education is concerned, Czech Republic is the closest to upper secondary education, denoting quite a low level of education, and it is far away from tertiary education.

Hungary and Slovak Republic are quite close to Czech Republic, denoting a similar behaviour; with respect to Czech Republic they are closer to tertiary education and share of gazelles, meaning that entrepreneurship is more dynamic.

4. Conclusions

From these preliminary analysis, it seems that the birth of new enterprises is influenced by quite a low level of education and by the percentage of SMEs which perform R&D.

On the other hand, in order to have an enterprise quick growth a higher level of education is needed.

We have also seen how MDS techniques can be used to analyse entrepreneurship across countries and in particular to identify countries with similar entrepreneurial characteristics.

Due to the very small number of observations, it is clear that we have to be very careful in drawing general conclusions. These first results need to be further investigated using more detailed data. In this respect we intend to analyse single country data and for this purpose we have already asked for the access to the Kauffman Firm Survey (KFS) confidential database. From this further investigation we hope to obtain results confirming for the US the conclusions we have drawn in this analysis. At the same time other aspects specific to the US economy may rise and shed further light on the phenomenon.

Sources

OECD, Education at a Glance, 2007
OECD, Market Regulation Database, October 2007
OECD, R&D Database, May 2007
OECD, Structural and Demographic Business Statistics (SDBS) Database

References

Ahmad, N. – Hoffman, A. 2008: A Framework for Addressing and Measuring Entrepreneurship. *OECD Statistics Directorate Working Paper*, STD/DOC(2008)2

Ahmad, N. – Seymour, R.G. 2008: Defining Entrepreneurial Activity: Definitions Supporting Frameworks for Data Collection. *OECD Statistics Directorate Working Paper*, STD/DOC(2008)1

Ahmad, N. – Gonnard, E. 2007: High Growth Enterprises and Gazelles, The International Consortium on Entrepreneurship (ICE), Meeting 22-23. February, 2007. Copenhagen, Denmark.

Borg, I. – Groenen, P. 1997: *Modern Multidimensional Scaling*. Springer, New York

Conway, P. – Janod, V. – Nicoletti, G. 2005: Product market regulation in OECD countries: 1998 to 2003. *OECD Economics Department Working Paper*, ECO/WKP(2005)6

Fairlie, R. W. 2009: Kauffman Index of Entrepreneurial Activity, 1996-2008, Erwing Marion Kauffman Foundation

Klapper L. – Raphael A. R. – Guillén M.F. 2008: Entrepreneurship and Firm Formation Across Countries.

Kruskal, J.B. – Wish, M. 1981: *Multidimensional Scaling*. Sage Publications, London.

Nicoletti, G. – Scarpetta, S. – Boylaud, O. 2000: Summary indicators of product market regulation with an extention to employment protection legislation. *OECD Economics Department Working Paper*, ECO/WKP(99)18

OECD /Eurostat 2007: *Eurostat-OECD Manual on Business Demography Statistics*, OECD, Paris

OECD/Eurostat 2008: *Measuring Entrepreneurship: A digest of indicators,* OECD, Paris.

Schiffman, S.S. – Reynolds, M.L. – Young, F.W. 1981: *Introduction to Multidimensional Scaling: Theory, Methods and Applications*. Academic Press, New York.

UNESCO 1997: *The International Standard Classification of Education: ISCED-97*. UNESCO, Paris.

Twostep cluster analysis: Segmentation of largest companies in Macedonia

Marija Trpkova[1] - Dragan Tevdovski[2]

One of the important procedures for segmentation and classification of the largest Macedonian companies is twostep cluster analysis. This clustering method is very efficient in classification of large data sets, has the ability to create groups using categorical and continuous variables and it is provided with automatic selection of number of clusters. These are all advantages of twostep analysis compared to the traditional clustering methods.

The goal of this paper is to present valuable application of the twostep cluster analysis in segmentation of the Macedonian companies. Every year, the Central Register of Republic of Macedonia and Euro Business Centre - Macedonia present a publication that reveals the 200 largest and most successful companies in Macedonia. In order to reveal the structure of the Macedonian companies, twostep cluster analysis is performed using the following continuous variables: total revenue in 2007, total revenue in 2006, earnings before taxes in 2007, revenue growth rate 2007/2006 and number of employees. Also, one categorical variable is included, type of industry.

The analysis successfully manages to create solution of four clusters or four different types of companies on the Macedonian market. The first type represents the most successful companies with significantly high revenues, earnings and stabile growth. These companies come from industries such as communications, electricity and manufacturing, and provide significant employment of the work force. The second type represents companies with relatively smaller revenues and earnings compared to the first type, but yet higher than the country's average. These are all manufacturing companies with steady growth. The third group has slightly smaller revenues and earnings than the second group, but the difference is that this group represents companies with high revenue growth rate, representing developing companies with significant potential. These are companies that mostly provide services, companies that provide telecommunication and transport, and also few companies from other industries. The last group represents the smallest companies from the analyzed 200 largest companies, having the smallest revenues, earnings and number of employees. These companies will develop further, but with much smaller rate than the companies in the third group. These are all companies that deal with retail and wholesale trade.

These findings are useful because mainly they provide the general structure of the largest Macedonian companies. For the potential foreign investors this analysis is an insight

[1] Marija Trpkova, Teaching and research assistant, Faculty of Economics, University "Ss. Cyril and Methodius" (Skopje)

[2] Dragan Tevdovski, MSc, Teaching and research assistant, Faculty of Economics, University "Ss. Cyril and Methodius" (Skopje)

to the most lucrative industries in the country. For the government, the presented results give information about which industries dominate in the most successful companies, in order to invest in their development through infrastructure, university education, tax relief and deduction of other expenditures.

Keywords: Twostep cluster analysis, log-likelihood distance measure, Schwarz Bayesian information criterion, segmentation

1. Introduction

In order to achieve the desired economic growth, Macedonian economy needs to attract foreign investments. The inflow of the foreign capital means growth of the gross domestic product, increase of the employment rate, improvement of the overall standard of living of Macedonian population and faster integration in the European Union. This paper represents an effort to create the general structure of the most successful Macedonian companies. For the potential investors it will act as indicator to the most lucrative industries in Macedonia and stimulus for increased inflow of foreign investments.

The twostep cluster analysis proves to be important procedure in clustering of the large data base consisted of 200 companies – observations with five continuous and one categorical variable.

Structure of the paper is following: the first part of the paper briefly explains the twostep cluster algorithm, determining the number of clusters and assignment of the observations in the most appropriate newly created clusters. The second part of the paper shows the application of the procedure in clustering of the 200 most successful Macedonian companies. Used data set and the empirical results are elaborated. Final conclusion about the findings of the analysis is given.

2. Twostep cluster analysis algorithm

Twostep cluster analysis is method of the statistical software package SPSS used for large data bases, since hierarchical and k-means clustering do not scale efficiently when n is very large (Garson 2009). This analysis can be used both for categorical and continuous variables, and has its application when there are categorical variables with three or more categories.

Twostep cluster analysis represents method that requires only one pass throughout the data. The process is consisted of two major steps: first step, where initial clustering of observations into small subclusters is performed and further on these subclusters are treated as separate observations. The decision whether the observation is joined in already formed cluster or a new cluster shall be formed is made on the bases on the distance criteria. The grouping of these new observations

is done by hierarchical cluster method. It is possible for the algorithm of the twostep cluster analysis to determine the number of clusters, or the number of clusters can be assigned previously. The second step is groping, where the subclusters are bases for the analysis, and they are grouped into the required number of clusters. Since the number of subclusters is significantly smaller than the number of observations, the traditional grouping methods are easy to be used. The method is more precise if there are more subclusters (Zhang et al. 1996).

In this analysis, if one or more variables are categorical, the log-likelihood distance measure is used, in such manner that the observations are grouped in the cluster which has the highest values of this measure, using a method developed by Meila and Heckerman (1998). If all of the variables are continuous, the Euclidean distance is used, so that the observations are grouped in the cluster that has the smallest Euclidean distance. SPSS algorithm uses a decrease in the log-likelihood distance measure for combining clusters as the distance measure because the log-likelihood method is compatible with categorical and continuous variables.

The procedure of twostep cluster analysis that uses log-likelihood distance measure assumes normal distribution for continuous variables and multinomial distribution for categorical variables. The twostep cluster analysis gives good results even if the normality assumption is not met. Another assumption of this analysis is that the sample is large (> 200).

The distance measure is needed in both steps, or in the step of the initial clustering and in the clustering step. There are two distance measures available, the first one is log-likelihood distance measure which represents the distance based on probability. The distance between two clusters is in a relation with the decrease of the value of the log-likelihood distance measure, when two clusters are joined in one (Banfield-Raftery 1993). While calculating the log-likelihood distance measure, normal distribution for continuous variables and multinomial for categorical variables is assumed. Also, the independence of the variables and independence of observations is also assumed. The distance between clusters R and S is defined as

$$d_{(R)(S)} = \xi_R + \xi_S - \xi_{(R,S)}$$

where

$$\xi_v = -N_v \cdot \left(\left(\sum_{k=1}^{K^A} \frac{1}{2} \cdot \log(\hat{\sigma}_k^2 + \hat{\sigma}_{v \cdot k}^2) \right) + \left(\sum_{k=1}^{K^B} \hat{E}_{v \cdot k} \right) \right)$$

and where

$$\hat{E}_{v \cdot k} = -\sum_{l=1}^{L_k} \left(\frac{N_{v \cdot k \cdot 1}}{N_v} \cdot \log\left(\frac{N_{v \cdot k \cdot 1}}{N_v} \right) \right)$$

where

K^A is the total number of the continuous variables in the analysis; K^B is the total number of the categorical variables in the analysis ; R_k is the interval or range of the

k continuous variable; N is the number of observations in the data base; N_k is the number of objects in k cluster; $\hat{\sigma}_k^2$ is the estimated variance of the k continuous variable for all data; $\hat{\sigma}_{Rk}^2$ is the estimated variance of the k continuous variable in the R cluster; N_{Rkl} is the number of objects in the R cluster, where k categorical variable takes the l category; $d_{(R)(S)}$ is the distance between the R and the S clusters; (R,S) is the index that represents cluster which is formed by joining of the clusters R and S (Chiu et al. 2001).

If the $\hat{\sigma}_k^2$ is ignored in the equation, the distance between the clusters R and S will be equal to the decreased value of the log-likelihood distance measure when two clusters are joined. The expression $\hat{\sigma}_k^2$ is given as a solution of the rising problem, if $\hat{\sigma}_{v\cdot k}^2 = 0$, by which undefined values for natural logarithm are reached. This problem occurs if the clusters have only one observation.

The other distance measure, the Euclidean distance, can be used only in a situation when all of the variables are continuous. The Euclidean distance between two points is clearly defined. The distance between two clusters is defined by the Euclidean distance between their centroids. The centroid of the clusters is defined as vector consisted of the means of all variables for a given cluster.

The procedure of the twostep cluster analysis begins with the first step, which is creation of initial cluster. This step uses method of sequential clustering. It analyzes the observations and decides if the given observation will join in one of the already formed cluster, or whether it will form a new cluster. This decision is based on the distance criteria.

2.1. Determining the number of clusters

When the process of clustering is started, the question is how many clusters should be formed. The answer depends on the data base. Characteristic of the hierarchical cluster analysis is to form a set of possible solutions from one pass throughout the date, with one, two, three or more cluster. K-means cluster algorithm has to be performed several times (each time for different number of clusters) so that a set of solutions is generated.

For automatic determination of number of clusters SPSS has developed the twostep procedure which is compatible with hierarchical cluster analysis. In the first step, *BIC* or Bayes information criterion or *AIC* or Akaike information criterion statistics is calculated for each different cluster solution with different number of clusters. In the second step, the initial estimate is improved by finding the highest distance increase between the two closest clusters during each stage in the hierarchical clustering.

The statistics BIC and AIC for R clusters is defined as

$$BIC_R = -2 \cdot \sum_{i=1}^{R} \xi_R + m_R \cdot \log(N)$$

$$AIC_R = -2 \cdot \sum_{i=1}^{R} \xi_R + 2 \cdot m_R$$

where

$$m_R = R \cdot \left\{ 2 \cdot K^A + \sum_{k=1}^{K} (L_k - 1) \right\}$$

where L_k is the number of groups in k categorical variable.

2.2. Assignment of the observations

The assignment of the observations into cluster is done by assigning the observations to the nearest cluster, if there is no transformation of the outliers. If there is an outlier transformation, then log-likelihood distance is used.

Let us assume that the extreme observations follow a normal distribution. Two separate likelihood functions are calculated, one when the observation is assigned to unstandardized cluster, and the other one to the nearest cluster of the unstandardized cluster. The observation then is assigned to the cluster that has the highest value of the log-likelihood function. The procedure is equal to assignment of the observation to the nearest cluster (which is not unstandardized) if the distance from the cluster is smaller than the critical value $C = \log(V)$ where $V = \left(\prod_k R_k \right) \cdot \left(\prod_m L_m \right)$. In rest of the cases, the observation is classified as outlier.

The object is assigned in the nearest cluster (which is not unstandardized) if the Euclidean distance is smaller than the critical value $C = 2 \cdot \sqrt{\frac{\sum_{l=1}^{K^A} \hat{\sigma}_{k \cdot l}^2}{K^A}}$, otherwise is classified as outlier.

Missing values are not allowed. The observations with missing values are excluded from the analysis.

3. Application of the twostep cluster analysis in clustering of 200 Macedonian companies

Twostep cluster analysis creates subclusters using hierarchical methods. If the analysis includes large data base, twostep cluster analysis is recommended, as well in situation when the categorical variables are included. The analysis performed with statistical software SPSS gives significant output, as well as variable importance charts.

3.1. Data set

The data set is consisted of 200 Macedonian companies. The data are provided from the Central registry of Macedonia and they refer to the most successful 200 companies. The main purpose of this analysis is to define clusters of companies on the base on the following continuous variables: total income in 2007, total income in 2006, income growth rate 2007/2006, earning before taxes in 2007 and number of employees. Also, there is one categorical variable, type of industry. The final clusters and their profiles will provide with the structure of the main groups of Macedonian companies. Analysis is performed with the statistical software SPSS 15.

3.2. Empirical results

The first step in the analysis is to examine the data and make the necessary transformations. The starting point is standardization of the variables. The twostep analysis performs standardization of the continuous variables, yet the final results are given in original values of the variables. The question about the outliers is solved by the Mahalanobis D^2 distance measure (Hair et al. 2005). This measure shows that less than ten companies may be potential outliers. The decision is to keep these companies in the analysis.

The correlation matrix calculates the Pearson coefficient of correlation between the variables and confirms that the correlation exists. This is expected, especially between the variables income and earnings. This analysis does not have a situation where several variables represent a factor, where if multicollinearity exists, it is necessary to reduce the variables for each factor or set of correlated variables. Because of this, the analysis can be continued, without fulfilment of this assumption.

The analysis includes five continuous variables and one categorical variable. If one or more variables are categorical, then a log-likelihood distance measure is used. If all variables are continuous, then the Euclidean distance is used. For determination of number of cluster Bayes or Akaike information criteria is used. The number of clusters also can be automatically assigned by the researcher. The analysis gives the following results.

Table 1 is automatic clustering table.

The presented statistics are values of the Bayesian criterion and value of BIC change for solutions with different number of clusters. When performing automatic clustering, different criteria are given, so that the best cluster solution is the one that has the lowest value of the Schwarz's Bayesian Criterion, or the lowest vale of Akaike information criterion. The statistical software SPSS chooses the solution that has significantly high value of BIC change and high value of ratio of distance measures. The performed simulation studies confirm that the ratio of BIC changes, which is combined criterion, givers better results than the individual values of Bayesian and Akaike information criterion.

In the presented example, the best solution is the one with 4 clusters, because this solution gives highest value for the ratio of distance measures and the lowest value of the Schwarz's Bayesian Criterion. The SPSS algorithm need not agree with the BIC criterion used alone, though it does in this example. When it differs, in essence the SPSS algorithm judges that the gain in information from having more than the number of clusters specified by BIC alone is not worth the increased complexity (diminution of parsimony) of the model. The researcher has the option to override this default and specify 6 or some other number of clusters (Garson 2009).

Table 1. Automatic clustering

Number of clusters	Schwarz's Bayesian Criterion (BIC)	BIC Change (a)	Ratio of BIC Changes (b)	Ratio of Distance Measures (c)
1	1272,747			
2	990,722	-282,026	1,000	1,839
3	884,535	-106,186	0,377	1,152
4	806,071	-78,464	0,278	2,168
5	825,571	19,500	-0,069	1,906
6	884,966	59,395	-0,211	1,212
7	952,054	67,088	-0,238	1,394
8	1029,408	77,354	-0,274	1,034
9	1107,607	78,199	-0,277	1,146
10	1189,023	81,416	-0,289	1,209
11	1274,238	85,216	-0,302	1,015
12	1359,726	85,488	-0,303	1,318
13	1449,533	89,808	-0,318	1,373
14	1543,036	93,503	-0,332	1,189
15	1638,112	95,076	-0,337	1,126

(a) The changes are from the previous number of clusters in the table.
(b) The ratios of changes are relative to the change for the two cluster solution.
(c) The ratios of distance measures are based on the current number of clusters against the previous number of clusters.
Source: Own creation

The Table 2 presents the distribution of the observations in the cluster, or the number of observations in each cluster. This is the first indicator of the size of clusters. Also, the number of the excluded observations is also given, or in this example 24 observations are excluded because they to not have sufficient data for the chosen variables so they can be grouped in of the clusters.

Table 2. Cluster distribution

Cluster	Number of observations	% of combined cluster	% of total
1	7	4,0%	3,5%
2	64	36,4%	32,0%
3	43	24,4%	21,5%
4	62	35,2%	31,0%
Combined clusters	176	100,0%	88,0%
Excluded observations	24		12,0%
Total	200		100,0%

Source: Own creation

The centroids Table 3 represent descriptive statistics for the continuous variables. The mean values for all continuous variables for each cluster are presented.

Table 3. Cluster centroids

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Total income in EUR (2007)	301.044.003	25.714.168	21.771.629	16.452.301
Total income in EUR (2006)	248.912.033	21.215.261	16.912.291	13.322.724
Income growth rate 2007/2006	27	33	6.119	129
Earnings before tax EUR (2007)	48.756.644	2.103.284	1.714.597	845.832
Number of employees (2007)	1.872	298	371	88

Source: Own creation

The interpretation of the clusters profiles shows that the first cluster has 7 companies. According to the cluster centroids, these are the largest and the most profitable companies, with the highest number of employees and steady growth rate. The main industries of these companies are manufacturing, electricity and

telecommunication. These are the largest and most successful companies in Macedonia.

The second cluster is the biggest and has 64 companies that have significantly lower income and earnings from the companies in the first cluster. These companies have steady growth and the number of employees is also lower than the number of employees in the first cluster. These are middle size companies with normal growth.

Table 4. Frequencies

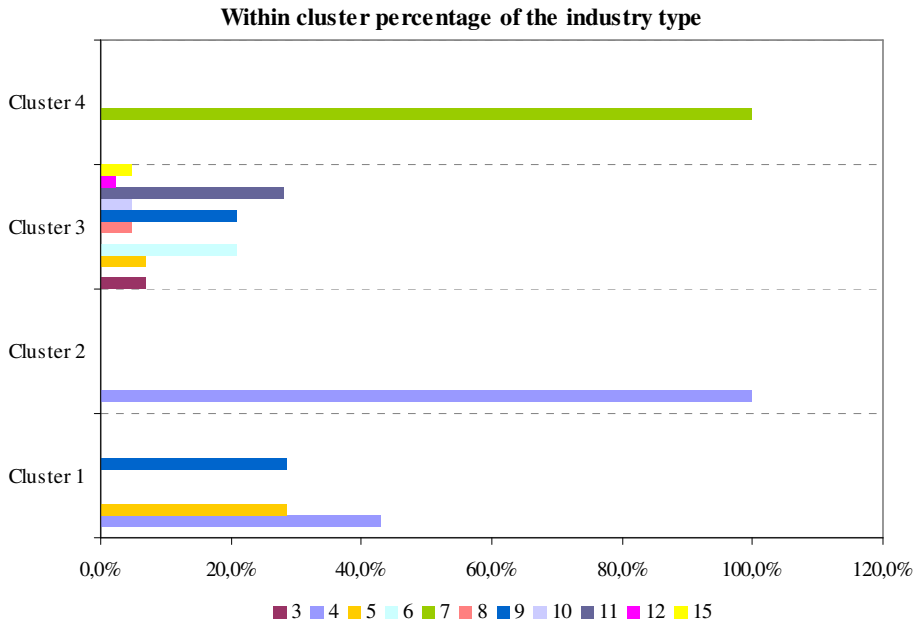
		Industry code										
		3	4	5	6	7	8	9	10	11	12	15
Frequency	Cluster 1	0	3	2	0	0	0	2	0	0	0	0
	Cluster 2	0	64	0	0	0	0	0	0	0	0	0
	Cluster 3	3	0	3	9	0	2	9	2	12	1	2
	Cluster 4	0	0	0	0	62	0	0	0	0	0	0
	Combined	3	67	5	9	62	2	11	2	12	1	2
Percent	Cluster 1	0	4,5	40	0	0	0	18,2	0	0	0	0
	Cluster 2	0	95,5	0	0	0	0	0	0	0	0	0
	Cluster 3	100	0	60	100	0	100	81,8	100	100	100	100
	Cluster 4	0	0	0	0	100	0	0	0	0	0	0
	Combined	100	100	100	100	100	100	100	100	100	100	100

Source: Own creation

The third cluster has 43 companies and has slightly lower income and earnings from the companies in the second cluster. The number of employees is greater than the number of employees in the second cluster. The main difference here is in the income growth rate. The companies in this cluster have very high growth rate, indicating that these are developing companies, with potential for further development in even more successful companies.

The fourth cluster has 62 companies, with lowest income, earnings and number of employees. It is important to say that these companies also have high income growth rate. This is a cluster with companies significantly smaller then the companies in the other clusters, yet since the growth rate is high, further growth can be also expected here.

Figure 1. Graphical presentation of the frequencies (in percentage)



Source: Own creation

The frequency Table 4 uses the descriptive statistics for categorical variables. For each variable separate table is created. In this table is clearly shown that in the first cluster the companies that have industry code 4 – manufacturing, industry code 5 – electricity and industry code 9 – telecommunications dominate. In the second cluster all companies have industry code 4 – manufacturing. In the third cluster there are companies from different industries, practically all industries are present here. In the fourth cluster all companies have industry code 7 – trade.

The industry codes are: 1 – agriculture, forestry and hunting, 2 – fishing, 3 – mining and quarrying, 4 – manufacturing, 5 – electricity, gas, water supply, 6 – construction, 7 – wholesale and retail trade, 8 – hotels and restaurants, 9 – transport, storage and communications, 10 – financial intermediation, 11 – real estate, renting and business activities, 12 – public administration and defence, compulsory social security, 13 – education, 14 – health and social work, 15 – other community, social and personal service activities.

Frequency table does not have data about the industries with codes 1, 2, 13 and 14. The reason for this is because these types of industries are not present in the companies from the observed data base, or they are excluded because of the missing data.

The frequency distribution in the cluster is also shown on a chart on Figure 1, so that the group structure of the categorical variable in each cluster is clearer. The distribution is shown in percentage.

Figure 2 shows the within cluster variation. There is one error bar chart for each categorical variable which shows the arithmetic mean for all clusters for that particular variable. Since a sample is used, there is 95% confidence interval. The first chart represents the mean for the variable total income in 2007. It is clearly shown that the first cluster has significantly higher value from the other clusters, which have approximately same value for this variable. The same situation is for the variable total income in 2006, earnings before tax and number of employees. Only the variable growth rate has significantly higher confidence interval in the third cluster than the other cluster, which has moderate growth.

SPSS offers another group of charts as an outcome in the twostep cluster analysis – charts that show the significance of the variables (Figure 3). For the continuous and categorical variables, on special charts, on the X axis the χ^2 value is given, and on the Y axis the particular variable. Bar lines that are longer than the critical value show that the variable is statistically significant in differentiation of the clusters. The charts are shown in Figure 2.

The first variable is the categorical variable Industry. The value of χ^2 shows that this variable is statistically significant for differentiation of the clusters, especially for clusters 3,4 and 2, while less important for cluster 1. The second variable, total income in 2007 and the third variable, total income in 2006 are variables that significantly contribute in differentiation of the clusters 4, 3 and 1, while they do not contribute in differentiation of the cluster 2.

The fourth variable, growth rate of the total income, has particularly significant influence in differentiation of clusters 2 and 1, small influence in differentiation of cluster 4, and no influence in differentiation of cluster 3. The variable earnings before taxes in 2007 significantly differentiates cluster 4 and 3, while it is insignificant for clusters 2 and 1. The last variable, number of employees in 2007 is significant only for differentiation of cluster 4, while there is no influence on clusters 3, 2 and 1.

Table 5 represents all 200 companies and their cluster membership.

4. Conclusion

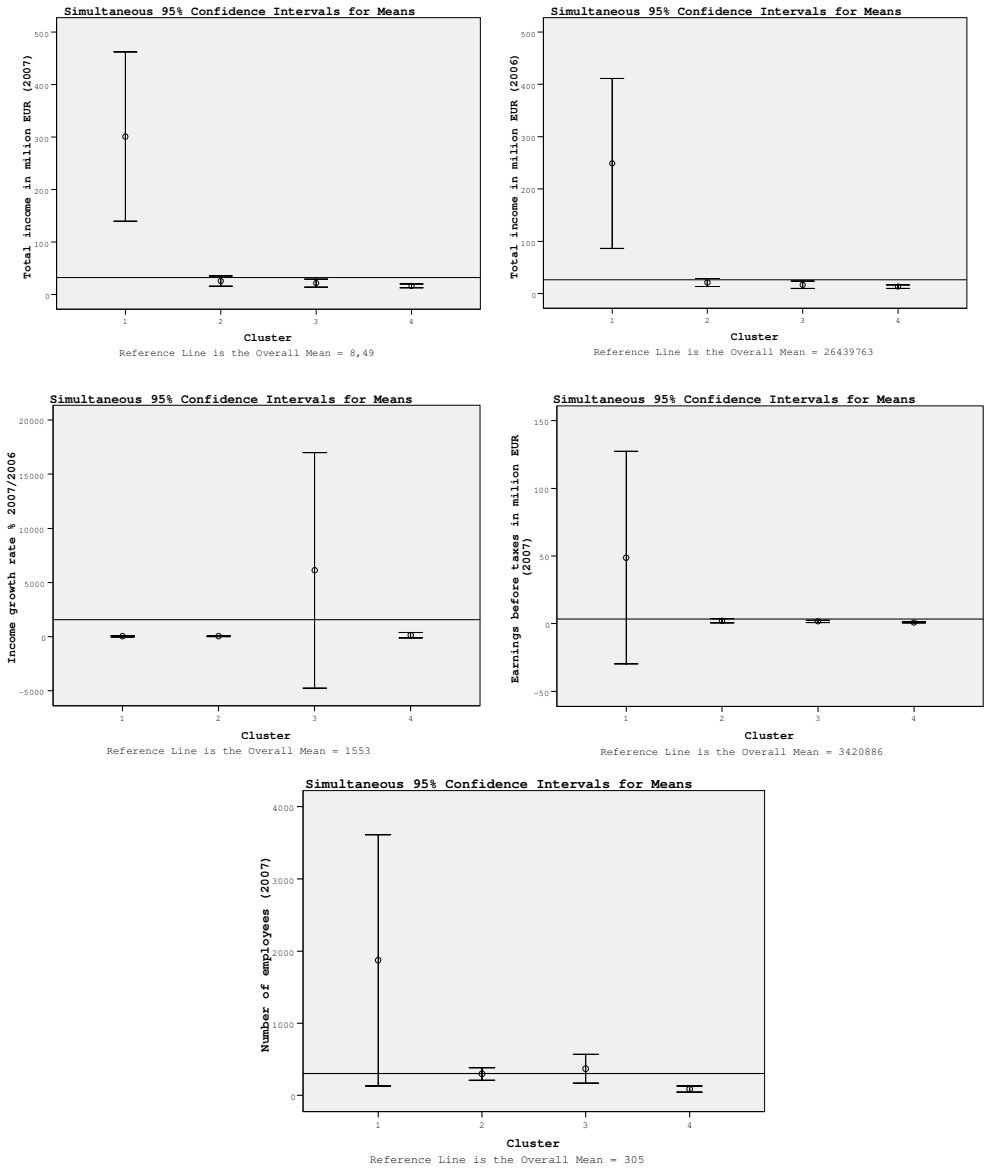
The main idea of this paper is to examine and to prove that the twostep cluster analysis is appropriate for segmentation of the Macedonian companies, so that the formed clusters are represents of the country's main company structure.

As presented in the empirical results, the four cluster solution seems to be a good represent of the main structure of the Macedonian companies. Distribution of

the observations in the clusters can be evaluated as satisfactory, since the number of companies in the clusters is relatively equal, except for the first cluster. The first cluster has small number of companies, yet these are the companies that are the base of the Macedonian company because they create a significant amount of revenues and also they provide the country with energy resources and telecommunication. Most of them were privatized from the state ownership and today they are joint stock companies partly owned by foreign investors. The second cluster is the largest cluster that includes some large companies and mostly middle sized companies. The industry of these companies is manufacturing. These are companies with steady growth, they have been present on the Macedonian market for number of years and can be qualified as firm companies. Except for the few large companies, most of the middle sized companies are owned by the domestic owners. The third cluster can be named as "the growing companies cluster" since here the income growth rate is significantly higher than the other clusters. Most of these companies are successfully companies that are more likely to grow in the years to come. Also, these companies employ significant part of the countries work force. The cluster is consisted of 43 companies. They come from different industries, yet mostly are from construction, transport, storage, telecommunications and other business activities. Investing in some of these companies may seem like a good decision, since most of these companies are likely to expand their business. The last cluster is consisted of the smallest, yet successful companies in wholesale and retail trade. Since the industry is trade, the earnings before tax are smaller than in clusters with companies from industries like manufacturing or services. Similar to the second cluster, majority of the companies are in domestic ownership, and can be classified as small and medium enterprises.

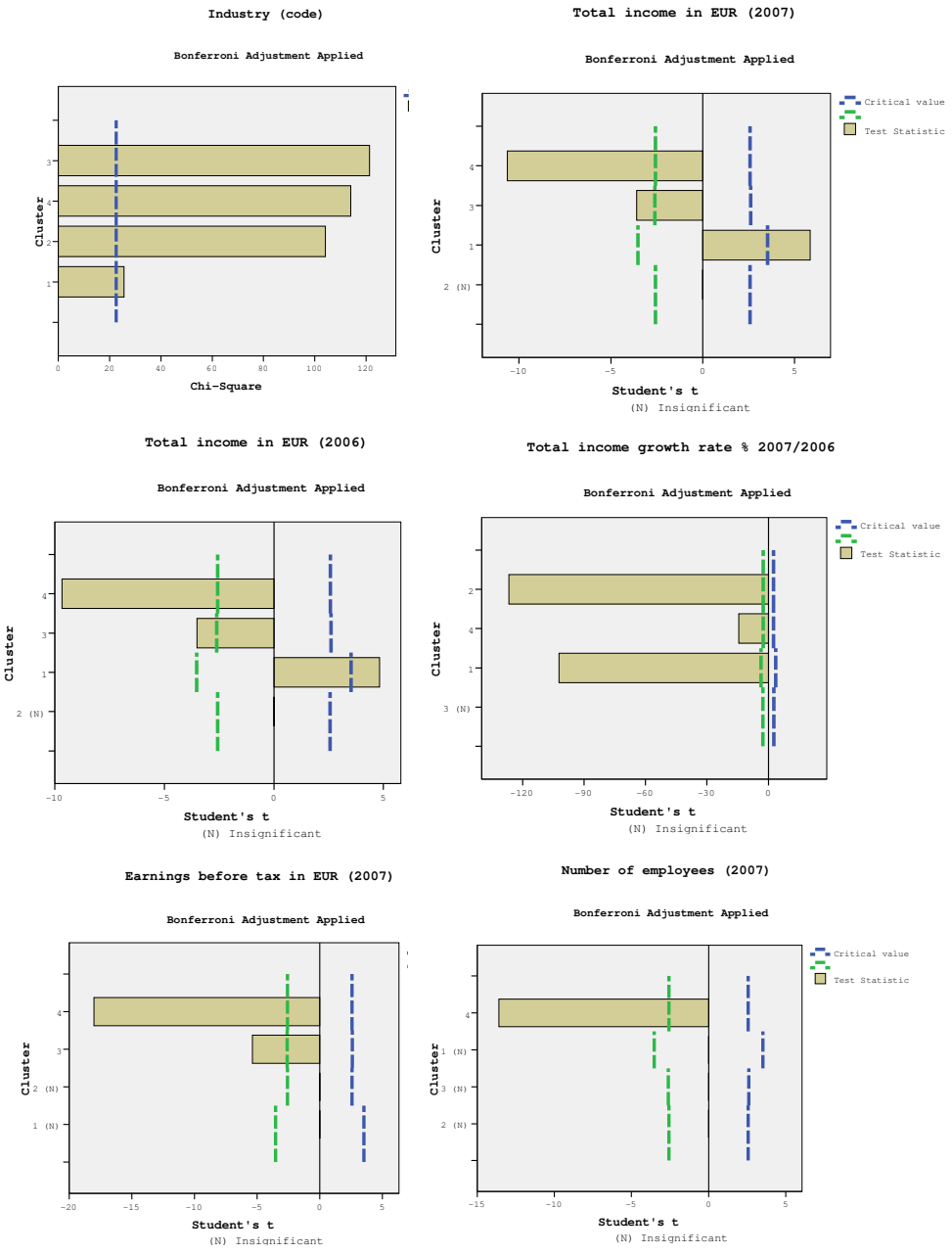
The given explanation of the empirical results proves that the twostep analysis is practical tool for segmentation of the large data base of 200 Macedonian companies. Findings of the analysis are useful because mainly they provide the general structure of the largest Macedonian companies. For the potential foreign investors this analysis is an insight to the most lucrative industries in the country. For the government, the presented results give information about which industries dominate in the most successful companies, in order to invest in their development through infrastructure, university education, tax relief and deduction of other expenditures. For further research, the cluster membership of the twostep cluster analysis can be used as dependent categorical variable in discriminant analysis.

Figure 2. Within cluster variation for all continuous variables



Source: Own creation

Figure 3. Variablewise importance



Source: Own creation

Table 5. Cluster membership of the companies

Company / Cluster		Company / Cluster	
Okta AD Skopje	1	Adrijus DOOEL Skopje	2
Feni industries AD Kavadarci	1	MZT Hepos AD Skopje	2
Makpetrol AD Skopje	1	ADE Skopsko pivo Tetovo	2
Makedonski telekom AD Skopje	1	Swisslion-Agroplod AD Resen	2
EVN Macedonia AD Skopje	1	Ekstra-Skopsko Kosel DOOEL Ohrid	2
T-Mobile Macedonia AD Skopje	1	Metro AD Skopje	2
Macedonian power plants JSC Skopje	1	Bas tuti friuti DOOEL Skopje	2
Arcelormittal Skopje (CRM) AD Skopje	2	Brako DOO Veles	2
Arcelormittal Skopje (HRM) AD Skopje	2	Fabika Karpos AD Skopje	2
Makstil AD Skopje	2	Mlekara Zdravje Radovo, Strumica	2
Usje AD Skopje	2	Metalopromet DOOEL Strumica	2
Pivara Skopje AD Skopje	2	Prototip DOOEL Skopje	2
Alkaloid AD Skopje	2	MZT Learnica AD Skopje	2
Igm-trade DOO Kavadarcki	2	Helmateks AD Strumica	2
Dojran stil DOO Dojran	2	Tondach-Makedonija AD Vinica	2
Skopski leguri DOOEL Skopje	2	Kiro Kucuk AD Veles	2
11 Oktomvri AD Kumanovo	2	Zdenka DOOEL Negotino	2
Tutunski kombinat AD Skopje	2	Fersped AD Skopje	3
Dil Petrol DOOEL Stip	2	Euro tabak DOO Skopje	3
Toplifikacija AD Skopje	2	Cosmofon AD Skopje	3
Brilijant DOOEL Stip	2	Granit AD Skopje	3
Silmak DOOEL Tetovo	2	Makedonski aviotransport AD Skopje	3
Swisslion DOO Skopje	2	Tec Negotino AD Negotino	3
Mlekara AD Bitola	2	NLB Lizing DOOEL Skopje	3
Zito luks AD Skopje	2	Knauf-radika AD Debar	3
Vinarska vizba Tikves AD Skopje	2	JP za stop. so stanben i deloven prostor Skopje	3
Bomex DOO Skopje	2	Terna AD Skopje	3
Pekabesko AD Skopje	2	Makosped AD Skopje	3
ZK Pelagonija AD Bitola	2	Makoten DOOEL Gevgelija	3
F.I. Vitaminka AD Prilep	2	EFT Makedonija DOOEL Skopje	3
EMO AD Ohrid	2	Sasa DOOEL Makedonska kamenica	3
Teteks AD Tetovo	2	Alma-m DOO Skopje	3
Mega DOOEL Skopje	2	Makedonska posta AD Skopje	3
EMO DOOLE Ohrid	2	JP Vodovod i kanalizacija Skopje	3
Prilepska pivarnica AD Prilep	2	Bucim DOOEL Radovis	3
Droga kolinska DOOEL Skopje	2	Beton AD Skopje	3
Vest DOOEl Bitola	2	Media print Makedonija DOO Skopje	3
Strumica tabak AD Strumica	2	Pexim DOOEL Skopje	3
Promes DOO Skopje	2	Mavrovoinzenerig DOOEL Skopje	3
Alayans uan Makedonija AD Kavadarci	2	Peas Macedonia Skopje	3
FI Blagoj Gorev JSC Veles	2	JP Komunalna higiena Skopje	3

MIK Sveti Nikole DOO Sveti Nikole	2	DS Iskra steel construction DOO Kumanovo	3
Imperijal-tabako AD Valandovo	2	Indo minerals&metals DOOEL Skopje	3
Kontihidroplast DOOEL Gevgelija	2	Neocom AD Skopje	3
Ideal sipka DOO Bitola	2	Makinvest DOO Skopje	3
Pucko petrol DOO Makedonski brod	2	Dzesas insaat tidzared i sanaji AS Podruznica SK	3
Komuna AD Skopje	2	Alfeks inzenering DOO Skopje	3
Riomk bomeks - refraktori AD Pehcevo	2	DGU Pelister Bitola DOO Bitola	3
TGS Tehnicki gasovi AD Skopje	2	On.net DOO Skopje	3
Evropa AD Skopje	2	Pakom kompani DOOEl Skopje	3
Leov kompani DOOEL Veles	2	Dauti komerc AD Skopje	3
Makprogres DOO Vinica	2	Mlaz AD Bogdanci	3
FHL Mermeren kombinat AD Prilep	2	Makedonija Turist AD Skopje	3
Pavor DOOEL Veles	2	Kiro D. Dandaro AD Bitola	3
Bunar petrol DOO Gostivar	3	Centro union DOO Skopje	4
Kvalitet-prom DOOEl Kumanovo	3	Gorenje DOOEL Skopje	4
Senker DOOEL Skopje	3	Podravka DOOEL Skopje	4
Pelagonija Inzinering DOOEL Skopje	3	Energomarket DOO Skopje	4
Publicis DOO Skopje	3	Tabako-promet BM DOOEL Valandovo	4
Int trejd DOOEL Kocani	3	Eurotrejd DOO Skopje	4
Lukoil Macedonia LTD Skopje	4	Montenegro DOO Gostivar	4
Tinex-mt DOOEL Skopje	4	Ka-dis DOO Skopje	4
Veroupulos DOOEL Skopje	4	Kolid kompani AS DOO, s. Kolesino Novo Selo	4
Porshe Makedonija DOOEL Skopje	4	German PX DOO Skopje	4
Gemak-trade DOOEL Skopje	4	Gamatroniks DOOEL Skopje	4
ZEGIN DOO Skopje	4	Automakedonija AD Skopje	4
Skopski Pazar AD Skopje	4	Marija treid DOO Veles	4
Euro aktiva DOO Skopje	4	Gross prom DOO Skopje	4
KAM DOOEL Skopje	4	Grosist DOOEL Bitola	4
Makautostar DOOEL Skopje	4	Agrokumanovo AD Kumanovo	4
Nelt st DOOEL Skopje	4	Swisslion Mak DOO Skopje	4
AD D-r Panovski Skopje	4	Agroefodia DOOEL Strumica	4
Pharmacy Zegin farm Skoopje	4	Kemo-farm DOOEL Skopje	4
Promedika DOO Skopje	4	Avto kuka DOO Skopje	4
Euro media DOO Skopje	4	Krka-farma DOOEL Skopje	4
Euroimpex DOO Skopje	4	Zito DOOEL Veles	4
KIK DOO Skopje	4	Mepso AD Skopje	*
Ekspanda DOOEL Skopje	4	Kameni most komunikacii AD Skopje	*
Makoil DOOEL Skopje	4	ADG Mavrovo Skopje	*
Automobile sk DOOEL Skopje	4	Zito vardar AD Veles	*
Elektroelement DOO Skopje	4	JP Makedonski sumi Skopje	*
Jaka 80 AD Radovis	4	Jaka tabak AD Radovis	*
Replek AD Skopje	4	Hypo-alpe-adria-lizing DOOEL Skopje	*
Ramstore Macedonia DOO Skopje	4	4 Noemvri AD Bitola	*

Filip Moris Skopje DOOEL Skopje	4	Tutunski kombinat - cigari DOOEL Prilep	*
Toyota avto centar DOOEL Skopje	4	Public transportation enterprise Skopje	*
Mako-market DOO Skopje	4	Haier Makedonija trejd DOOEL Skopje	*
Merkur Makedonija DOO Skopje	4	JP Makedonijapat Skopje	*
Makedonija Lek DOO Skopje	4	Germanos Telekom AD Skopje	*
Mi-da motors DOO Skopje	4	Lek DOOEL Skopje	*
Libra 1 AG Skopje	4	Lotarija na Makedonija AD Skopje	*
Euromilk DOO Skopje	4	JP Makedonska radio televizija Skopje	*
Tediko super DOOEL Skopje	4	TCG learnica DOOEL Ohrid	*
Internacional food bazar DOO Skopje	4	AD Negotino Negotino	*
Avtonova DOO Skopje	4	Prima.mk DOO Skopje	*
Rudine-mm DOO Skopje	4	Fruktal mak AD Skopje	*
JUS MB DOO Skopje	4	Fabrika za kvasec i alkohol AD Bitola	*
Makkar DOO Skopje	4	SAF Energiferzorgungslezngen GmbH podr. SK	*
Inter tobako DOOEL Skopje	4	Alfa kopi DOOEL Skopje	*
Kola DOOEL Skopje	4	MIA Beverages DOO Skopje	*

** Excluded observations due to the missing data*
Source: Own creation

References

Banfield, J. D. - Raftery, A. E. 1998: Model based Gaussian and non Gaussian clustering. *Biometrics,* 49, 803-821. p.

Chiu, T. - Fang, D.- Chen, J. - Wang, Y. - Jeris, C. 2001: A Robust and Scalable Clustering Algorithm for Mixed Type Attributes in Large Database Environment. *Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining,* SanFrancisco, CA: ACM, 263–268. p.

Euro Business Centre, Central Registry of Republic of Macedonia, 2008: *200 Largest*. Euro Business Centre.

Hair, J. F. - Black, B. - Babin, B. - Anderson, R. E. 2005: *Multivariate data analysis*. Upper Saddle River, NJ , Prentice Hall.

Garson, D. 2009: *Cluster analysis from Statnotes: Topics in Multivariate analysis* retrieved from http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm.

Meila, M. - Heckerman, D. 1998: *An experimental comparison of several clustering and initialization methods.* Microsoft Research Technical Report MSR-TR-98-06.

SPSS 2007: *SPSS 16.0 Command syntax reference*. SPSS Inc., Chicago.

Zhang, T. - Ramakrishnon, R. - Livny, M. 1996: BIRCH: Method for very large databases. *Proceedings of the ACM. Management of Data,* 103–114. p. Montreal, Canada.

Different Clustering Techniques – Means for Improved Knowledge Discovery

Olivera Grljević[1] - Saša Bošnjak[2] - Zita Bošnjak[3]

Application of different clustering techniques can result in different basic data set partitions emphasizing diversified aspects of resulting clusters. Since analysts have a great responsibility for the successful interpretation of the results obtained through some of the available tools, and for giving meaning to what forms a qualitative set of clusters, additional information attained from different tools is of a great use to them.

In this article we presented the clustering results of small and medium sized enterprises' (SMEs) data, obtained in DataEngine, iData Analyzer and Weka tools for intelligent analysis.

Keywords: Data mining, clustering, DataEngine, iData Analyzer, Weka

1. Introduction

The idea of *Knowledge Discovery in Databases* (KDD) is to search for relations and global schemes that exist in large databases and are hidden in the vast amount of data. *Data mining*, as the part of KDD, is the process of using one or more computational techniques in automated search for hidden information and relationships among data. As such, it represents indivisible part of qualitative research. Knowledge discovered through different data mining methods and techniques reveal behavioral patterns, profiles of entities, and similar regularities in data. Using solely statistical methods, qualitative data model can not be built.

[1] Olivera Grljević, MSc, University of Novi Sad, Faculty of Economics Subotica, Serbia
[2] Saša Bošnjak, PhD, associate professor, University of Novi Sad, Faculty of Economics Subotica, Serbia
[3] Zita Bošnjak, PhD, full professor, University of Novi Sad, Faculty of Economics Subotica, Serbia

Acknowledgement: This paper is the result of a research on the project titled: "Comparative Advantages of Intelligent Data Analysis Methods in Strengthening the Sector of Small and Medium Sized Enterprises", No. 114-451-01092/2008-01, funded by the Ministry of Sciences and Technology Development of Province of Vojvodina, Republic of Serbia.

Besides large databases, sophisticated algorithms are needed, which are subject of knowledge discovery in databases.

As proven by now (Liao- Triantaphyllou 2008), (Harrison-Llado 2000) each clustering algorithm, sometimes even the same algorithm applied several times on the initial dataset, can result in different basic dataset partitions, putting an accent on a specific aspect of the resulting clusters. Apart from diverse outputs, clustering algorithms use different visualization techniques to represent the derived clusters, which enable better insight into their structure and grouping relationships of similar entities. Furthermore, they can denote cluster centers, typical and least typical representatives of clusters, etc.

The selection of a subset of attributes of the database for clustering data, as well as the determination of the most adequate number of clusters, is under subjective appraisal of analysts. Furthermore, they have a great responsibility to carry out the interpretation of the results gained through some of the available tools successfully, and to give meaning to what forms a qualitative set of clusters. Consequently, additional information attained from different tools that support clustering techniques is of great use in clusters shaping.

Collecting and compounding various information about defined clusters, contributes to qualitative decision making on optimal cluster number and elements that constitute them. Consequently, to obtain as qualitative results as possible, and to facilitate cluster interpretation, analysts should combine different tools in the process of data clustering.

In our paper, we described a *composite approach* that implies diversity of tools and obtained results that significantly simplify the work of analysts in knowledge discovery, helps the interpretation of results, and facilitates the derivation of detail and clear conclusions. We present the results of clustering small and medium sized enterprises' (SMEs) data in Vojvodina province using *DataEngine, iData Analyzer* and *Weka tools* for intelligent analysis. Each tool supports a different clustering algorithm.

2. Techniques used in the empirical examination

The broader goal of our research was to determine discriminators between successful and less successful enterprises, and to distinguish the profile of businesses that will succeed in their goals from those that are likely to fail. These tasks are classification and clustering tasks, respectively. (Bošnjak et al. 2009) provides more detailed presentation of these problems. In this article we presented only the results of clustering techniques utilization, since they are common to all three tools we used, and are in compliance with the goals of our research.

In Komem- Schneider (2005), Bratko et al. (1998), Jiawei-Kamber (2001) it is defined that clustering is a process of grouping feature space vectors into classes in

the self-organized mode. Cluster is a group of points in a multi-dimensional space. The points aggregated in such a way are closer to each other and to their "cluster center" than they are to the centers of other groups. Within our research, we have created data models by c-means algorithm, improved fuzzy c-means algorithm and Kohonen neural networks for clustering tasks.

C-Means algorithm is a prototype-based, partitioning technique that attempts to find a user-specified number of clusters (c), which are represented by centroids. Centroid is usually the mean of a group of points and is typically applied to objects in a continuous n-dimensional space, (Tan et al. 2006), (Witten-Frank 2005). It is a very simple and fast algorithm. Since c-means requires that the user knows the exact number of clusters in advance, and usually this number is not obvious, so determining the initial value of c is a major difficulty in using this algorithm. Furthermore, a lack of explanation requires additional analysis by a supervised learning model. In a crisp C-means algorithm, each entity belongs to only one cluster, not being the case in numerous real world situations, and hence the algorithm is facing a limited usability.

One improvement of the C-means clustering algorithm incorporates the theory of fuzzy sets, resolving the single-membership problem by measuring the degree of membership of all entities to each cluster, by the membership function. However, the fuzzy C-means (FCM) algorithm still has several drawbacks that influence its performance. (Binu et al. 2009) it is stated that "the main drawback is from the restriction that the sum of membership values of a data point x_i in all the clusters must be one, and this tends to give high membership values for the outlier points." The second limitation refers to the fact that membership of a data point in one cluster is directly related to its membership in other clusters, and also the partial membership of all data members moves clusters' centers towards the center of all data points, producing sometimes unrealistic results. Consequently, additional information attained from different data analysis tools that support clustering techniques is of great use in clusters shaping.

3. Data mining tools overview

DataEngine (DE) software tool for intelligent data analysis is a very powerful tool that facilitates knowledge discovery in data. It combines statistical methods with neural networks technology, both supervised and unsupervised learning models, and fuzzy technology. Intelligent technologies DE supports are well proven in business, technology and academy work. In DE all data processing steps can be automated by graphical macro language and all models developed in DE can be incorporated into user's own programs (if they are built as Dynamic Link Libraries, for instance).

DE uses the Fuzzy C-Means (FCM) algorithm for partitioning a collection of points into a number of clusters. These data points are represented as feature vectors

and are describing objects. The objects within a cluster show a certain degree of closeness or similarity. Objects are assigned to each cluster with a corresponding membership degree. The algorithm is using validity criteria to determine number of clusters in the data.

(Roiger- Geatz 2003) it is stated that "The iData Analyzer (iDA) provides support for business or technical analyst by offering a visual learning environment, an integrated tool set, and data mining process support". iDA consists of a preprocessor for improving the quality of data, three data mining tools: unsupervised clustering, supervised learning and neural networks, and a report generator. iDA is an Excel add-on, so the user interface is Microsoft Excel. It uses first three rows of a spreadsheet to store the information about individual attributes. In this way, it states if the attribute has categorical or numerical value, if it should be used as input in model building or as an output attribute. There is also a possibility to declare certain attributes as unused or display-only, when they would not be used for building a model. Each column in MS Excel spreadsheet can represent an individual attribute.

The essential limitation of commercial version of iDA is that it can work with a single MS Excel spreadsheet, which allows maximum of 65536 rows and 256 columns. The version of iDA, which we have used, has even greater limitation regarding the dataset size – no more than 7000 data instances can be mined with this tool. The maximum size of an attribute name or value stored in one cell is 250 characters. The last limitation is that RuleMaker in iDA will not generate rules if the number of derived classes exceeds 20.

An exemplar-based data mining tool (ESX), which builds a concept hierarchy to generalize data, can, as stated in Roiger- Geatz (2003), "help create target data, find irregularities in data, perform data mining, and offer insight into the practical value of discovered knowledge". ESX will not make statistical assumptions about the nature of mined data. Furthermore, it can emphasize certain inconsistencies and unusual values in dataset. If ESX is performing supervised classification, it can provide information about those instances and attributes which could classify in the best fashion new instances of unknown origin. When performing unsupervised clustering, ESX incorporates a globally optimizing evaluation function that encourages a best instance clustering. In contrary to DataEngine, iDA can work with both categorical and numerical data values.

Waikato Environment for Knowledge Analysis - Weka is suite of Java class libraries and it implements many acknowledged machine learning and data mining algorithms. In contrary to DE and iDA, algorithms in Weka can be applied either directly to a dataset or can be called from Java code. It contains tools for preprocessing, classification, regression, clustering, association rules and visualization. It is also suited for developing new machine learning schemas.

Pros for using Weka tool are the following: it covers the entire machine learning process, it facilitates comparison of the results of different algorithms implemented, it accepts one of the most widely used data formats as input – ARFF

format, there are flexible APIs for programmers, and customization possibilities. Weka has also some deficiencies: it requires Java Virtual Machine to be installed for its execution, and visualization of mining results is not possible.

Weka tool implements clustering methods as C-Means, EM, Cobweb, X-means, FarthestFirst, and others. We decided to use simple k-Means algorithm as it is one of the oldest and most widely used clustering algorithms. We decided to use the simple c-means clustering algorithm, as it is available in all three selected data mining tools and is generally in wide use.

4. Data understanding

The goal of our research was to discover knowledge hidden in small and medium sized enterprises' (SMEs) data, by means of intelligent data analysis and in that way to support the development of this sector. The SMEs data were provided by four Regional Agencies for the Development of Small and Medium Sized Enterprises and Entrepreneurship from province of Vojvodina. The data was collected in 2006. by means of the questionnaire these Agencies provided.

The questions in the questionnaire were divided into two groups. The first group aimed to collect general enterprise data. The second group of data was formed by answers of individual enterprises to the questions related to business itself, technical, technological and financial aspects, market conditions and distribution, administrative and legislative conditions, human resources, business connectivity, and the need for non-financial services.

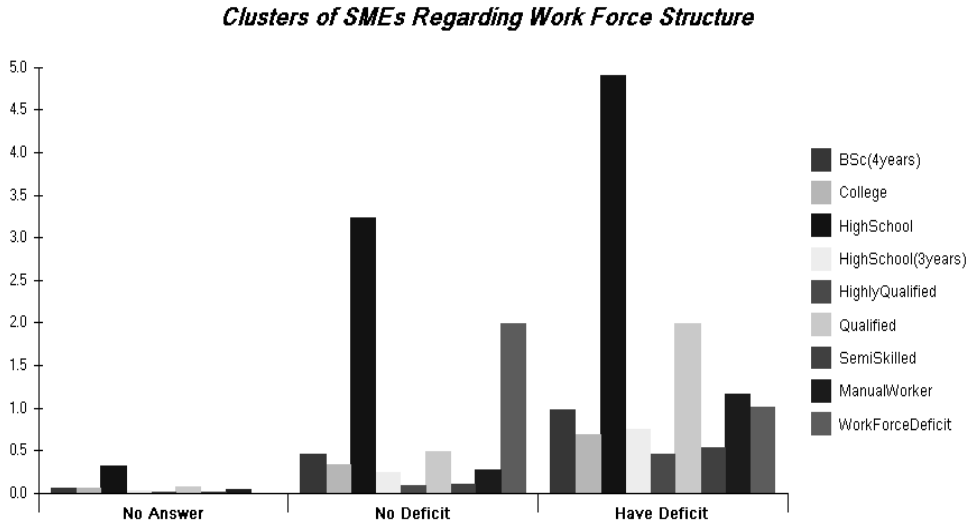
The final data collection consists of 2365 records on SMEs in the province of Vojvodina. Each data record is described with more than one hundred attributes. The data was originally stored in MS Access format and contained many missing data. Therefore, there was a need for qualitative data transformation into a format required by each data analysis tool we used in our research. Also, in the data preprocessing phase, many of initial attributes were removed from further analysis (data preprocessing is described in more detail (Grljević- Bošnjak 2008).

The resulting set of data was divided into subsets, and different tools, data mining methods and techniques were used for their analysis. In this paper we presented the data analysis results, using different clustering techniques. At this point, it is essential to emphasize the fact that the quality of collected data was poor and that we faced many challenges during the data mining. Consequently, there are some limitations in applicability of revealed knowledge (these challenges and limitations are described in more detail (Bošnjak et al. 2009).

5. Data analysis

The analyses we have carried out, and the results presented in this article, refer to human resources data. They consist of data concerning an employee's qualification structure (B.Sc level, College, High School – 4 years, High School - 3years, Highly Qualified, Qualified, Semi-qualified, Without qualification) and the deficit of adequate work force, experienced by SMEs (facing/not facing such a deficit). Firstly we developed a clustering model which divides SMEs according to the structure of work force. DE tool offers a possibility of cluster analysis, where we used the partitioning coefficient as a validity measure to determine the best number of clusters. For the same purpose, other two validity criteria can be used: proportion exponent and classification entropy. These are three known criterions by which fuzzy clustering can be judged. Also, they can be presented in a form of graph. We inspected these criteria, putting in relationship the partitioning coefficient (pc) and the classification entropy (ce). As stated in (MIT GmbH, 1997) both of these validity criteria tend towards monotone behavior depending on the number of clusters. Therefore, to determine the optimal number of clusters (c) we had to look for the number of clusters at which these values have a kink, so called "elbow criterion" (the best number of clusters according to this criteria was 3). DE offers an additional functionality, called cluster analysis that automatically determines the optimal cluster number. According to cluster analysis, the optimal number of clusters was also 3. Combining these two possibilities, we found that the best partitioning of the SMEs data regarding work force structure and work force deficit an enterprise is facing with, is achieved with three clusters. Fig. 1 represents these SMEs groups in DE.

Figure 1. DataEngine clustering model regarding the work force deficit



Source: Own creation

Analyzing the resemblance score, as a main indicator of successfulness of clustering process in iDA tool, and the goodness of the model developed, we revealed that usage of clustering technique in iDA to partition the same data, resulted in the same optimal cluster number. Furthermore, we used this knowledge to set the initial number of clusters in Weka. Although, each tool uses different clustering algorithm and therefore a structure of obtained clusters is also different, we were able to gain additional information analyzing the defined clusters in each tool. This information was very valuable for better understanding of SMEs structure regarding the work force and problems enterprises cope with, and for determining other relationships.

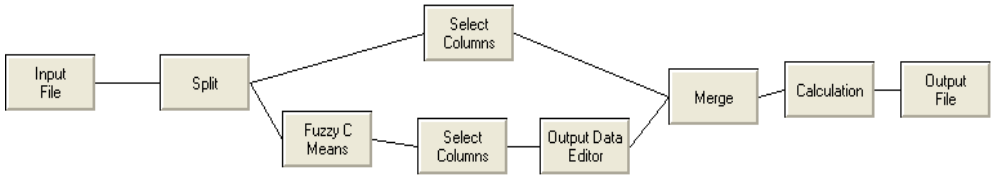
Out of those enterprises that cooperated and gave answers to the questions, majority stated that they do not have a work force deficit. Also many SMEs did not provide an exact number of employees distributed amongst different levels of qualification. Given the previously stated fact, the generated cluster with all ranks ranging around zero (Fig. 1) is fully justified. Each tool we have used generated one such cluster and two others. One of them consists of SMEs facing work force deficit, while the other contains SMEs that declared they do not have such a deficit (rank 1 and 2 on Fig. 1, respectively).

Additional information gained in iDA, that were not available in other two tools, refer to typical representatives of each cluster. iDA provides the list of all

instances belonging to one cluster ranked from the most typical for that cluster to the least typical, with associated typicality scores. The most typical representatives of cluster denoted NoDeficit are those enterprises that employ one employee with BSc level (four years study) of qualification, one employee with academic level of education and five or six employees with high school level of qualification. The most typical representatives of cluster HaveDeficit are those enterprises that employ one employee with BSc (four years study), and one with academic level of education, two to five employees with high school, and one highly qualified, and one qualified worker. We could conclude that, according to the most typical representatives of each cluster, a relationship between employee's level of education and deficit of qualified workforce is not relevant.

DE offers the possibility to automate data processing steps using graphical macro commands which are in a form of function blocks. Function blocks are placed on a card that allows their easy configuration and connection. Each card should contain at least one input and one output in the form of data or a graph. The processing steps are placed between input and output. We used this DE tool to inspect in more detail the structure of expressed deficit related to the level of qualifications. Fig. 2 shows a card that classifies all data according to the determined optimal cluster number, as we previously concluded it as three. After the correct classification of input data, this card will allow us to select several columns which will serve as input to macro block called calculation. This macro block provides the

Figure 2. Automated selection of SMEs with work force deficit



Source: Own creation

code that will select only those SMEs that surely have work force deficit, and data of such SMEs, together with their industry code, will be placed into in beforehand prepared data file.

The goal was to inspect in more detail the work force deficit across different industry branches. The knowledge discovered revealed that 108 SMEs were facing work force deficit. Additional analysis of these data led us to the conclusion that SMEs come from very wide range of industry sectors, 60 different ones in total. The

final conclusion we have driven was that the observed deficit was not sector specific.

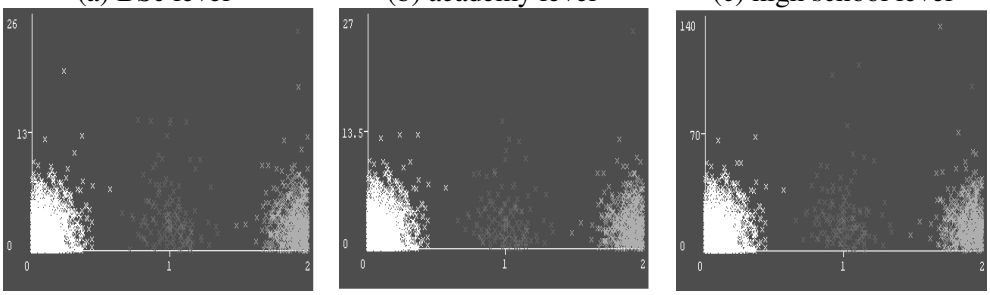
After such a conclusion, we were peculiarly interested in relationships between work force deficit and different towns or municipalities in province of Vojvodina. Therefore, we did some further analysis. The presented card was modified in such a way that additional information was added to the already extracted and grouped ones. This information refers to the zip codes of towns where SMEs facing deficit located their businesses. In this way we gained better insight into desired relationships. The results show that majority of SMEs having work force deficit are from West-Bačka and South-Bačka administrative districts, and mainly from Novi Sad and Sombor municipalities. Therefore, we could conclude that, as we have suspected, the region and the deficit of work force were related.

As Weka offers some irreplaceable visualization functionalities, we used them to inspect further human resources aspect of SMEs. Figure 3 consists of three plots marked with (a), (b) and (c), displaying a frequency of occurrence of answers within clusters. Such presentation of cluster structure was not possible in other two tools. X axis represents a work force deficit, while Y axis represents the number of employees. This is displayed for 3 different levels of qualification: bachelor, academic and high school level of education as three highest levels. Other levels of qualification are omitted from this figure. The dots on graphs (a), (b) and (c) around (0,0) are those employees that answered with HaveDeficit, the dots in the middle of each graph represents those employees that answered with NoDeficit, and the cluster of dots on the right hand side represents employees that have not provided any answer. As can be seen in Fig. 3, SMEs that have fewer employees commonly face the deficit of adequate work force, since marks are concentrated around zero. The same trend can be observed when we visualize other levels of qualification with remarked deficit.

Generally speaking, the structure of work force in SMEs in province of Vojvodina, according to the level of qualification, is devastating. Probably one of the main reasons for this is the fact that SMEs are mainly family businesses that are inherited or developed irrespectively to the level of qualification of their owner(s) or employees. This is what led SMEs to a situation where most of directors are highly uneducated people (table 1).

Figure 3. Weka Cluster Visualization

(a) BSc level (b) academy level (c) high school level



Source: Own creation

Table 1. Level of SMEs directors' education

Director – education	
PhD	45
MSc	37
BSc (4 years)	346
Academy	237
High school (4 years)	77
High school (3 years)	809
Primary school	146
Other	17
No answer	628

Source: Own creation

Table 2 illustrates a number of employees according to their qualification structure. It shows that SMEs in province of Vojvodina mainly employ workers with high school level of education.

SMEs in province of Vojvodina are facing many different problems in everyday business, such as lack of available funds, complex administrative and legislative regulations, disharmony with standards, insufficient market information, insufficient information on technologies. Results of data analysis about problems SMEs cope with, innovations they have conducted in the previous two years, over - aging of fixed assets, percentage of capacity utilization, and ownership structure are presented in more details (Grljević- Bošnjak 2009).

Table 2. Employees qualification structure

Number of employees according to qualification structure	
PhD	679
Academy	526
High school	3962
High school (3 years)	550
Highly qualified	238
Qualified	1257
Semi-qualified	299
Unqualified	717

Source: Own creation

6. Conclusion

In this article, we described the application of different clustering techniques of small and medium sized enterprises, which could support the development of SMEs sector. The results presented refer to work force data, and unavailability of qualified work force, and human resources development.

During the research and analysis of data we concluded that a composite approach to data analysis process, that implies diversity of tools could not help in achieving each and every data mining goal. Despite this fact, we managed to take advantage of the utilization of three tools – DE, iDA, and Weka, when clustering tasks are in question. Each tool has added additional information to the previously discovered knowledge. We presented these results in short in this paper, while additional results are available (Grljević- Bošnjak 2009).

The conducted analysis and the results presented in this paper are merely the starting point for further analysis. Additional analyses are needed to reveal, if possible, an exact reason of work force deficit. Agencies for the Development of Small and Medium Sized Enterprises and Entrepreneurship could use the results we described in this article, to create more successful employment policies and to maintain balance between supply and demand on work force market.

References

Binu, T. – Raju, G. – Sonam, W. 2009: A Modified Fuzzy C-Means Algorithm for Natural Data Exploration. *Proceedings of world academy of science, engineering and technology,* Volume 37, January 2009, pp. 478-481.

Bošnjak Z. – Grljević, O. – Bošnjak, S. 2009: *CRISP-DM as a Framework for Discovering Knowledge in Small and Medium Sized Enterprises*. Scientific Bulletin of „Politehnica" University of Timisoara, Romania, Transactions on Automatic Control and Computer Science, Vol 10. Fasc. /2010.

Bošnjak, Z.- Bošnjak, S.- Stojković, M. 2005: *Application of Fuzzy Clustering for Searching Trends in Data – The Public Transport Company in Subotica Case Study.* Proceedings of EUROFUSE 2005, 15th -18th Jun, Belgrade, Serbia.

Bratko, I. - Kubat, M. – Michalski, R.S. 1998: *Machine Learning and Data Mining: Methods and Application*. John Wiley & Sons Inc, New York.

Cahlink, G. 2000: *Data Mining Taps and Trends*, Government Executive Magazine, http://www.povexec.com/tech/articles/1000managetech.html, Oct. 1, 2000.

MIT GmbH, 1997: *DataEngine Tutorials and Theory*, Aachen, Germany.

MIT GmbH, 2008: *DataEngine – User Guide*, Aachen, Germany, available at http://www.dataengine.de/

Grljević, O. – Bošnjak, Z. 2008: *Primena CRISP-DM Metodologije u Analizi Podataka o Malim i Srednjim Preduzećima (CRISP-DM Methodology Utilization in Preprocessing Small and Medium Sized Enterprises Data).* Book of proceedings of XXXV symposium on OR, SYM-OP-IS, Septembar, Soko Banja, Serbia, pp. 275-279.

Grljević, O. - Bošnjak, Z. 2009: *Combining different Clustering Techniques for Improved Knowledge Discovery.* Central European Conference on Information and Intelligent Systems, CECIIS 2009, Zagreb, Croatia.

Harrison, P.G. – Llado, C.M. 2000: *Performance Evaluation of a Distributed Enterprise Data Mining System Source*. Lecture Notes In: Computer Science, Springer-Verlag, London, Vol. 1786, pp. 117-131.

iData Analyzer, available at http://www.infoacumen.com

Jiawei, H. –Kamber, M. 2001: *Data Mining Concepts and Techniques*. Morgan Kaufman Publishers, San Francisco.

Komem, J. – Schneider, M. 2005: *DataEngine Tools for Inteligent Data Analysis and Contro.,* Data Mining and Knowledge Discovery Handbook, Springer Science+Business Media Inc., pp. 1371-1377.

Liao, T.W. – Triantaphyllou, E. 2008: *Recent Advances in Data Mining of Enterprises Data: Algorithms and Applications*. Series on Computer and Operations Research, Vol. 6.

Nemati, H.R. – Barko, C.D. 2003: *Organizational Data Mining: Leveraging Enterprise Data Resources for Optimal Performance*. Idea Group Inc (IGI).

Roiger, J.R. –Geatz, M.W. 2003: *Data Mining: A Tutorial – Based Primer*. Addison Wesley, USA.

Tan, P.-N. - Steinbach, M. – Kumar, V. 2006: *Introduction to data mining*, Addison Wesley, USA.

Waikato Environment for Knowledge Analysis. available at http://www.cs.waikato.ac.nz/ml/weka/

Witten, I.H. – Frank, E. 2005: *Data Mining: Practical Machine Learning Tools and Techniques.* Elsevier Inc., San Fransisco.

Some Empirical Evidences about the Biphasic Action of e-Government Processes

Massimo Pollifroni[1]

The aim of the paper is to investigate - by a Business Economics approach - about the potential correlation between two clusters (or variables): innovation and ethical behaviours related to the life standards in a country. The first cluster (innovation) includes Information Communication Technologies (ICT), Research & Development Expenditure, Education Investment, (etc.); while the second one (ethical behaviours) contains elements such as ethical values, the observance of the law, merit rating system, (etc.).

Inside the first cluster are located the e-government processes that the paper systematizes in the Introduction. The central part of the paper shows the potential correlation between the two clusters by an empirical research concerning the European Union (EU) countries area.

The final pages of the paper are dedicated to comment the research result that shows the biphasic action of e-government processes: on the one hand these processes represent a right way to introduce efficiency and effectiveness in the public sector management (short period), on the other hand e-government applications can have a useful effect on the ethical shared behaviours (long period).

Keywords: Business Economics, e-Government, Ethics, Innovation

1. Introduction

The current feeling of mistrust towards the system-company can be attributed in part to problems related to so called "lacks of a business ethic approach". In particular, the behaviour of some companies, in current models of government to denature the target given by: "(…) *the production of goods or services to economic conditions, conditions for which the company has durability as a function of capacity, its existence would consolidate over time, the instrumental function which characterizes the company as part of the fulfilment of human needs* (…)" (Ferrero 1965).

The business, therefore, seems - today as yesterday - to direct its attention towards a long-term and strategic approach, making converge values and principles

[1] Massimo Pollifroni, Associate Professor, University of Turin, Faculty of Economics, Department of Business Administration (Turin).

of an ethical nature. The concept of Government itself and the methods used by public companies, presents a number of critical issues (Savoie 1995), such as:

1. the risk of "*self-reference*" of the model and the related accounting records, reporting or programmatic reports produced;
2. the current "*imbalance*" in terms of informative spaces available (representing, for example, the availability of documents on web sites), reserved to the budgets of sustainability (such as the social, the environmental, etc.) compared to traditional financial statements (such as, for example, the budget, the budget report, etc.);
3. the weakness of actions on the concept of "*sustainability*": at the local side is referred to the "*social and environmental sustainability*" (strongly emphasized in the financial sustainability, it as has been told) and on the "*financial*" (poorly reflected in traditional financial statements), both attributable to a single issue: that of "*sustainable development*" of a community guided by a model reference value of improved and oriented to the ethical business;
4. (etc.).

Inside the Italian system, for example, we tried to remedy in time to these critical issues through a cyclical and often massive regulatory intervention that, in fact, led to inefficient outcomes, or in some cases, insufficient demand, which then arises spontaneously from the reasoning, whether there are other ways to improve these imbalances, as an alternative to that offered by the pure regulatory action (Puddu 2008).

The alternative way, to be followed for the improvement of the ethical model, could be represented by innovation, a cluster that includes Information Communication Technologies (ICT), Research & Development Expenditure, Education Investment, (etc.): inside the public sector management the technological component of innovation gives rise to the related concepts of e-government and e-governance (or e-democracy).

The concept of e-government (or e-administration) is referred to the use of modern Information and Communication Technologies (ICT) linked to the development of electronics and the Internet in the modernisation process of the Public Administration (Pollifroni 2005). The different processes of e-government may be analysed with reference to the various models, that the Public Institution may adopt during the modernisation process of the structure. The different e-government models are:

1. G2C model (Government to Citizen model): this model concerns the activities carried out by the Public Institution towards citizens (for example to build Institutional Portal Web and to provide Internet on line services such as the presentation of the Individual Tax Return in electronic format, or the application of electronic documents by the Registry Offices, etc.).

2. G2B model (Government to Business model): this model concerns the activities carried out by the Public Institution towards business companies (for example to provide Internet on line services such as the presentation – in electronic format - of the following documents: Income Tax Return, Annual Report, etc).
3. B2G model (Business to Government model): this model concerns the activities carried out by the Public Institution towards external supplier (for example e-procurement activities, e-auctions on line, etc.; in Italy these activities are made by Consip S.p.A., a Public Company of the Italian Treasury Department).
4. G2E model (Government to Employees model): this model concerns the activities carried out by the Public Institution towards employees (for example to provide Internet on line services such as e-learning activities).
5. G2G model (Government to Government model): this model concerns the activities carried out by the Public Institution towards other Domestic Public Institution (electronic integration between several Departments or between Central and Local Public Institution) or towards other International or Foreign Public Institutions (for example intelligence activities, International Co-operation actions, etc.).

The development of the e-government processes (conditioning processes or causes) determines an improvement in the governance processes of the Public Institution that - using highly technological solutions - are named e-governance processes (conditioned processes or effects) (Haque 2001; Osborne et al. 1992).

Consequently, the e-governance is the second aspect of technological innovation applied to Public Administration processes: that is to say the possibilities for the improvement of the democratic participation processes offered by the new technologies (Pollifroni 2005). The digital revolution multiplies the individual's possibilities of communication and interaction in an exponential fashion, making it possible to re-launch the classic idea of the individual at the centre of the "*Res Publica*" (Kettl 2000). These e-governance processes [also called digital democracy (or e-Democracy)] include, for example:

1. direct participation of the employees to the internal decision of the Public Institution: these processes influence the internal governance with activities, e.g., of internal electronic poll, also called e-Decision;
2. direct participation of the citizens to the political choices: these processes influence the external governance of the Public Institution by e-Voting activities.

The innovation activity can be defined as scientific, technological, organizational, financial or commercial implementation of processes. Innovation is a key word now in the economic system and in the social world: innovate is often synonymous with success, to survive in the market and to gain competitive advantage (business sector). By investing in this process, the government shows more and more interest

towards the Information Technology and management techniques that can ensure an higher level of control, efficiency and quality of services, compared with increasingly rigid spending constraints and changes in the economy growing very fast, more and more demanding requirements from the user-citizen (public sector) (Bajjaly 1998; Werlin 2003).

Governments and Public Institutions in recent years have been directed toward more sophisticated methods, which provide tools for increasingly complex data analysis and stringent reporting capabilities and more sophisticated, bringing out in this way also for the public sector, by applying tools able to support the strategic control and decision making process (Brown et. al. 1998).

The development of such technology called e-government, passes thought the process of computerization of public administration and together with shares of organizational change, can handle the documentation and helping to manage processes with digital systems through the use of information and communication technologies (ICT) (Cantino 2005). This will optimize the work of the organizations and provide users (citizens and businesses) faster and new services, using, for example, the websites of the authorities concerned. Following this address, national and regional governments of the most technologically advanced countries (North America, Japan, European Area) have initiated strategic plans to guide the transition and to accelerate steps to force the diffusion of ICT in the public sector.

In recent years, in addition to the implementation and development of technological innovation, has developed a parallel process of attention to ethics, as a related discipline (Maesschalck 2004); some studies have sought to show how innovation is able to influence the ethical behaviour, triggering a virtuous circle, (such as the city satisfied with the service), to monitor and encourage ethical behaviours in several fields, such as: tax evasion control, observance of the law, reengineering a public merit rating system, (etc.) (Neilsen 1995).

The aim of this research is, therefore, found in the verification of the existence of a possible correlation between the indicators that measure the level of innovation (independent variable) and ethical behaviours (dependent variable): while these two seemingly independent aspects, if the outcome would be positive, it could be said that innovation in the public management can be a valuable tool for improving its ethical model.

2. Research Methodology

2.1. Path research of structural indicators

To achieve the goal described above, two baskets of indicators have been identified:

1. the first basket (basket of innovation indexes) is the Summary Innovation Index (SII), that is an arithmetic weighted average of 33 innovation indexes (data sources: European Commission/Eurostat);
2. the second basket (basket of ethical indexes) includes the following seven ethical indexes: 1) AEI Standard Ethics (data source: Agenzia Europea di Investimenti Standard Ethics); 2) Corruption Perception Index (CPI) (data source: Transparency International); 3) Control of corruption (data source: World Bank); 4)Voice and accountability(data source: World Bank); 5) Government effectiveness (data source: World Bank); 6) Political stability and absence of violence (data source: World Bank); and 7) Regulatory quality (data source: World Bank).

Each index has presented the following characteristics:

1. availability for the period 2003-2007;
2. applicability to almost all of the 27 European Union countries;
3. representativeness of the country;
4. possibility of comparison between them.

The research of the indicators was carried out by consulting the data sources offered by the following international bodies: European Commission, Eurostat, Transparency International, AEI (Agenzia Europea di Investimenti) Standard Ethics and World Bank. The paper continues with a brief presentation of the indices identified.

2.2. Presentation of the basket of innovation indexes

The basket of innovation indexes includes the Summary Innovation Index (SII), that is an arithmetic weighted average of 33 innovation indexes (data sources: European Commission/Eurostat). The indicator is composed of a basket of sub-indicators that vary over time.

This composite index measures the "innovation performance" through three innovation inputs [A1) drivers of innovation, A2) creation of new knowledge, A3) innovation and entrepreneurship] and two innovation outputs [B1) applications, B2) intellectual property]: the sub-indicators considered for the purposes of this study have the characteristics specified below.

A1) Drivers of innovation (7 indexes).

1. Graduates in science and engineering per 1,000 population (age group 20-29 years) - S & E graduates (% of population aged 20-29): this indicator brings together university graduates in science, physics, mathematics, statistics, computer science, engineering, architecture with the population under study, between 20 and 29 years (included).
2. Population with tertiary education in the field (age 25-64) - Population with tertiary education (% of population aged 25-64): this indicator brings

together the number of people in age group 25-64 formed for the tertiary sector, with the entire population in that range of reference.

3. Rate of broadband penetration (number of broadband lines per 100 inhabitants) - Broadband penetration rate (number of broadband lines per 100 population): this indicator brings together the number of broadband lines with the total population.
4. Participation in a long training period (age 25-64) - Participation in life-long learning (% of population aged 25-64): this indicator brings together the people taking part in a formation of long-term with the entire population within the age group 25-64.
5. Level of education achieved at a young age (% of population aged 20-24 years who have completed university) - Youth education attainment level (% of population aged 20-24 having completed at least upper secondary education): this indicator brings together people aged between 20 and 24 years who have completed university, with the entire population in that age range.
6. Internet Access or domestic - Level of Internet access of households: it indicates the ratio between the number of homes with Internet access and the total case.
7. Share or SMEs with a website - Level of Internet access of enterprises: it indicates the ratio between the number of SMEs with a website and the total number of SMEs.

A2) Creation of new knowledge (6 indexes).

1. Public expenditure on research and development (% of GDP) - Public R & D expenditures (% of GDP): this indicator has been extrapolated from the Eurostat database and shows the expenditure on research and the development level as a percentage of total GDP of each country of the European Union.
2. Private expenditure on research and development (% of GDP) - Business R & D expenditures (% of GDP): this indicator brings together all the expenditure in R & D performed by private sector (industry and services), with the GDP.
3. Share of R & D in medium-high and high technology (% of expenditure in R & D in Industry) - Share of medium-high-tech and high-tech R & D (% of manufacturing R & D expenditures): this indicator brings together the expenditure in R & D for high-and medium-high technology industry, with total spending on industrial R & D.
4. Proportion of firms that receive public funds for innovation - Share of enterprises receiving public funding for innovation: this indicator brings together a number of innovative firms that receive public funds, with the total number of firms.

5. University R & D financed by the private sector - University R & D expenditures financed by business sector: this indicator brings together the expenditure in R & D in universities, with total expenditure in R & D university, highlighting the degree of cooperation between public and private.
6. Share of venture capital investments in High-tech venture capital (% of venture capital invested): this indicator brings together the investment of venture capital in high-tech, with total investments of venture capital. Investment of venture capital in high-tech refers to the following areas: computer science, electronics, biotechnology, medicine, industrial automation and financial services.

A3) Innovation and entrepreneurship (6 indexes).

1. Industrial products and services, created in SMEs (% product and service): this indicator is the sum of all products / services created by SMEs in innovation activities (for businesses to innovate means both producing knowledge by them self, or producing it by collaborating with other firms), with the total number of products / services generated by SMEs.
2. Proportion of Early-stage venture capital (% of GDP): this indicator measures the dynamism in creating new business.
3. SMEs innovating in cooperation (% product and service): this indicator measures the flow of knowledge and between enterprises and between public research and enterprises.
4. Expenditure on innovation - Innovation expenditures (% of turnover): this indicator links total expenditure on innovation by all firms producing goods or providing services, with the total turnover generated from goods / services.
5. ICT expenditure (% GDP) - ICT expenditures (% of GDP): this indicator links the total expenditure in Information and Communication Technology (ICT), with the GDP.
6. Share of SMEs that do not change on a technical level - SMEs using non-technological change (% of SMEs) : this indicator considers the companies that do not implement technical improvements, new facilities and do not change the design of at least one product.

B1) Applications (7 indexes).

1. Employees in high-tech services (% of the workforce) - Employment in high - tech services (% of total workforce): this indicator brings together people working in areas of high-tech services (post and telecommunications, information technology including the development of software and services for R & D), with the total workforce in all industries and services.

2. Employed in the production of high-or medium-high technological content (% Labour Force) - Employment in medium/ high and high - tech manufacturing (% of total workforce): this indicator brings together the number of employees in the production of products of high or medium-high technological content (chemical, machinery, office equipment, telecommunications, precision instruments, automobiles, aerospace and other transport equipments) with the total workforce.
3. Exports of high technology products as a share of total exports: this indicator measures the competitiveness of the European Union in commercialising the results of research and development and innovations on international markets.
4. Sales of new products (% of sales) - Sales on new market products (% of turnover): this indicator brings together the revenue generated from the sales of new or improved products, with the total turnover.
5. Sales of new products for the firm, but not new to the market (% of turnover): this indicator brings together the revenue generated from new products considered by some businesses but not regarded as such by all the companies on the market, compared with the total turnover.
6. Value-added in high-tech manufacturing (% of manufacturing value-added) : this indicator brings together the value added industrial production in five high-tech sectors (pharmacy, office equipment, telecommunications equipment, aerospace), with the total value added of the manufacturing sector.
7. SMEs Rate of volatility (sum of birth rate and death rate): this indicator links the rate of volatility, with the total number of SMEs; the rate of volatility interprets business dynamism and the contribution given to increase productivity. A high degree of volatility indicates a capability to adapt to changes.

B2) Intellectual property (7 indexes).

- European habitants: this indicator brings together the number of high-tech patents validated by the European Patent Office, with the total population.
- American habitants. (New) USPTO high- tech patents: this indicator is the U.S. equivalent, of the above described for Europe.
- EPO patents: this indicator brings together the number of patents approved by the European Patent Office (EPO) with the total population.
- USPTO patents per million Americans: this indicator brings together the number of patents approved by the U.S. Patent Office (USPTO) with the total population.
- New Triadic patent families per million population: this indicator brings together the number of patents of the “triad”, with the total population. A patent is the triad if and only if it was lodged with the European Patent Office (EPO), the Japanese Patent Office (JPO) and the U.S. Patent and Trademark Office (USPTO).

- Number new domestic community trademarks (CTM) per million population: this indicator brings together the number of new trade marks, with the total population.
- Number of (new) domestic community industrial designs per million population: this indicator brings together the new design community, with the total population.

2.3. Presentation of the basket of ethical indexes

The second basket (basket of ethical indexes) includes the following seven ethical indexes:

- AEI Standard Ethics (data source: Agenzia Europea di Investimenti Standard Ethics). Evaluations in terms of ethical Rating (national or regional) have as a reference the concept of Ethics and Social Responsibility issued according to parameters set by international bodies like the UN, OECD and the European Union. The final evaluations of the EEA Ethics Standards are expressed in the form of a rating to eight levels (EEE, EEE-, EE+, EE, EE-, E+, E, E-). The rating is the result of statistical and scientific activity carried out with the intention of photographing the world of business in relation to ethical principles promoted by large international organizations.
- Corruption Perception Index (CPI) (data source: Transparency International). The index of perceptions of corruption in English Corruption Perception Index (CPI) is an indicator published annually since 1995 by Transparency International ordering the countries of the world on the basis of the level that the existence of corruption is perceived among public and political office.
- Control of corruption (data source: World Bank). The indicator provided by the World Bank measures the ability of the political, legal and judicial systems to prevent and combat corruption.
- Voice and accountability (data source: World Bank). This index provided by the World Bank measures the degree of civil liberties and political rights and influence of the effective population in the election of political leaders, so far, to the level of independence of the media from political pressure.
- Government effectiveness (data source: World Bank). The indicator published by the World Bank that measures the quality of public services, the credibility of the Government on the measures to be implemented, the quality of the bureaucracy and the independence of civil servants from political pressure.
- Political stability and absence of violence (data source: World Bank). The index published by the World Bank, which measures the perceptions of the likelihood that destabilize the government or be removed by unconstitutional or violent means, including domestic violence and terrorism.
- Regulatory quality (data source: World Bank). Indicator published by the World Bank, which measures the ability of the government in formulating and

implementing policies that can enable and promote the development of the private sector.

3. Data Standardization and presentation of correlation data-results

In order to compare these indexes, their values have been standardized, and traced back to a single scale in terms of cents: the process used is explained below.
Innovation Indicators. Summary Innovation Index (SII) Standardization was obtained by multiplying by 100 the original data, according to the following proportion:

Since the original: Given standardized (x) = 1:100.

Ethics Indicators.

1. AEI Standard Ethics. Cents in the conversion of this quality indicator is obtained through the following conversion scale: EEE=100; EEE-= 85.71428571; EE + =71.42857143; EE=57.14285714; EE-=42.85714286; E +=28.57142857; E=14.28571429 and E-=0.
2. Corruption Perception Index (CPI). The indicator in question is represented by a scale from 0 to 10, its conversion into cents was realized through the following proportion: since the original: Given standardized (x) = 10:100.
3. Control of corruption. 4) Voice and accountability. 5) Government effectiveness. 6) Political stability and Absence of Violence. 7) Regulatory quality. The five indicators of the World Bank are expressed on a scale whose values range from -2.5 to +2.5. Cents in the conversion has been obtained through the following conversion scale: since normalized (x) = (as original + 2.5) * 20.

For the purposes of this work, the calculation of the correlation was obtained by the following indicators:

- the independent variable “Innovation”: the indicator is calculated as a result of several sub-indicators and corresponds to the Summary Innovation Index;
- the dependent variable “Ethics”: the data used is the value that results from the average of the basket composed of the seven indicators described above;
- the values that derives from the process of normalization of the original data bases.

The following tables and charts show, year by year, the results of the research.

Table 1. Calculation of the correlation between "Innovation" (independent variable: x) and "Ethics"(dependent variable: y) – Year: 2003

Nations	**x**	**y**	**(x – mx)**	**(y – my)**	**$(x – mx)^2$**	**$(y – my)^2$**	**$(x – mx)^2$ $(y – my)^2$**
Austria	47,00	81,67	8,48	9,46	71,94	89,42	80,20
Belgium	51,00	79,73	12,48	7,51	155,79	56,45	93,78
Bulgaria	20,00	51,29	-18,52	-20,92	342,94	437,77	387,46
Cyprus	29,00	67,90	-9,52	-4,32	90,60	18,63	41,09
Denmark	68,00	89,57	29,48	17,36	869,16	301,20	511,65
Estonia	35,00	69,27	-3,52	-2,95	12,38	8,70	10,38
Finland	69,00	89,67	30,48	17,46	929,12	304,73	532,10
France	48,00	74,56	9,48	2,34	89,90	5,49	22,21
Germany	59,00	79,33	20,48	7,11	419,49	50,60	145,70
Greece	26,00	63,43	-12,52	-8,78	156,71	77,16	109,97
Ireland	50,00	79,84	11,48	7,63	131,82	58,18	87,58
Italy	32,00	65,29	-6,52	-6,93	42,49	47,99	45,15
Leetonia	16,00	61,47	-22,52	-10,75	507,08	115,56	242,07
Latvia	23,00	64,53	-15,52	-7,68	240,82	59,03	119,23
Luxemburg	50,00	84,15	11,48	11,93	131,82	142,33	136,97
Malta	27,00	75,48	-11,52	3,26	132,68	10,65	-37,59
Netherlands	50,00	85,07	11,48	12,86	131,82	165,29	147,61
Poland	21,00	58,36	-17,52	-13,85	306,90	191,92	242,69
Portugal	21,00	73,23	-17,52	1,02	306,90	1,03	-17,79
United Kingdom	57,00	81,76	18,48	9,54	341,57	91,05	176,35
Czech Republic	32,00	63,49	-6,52	-8,73	42,49	76,16	56,89
Romania	16,00	45,98	-22,52	-26,24	507,08	688,39	590,82
Slovakia	23,00	60,22	-15,52	-12,00	240,82	143,91	186,17
Slovenia	32,00	68,43	-6,52	-3,78	42,49	14,31	24,66
Spain	32,00	75,27	-6,52	3,06	42,49	9,34	-19,92
Sweden	82,00	88,97	43,48	16,76	1890,64	280,73	728,53
Hungary	24,00	66,49	-14,52	-5,73	210,79	32,80	83,15
European Average	38,52	72,22	====	====	310,69	128,85	175,08
Correlation Index				0,88			

Source: own creation

Table 2. Calculation of the correlation between "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2004

Nations	**x**	**y**	**(x – mx)**	**(y – my)**	**$(x – mx)^2$**	**$(y – my)^2$**	**$(x – mx)^2$ $(y – my)^2$**
Austria	46,00	82,36	7,63	10,46	58,21	109,43	79,81
Belgium	49,00	78,59	10,63	6,69	112,99	44,75	71,11
Bulgaria	21,00	52,01	-17,37	-19,89	301,73	395,60	345,49
Cyprus	29,00	65,40	-9,37	-6,50	87,80	42,22	60,89
Denmark	66,00	90,14	27,63	18,25	763,40	332,89	504,11
Estonia	34,00	69,77	-4,37	-2,13	19,10	4,54	9,31
Finland	68,00	89,76	29,63	17,86	877,91	319,01	529,21
France	48,00	75,33	9,63	3,43	92,73	11,78	33,05
Germany	59,00	79,73	20,63	7,83	425,58	61,35	161,58
Greece	26,00	62,80	-12,37	-9,09	153,03	82,70	112,50
Ireland	49,00	79,53	10,63	7,63	112,99	58,25	81,13
Italy	33,00	63,58	-5,37	-8,32	28,84	69,27	44,70
Leetonia	16,00	60,20	-22,37	-11,70	500,43	136,83	261,68
Latvia	24,00	63,30	-14,37	-8,60	206,51	73,92	123,55
Luxemburg	50,00	84,09	11,63	12,19	135,25	148,64	141,78
Malta	27,00	73,63	-11,37	1,74	129,29	3,01	-19,74
Netherlands	49,00	84,93	10,63	13,03	112,99	169,84	138,53
Poland	21,00	56,42	-17,37	-15,48	301,73	239,56	268,85
Portugal	24,00	71,75	-14,37	-0,15	206,51	0,02	2,17
United Kingdom	57,00	82,22	18,63	10,32	347,06	106,46	192,22
Czech Republic	33,00	62,72	-5,37	-9,18	28,84	84,27	49,30
Romania	15,00	46,55	-23,37	-25,35	546,17	642,47	592,37
Slovakia	22,00	60,73	-16,37	-11,16	267,99	124,62	182,75
Slovenia	34,00	68,40	-4,37	-3,50	19,10	12,23	15,29
Spain	31,00	74,27	-7,37	2,38	54,32	5,64	-17,51
Sweden	80,00	88,97	41,63	17,07	1733,03	291,51	710,78
Hungary	25,00	66,00	-13,37	-5,89	178,77	34,74	78,81
European Average	38,37	71,90	====	====	288,97	133,54	176,06
Correlation Index				0,90			

Source: own creation

Table 3. Calculation of the correlation between "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2005

Nations	**x**	**y**	**(x – mx)**	**(y – my)**	**$(x – mx)^2$**	**$(y – my)^2$**	**$(x – mx)^2 (y – my)^2$**
Austria	48,00	82,10	9,44	10,96	89,20	120,06	103,48
Belgium	49,00	77,33	10,44	6,19	109,09	38,26	64,61
Bulgaria	20,00	52,09	-18,56	-19,05	344,31	362,93	353,50
Cyprus	30,00	66,07	-8,56	-5,08	73,20	25,78	43,44
Denmark	65,00	88,80	26,44	17,66	699,31	311,72	466,89
Estonia	35,00	69,50	-3,56	-1,64	12,64	2,70	5,85
Finland	65,00	88,87	26,44	17,73	699,31	314,30	468,82
France	48,00	75,44	9,44	4,30	89,20	18,49	40,61
Germany	59,00	80,22	20,44	9,07	417,98	82,29	185,46
Greece	26,00	61,98	-12,56	-9,17	157,64	84,08	115,13
Ireland	50,00	80,50	11,44	9,36	130,98	87,55	107,09
Italy	33,00	60,13	-5,56	-11,01	30,86	121,22	61,17
Leetonia	17,00	60,77	-21,56	-10,38	464,64	107,70	223,70
Latvia	24,00	63,43	-14,56	-7,71	211,86	59,46	112,24
Luxemburg	53,00	82,63	14,44	11,49	208,64	131,97	165,93
Malta	28,00	71,77	-10,56	0,62	111,42	0,39	-6,57
Netherlands	49,00	83,64	10,44	12,50	109,09	156,25	130,55
Poland	22,00	54,69	-16,56	-16,45	274,09	270,63	272,35
Portugal	23,00	71,60	-15,56	0,46	241,98	0,21	-7,14
United Kingdom	56,00	80,27	17,44	9,13	304,31	83,33	159,24
Czech Republic	33,00	61,96	-5,56	-9,18	30,86	84,30	51,01
Romania	16,00	47,07	-22,56	-24,08	508,75	579,82	543,12
Slovakia	23,00	61,89	-15,56	-9,25	241,98	85,62	143,93
Slovenia	34,00	67,67	-4,56	-3,48	20,75	12,09	15,84
Spain	32,00	73,67	-6,56	2,53	42,98	6,39	-16,58
Sweden	78,00	87,03	39,44	15,88	1555,86	252,31	626,54
Hungary	25,00	63,82	-13,56	-7,32	183,75	53,65	99,29
European Average	38,56	71,14	====	====	272,77	127,91	167,76
Correlation Index						0,90	

Source: own creation

Table 4. Calculation of the correlation between "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2006

Nations	**x**	**y**	**(x – mx)**	**(y – my)**	**$(x – mx)^2$**	**$(y – my)^2$**	**$(x – mx)^2$ $(y – my)^2$**
Austria	48,00	82,24	8,85	11,03	78,36	121,74	97,67
Belgium	48,00	77,67	8,85	6,46	78,36	41,76	57,20
Bulgaria	22,00	52,09	-17,15	-19,12	294,06	365,46	327,82
Cyprus	32,00	67,40	-7,15	-3,81	51,10	14,52	27,24
Denmark	64,00	89,37	24,85	18,16	617,61	329,82	451,33
Estonia	37,00	70,73	-2,15	-0,48	4,61	0,23	1,03
Finland	67,00	88,79	27,85	17,58	775,73	308,94	489,54
France	48,00	74,84	8,85	3,63	78,36	13,20	32,17
Germany	59,00	80,24	19,85	9,03	394,10	81,61	179,34
Greece	25,00	61,52	-14,15	-9,69	200,17	93,95	137,13
Ireland	49,00	80,42	9,85	9,21	97,06	84,74	90,69
Italy	33,00	58,75	-6,15	-12,46	37,80	155,30	76,62
Leetonia	18,00	62,40	-21,15	-8,81	447,24	77,63	186,33
Latvia	26,00	62,47	-13,15	-8,74	172,87	76,45	114,97
Luxemburg	57,00	82,86	17,85	11,65	318,69	135,73	207,98
Malta	29,00	72,57	-10,15	1,36	102,98	1,84	-13,76
Netherlands	48,00	83,27	8,85	12,06	78,36	145,50	106,77
Poland	23,00	54,21	-16,15	-17,00	260,76	289,09	274,56
Portugal	25,00	70,00	-14,15	-1,21	200,17	1,46	17,08
United Kingdom	55,00	82,04	15,85	10,83	251,28	117,37	171,74
Czech Republic	34,00	62,88	-5,15	-8,33	26,50	69,45	42,90
Romania	17,00	48,67	-22,15	-22,55	490,54	508,30	499,34
Slovakia	24,00	61,55	-15,15	-9,66	229,47	93,35	146,36
Slovenia	36,00	68,97	-3,15	-2,24	9,91	5,03	7,06
Spain	32,00	70,36	-7,15	-0,85	51,10	0,73	6,09
Sweden	76,00	87,17	36,85	15,96	1358,06	254,75	588,19
Hungary	25,00	64,02	-14,15	-7,19	200,17	51,70	101,73
European Average	39,15	71,21	====	====	255,76	127,39	163,89
Correlation Index				0,91			

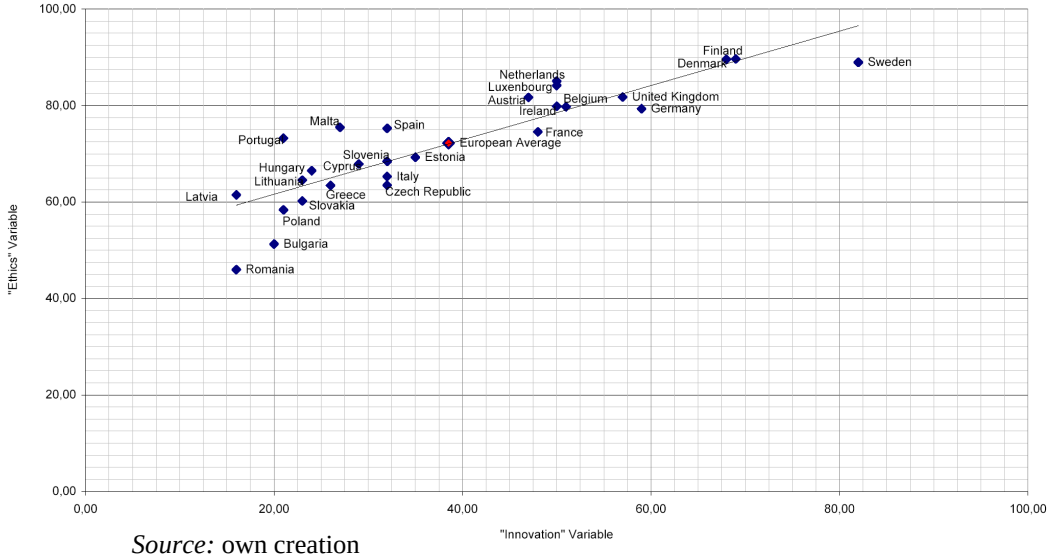
Source: own creation

Table 5. Calculation of the correlation between "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2007

Nations		**x**	**y**	**(x – mx)**	**(y – my)**	**$(x – mx)^2$**	**$(y – my)^2$**	**$(x – mx)^2 (y – my)^2$**
Austria	48,00	82,36	9,00	11,01	81,00	121,15	99,06	
Belgium	47,00	77,27	8,00	5,92	64,00	35,06	47,37	
Bulgaria	23,00	52,15	-16,00	-19,20	256,00	368,69	307,22	
Cyprus	33,00	67,23	-6,00	-4,12	36,00	16,96	24,71	
Denmark	61,00	89,34	22,00	17,99	484,00	323,67	395,80	
Estonia	37,00	70,37	-2,00	-0,99	4,00	0,97	1,97	
Finland	64,00	87,44	25,00	16,09	625,00	258,96	402,31	
France	47,00	74,24	8,00	2,89	64,00	8,37	23,14	
Germany	59,00	80,04	20,00	8,69	400,00	75,56	173,85	
Greece	26,00	61,12	-13,00	-10,23	169,00	104,74	133,04	
Ireland	49,00	80,99	10,00	9,64	100,00	92,84	96,35	
Italy	33,00	60,31	-6,00	-11,05	36,00	122,02	66,28	
Leetonia	19,00	58,67	-20,00	-12,69	400,00	160,94	253,72	
Latvia	27,00	62,37	-12,00	-8,99	144,00	80,73	107,82	
Luxemburg	53,00	83,69	14,00	12,34	196,00	152,21	172,72	
Malta	29,00	72,27	-10,00	0,91	100,00	0,84	-9,15	
Netherlands	48,00	84,22	9,00	12,86	81,00	165,48	115,77	
Poland	24,00	59,40	-15,00	-11,95	225,00	142,76	179,23	
Portugal	25,00	69,75	-14,00	-1,61	196,00	2,58	22,48	
United Kingdom	57,00	81,27	18,00	9,92	324,00	98,42	178,58	
Czech Republic	36,00	62,79	-3,00	-8,56	9,00	73,28	25,68	
Romania	18,00	49,58	-21,00	-21,77	441,00	474,05	457,23	
Slovakia	25,00	62,16	-14,00	-9,19	196,00	84,44	128,65	
Slovenia	35,00	68,93	-4,00	-2,42	16,00	5,85	9,67	
Spain	31,00	70,10	-8,00	-1,25	64,00	1,56	10,00	
Sweden	73,00	88,43	34,00	17,08	1156,00	291,61	580,61	
Hungary	26,00	62,99	-13,00	-8,36	169,00	69,90	108,69	
European Average	39,00	71,35	====	====	223,56	123,47	152,33	
Correlation Index				0,92				

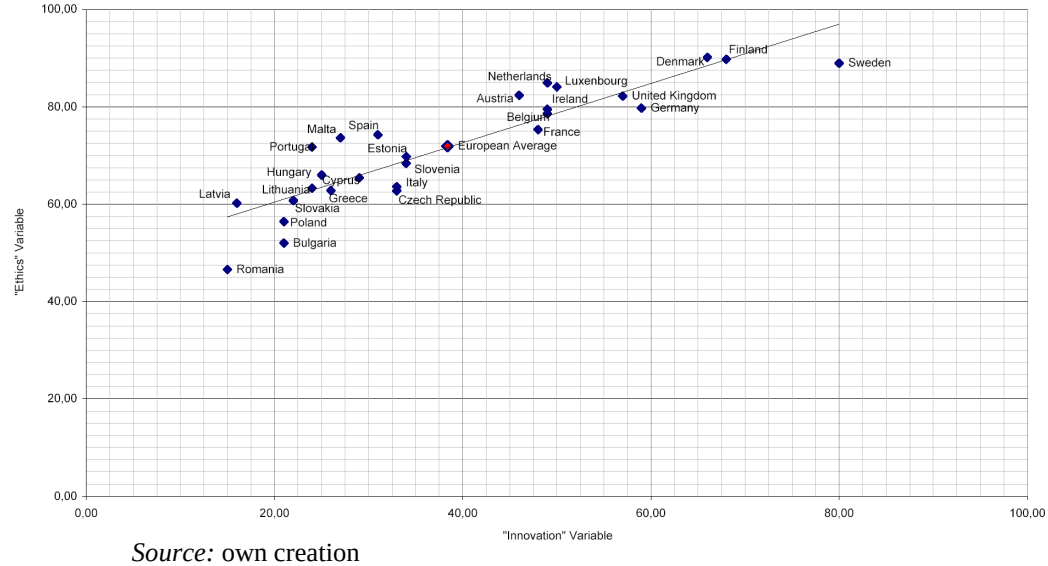
Source: own creation

Figure 1. Scatter chart and trendline concerning the two variables "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2003



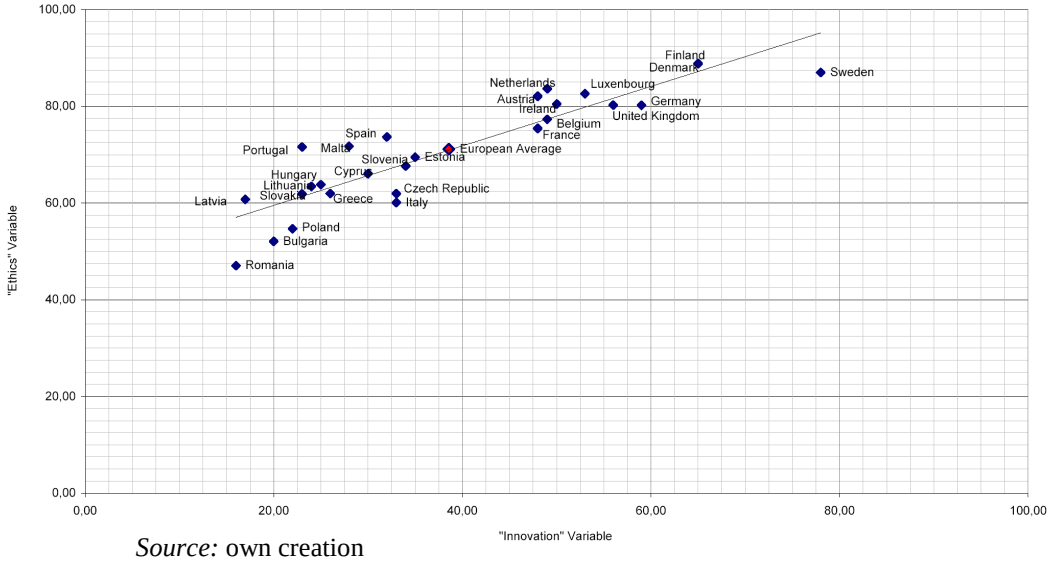
Source: own creation

Figure 2. Scatter chart and trendline concerning the two variables "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2004



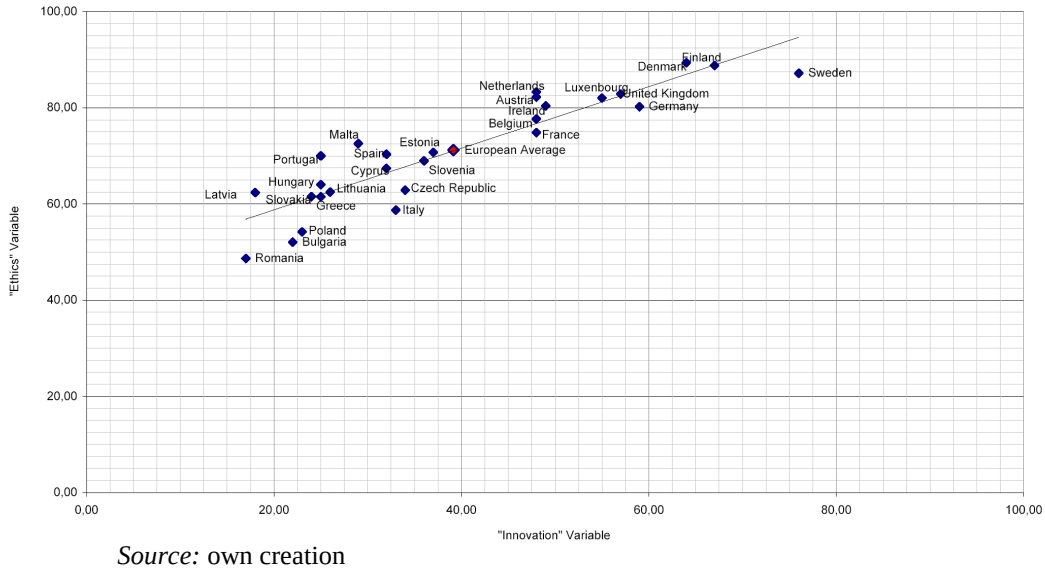
Source: own creation

Figure 3. Scatter chart and trendline concerning the two variables "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2005



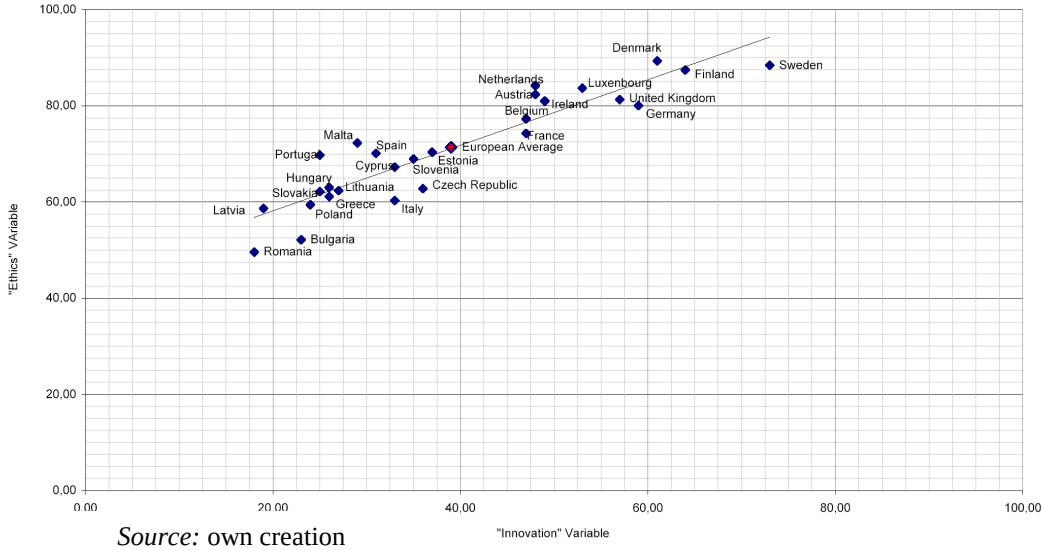
Source: own creation

Figure 4. Scatter chart and trendline concerning the two variables "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2006



Source: own creation

Figure 5. Scatter chart and trendline concerning the two variables "Innovation" (independent variable: x) and "Ethics" (dependent variable: y) – Year: 2007



Source: own creation

4. Conclusion

The contribution of this research has had, as prerequisite, the identification in the current processes for improvement and development of models of government of the crucial role represented by the share of the underlying reference model value, measured by ethical parameters.

In the model, the issue of governance and their criticality, has been pressing an action that often, as we have already registered, leading to inefficient results, or in some cases, insufficient demand, born spontaneously the reasoning above is whether there are other ways in addition to that legislation, the improvement of these imbalances: the alternative way (followed in this study) was designed to measure the level of innovation, cluster where the e-government processes are located.

According to the empirical evidence outlined above, taking place within the community, could be a positive measure between the two variables: innovation and ethical behaviours of a country (which ranges from 0.88 and 0.92 for the period 2003-2007).

The data-results show that in countries where the economic system is more oriented towards innovative practices (e.g., Sweden, Finland and Denmark), there are also the highest ethical standards. In conclusion, therefore, it is possible to state that the implementation of the component of innovation (a cluster that includes

Information Communication Technologies (ICT), Research & Development Expenditure, Education Investment, etc.) is one way to improve the ethical behaviour of a country, consequently overcoming the limitations and weaknesses produced by the mere regulation.

In conclusion the research results could shows the biphasic action of e-government processes:

- on the one hand these processes represent a right way to introduce efficiency and effectiveness in the public sector management (short period analysis);
- on the other hand e-government applications can have a useful effect on the ethical shared behaviours, such as tax evasion control, observance of the law, reengineering a public merit rating system, (etc.) (long period analysis).

References

Bajjaly, S. T. 1998: Strategic information systems planning in the public sector. *American Review of Public Administration*, 28, 1, 75-85. p.

Brown, M. – Brudney, J. L. - Maureen M. 1998: Public sector information technology initiatives. *Ad-ministration & Society*, Vol. 30, 4, 421-442. p.

Cantino, V. 2005: *Management Information System*. McGraw-Hill, Milan.

Ferrero, G. 1965: *Le determinazioni economico-quantitative d'azienda. Introduzione: le determinazioni conoscitive di ricerca aziendale nei caratteri connaturali al loro oggetto*. Giuffrè, Milano. 88-90. p.

Haque, A. 2001: GIS Public service, and the issue of democratic governance. *Public Administration Review*, 61, 259-266. p.

Kettl, D. 2000: *The global public management revolution: A report on the transformation of governance*. Brookings, Washington, DC.

Maesschalck, J. 2004: The impact of new public management reforms on public servants ethics: Toward a theory. *Public Administration: An International Quarterly*, Vol. 82, 465-489. p.

Neilsen, R. P. 1995: Alternative Managerial Responses to Ethical Dilemmas. *Planning Review*, Vol. 43, 24-29. p.

Osborne, D. – Gaebler, T. 1992: *Reinventing government*. Addison-Wesley, Reading, MA.

Pollifroni, M. 2005: The 'S-EPI MODEL: a Theoretical Model that links the E-Government processes to the Instruments of the Public Sector Social Responsibility. *Proceeding of the International Workshop e_Government'05 (e_GOV'05)*, Brunel University, 13 September 2005, West London (UK). 12-13. p.

Puddu, L. 2008: *Bilancio IPSAS: Sistema Integrato di Rilevazioni e Principi Contabili Pubblici*. RIREA, Rome.

Savoie, D. J. 1995: What is wrong with the new public management. *Canadian Public Administration*, Vol. 38, 112-121. p.

Werlin, H. H. 2003: Poor nations, rich nations: A theory of governance. *Public Administration Review*, Vol. 63, 329-342. p.

Sampling techniques for sampling units with different size

Mónika Galambosné Tiszberger[1]

Usually it is a not too difficult problem to select efficient sample from a population, which includes units with different size. Stratified sampling might be a proper solution for this matter. The situation is getting more complicated if the statistician has to observe more characteristics of the unit, and these characteristics behave in various ways. Homogeneous strata cannot be created from every point of view. The field of my concrete research is the livestock surveys in Hungary.

If we "only" keep in mind the official requirements of the European Union, data in 6 categories of different livestock have to be provided by given reliability level. It means in practice, that one unique sample has to be worked out to suit these requirements. In Hungary there are more than 600 thousand private holdings, of which more than 50 percent raise usually more kinds of livestock. There are very small households, producing mainly for their own consumptions, and there are also huge units, which produce one company scale. The deviation of the indicators is extremely high in most of the cases.

In my research I attempt to see what would be the optimal solution from every point of view through working out different sampling schemes (simple random, stratified, and concentrated, mixture of these). In the presentation I would like to show the results and the final outcome of the work.

Keywords: agriculture, sampling techniques, livestock survey

1. Introduction

Statistical data concerning the livestock have been collected in Hungary since more than one hundred and fifty years. At the beginning the livestock was surveyed concurrently with the census of the population. From 1884 individual livestock surveys were carried out, while from 1895 livestock has always been part of the recurring agricultural censuses. (Laczka 2000) Since 1957 representative surveys of private farms have been conducted whereas all agricultural enterprises have always reported on their livestock. At the beginning the frequency of surveys was three monthly; until 2008 they were conducted in every four months (1 April, 1 August, 1 December); currently twice in a year (1 June, 1 December). The surveys cover all kind of animals. Full-scope observation applies in case of the agricultural enterprises; on the other hand there is sample survey for private holdings and households. After 1986 the breakdown of main animal species including the cattle stock by breed was surveyed following the Agricultural Census in 2000 (AC 2000).

[1] Mónika Galambosné Tiszberger, PhD student, University of Pécs, Faculty of Economics, Doctoral School of Regional Policy and Economics (Pécs)

In the international literature (to leave out of consideration the general literature dealing with statistics and sampling theory), in the topic of agriculture livestock is not in the focus of the interest. Crop statistics and area frame sampling in connection with soil and crops are the fields, which are worked out. Livestock data in most of the developed countries are coming from administrative data, so the way of collection and the sampling procedure is out of scope.

Before looking up details about the sampling techniques, it is necessary to summarize, what is the population that needs to be surveyed in case of agricultural statistics. There are two main groups involved in agricultural activity:

Agricultural enterprises: Every enterprise engaged in any agricultural activity, regardless its size. It is a business unit with or without legal entity excluding private entrepreneurs and private holdings.

Private holding: A technically and economically stand-alone production unit involved in agricultural activity, or holdings operated by private entrepreneurs, that used

- productive land (arable land, kitchen garden, orchards, vineyard, grassland, forest, reed, fishpond) of at least 1500 m^2 area, or
- orchard or vineyard of at least 500 m^2 area, or
- 100 m^2 land area under cover, or
- 50 m^2 of mushroom area during the reference year, or
- had a livestock consisting at least of
- one large animal (such as cattle, pig, horse, sheep, goat, buffalo), or
- 50 heads of poultry (such as hens, geese, ducks, turkeys, guinea fowls), or
- 25 heads each of rabbits, furred animals, pigeons, or
- 5 bee colonies

on the reference date of the survey.

The share of agricultural enterprises and private holdings in the agricultural value produced by the country is about equal, 50-50 percent. However the distribution of their number is not so balanced. There are about 7 700 active agricultural enterprises and more than 618 000 private holdings according to the latest Farm Structure Survey (FSS) in 2007. It means, that in number the share is 1-99 percent. Obviously the group of private holdings can be observed only through appropriately selected samples. As it is, I will deal only with private holdings throughout this article.

The aim of this paper is to present the beginning of a way, which tries to create an effective sample method to fulfil the EU requirements for the livestock surveys. The whole research is based on the data collected in FSS 2007. I have the opportunity to use the database at the regional office of the Hungarian Central Statistical Office (HCSO) in Pécs. As an important topic, the recent sampling techniques used by the HCSO will be introduced. A chapter will show the nature and characteristics of the Hungarian animal husbandry and the specialties of the different species. Then I have to summarize the requirements for the accuracy of the livestock data, which is determined by EU legislation, and which will be the bottleneck of the re-

search. I will present the analyses concerning the results of using simple random sampling by taking different aspects of the population, and the first trials on stratification. The work is not finished yet, so only partial results are indicated.

2. Livestock of Hungary

In Hungary the two main livestock types are cattle and pigs. Poultry is also important, but from the point of view of this paper it is not included in the analyses. In Table 1 I present the main data on livestock from the year 2007. In case of cattle and pigs, most of the livestock is kept by enterprises, which specialized for animal husbandry, and have more heads in average, kept by more efficiency. Sheep and goats are in the hand of private holdings. Sheep shows a very impressive average stock, but later on we will see, that the distribution of the livestock is not too fortunate from the point of view of sampling. Goats are in an even worse situation, as the average stock is quite small. Both of these two latter types show high deviations, as we will see in Chapter 7.

Table 1. Main data of livestock in Hungary, 2007

Livestock	Country livestock total, heads	Livestock of agricultural enterprises, heads (percentage of total)	Livestock of private holdings, heads (percentage of total)	Livestock keeper private holdings, holdings	Average livestock in private holdings, heads
Cattle	705 077	485 250 (69)	219 827(31)	18 907	11,63
Pig	3 871 147	2 603 958 (67)	1 267 189(33)	281 930	4,49
Sheep	1 232 005	172 660 (14)	1 059 345(86)	21 468	49,35
Goat	67 271	2 872 (4)	64 399(96)	15 380	4,19

Source: FSS 2007 and own calculation

I would like to give a picture about the structure of the Hungarian livestock not only by its size, but the value they represent. It is obvious that livestock heads cannot be added up. One might use the national livestock unit – which is an equivalent of the total livestock used for aggregation of various species of different genders, ages, equal to one or more animals of 500 kg live weight – to compare the amount of animals. But as the base of this indicator is the weight of the animals, from economic aspect it is rather meaningless. The best indicator would be the standard gross margin[2] (SGM), as it is an indicator, worked out by Eurostat, and used by every member state, but unfortunately, at present time I have no details about this at

[2] The SGM is equal to the unit production value of products and services net of variable costs.

the level of livestock, only at the level of the holdings. Gross production value[3] of agricultural products is the indicator of HCSO to measure agricultural value. According to the available data of 2000, the composition of the 4 types this article works with shows that pigs have a share of more than 50 percent. Cattle follow them, which represents about 1/3 of the value. Sheep and goat together gives only a bit more than 10 percent of the value produced by these 4 types on animals. These facts underline the original statement at the beginning of this chapter, so the two most import livestock are cattle and pigs.

It is also important from the point of view of sampling to see the distribution of livestock holdings and of the livestock itself of the different animal species.

- 92 percent of the pig keeper holdings have less than 11 pigs, and they give almost half of the livestock. It means that if we would be able to cover the remaining 8 percent – which represents not too many – of the holders, it would give the other half of the livestock. The skewness of the distribution is notable.
- In case of cattle, holdings having 10 or less animals constitute almost 80 percent of the cattle keepers, but have a share of the livestock of only 28 percent. The distribution in the higher sections shows various pictures.
- Sheep stock represents more concentrated production. 15 percent of the biggest holders (having more than 50 heads) give more than 80 percent of sheep. Unfortunately the remaining holdings are very diverse from the aspect of the size of their livestock. The distribution of the livestock is skewed, but the opposite way as we saw in case of pigs.
- Goats are the most special ones. The mode of the held heads is 2 and 3. It means that statistician cannot really gain from the observation of bigger keepers, because there are only a few of them, and they represent just a few percent of the total production of goats.

Altogether the distribution of the different species shows different nature. It would cause problems during stratification and in the situation where I would like to create a combined sampling plan, which works for every species.

3. Sampling method today

Regular surveys mean those implemented every year (survey on sown area, annual production, crops, etc.) or even several times within a year (livestock). Surveyors

[3] Value of agricultural products produced in the framework of agricultural production in a certain time period, irrespective of whether those were produced in a unit in agricultural or in other branch. This value includes the value of two main branches of agricultural: crop production, animal husbandry. Gross production value of animal husbandry includes the production value of breeding (live born animals), value of livestock change and weight growth and values of products and by-products from animal husbandry. Gross production value is the sum value of total amount of produced products multiplied by average prices determined for each utilisation types.

visit the selected private holdings to fill in the questionnaires (face-to-face interviews). The sample frame is based on the AC 2000, and it is updated with the Census on vineyards and orchards' (2001) and the farm structure surveys' results. The frame population is divided into two groups according to the size of the holding. A smaller group, declared as large holdings, is selected on a full scope base, according to my initiatives, and they receive the questionnaire by mail. This preferential group is selected through natural figures like the size of livestock or land area. (Altogether about 1500 private holding belong to this group. Out of it 700 are livestock holdings.) The specific thresholds in the selection of this group were adjusted according to their share, paying attention to the financial possibilities. This kind of "take all" philosophy in case of large holdings results in more reliable figures, as a notable part of the production of private holdings is observed without sampling errors. Another advantage of my initiative was to introduce the data collection by mail to the respondents, and start a process to generalize and make this form of surveys acceptable among them.

The rest of the holdings (more than 900 thousand holdings as a frame population) form the base for sampling. A universal sample is used to keep the surveys' budget on a cost-effective level. It means that for all of the above-mentioned surveys only one sample is selected, and there are no special sample population for the different types of surveys.

A two-stage, concentrated, stratified sampling technique is applied. Organisational and financial reasons made it necessary to use concentration, which is theoretically not optimal for minimizing the sampling error with a given sample size. However, the face-to-face interviews are much more economical and faster if the surveyor has to visit holdings within a small district (a part of a settlement), instead of travelling kilometres to find the different data suppliers. It is also easier and more effective to organise and manage a smaller number of surveyors within the regions.

The sample selection implemented through the following steps: In the first stage every 9th survey district is selected randomly, stratified by county (on this NUTS III level there are 19 counties in Hungary). These districts are the primary sampling units (PSU). In these selected districts 2 strata are determined:

- Stratum "A": all holdings exceeding at least one of the following thresholds: 5 cattle, 10 pigs, 26 sheep, 100 chickens, 100 ducks, 100 turkeys, 26 geese, 25 bee colonies, 5 ha arable land, 1 ha vineyard or 1 ha orchard.
- Stratum "B": holdings not exceeding the thresholds mentioned above.

The secondary (of final) sampling units (SSU) are the private holdings within the selected PSUs. Every holding is selected in stratum "A", and randomly every 4th is in the sample in case of stratum "B". It results about 40 thousand holdings in the sample population. It means a sampling rate of about 4 percent. The size can be easily adjusted to financial possibilities or quality requirements by changing the sampling rate of either PSUs or SSUs in stratum "B". (Previously every 8th district was in the sample and the sampling rate was 33 percent in stratum "B".) New sample is

selected after the bigger surveys like censuses of farm structure surveys. It means that one particular private holding will be a data supplier for 3-4 years, and then the sample is refreshed.

The main problems with this sample design, as I see, are the following:

- Counties, or in the future regions as strata are necessary, because the main figures are published at county – in the future at regional – level, but obviously they do not form homogeneous groups. This aspect is an obligation (which is problematic, but it can not be "solved").
- The stratification on the second level does not serve the aim of building homogenous groups as well, because of the large number of variables used. Thus the variances of the different variables within the groups are still high.
- Concentration effects worse sampling errors with the same sample size compared to simple random sampling without concentration. However, I must admit, as long as there are face-to-face interviews, this part unfortunately cannot be modified in practice.
- The universality of the sample is the weakest point as I see. The effectiveness of the sampling cannot be sufficient if it has to cover so many topics and characteristics.

4. The requirements

After the coming census in 2010 new sample selection would be necessary. It would be a convenient solution to continue the method used in the previous 10 years. However, there is a new legislation, which declares clearly the allowed maximum relative standard errors for the main livestock types. This EU regulation[4] from the year 2008 requires the following relative standard error by 68 percent probability level for the country totals:

[4] Regulation (EC) No 1165/2008 of the European Parliament and of the Council of 19 November 2008 concerning livestock and meat statistics and repealing Council Directives 93/23/EEC, 93/24/EEC and 93/25/EEC.

Table 2. Maximum relative standard errors according to EU legislation

Livestock	Maximum relative standard error
Cattle	5%
Cow	5%
Pig	2%
Sheep	2%
Goat	5%

Source: Regulation (EC) No 1165/2008 of the European Parliament and of the Council

As sampling techniques are used only in case of private holdings, the relative standard error (coming from the nature of sampling) can be concerned as 0. It means, that by taking into account the distribution of the livestock totals per holding type, we can define the maximum relative standard errors for the livestock total of the private holdings (which will be obviously higher, less strict at the end). I have collected the necessary information for these calculations from the FSS 2007.

Table 3. Data of the main livestock, 2007

Livestock	Country livestock total	Livestock of agricultural enterprises	Livestock of private holdings	"Healing" weight of the livestock total of enterprises	Maximum relative standard error for the livestock total of private holdings
Cattle	705 077	485 250	219 827	0,312	**16,04%**
Cow	322 369	225 477	96 892	0,301	**16,64%**
Pig	3 871 147	2 603 958	1 267 189	0,327	**6,11%**
Sheep	1 232 005	172 660	1 059 345	0,860	**2,33%**
Goat	67 271	2 872	64 399	0,957	**5,22%**

Source: FSS 2007 and own calculation

The "healing" weight of the livestock total of agricultural enterprises comes from the following relationship (Ay 1976):

$$V = \sqrt{\frac{X_{ae}^2 \times 0 + X_{ph}^2 \times V_{ph}^2}{X^2}}$$

where:
V_{ph}: relative standard error of private holdings
V: required relative standard error
X_{ae}: total of agricultural enterprises' livestock
X_{ph}: total of private holdings' livestock
X: country total (livestock)

As the required error limits are given to the country totals of the main livestock types, the maximum relative standard error allowed for the livestock total of the private holdings can be calculated by rearranging the above-mentioned equation. (The results are in the last column of Table 3.)

$$V_{eg} = \sqrt{\frac{V^2 \times X^2}{X_{eg}^2}}$$

Apparently the estimations of the private holders' livestock have more latitude in those cases where the livestock is held mainly by the agricultural enterprises. The biggest challenge is stated for the number of sheep and goat. The modest size of the livestock kept by the enterprises does not really ease the strict requirements.

5. The base of the research

It would have been really nice to do every calculation and simulation on an up-to-date population of agricultural private holdings, but for the nature of statistical work, it is off course impossible. The latest full scope survey has been carried out in 2000.

Naturally the database of the registered 958 534 private holdings is available at HCSO, but the agricultural sector shows rapid changes in the last 10 years (35 percent of the private holdings had disappeared), this database is not good enough to simulate the current situation. On the ground of these reasons, I decided to be satisfied with a smaller part of the population, but with more recent data. The latest "bigger" survey has been carried out in 2007 (FSS 2007). The sampling rate was about 18 percent. Concentrated, one-stage sampling design had been worked out for this FSS. Every 6th survey district was selected by simple random within the regions (NUTS II level). During the survey, the enumerators looked up the whole district. It means, that they have asked every household about their agricultural activity, and not only those, who were on their list. It gave us the opportunity to collect information about the new agricultural holdings (in the selected districts) as well, and not only about those who were previously surveyed. In my research I use this database of FSS 2007, which includes almost 111 thousand private holdings. The size of the sample and the way of selection ensure that this sample population shows practically the same distributions according to the different variables as the frame population, even on regional level (NUTS II). To justify these statements, I investigated the distribution of the sample population by regions, and compared it to the official results of FSS 2007. The differences of the proportions are under 3.7 percentage points, and

the sample rate is also similar in the regions. According to these results, I believe, that general conclusions can be drawn[5] from the analyses of the sample population.

The final aim of the calculations is to work out an optimal sampling design. By optimal I mean to stay within the resent sample size (30-40 thousand private holdings), and ensure the fulfilment of the EU requirements about the maximum relative standard errors for this 5 specific livestock. As I work according to theoretical possibilities, I would handle the costs and management of the survey as secondary elements. During my work in HCSO I had to work within the financial and organisational frames, so I have never had the opportunity to try the statistical theory in practise. Naturally I hope, that HCSO would be able to build in my results into the sampling designs of the future.

6. Special features of livestock

In general the livestock holdings keep 1,86 animal types at one time. (It includes every type of livestock, in agricultural meaning.) It is good news from the point of view of sampling, because it means that the sum of the necessary sample size by animal species would give a good approximation, as most of the holdings keep only one of the 5 animals. It also predict, that stratification by animal types would be part of the optimal solution.

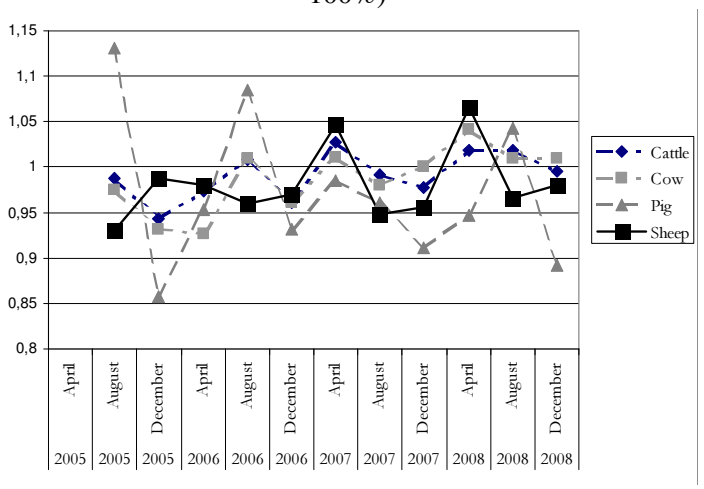
While we plan a sample design, there is another important aspect to take into account: how often does a livestock holding change its activity (change the livestock type, or close the business). Off course the stable livestock holdings would be ideal from the point of view of sampling. If they remain unchanged in time, the stratification and the sampling errors can be predicted in a more effective way. To analyse this question, it would be nice to see the whole life circle of at least some of the private holdings. Full scope information is available only from the year 2000. In the following years only samples had been surveyed. The FSS 2003 can be one more guideline because the holdings above 1 ESU[6] had been observed on full scope base. But altogether the behaviour of the holdings can be followed only in very few cases. The estimated results of my own calculations show, that between 2000 and 2007 the permanent livestock holdings give 70-90 percent of the total. In my understanding these numbers suggest that the livestock keepers are a quite stable group. It means that an optimal sampling design would be efficient in a longer period.

[5] In the formulas I use capital letters in most of the cases, because, as it is mentioned earlier, I regard the FSS 2007 as a whole frame population. So the calculated figures – mean, deviation, etc. – are understood for the entire population, and not as a sample variable.

[6] The economic size of the holding is determined on the basis of the total SGM value of products and activities of the holding and expressed in European Size Unit (ESU), where one ESU worth of SGM is equal to 1200 €.

Besides these aspects, there is one more factor that the statistician has to be aware of. Animal husbandry is a seasonal phenomenon. The size of the livestock is different in the to sampling period within the year. (As an illustration of the seasonal characteristic, I present Graph 1.) It is natural, but during the work we must pay attention to this fact.

Graph 1. Seasonal changes of livestock of private holdings (previous season = 100%)



Source: www.ksh.hu (Agricultural long time series and censuses)

7. Background and possible methods of sampling

There are basically two sources of the error in sample surveys. One of them is the random sampling error, which comes from the fact that only a part of the whole population is observed. The other source is the bias, which comes from the biased estimator. Sampling error cannot be avoided, but the bias can be overcome by using unbiased estimator. (unbiased = the estimator's expected value equals the variable that wished to be estimated). Simple random sampling has the advantage – contrary to the non-random samples – that the error limit of the estimations for the population can be calculated by exact methods from the sample itself. It can be measured by mathematical calculations and through the "arbitrary" increase of the sample size it might be decreased for the needed level.

The other group of errors is the branch of non-sampling errors. These may occur in different ways during an observation:

- Uncertain information of the populations (the level of coverage is under 100 percent).

- Respondents provide uncertain data about themselves.
- "Clerical error" during the survey, or the data entry.

It seems quite obvious, that using administrative data sources, undercoverage can be controlled and corrected in the registers. Unfortunately there are fields, where this kind of source is unavailable or includes false, not up to date information. In case of the other two types, the control rules, blind checks, educated enumerators might help to avoid the mistakes, but it would be unreasonable to expect 100 percent elimination of these biases. During the practical part of the surveys, these items must be kept in mind, but as such mistakes are hard to calculate, in my research I do not have the opportunity to cover them.

As I consider the sampling error as the strictest condition, I started my work by the analyses of the deviations of the different livestock species. Although the final aim is to estimate the sum of the values, to keep the calculations simple I will do the work first for the averages. From the point of view of relative deviation and the necessary sample size, the results would be the same, anyway.

I have used the formula of the simple arithmetic mean to calculate the average livestock of one holding:

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

where:
X_i: total of private holdings' livestock
N: population size (holdings)

The formula of the deviation:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} \left(X_i - \overline{X}\right)^2}{N}}$$

The formula of the relative deviation:

$$RD = \frac{\sigma}{\overline{X}}$$

These are simple formulas, however the calculations are a bit more complicated in the practise. The livestock table of HCSO database includes only those

holdings, which had kept livestock at the reference date of the survey (1st December 2007). It means, that additional calculations are necessary to reach the variables (average, deviation) for the entire population. The number of holdings not keeping animals is important, because they are part of the population (N), and their group will decrease the average, and most probably will increase the deviation. To make this point more visible, I created three different tables (Table 4, 5 and 6) with the same indicators in them. The three different tables concerns three different aspects of the population.

Table 4. Data of the population

Livestock	Average livestock of 1 holding	Deviation	Relative deviation	Number of holdings
Cattle	0,25	2,48	981,6%	110 949
Cow	0,11	1,22	1 138,4%	
Pig	2,04	8,57	420,5%	
Sheep	1,34	18,83	1 407,6%	
Goat	0,10	1,78	1 868,3%	

Source: own calculation

Table 5. Data of livestock holdings

Livestock	Average livestock of 1 holding	Deviation	Relative deviation	Number of holdings
Cattle	0,35	2,92	831,0%	79 837
Cow	0,15	1,44	964,2%	
Pig	2,83	9,99	352,8%	
Sheep	1,86	22,17	1 192,9%	
Goat	0,13	2,09	1 584,0%	

Source: own calculation

Table 6. Data of livestock holdings with the specific livestock species

Livestock	Average livestock of 1 holding	Deviation	Relative deviation	Number of holdings
Cattle	8,13	11,57	142,3%	3 447
Cow	4,89	6,78	138,6%	2 438
Pig	4,41	12,18	276,4%	51 293
Sheep	38,56	93,72	243,0%	3 848
Goat	3,87	10,67	275,6%	2 725

Source: own calculation

We can see, that the relative deviations are getting smaller by narrowing the aspect of the population. It is also visible, that the degree of the decrease is relatively small between Table 4 and 5, but we can get much more homogeneous groups by handling the livestock holdings having the specific livestock specie separately. Unfortunately the best results (Table 6) would be the less useable in practise. I would need one sample to estimate every species, and the handling of different populations within one process would be rather insolvable. The other problem is that there are overlapping among the groups.

7.1. Simple random sampling

If we know the deviation of the variables, we can easily calculate the necessary sample size to fulfil the requirements. I started the analyses by simple random sampling. Simple random samples require known probabilities (non zero) of selection for every element. In this case we get the necessary sample size by using the following formula[7]:

$$n = \frac{\sigma^2}{\Delta^2 + \frac{\sigma^2}{N}}$$

where

- n: sample size (holdings)
- Δ: accepted maximum error limit – $V_{ph} \times \overline{X}$ (heads)

I have calculated the necessary sample sizes for the different aspects of the population (entire, livestock holdings, livestock holdings with the specific livestock

[7] The probability level is 68%.

type). It seems quite obvious, as the relative deviations decrease thanks to the narrowed concept of the population that the necessary sample size would decrease as well. As I work with a research database, I think the sampling rate will illustrate the differences and the volume much better than the number of the holdings. In every case in the tables of the article I divide the necessary sample size by the original entire population (n/N), to make the comparison possible. The sampling rates can be seen in Table 7.

Table 7. Necessary sampling rate from the total population

Livestock	Total population	Livestock holdings	Livestock holdings with the specific livestock type
Cattle	3,27%	2,36%	0,07%
Cow	4,05%	2,94%	0,06%
Pig	4,09%	2,92%	1,77%
Sheep	76,69%	70,26%	2,56%
Goat	53,59%	45,35%	1,24%

Source: own calculation

In case of cattle and pigs, the sampling rates seem to be acceptable even at the first stage. However if we take a look at sheep and goats, we can see that only the last column of Table 7 gives us acceptable sampling rates. In the first and second case the sampling rates of these types are around of even above 50 percent. Such a sampling rate does not make any sense to use in a real survey. Hunyadi (2001a) introduces a similar example, where he looks for an optimal sampling design of the average production of 3 types of wheat. In the example we find that wheat "A" has a relatively small weight in the total production but with relatively high deviation. In this case, because of the heterogeneity of the produced amount, extremely high sampling rate would be "optimal". At the same time, as its proportion of the total production is quite small, Hunyadi suggests to be satisfied with smaller sample in these cases, and let the deviation (sampling error) remain high, to avoid the unnecessarily big sample size. In our case of livestock surveys, the translation of this example would mean that as in the Hungarian livestock sheep and goats are not so important, and their share is quite small (compared to cattle and pigs), and the livestock is very heterogeneous, we should be satisfied with higher sampling errors during the estimation of the total of these two types. As a result, we would be able to keep the sample size within reasonable limits. Sure enough, but unfortunately the EU legislation clearly defines the required maximum sampling errors, and do not take into account

the importance and proportion of the given livestock within the country's livestock production. So we must investigate other solutions.

7.2. Stratification

The next classical topic of sampling theory is stratification. Stratification has the precondition of knowing the entire population from the point of view of the stratification characteristic. It is also important, that every item of the population is classified into only one stratum, and every item can be classified. In an optimal case the best grouping variable would be the one that we are about to observe. Obviously from this variable we do not have complete information, as the final aim is just to collect it. (Cochran 1977, Kish 1995) In a general agricultural survey, which is multipurpose and the aim of the observations is not just a few indicators, it is a real problem to find good variables for stratification. (Kish 1989) In case of livestock surveys it is especially difficult to find a good variable, which will divide the entire population into distinct subpopulations. Holdings, who keep more types in parallel, might be part of more subpopulations if we use livestock type as the base of stratification. This is the reason why in practise we usually formulate more conditions at once to create distinct groups of smaller and larger holdings within the population. By using more variables for stratification there would be no overlapping of the subpopulations, but the original aim of stratification seems to be lost. Namely, we plan to use stratification to decrease the variances of the sample estimates. The original consideration standing behind stratification is to divide the heterogeneous population into homogeneous subpopulations. However if a holding would be large according to the number of its cattle, it is quite probably, that it would belong to the smaller holdings from the point of view of pigs or sheep. In practise it is general, that holdings are specialized in something, and not keep every type of livestock in big amount. So it results in high deviations for every variable, because the original aim of creating homogeneous groups cannot be reached. The deviation of the variables does not really decreased by this multivariable stratification. This is one part of the problems.

7.2.1. Current stratification

The stratification, which HCSO currently applies, is introduced in Chapter 3. To give the reader an idea of how does these aspects classify the population I present the distribution of holdings by the current stratification, giving extra information on livestock and land use aspects within the strata. Looking at Table 8 we see, that most of the holdings (86,9 percent) remained in one stratum, stratum "B", the smaller holdings. So we must conclude that homogeneous groups are not created. The share of the different aspects, namely livestock and land use, also confirms remained heterogeneity. Only 2,7 percent of the holding can be regarded as "large" holding by

both branches of agriculture. The rest 10,3 percent is specialized in one of them, so putting them into one stratum will not result in homogeneity.

Table 8. Distribution of holdings according to different aspects of stratification, 2007

Strata	„A" according to livestock	„B" according to livestock
„A" according to land use	2,7%	6,2%
„B" according to land use	4,1%	86,9%

Source: own calculation

If we look at the values behind the holdings (Table 9), we find that the produced standard gross margin shows a bit better distribution by stratum "A" and "B", but still, more than half of the production is kept in one group.

Table 9. Distribution of SGM according to different aspects of stratification, 2007

Strata	„A" according to livestock	„B" according to livestock
„A" according to land use	6,7%	9,6%
„B" according to land use	30,9%	52,9%

Source: own calculation

Another important point about the recent stratification method is its change in time. It was mentioned earlier, that the elements of the population could be divided into the subpopulations only if we know the value of their variables in advance. Let us take a look at the current stratification system in 2000 and 2007. From Table 10 we can conclude that the proportions of the strata has changed a lot in time, and holdings has moved from one stratum to the other in a lot of cases. It means, that the agricultural activity is still changing rapidly in Hungary and the classification of holdings by 7-year-old data results in misleading proportions. This aspect arises the thought of post stratification or two-phase sampling, as possible solutions in the future. *(The analyses of these two further methods are not part of this paper, but of later research.)*

Table 10. Change of stratification in time

	Based on AC 2000	**Based on FSS 2007**
Stratum	**Number of private holdings**	
"A"	29 805	17 547
"B"	60 073	101 038
Non-classified[8]	22 152	-

Source: own calculation

7.2.2. Stratification by regions

The next part is the required territorial breakdown of the provided data, which is very import for the domestic data users. At least regional (NUTS II) level data is required for the main livestock. It means for the statistician, that region would be a stratifying variable no matter of its efficiency in creating homogeneous groups. The 7 regions of Hungary are subpopulations as an obligation. The statistician must start the whole work within these frames. Additional stratification can be used, off course, but we have to balance the number of strata within reasonable limits. It would be inefficient to create too many subpopulations (including only a few holdings). As the stratification aspects are multiplied, we do not have the option to come out with too many ideas.

After computing the deviations for the different regions, I had the possibility to analyse the necessary sample sizes for a stratified sample. I did the calculations by conditioning simple random selection within the subpopulations (regions). I have used the following formula of relative sampling error[9] for stratification by proportionate allocation:

$$V_{RRA} = \sqrt{\sum_{j=1}^{7}\left(\frac{N_j}{N}\right)^2 \times \frac{\sigma_j^2}{n_j}\left(1-\frac{n_j}{N_j}\right)}$$

where[10]

- j: regions
-

It was not possible to rearrange the formula, and give an exact equation for the sample size in this case. So during the calculations I used iterations to see, what would be the necessary sample size. The results can be found in Table 11, where I

[8] Non-classified holdings did not exist in the year 2000.

[9] These general formulas of arithmetic mean, deviation, sampling error in case of simple random and stratified sampling can be find in several books, like: Kish (1995), Hunyadi-Vita (2003) or Pintér-Rappai (2007).

[10] In the formula I use the known deviation, as I work from the database of the entire population.

have included the former results to be able to compare the sampling rates. Unfortunately we find similar, or even higher percentages for the stratification by proportionate allocation as by the above mentioned simple random sample of the entire population. It means that the obligation of applying regions as strata did not increase the efficiency at all. Because of the heterogeneous regions, this territorial breakdown does not have good characteristics from the point of view of stratification. So, although it is not a surprise, this result is quite sad. We have already used partly one of our weapons (stratification) to make the sampling method more effective, but we gained basically nothing.

Table 11. Necessary sampling rates in different aspects

Livestock	**Population**	**Stratification by regions (proportionate)**
Cattle	3,27%	3,3%
Cow	4,05%	4,3%
Pig	4,09%	4,1%
Sheep	76,69%	77%
Goat	53,59%	54%

Source: own calculation

I did not give in at this point. I tried another type of allocation to see if it is going to show better results for the regional stratification. If the aim is to decrease the sample size and keeping the sampling error at the same level, Neyman optimal allocation by using the deviations might lead us for better solution. The essence of this type of allocation is to have higher sampling rate in the more heterogeneous regions to ensure more accuracy at the end. The following formula shows the allocation of the sample size:

$$n_j = n \times \frac{N_j \sigma_j}{\sum_{j=1}^{7} N_j \sigma_j}$$

The sampling error is the same, as it was in case of proportionate allocation. In this case I applied iterative approach, by changing the sample size (n) until I have reached the necessary sampling error.

Neyman optimal method gives better allocation design, if the deviation of the variables shows big variety among the regions. It is only true for sheep and goats. Regions have very similar deviations for pigs and cattle. Table 12 justify these facts. Compared to the proportionate allocation Neyman optimal allocation decreases the sampling rate in notable way only in case of sheep and goat. However from these

results we must conclude that Neyman optimal allocation overall is better than proportionate.

Table 12. Stratification by regions with different allocation strategies, necessary sampling rates

Livestock	Proportionate	Neyman optimal
Cattle	3,3%	3,2%
Cow	4,3%	4,2%
Pig	4,1%	3,7%
Sheep	77,0%	66,5%
Goat	54,0%	44,0%

Source: own calculation

8. Conclusions

From the article, it is quite clear that sampling of units with different size and nature is very difficult. The diverse distribution of the animal species makes it hard to harmonize the information into one sample. The high values of deviation require too high sampling rates to reach good quality estimations. The narrowed aspects of the population gave promising results. The handling of livestock holdings as the frame populations gives better starting as taking every agricultural private holding. The conditions of the work are also very strict. The necessary sampling errors are probably too small for sheep and goat, and the requirements of Eurostat do not take into account the importance of the specific livestock species in the different countries. The necessity of regional data also restricts the work of the statistician. Regions do not create homogeneous strata, but give an additional extra aspect for the stratification.

Administrative data sources would ease every problem, but so far in the field of agriculture these sources do not have the appropriate quality for statistical purposes.

The final aim of my research is not reached yet. The results so far show that the working out of different stratification plans has to be continued. The distribution of the animal species suggests that stratification by size would be effective, but the thresholds and the way of building distinct groups are still questions to answer. Post stratification and two-phase sampling are also techniques that have to be investigated to the matter in hand.

I plan to finish the research before the census of 2010, to be able to offer the results for practical application to the sample selection of 2010 at the Department of Agriculture of HCSO. It would be also important to widen the scope of the investi-

gation into land use. The results of the coming census will give new database for the further work.

References

Ay, J. 1976: A *mintavételes állatösszeírások módszertani kérdései.* Kandidátusi értekezés, Budapest. 150 p.

Cochran, W. G. 1977: *Sampling Techniques.* Wiley & Sons, Inc., New York. 428 p.

FAO 1995: "*Programme for the World Census of Agriculture 2000*", Statistical Development Series 5, Rome.

Hunyadi, L. 2001a: *A mintavétel alapjai.* SZÁMALK Kiadó, Budapest. 96 p.

Hunyadi, L. 2001b: *Statisztikai következtetéselmélet közgazdászoknak.* Központi Statisztikai Hivatal, Budapest. 483 p.

Hunyadi, L. – Vita L. 2003: *Statisztika közgazdászoknak.* 2. kiad. Központi Statisztikai Hivatal, Budapest. 770 p.

Kish, L. 1989: *Sampling methods for agricultural surveys.* Food and Agriculture Organization of the United Nations, Rome. 261 p.

Kish, L. 1995: *Survey sampling.* Wiley & Sons, Inc., New York. 643 p.

Laczka, S. 2000: *Mezőgazdasági összeírások Magyarországon, 1895-2000* = Statisztikai Szemle 78. évf. 4. sz. 283-289. p.

Pintér, J. – Rappai, G. (editors) 2007: *Statisztika.* Pécsi Tudományegyetem Közgazdaságtudományi Kar, Pécs. 508 p.

Regulation (EC) No 1165/2008 of the European Parliament and of the Council of 19 November 2008 concerning *livestock and meat statistics* and repealing Council Directives 93/23/EEC, 93/24/EEC and 93/25/EEC.

www.ksh.hu (10.08.2009 – 31.12.2009)

Producer price forecasting in beef cattle sector

Béres Dániel[1] *– Mészáros Kornélia*[2]

The goal of this paper is to introduce a model which creates a system by using a chain of simple statistical methods. This model is able to give an approaching estimation from the inputs' price changes to the prices of the output(s) (which of the inputs have the biggest effect on the output). By this way we are able to define the measure of the risk of the entrepreneurs, the companies or even the agricultural producers. The defined risk factors serve as a basis of the later analyses where the decision makers can classify these risk factors to choose out the best methods to risk management. The model is tested on beef cattle sector where the authors are making an experiment to explain the changes of the beef producer price with the price changes of the predefined key input factors.

Keywords: beef cattle sector, forecasting model, risk management, price forecasting

1. Introduction

Future and uncertainty: these two words are often used together in one sentence not only in Hungarian language. It is no accident because the **future** always contains a kind of **uncertainty** which is in fact a **risk factor** for the economic organizations. All the risk factors must be managed to avoid inadequate operation that may endanger the continuous course of business namely the **liquidity, good standing, and profitability**.

The mine of methods and procedures that could be effectively used during risk management is fairly large and well documented but many of these methods are based on multivariate statistical analysis. This means that the circle of people who are able to use these methods is limited partly because of the lack of knowledge partly because the statistical softwares are very expensive.

Our aim was to **create a model** that is **simple** and **easy to understand** and based on the calculations of descriptive statistics. With the chain of these simple statistical methods we could define the risk factors of the operation of a corporate and this way it will be possible to respond to the market changes before it would be

[1] Dániel Béres, PhD student, Szent István University GTK PSZI (Gödöllő)
[2] Kornélia Mészáros, PhD student, Szent István University GTK VGSZI (Gödöllő)

realized. The model is tested on cattle beef sector where the factors of the producer price were examined.

2. Material and Methods

Three things what are usually examined by strategic planners before starting their job: (1) what happened in the past (2) what is the present situation like, and finally (3) what are the expectations like. The first two questions could be answered relative easily thus the stress is on the future namely the **expectations**. As it was mentioned in the introduction the future means uncertainty and risk that must be managed by the economic organizations.

Information about the figures of the past could be easily collected from the different **databases** (internal and/or external databases). All forecasting systems are based on these databases. The first question is: how long should be the time series namely how many figures do the calculations need. There is no unequivocal answer: it always depends on the nature of the examined occurrence but it could be said in general that **the longer time series are accessible for the calculations the more complex conclusions could be drawn**. It must be considered that the figures need (1) to refer to similar time, (2) to contain similar time intervals, and (3) to have similar content (Szűcs 2004).

The inputs of the model are the **prices of input factors** and the sum of the input usages in the final product. According to it the cost of production was calculated as it shown below (1).

$$\sum_{i=1}^{n} P_{x_i} * z_{x_i} \tag{1}$$

where: x_i means: the inputs
$P_1, P_2, P_3 \ldots P_n$ mean: prices per input unit
$z_1, z_2, z_3 \ldots z_n$ mean: the sum of the input usages per final product

In order to the model could explain the price changes of the final product with the price changes of the inputs it is essential to determine that how an input price changing influence the final cost of production.

$$\gamma_i = \frac{\sum_{i=1}^{n} P_{x_i} * z_{x_i} + \Delta P_{x_i} * z_{x_i}}{\sum_{i=1}^{n} P_{x_i} * z_{x_i}} \tag{2}$$

where: γ_i is the influence of the input price changing on the final cost of production (elasticity)

The new production cost will be equal to the product of former production cost and γ_i.

$$CoP_1 = \gamma_i * CoP_0 \quad (3)$$

where: CoP is Cost of Production

If more than one input price is changing and there is no any significant and professionally justified relation between them then **linear regression model** could be effectively used. This means that the original cost of production by influence of the price changes of the inputs are multiplied and this way the new production cost will be got.

The situation is a little different when there is relation between two independent variables (dependent variable is only the cost of production). In this case the formula number (3) must be corrected as it could be found in formula (4).

$$CoP_1 = \gamma_i * CoP_0 * (1 + r_{ij}) \quad (4)$$

where: r_{ij} is the correlation coefficient between two independent variables (only in the case of significant and professionally justified relations)

If more than one input price is changing than the changing of production cost will be equal to the sum of subtractions of the new calculated and original production cost (5).

$$\Delta CoP = \sum_{i=1}^{n} (\gamma_i * CoP_0 * (1 + r_{ij}) - CoP_0) \quad (5)$$

Risk factors:

- The larger the input usage ratio during production the smaller changes in input prices may cause significant changes in production cost
- Volatility of input prices

Input usage in a final product is signed by z. This data could be got from the calculation of the cost of production related to the basis period and the prices of the inputs and its changes could be extracted from the informatics systems containing figures from the past.

Every input factor's price is following a kind of **trend** that contains a lot of information. An average price value could be calculated for each period which of course have an own volatility too. If we fit a trend line to this average values a price changing tendency will be shown. The volatility shows how exact the trend line is (for instance: is there seasonality or not). Consider all have been mentioned before

we could apply **interval estimation** to predict the **new production cost at a certain level of probability**.

Naturally not all of input factors' price should be predict this way only those that have significant influence on the final production price. (When relatively few independent variables are determining the dependent variable it is worth to do this estimation to every variable.)

Finally it is very important to take note of **seasonality**. If it appears to be it is needed to handle.

The steps of estimating production cost:

1st determination of inputs which influence the production cost and data collection
2nd check on multicollinearity
3rd determining of the values of γ_i
4th determining of the risk level of the input factors and select the most risky ones that should and could be managed.
5th calculating and representation of the periodic average input prices, fitting trends – volatility, seasonality
6th interval estimation
(7th changing the price of the final product to keep the contribution ratio)

3. Results

The model is tested on **beef cattle sector**. The examined period contains figures from 1 January 2004 to 31 August 2009. **The input factors and the bounds for the model are the followings**:

Identified input factors are (1) maize, (2) hay, (3) wage, and (4) other expenses.

The examination is based on the databases of **AKI[3] Pair** and a farm located in Hortobágy (Kovács Szabolcs "Zöldvonal" agricultural entrepreneur). The entrepreneur is breeding beef cattle. The territory of the farm is 160 hectare that is averagely 4-5 golden crown[4] per hectare is divided into two parts. Half of this is used as grazing ground the rest as meadow. The meadow is allowed to scythe once a year (early July) by Hortobágyi National Park. The cattle used to be kept in the grazing ground from May to end of August (the 80 ha is divided into 4 parts – each part contain 20 ha) then from September until November they are kept in meadow.

[3] Research Institute of Agricultural Economics

[4] All land in Hungary has a golden crown value per hectare (originally per cadastral yoke) which is periodically related to the real unit of currency (HUF).

During winter period the cattle is kept in cowshed. Thus the summer grazing period last 215 days long, the winter one lasts 150 days long.

The cattle consume maize continuously. It's demand 0.2 tons/animal/year. As the consumption is continuous the purchase of maize is similar to it. Demand for hay is appearing only in the winter period: 12 bales/animal/year. The price of one bale is depending on the weather. In those years when it was no drought the price of a bale is about 3,000 HUF/bale. When it was drought the price is much higher namely 5,000 HUF/bale. The examination was supposed that the bales are purchased at the beginning of the winter period, the farmer breeding 80 cattle, and 2 member staff is needed to take care of the animals. The average monthly earnings in the sector of agriculture are low so the cost of one employee was calculated according to minimum wage. Because of the non-detailed data of other expenses (vet cost, public utilities cost, etc.) we used the data provided by an AKI study (Beládi–Kertész 2007) where other expenses are determined as animal/year.

After collecting the required information could the second step come namely the scope of **multicollinearity** problems. Here we examined that how the chosen variables (independent variables) – the input prices – influence on each other, and the dependent variable (the price of production or here: producer price). As a result of the data examination no multicollinearity was found the inputs influenced only the (final) producer price. (It was also checked by SPSS program.) Thus the producer price could be calculated according to the formula (6).

$$CoP_l = \sum_{i=1}^{n} \gamma_i * CoP_0 \qquad (6)$$

Determining the γ_i values is the third step namely how the final producer price was changed averagely when there had been one percent change in the price of the inputs.

Table 1. The γ_i values of the inputs

Input	γ_i
price of maize	0,0675%
price of hay[5]	0,6622%
wage cost	0,1771%
other expenses	0,3123%

Source: own creation

The most significant influence has the hay and the least significant influence has the maize price changing on the final producer price. From this

[5] only in winter period

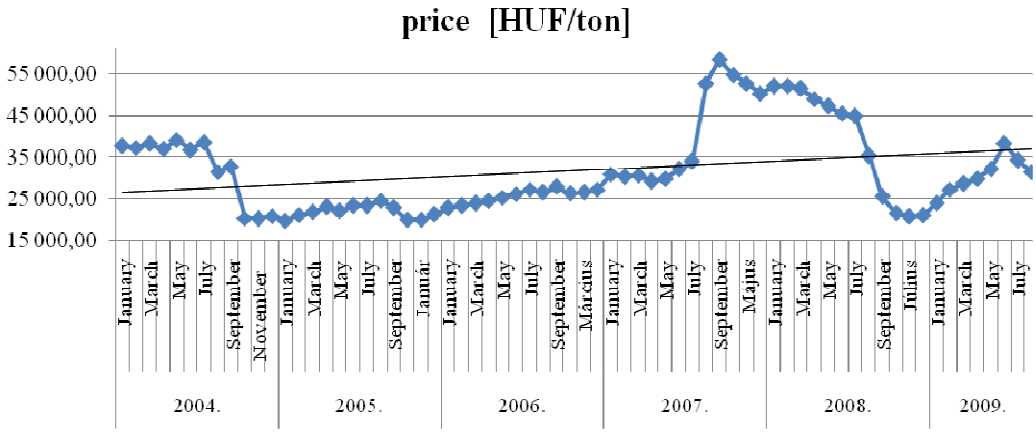
viewpoint the most risky input is the hay and especially the price of it should be followed during pricing then could come the other inputs as well.

The price of the hay depends on the weather mostly. The price of the bales is determined according to the yearly precipitation. As the hay purchasing used to be one time per year (before the winter period) then the cost of it is also appearing then. It comes from the previous that the producer knows how the producer cost will be change thus the producer is able to change the sale price to avoid the margin decreasing.

In this case the other expenses data are aggregated figures. The further breakdown of the numbers is indispensable in real situations. The wage cost is relatively inelastic because it changes only a few times per year thus the producers are able to manage the risk coming from the increasing of the wages easily in time. During the calculations the minimum wage was used which is changing only one time every year – determined by the government – so the producers have enough time build the grown cost into the cost structure.

Although the price changing of the maize had the least significant effect on the final producer price it is good to demonstrate the essence of the model. The **price of the maize had a big volatility in the examined period** (from January 2004 to August 2009) as it could be seen in the diagram below (Figure 1).

Figure 1. Variation in maize price (2004-2009)

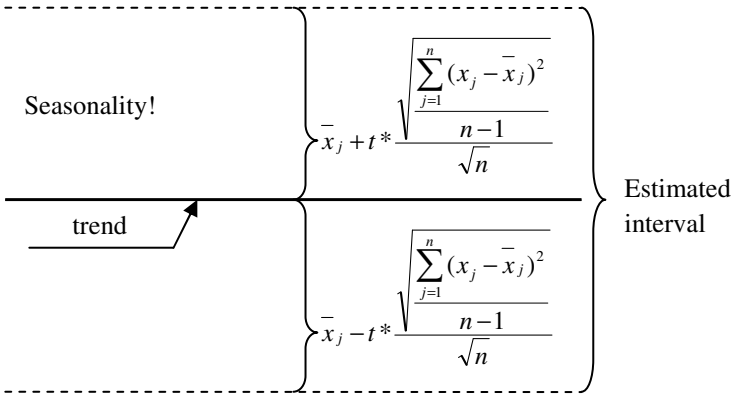


Source: AKI

The diagram shows that the (linear) trend is slowly increasing but there were periods when there were bigger changes in the price. It seems at first sight that the maize price is changing only long term but the cause of it is that the price of the maize is higher when it was draught and lower when was not. It follows from this

that these **different draught periods must be handled separately**. If we do not do this the estimated interval will be too wide and the analysis will be worthless.

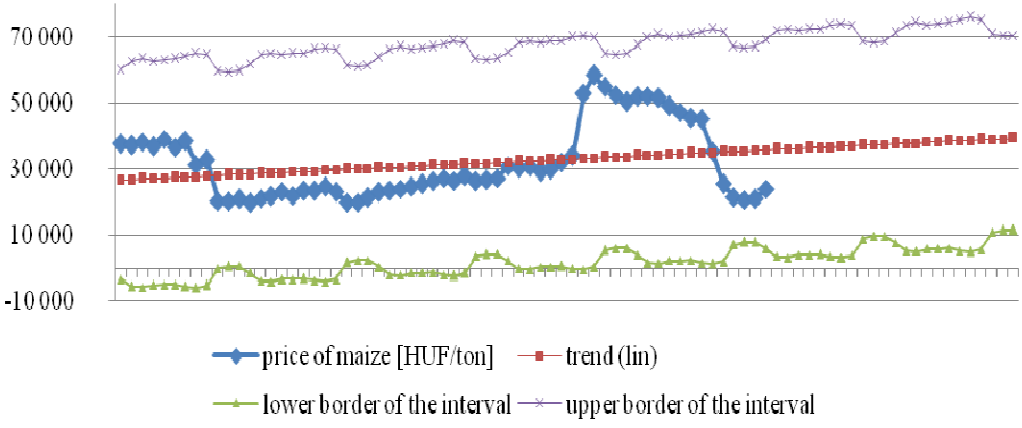
Figure 2. Calculation of estimated interval



Source: own creation

The width of the estimated interval is averagely 66,068 HUF if there is no separation between the periods (see Figure 2). This means an average 33,034 HUF difference from the trend both positive and negative direction. Considering that the price of one tone of maize is averagely 24,108 HUF in non-drought period it means that in extreme situation the price of the maize could increase by more than 100%.

Figure 3. Estimated interval without separation of the periods

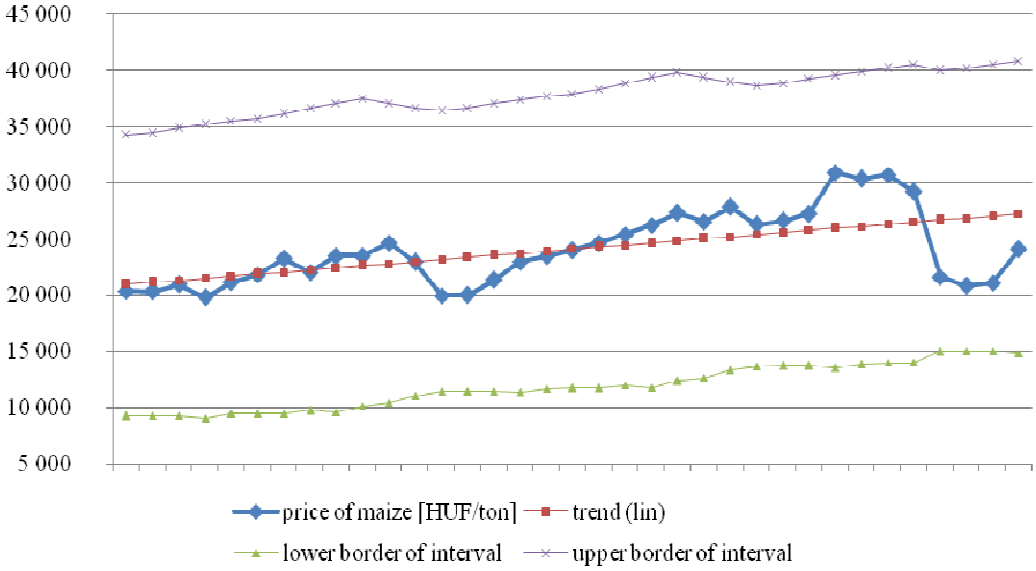


Source: AKI and own calculations

It could be seen from the interval borders that **the price of the maize is changing significantly one month to another – it refers to seasonality**.

When separated the non-drought and drought periods the values of the transition period were not considered because the data from these periods may cause distortions (wider intervals). Besides, the non-drought periods were examined together (Figure 4).

Figure 4. Estimated interval for non-drought period



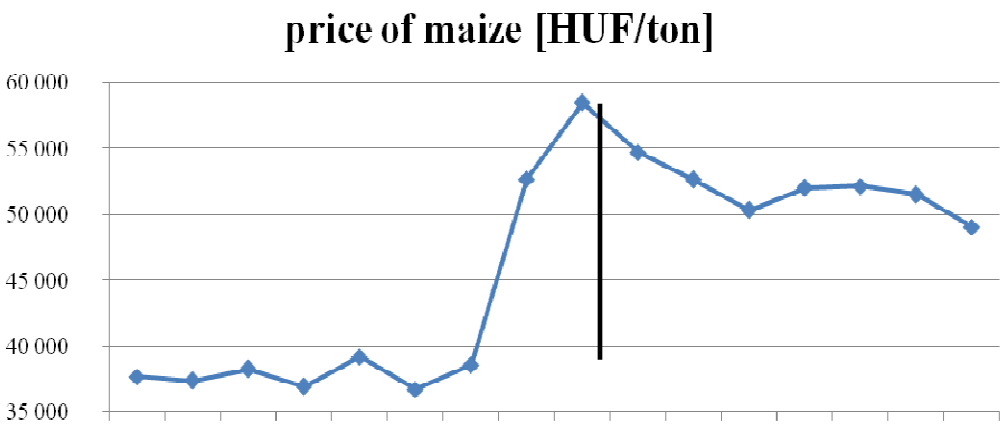
Source: AKI and own calculations

The average width of the interval now is only 26,147 HUF[6]. The interval is approximately 3 times narrower than it was before separation of the periods.

Unfortunately the data from the drought period were not enough to make such an analysis. The protocol is almost the same. The little different is that the gravity of the drought is not similar in every year so this must be examined separately as well. Figure 5 shows it us well.

[6] t value was 0.5. When it was 1.96 the width of the interval was 102,496 HUF.

Figure 5. Differences between drought periods



Source: AKI and own calculations

On the left side of the diagram (Figure 5) we could find **mild-drought period** while **hard-drought period** is on the other side. The same estimated interval was used for both two periods.

The following table contains the upper and lower interval borders fitted to the trend of non-drought period.

Table 2. Interval borders of non-drought period

Interval border of non-drought period		
	lower	**upper**
January	12,192	38,270
February	11,691	37,743
March	11,742	37,962
April	11,742	38,342
May	10,871	37,813
June	10,709	38,248
August	11,535	38,250
September	12,207	37,819
October	12,342	37,343
November	12,392	37,573
December	12,396	37,950

Source: own creation

The average price of the mild-drought period is 37,851 HUF while the hard-drought period has an average price of 52,628 HUF. The estimated interval (26,147

HUF) was fitted to them.[7] If the price of the maize in the basis period had been calculated according to non-draught period then the price of it could increase by 54% if there was hard-drought in the present period. If the basis was mild-drought-year then this number is 35%. If both two years have hard-drought then maximum 25% increasing is expected. It is true of course in the case of decreasing of the prices as well.

One percent change in maize price resulted 0.0675% change in producer cost. The following table (Table 3) contains the maximum effects of the maize price changing on producer cost.

Table 3. Maximum effect of the maize price changing on producer cost

If the former period was	
non-drought	3.6450%
mild-drought	2.3625%
hard-drought	1.6875%

Source: own creation

According to the table (Table 3) the most significant effect could be done if the former period was non-drought but it must be considered that all three periods have the same (non-drought) interval width. It means that **in worst case the price changing of the maize will result maximum 3.645% increase in the producer cost**.

4. Conclusions

The price of hay shows the largest γ_i value and the price of maize shows the lowest one. It means that the hay has the most significant effect on the producer price but it is always generated by the weather.

The average width of the interval is 66,068 HUF which means that the price could deviate from the trend with the half value of it in both sides. Considering that the average price of the maize in non-drought period was 24,108 HUF per ton this means that in extreme situation the price of the maize could be multiplied.

We made calculations: one to the drought and one to non-drought period. The width of this new interval is now only 26,147. In mild-drought period the average price was 37,851 in hard-drought it was 52,628.

One percent price change of the maize was made average 0.0675% effect on the producer price. This whole means that in worst case the producer price will be increased by 3.645%.

[7] At mild-drought period the interval fitting contains a period from January to July.

Finally, the study demonstrated that the price of the output had influenced considerably by changes in hay price while it could not been said about the price of maize.

References

Béládi, K. – Kertész, R. 2007: A főbb mezőgazdasági ágazatok költség- és jövedelemhelyzete 2006-ban a tesztüzemek adatai alapján. *Agrárgazdasági és Kutató Intézet,* 2007/7, 133-137. p.

Kerékgyártó, Gy. – Mundruczó, Gy. – Sugár, A. 2003: Statisztikai módszerek és alkalmazásuk. AULA, Budapest, 95-124., 441-486. p.

Sajtos, L. – Mitev, A. 2007: SPSS Kutatási és adatelemzési kézikönyv. Alinea Kiadó, 91-136. p.

Szűcs, I. (szerk.) 2004: Alkalmazott Statisztika. AgroInform Kiadó és Nyomda Kft., 181-210. p.

www.aki.gov.hu
https://pair.aki.gov.hu/pair-public/index.jsp

Price transmission in Serbian milk commodity chain

Rade Popovic[1] *– Boris Radovanov*[2]

Focus in the paper is on price transmission at Serbian dairy market. It analyzes milk price transmission for three products: pasteurized milk, white yoghurt and cheese. Those three products types participated in formal milk market with 65% of processed row milk. Monthly data are collected from three levels of milk chain: farm level, processors and retailers. Three step approach based on: intensity, asymmetry and time lags in price transmission was applied. Distinction of two market levels was made, first between farmers and processors and second among processors and retailers. Additionally price transmission from world to nationally market was examined.

Keywords: price transmission, milk and milk products, milk supply chain, market structure

1. Introduction

Milk production is one of the most important agriculture branches, with 11% gross agriculture output share. Traditionally it is based on over 200.000 mainly small family farms. Family farms produced 91.3% of milk intended for human use in 2008, while rest was supplied by non-family farms. Total milk production was stable in last 10 years. Milk market is separated on formal, where dairy plants process milk, and informal, where milk is used for consumption on dairy farms and for direct sale. Formal milk market is developing in recent years. The most responsible for that positive trend is dairy processing sector. Successful privatization and huge investments in reconstruction and new capacities enable faster development of formal market. In 2008 it reached 53.8% of total milk production intended for human use. Now in Serbia operate 193 dairy processing firms, from which the biggest one Danube Food Group control 43% of formal milk market. Four biggest

[1] Rade Popovic, University of Novi Sad, the Faculty of Economics Subotica (Serbia)

[2] *Boris Radovanov,* University of Novi Sad, the Faculty of Economics Subotica (Serbia)

Acknowledgement: This study was conducted within a project of Ministry for Science and Technological Development of the Republic of Serbia, no. TR-20008A.

dairy companies have 61% of market share. Process of concentration in dairy processing industry is ongoing.

Main group of dairy products in structure of formal milk market (Popovic 2009), measured in milk equivalents, was fermented milk products with 35% share in 2008. On second place was liquid milk with 34% share, from which UHT milk count 17% and pasteurized milk another 17%. Other products like: cheese, butter, cream, milk powder, cheese spreads etc. participated together with 31%.

Retail sector is composed from several big domestic and foreign retail chains and small retail shops. Process of concentration in retail sector is started earlier than in dairy processing industry. Until now, retail sector still didn't take strategic steps in this sector to ensure stability and competitive milk prices through own labels. In recent period, retailers practiced market power with pushing dairy processors to ensure longer credit period (90 and more days). That situation caused negative boomerang effect on other levels of dairy supply chain (farmers and processing industry) in few months period during 2009.

Trend of decreasing farmers' share in consumer expenditure is result of faster increase of productivity in agriculture production than in processors or retailer level. Such trend becomes obvious also in Serbian dairy market over previous period. But, in some cases it could be result of abuse of market power by processing or distribution trades (Bunte 2004). Downstream industries may practice market power by achieving lower purchase prices or higher margins, or both. If processors or retailers are able to practice market power they tend to increase difference between consumer and farmer price and reduce farmers' share in consumer expenditure. In recent years market power in dairy sector was one of the major policy concerns of Ministry of agriculture forestry and water management of Republic Serbia.

Price transmission is one of often discussed issues in agriculture economics. Between all agricultural products, in empirical studies, price transmission analyses of milk are the most often (Cotterill R., McCorriston S., Conforti P., Lechanova I., Novak P., etc.). Price transmission means, the way how prices at one level of product chain react to the price changes at another level. Market power may explain that price changes at one level are not transmitted to other levels (Bunte 2004). Except market power imperfections in price transmission can be caused by adjustment costs, profit maximizing inventory management, non-linearities in supply and demand. Between all two main causes are dominating in literature: non-competitive markets and adjustment costs (Meyer-Cramon 2002). McCorriston in his paper from 2002 urges that empirical studies with highly disaggregated data are becoming the main approach for gauging the impact of competition in multi-products industries with oligopolistic structure.

Dairy market is one of the most regulated food markets in vast countries. Recent changes in world milk prices had similar trends in Serbian market and affected mostly farmers and consumers. The additional aim of the article is to

analyses transmission of dairy world market price to Serbian market in farm gate prices using appropriate methodology.

2. Methodology

Applied methodology in the paper is aimed to research phenomena of price transmission in milk commodity chain in Serbia, distinguishing product with low value added (pasteurized milk) and product with high value added (yoghurt and cheese). Developed markets food commodity chains are usually consisted from five levels: farmers, wholesales, processors, distributors and retailers. Milk commodity chain in Serbia consist only three levels: farmers, processors and retailers. Processors are the most responsible for vertical coordination in chain. They collect and transport milk from farmers, process and distribute milk products to retailers. Therefore monthly data used in analysis are collected from three levels in period January 2007 – December 2009.

Price transmission is conducted subsequently between farmers and processors level and between processors and retailers level for chosen milk products. Analysis of price transmission is based on three step approach proposed by Lechanova in 2005.

In first step were analyzed coefficients of price transmission elasticity (EPT) as the basic measure of price transmission intensity. If two partial markets of dairy supply chain are denoted as *i* and *j* coefficient of price transmission can be defined as:

$$EPT_{ij} = \frac{\frac{\partial p_j}{P_j}}{\frac{\partial p_i}{p_i}}$$

Sequence of parameters *i* and *j* determines direction of assessed process of price transmission. Coefficient of EPT_{ij} explains by how much price of j^{th} level will change if the price at i^{th} level changes by 1%. With assumption of fixed proportion of technology, and if for example, row milk participated with 50% in total costs of processing one milk product, than transmission should be 0.5 (McCorriston 2002).

In second step, with focus on subsequent market level price differences, asymmetry of price changes were tested on basis of regression models (multiple regressions). With distinction of positive and negative price changes models has following forms:

$$\Delta P_{jt} = A^{+} + \sum_{l=1}^{k} B_l^{+} \cdot \Delta P_{it}^{+} \qquad \Delta P_{jt} = A^{-} + \sum_{l=1}^{k} B_l^{-} \cdot \Delta P_{it}^{-}$$

where:

ΔP_{jt} is downstream price changes,

$\Delta P_{it}^{+} = P_{it} - P_{it-1} > 0$ is positive price change at i^{th} level of commodity chain and

$\Delta P_{it}^{-} = P_{it} - P_{it-1} < 0$ is negative price change at i^{th} level of commodity chain.

The correlation intensity of positive respectively negative price differences can be evaluated by correlation coefficient.

In third step are analyzed the impact of time delay on the transmission of price changes between subsequent market levels. Monthly price differences, for chosen dairy products, at all market levels are used as data for analysis. Time delay is tested for 1, 2, 3 and 4 month period by determination coefficients.

As result of simplification some impediments in research data are unavoidable. They could be found in use of monthly prices instead of daily or weekly prices, excluding value added tax (VAT) from analysis, non accessible contract elements between processors and retailers (period of payment, confidential rebates), etc.

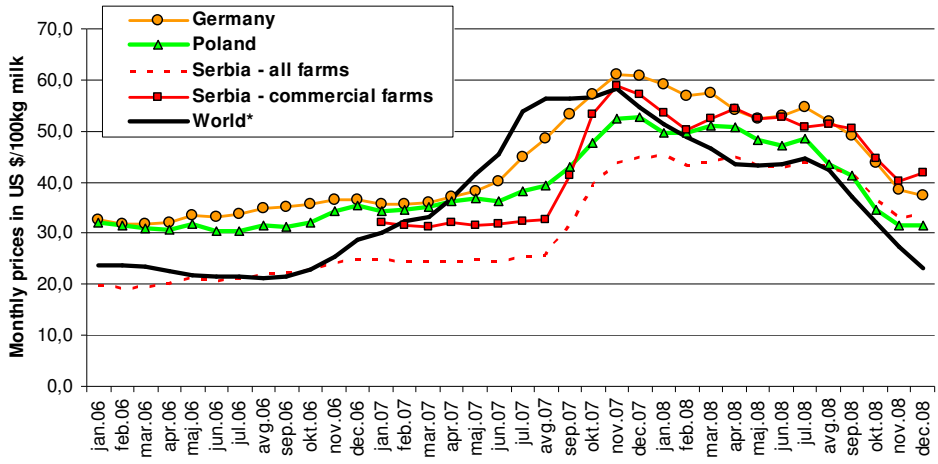
3. Results

World milk market experienced the highest milk price volatility in period end of 2006 to 2009 that are unrecorded in last three decades. Trend of world farm gate milk prices was followed by certain degree of national prices. Every country was affected by the milk price increase in period of 2007 and price decrease in 2008 year. The difference between them is the time lag and the force with which the national price responded to the world market price signals (Hemme et al. 2009).

In case of Serbia (Figure 1) world milk price variations were followed with time lag of 9 months for all farms, and 6 months for group of commercial farms[3]. At same period in Germany and Poland national farm gate milk prices reacted faster on world trend, one and four months respectively. Some possible reasons for big time lag on Serbian market could be: Serbian milk market is not well integrated in world market, and farmers in Serbia have lower market power than in other two countries. In case of commercial farms first reason is not possible because eventually when milk prices starts to rise they reach level and remain on higher level of world milk prices. Second reason is more acceptable, since there no dairy cooperatives.

[3] Commercial farms include family and non family farms with 10 and more cows in herd.

Figure 1. Farm gate milk prices in Serbia, Poland, Germany and World in period 2006-2008



Source: IFCN and own research

In Figure 2 is depicted price development for pasteurized milk, yoghurt and cheese respectively in three year period from January 2007 to December 2009. For yoghurt and cheese conversion rates are applied to provide data comparability. Conversion rate for yoghurt is 1.2 liter of raw cow milk. For feta type cheese conversion rate is 5 liters of cow milk, but because in analysis were included cheese packaging of 500 g, allotted amount of cow milk is 2.5 liters.

From visual insight in price developments from three milk products it's possible to make some conclusions for first and second market level. Prices between farmers and processors have partially similar variation only in case of pasteurized milk. For yoghurt and cheese processors' prices followed farmers prices only in 2007, after that they became rigid and insensitive for pattern of farmers' prices. Milk as main input for those products, which dominate in their total cost, should have the main impact of processors' prices. In analyzed period energy costs in terms of fuel for vehicles had similar variations, but share of this cost usually don't exceed 5% in milk products cost.

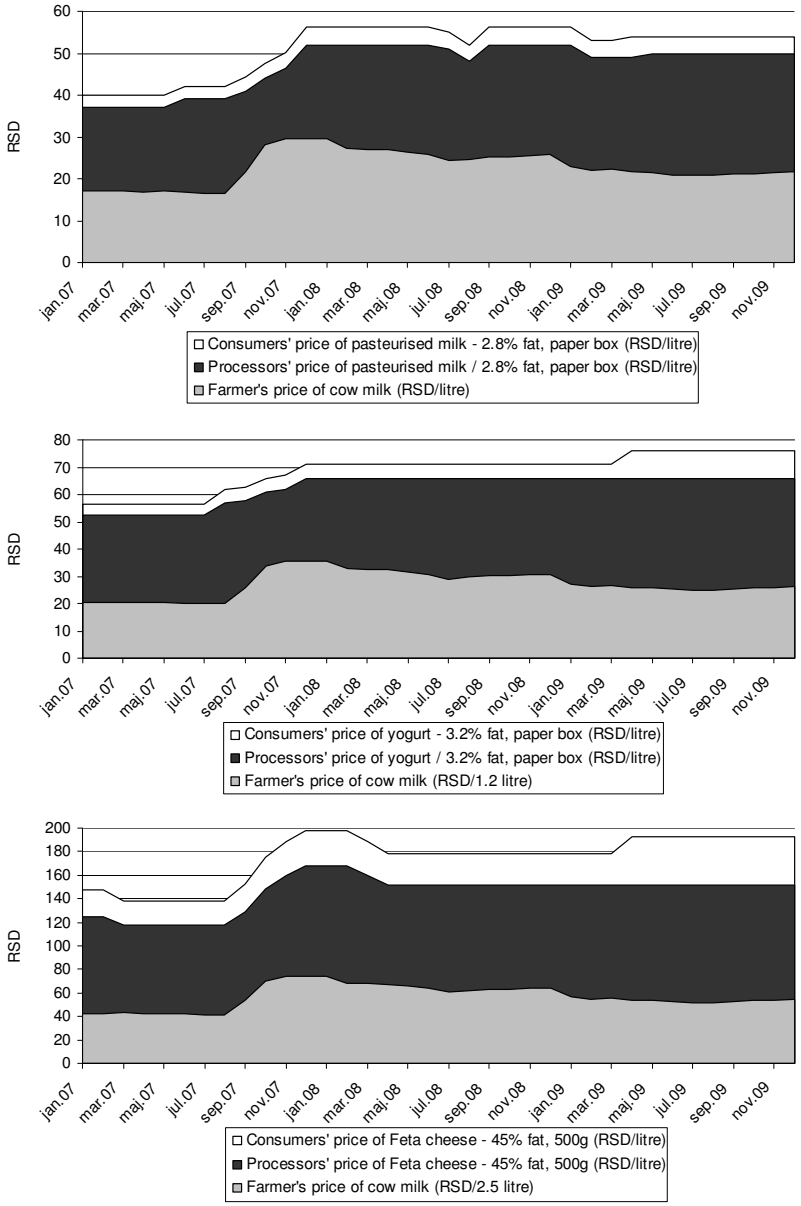
On the market level between processors and retailers it is obvious that retailers' prices well followed movement of processors' prices. Retailer prices do not absorb any processor price change and react immediately in same direction. Retailer share (or retailer margin) in consumer price is probably bigger than data in figure 2 presents. Difference lays is in confidential rebates, which are not visible out of processor – retailer contract. On this conclusion refers consumer price movement for yoghurt and cheese in May 2009. Contract relationships between processor and retailer changed at that time and probably rebates were decreased. On basis of new

contract retailer decide to increase their margin, which previously was substituted by bigger rebate.

Structures of consumer prices for three milk products can be analyzed by shares of supply chain participants. Average share of farmer in consumer price is highest in case of pasteurized milk 44.3% in three years period. In 2009 farmer share for this product reduced to 39.98%. In case of other two milk products with higher value added, average farmer share in consumer prices was 39.2 and 32.3% for yoghurt and cheese respectively. Similarly, as for pasteurized milk in 2009, farmer share reduced on 34.6 and 28.4% respectively.

According to presented data, average shares of processors in consumer prices for: pasteurized milk, yoghurt and cheese were 48.3, 51.7 and 50.8%. In practice those shares were certainly lower, because of rebates and credit period that processors allowed to retailers.

Figure 2. Price volatility in period 2007 – 2009 (farm gate, processor and consumer prices) in milk supply chain



Source: own research

First step of price transmission analysis, according to applied methodology, is calculating intensity of price transmission by EPT coefficients. Results are presented

in Figure 3. To understand correctly EPT coefficients it is necessary to define share of milk cost for each milk product. With assumption that processor price cover all economic costs of milk products, average shares of row milk cost are: 48% for pasteurized milk, 43% for yoghurt and 39% for cheese.

At market level between farmer and processor EPT coefficients indicates elastic (over-shifting) price transmission for pasteurized milk and cheese. In other words, 1% increases of row milk price results with 0.56% price increase of pasteurized milk and 0.66% for cheese, while full transmission should be 0.48 and 0.39% respectively. Some of explanation for over-shifted price transmission should be price increase of other inputs or market power. For yoghurt it's observed inelastic EPT coefficient.

Established EPT coefficients on market level between processor and retailer indicates price transmission elasticity for all three products respecting shares of processors in consumer prices. Processor prices changes of milk products are well transmitted to retail level.

Figure 3. Matrixes of coefficients of price transmission elasticity for: pasteurized milk, yoghurt and cheese

EPT milk	Farm price of cow milk	Processors` price of milk	Consumers` price of milk
Farm price of cow milk	1	0.56	0.56
Processors` price of milk	1.06	1	1.00
Consumers` price of milk	1.06	0.99	1

EPT white yogurt	Farm price of cow milk	Processors` price of white yogurt	Consumers` price of white yogurt
Farm price of cow milk	1	0.39	0.29
Processors` price of white yogurt	1.18	1	0.88
Consumers` price of white yogurt	1.19	0.97	1

EPT cheese	Farm price of cow milk	Processors` price of cheese	Consumers` price of cheese
Farm price of cow milk	1	0.66	0.51
Processors` price of cheese	1.22	1	0.73
Consumers` price of cheese	1.96	1.38	1

Source: own research

The most of empirical studies focused on price changes transmission between subsequent market levels revealed that positive price changes are much faster and in bigger volume are transmitted on subsequent commodity chain than negative price changes. In table 1 are presented correlation coefficients for positive and negative price changes on two subsequent market levels.

On first market level, related to farmers and processors, positive price changes are positively correlated for all three milk products. Negative price changes are positively correlated for pasteurized milk and negatively for cheese, which means when farm milk price decrease price of cheese are increased. Correlation of negative price changes don't exist in case of yoghurt, because during three year period processor price of yoghurt just increased. Taking together positive price changes of row milk are transmitted in greater extend than negative price changes.

At second market level of dairy chain between processor and retailer both positive and negative price changes has high value of positive correlation coefficients. Price decrease of processors product are fully transmitted to consumers

and price increases are almost fully transmitted. In case of yoghurt and cheese price increase are in small extend absorbed with consumers prices, probably because of stocks in markets.

Table 1. Correlation of positive and negative price changes

Article	First market level	Second market level
Pasteurized milk		
Price increase	43,56%	98,60%
Price decrease	9,27%	100,00%
Yogurt		
Price increase	38,11%	83,36%
Price decrease	0,00%	0,00%
Cheese		
Price increase	89,12%	89,30%
Price decrease	-8,25%	100,00%

Source: own research

Third step of analysis is time delays in output prices reaction to changes in input prices. Time delays are confirmed on first market level for all three products in periods of 1 and 2 months. On second level time delay exists in one month period only in case of cheese. It is because of characteristics of product which can be used and kept in storage in longer period. For other two milk products it is not possible because of their short periods for use which are for pasteurized milk 3 days and for yoghurt 15 days. The highest coefficients of determination are on second market level for non delays.

Table 2. Determination coefficients for examined time delays

Article	Time delay (in months)				
	1	2	3	4	No delay
Pasteurized milk					
First market level	38,12%	23,57%	6,92%	0,54%	17,22%
Second market level	2,50%	1,42%	1,36%	0,01%	98,06%
Yogurt					
First market level	16,39%	22,17%	8,53%	0,31%	15,46%
Second market level	2,66%	14,79%	0,10%	4,84%	69,48%
Cheese					
First market level	53,10%	21,05%	1,14%	0,65%	51,67%
Second market level	37,93%	7,61%	0,02%	5,04%	84,89

Source: own research

4. Conclusions

Dairy supply chain in Serbia is moving toward successive oligopsony (dairy processing and retail industry). Since 2004, when privatization of dairy processing firms started, Serbian dairy industry experienced dynamic development. Concentration rate in dairy industry was CR4 = 61% in 2008. The four biggest dairy processing companies are: Danube food group BV, Mlekara Sabac, Somboled and Mlekoprodukt. Concentration process in food retail industry, as the dominant channel for dairy food distribution, resulted with several domestic and foreign retail chains. While concentration rates are important in describing an industry, high concentration rate is not evidence that companies behave anti-competitively.

Empirical results of price transmission suggest that there is over-shifted elasticity, that's mean when farmer milk price increase for 1% processor and retailer prices increase more than 1%. Price asymmetry in dairy commodity chain in Serbia is also confirmed. Processor' and retailers' prices are more likely to increase if farmer price increase, than to decrease if farmer price decrease. Also time lag analysis revealed that on second market level wasn't time delay, and on first market level between farmers and processors time lags exists for 1 and 2 month periods. In other words positive price changes are much faster and in bigger volume are transmitted on first commodity level than negative price changes. Revealed results infer some signs of market power of processors. Practicing market power from participants at any level of commodity market chain leads to reduction of competitiveness and other performances of whole food chain.

References

Bunte, F. 2004: Pricing and performance in agri-food supply chains, In Proceedings of the Frontis workshop on quantifying of the agri-food supply chain. In ed. Onderstejin K. et al., Wageningen, Netherlands 22 – 24 October 2004.

Conforti, P. 2004: Price transmission in selected agricultural markets. FAO commodity and trade policy research working paper No. 7.

Cotterill, R. 2002: Who benefits from deregulated milk prices: the missing link in the marketing channel. Proceedings, University of Connecticut.

Hemme et al. 2009: IFCN Dairy Report 2009, International Farm Comparison Network, IFCN Dairy Research Center, Kiel, Germany.

Lechanova, I.-Novak P. 2006: Price transmission in Czech milk commodity chain, In: Annals of the Polish association of agricultural and agribusiness economists, Poznan, Poland, Vol. 8, No 6, 108-112. p.

McCorriston, S. 2002: Why should imperfect competition matter to agriculture economists? European Review of agriculture economics, Oxford University Press. Vol. 29, 349-371. p.

Meyer, J.-Von Cramon, T. S. 2002: Asymmetric price transmission: a survey: paper at the 10th EAAE conference in Zaragoza. [http://www.jochenmeyer.de/pdf/meyer-cramon.pdf].

Novak, P. 2007: Some approaches to the analysis of market structure's impact in milk commodity chain. 104th EAAE-IAAE Seminar Agricultural Economics and Transition, Corvinus University of Budapest, Hungary.

Peltzman, S. 2000: Prices rise faster than they fall. *Journal of political economy*, The University of Chicago Press. Vol.108, No 3, 466-502. p.

Popovic, R. 2007: Packaging types of milk and dairy products on Serbian market, *Food industry – Milk and dairy products*, Faculty of Technology Novi Sad, Serbia, Vol. 18, No. 1-2, 65-69. p.

Popovic, R. 2009: Structural changes on Serbian formal milk market, *Food industry – milk and dairy products*, Belgrade, Vol. 20, No. 1-2, 7-12. p.

Popovic, R. 2009a: Effects of market structure changes on dairy supply chain in Serbia, Paper presented at Agribusiness and Agro-industries development in Central and Eastern Europe, A joint FAO – IAMA Workshop at International food and agribusiness management association, 19th Annual World Forum & Symposium, page. 113-124, Budapest 20-21 Jun 2009.

Simulation model on optimizing the sowing structure of precision plant production

Enikő Lencsés[1] *– Dániel Béres*[2]

During the past decade, many people deal a lot with the Hungarian agriculture, its views and opportunities in the future. In order to the Hungarian agriculture be competitive on the European market it is needed to be able to follow the market motions and its changes. To do this, it needs making investments on certain areas which requires capital. The agricultural producing can only be competitive if the farmers keep the environmental viewpoints and the sustainable farming with an eye.

The precision cultivation can be one of the implement of the so many voiced sustainable development at the field of agriculture. The precision cultivation requires surplus expenditures (purchase devices, operating the devices, etc.) but it has advantages too (yield increase, decreasing of material costs and yield insecurity, etc.). The comparison of the surplus expenditures and surplus yields serves as a basis of a complex economical analysis where not only the costs and revenues but the sowing structure changes are also appearing. The aim of this paper is to determine an optimal sowing structure for a 250 ha farm which provides the highest income with the technology of precision plant cultivation.

Keywords: sustainable agriculture, precision cultivation, simulation

1. Introduction

Nowadays, there could be heard a lot about the environmental protection, **environment friendly agriculture and sustainable growth**. The **precision agriculture** is a farming method which takes part in sustainable development. (Swinton 1997) This was the main reason why the precision agriculture is on the focus of this paper.

The main tasks of the modern agriculture are the **efficient utilization of the resources**, integrating the biological processes and regulating mechanisms of the production where it is possible and through this, confirm the cost-effectiveness of the agricultural manipulation, preserve agricultural human resources and retain

[1] Enikő Lencsés, PhD student, Szent István University, Vállalatgazdasági és Szervezési Intézet (Gödöllő)
[2] Dániel Béres, PhD student, Szent István University, Institute of Finance and Accountancy (Gödöllő)

living-standard of provincial society. (Barkaszi et al. 2006, Csiba et al. 2009, Sándor et al. 2009)

Agriculture needs to face the challenge that it should produce the food for greater population on smaller field all over the world. The **site-specific (precision farming) technology** which **optimizes inputs** (fertilizer, herbicide, pesticide, etc.) on parcel-level might be a solution for this problem. Due to the site-specific optimizing this technology increases the yield and decreases the environmental damages. (Batte 1999, Székely et al. 2000, Takácsné 2003, Takács–Barkaszi 2006, Kis–Takácsné 2006, Pecze 2008)

The environmental debit of the production could be decreased for example by **precision weed-management technology** that results cost saving (only those parcels are treated that contain weeds). The amount of the savings which comes from the site-specific treatment is different according to the various researches (between 20% and 60%). (Leive et al. 1997, Batte 1999, Luschei et al. 2001, Takács-György et al. 2002, Reisinger 2004)

The parameters of the soil are: (1) the features of ground, water- and nutrient supply, (2) injuries, and (3) yield. These factors show the heterogeneity of the field. The soil is handled in **precision farming technology (PFT)** as a heterogenic unit which influences positively the success of farming by the meaning of site-specific treatment. The more detailed information we have about the heterogeneity the better treatment could be realized with site-specific treatment. (Weiss 1996, Pecze-Horváth 2004, Reisinger 2004, Csathó et al. 2007)

The PFT could not be applicable completely for every crop. For instance in the case of sunflower production the problem of yield-measure is not solved, while in the case of maize production every technology elements are applicable by site-specific method. (Table 1)

Table 1. Applicability of the precision plant production elements in different plant culture

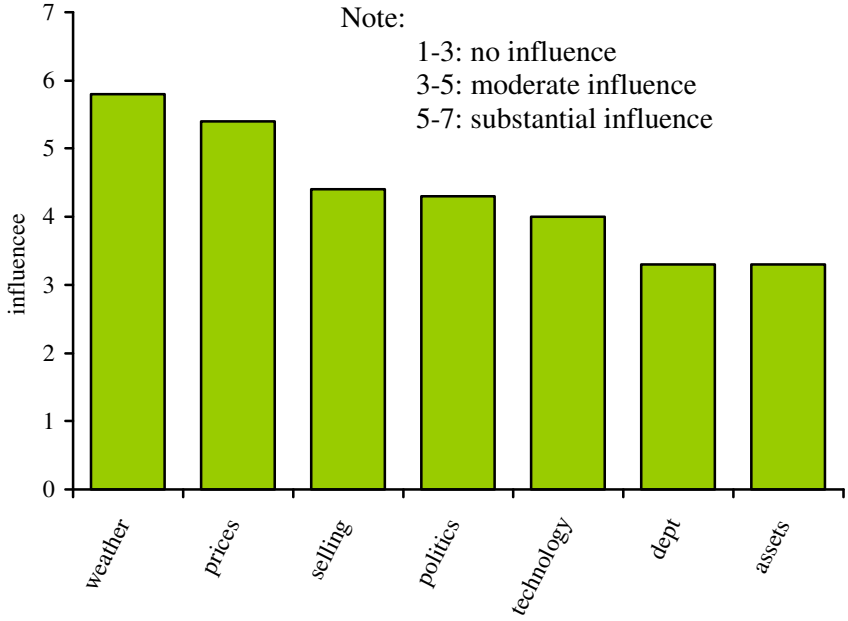
	Precision soil-sampling	**Yield mapping**	**Differential fertilization**	**Precision weed management**
Winter wheat	+	+	+	+
Maize	+	+	+	+
Sunflower	+	-	-	+
Alfalfa	+	-	-	-
Potato	+	-	-	-
Green bean	+	+	+	+
Soya	+	+	+	+
Colza	+	-	-	-

Source: Pecze 2006 and own creation

The international literature of PFT is considerably wide. The center of the research of Weiss, Lowenberg-DeBoer and Boehlje is the microeconomic questions of PFT especially the classic production economic analysis. Due to employing this technology smaller and smaller farm size could realize **profit**. Kalmár et al. argued in a study (in 2004) that this technology is viable on the farm-size that includes more than 1,000 ha. Kovács and Székely claimed in 2006 that 250 ha are enough to viability. According to the latest researches this number could be 206 ha depending on the sowing structure. (Kalmár et al. 2004, Kovács–Székely 2006, Takács-György 2007)

The **sowing structure and PFT** are not the only key factors of success because agribusiness has many factors of risk as well. According to Székely and Pálinkás the **most common risk factors** are: (1) production risk, (2) market (price) risk, (3) financial risk, (4) institutional risk and (5) personal risk. Their research which was made in 6 EU members namely Hungary, Poland, Holland, Spain and Germany claimed that the most significant risk factor is the weather- and natural risk (production risk). The volatility of the prices had only the second place in the ranking list (Figure 1). (Székely–Pálinkás 2008)

Figure 1. Ranking of risk factors by influence on production



Source: Székely–Pálinkás 2008

Technological development could be one aspect of **risk management**. The technology of precision plant production may lead to savings by site-specific treatment that saves material costs and exploitation of yield potential that improves yield security.

2. Data and methods

The main aim of this paper is to determinate an **optimal sowing structure for a 250 ha farm** which provides the highest income with the technology of **precision plant production**. The examined period is 10 year long. Main conditions of the simulation model are:

- Stipulations of **corp rotation**[3] and **intercropping**[4]: winter wheat and maize cannot be sawed in the same soil for 2 years. This number is 6 years for sunflower and 4 years for colza.
- **Weather conditions**: during the examination the model supposed that in 70% of the cases there were non-draught period and in the rest 30% there were draught period.
- **Input prices** (seeds, corp protection chemicals, artificial fertilizer) were changed according to the weather conditions.

Maximizing the gross margin is the decision criterion during optimizing the sowing structure.

The used data relating to the input costs, expenses and incomes come from the database of AKI (Research Institute of Agricultural Economics). The following changes were made on these figures: the costs of the seeds (-4%), the artificial fertilizer (-15%) and the crop protection chemicals (-10%) were decreased – the latter one is true only for those corps that has wide row spacing (e.g. sunflower and maize). Besides, the expenditures connected with the machinery were raised by 20%.

Average costs and values of production data were determined separately for non-draught and draught periods and for each corps according to the data of the period 2000-2006. The simulation model uses these figures considering the standard deviations namely the value of randomized data could be somewhere between the maximum and minimum marginal values (Table 2).

[3] temporal diversification
[4] spatial diversification

Table 2. Marginal values of operating expenses and sales price (data in HUF)

non-draught period	winter wheat		maize		sunflower		colza	
	min	max	min	max	min	max	min	max
cost of seed	10 752	12 965	14 839	20 143	10 248	14 444	7 258	12 414
cost of artificial fertilizer	12 198	15 741	13 593	15 518	6 127	11 997	14 742	17 889
cost of corp protection chemical	8 841	10 816	8 280	9 293	9 640	10 388	11 253	15 901
cost of machinery	24 007	29 342	29 361	32 165	22 217	33 767	22 186	30 442
sales price	21 000	25 200	19 427	24 445	48 470	58 125	48 359	55 744

draught period	winter wheat		maize		sunflower		colza	
	min	max	min	max	min	max	min	max
cost of seed	10 369	11 228	15 009	16 471	10 988	11 842	7 842	10 447
cost of artificial fertilizer	11 797	11 871	12 191	14 674	6 494	8 182	12 769	16 033
cost of corp protection chemical	9 137	9 300	8 380	9 667	10 422	10 559	11 940	16 631
cost of machinery	20 580	28 726	26 004	36 416	26 052	33 703	20 900	32 515
sales price	21 499	31 880	19 433	33 007	49 692	65 526	50 930	55 841

Source: own creation

Figure 2. The used formulas during interval determination

average: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ standard deviation: $\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$

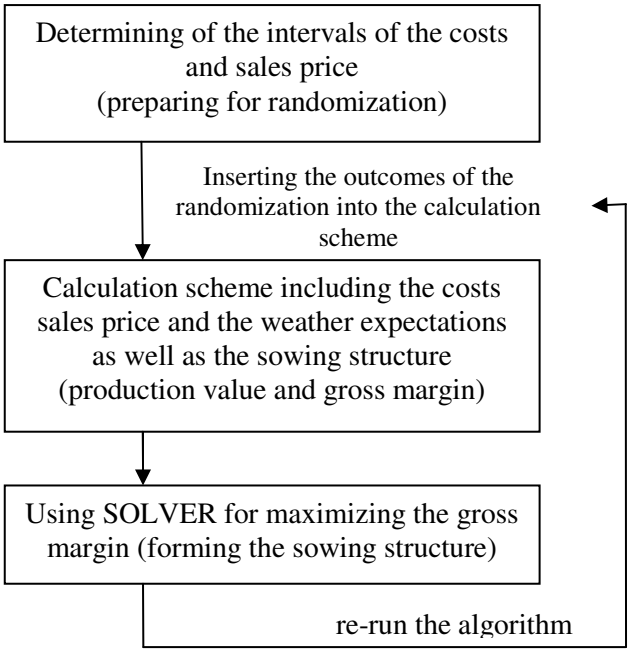
min: $\bar{x} - \sigma = \frac{\sum_{i=1}^{n} x_i}{n} - \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$

max: $\bar{x} + \sigma = \frac{\sum_{i=1}^{n} x_i}{n} + \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$

Source: own creation

The following flow-chart shows how the simulation model works. (Excel 2007 program was used for calculations.)

Figure 3. Flowchart of the model



Source: own creation

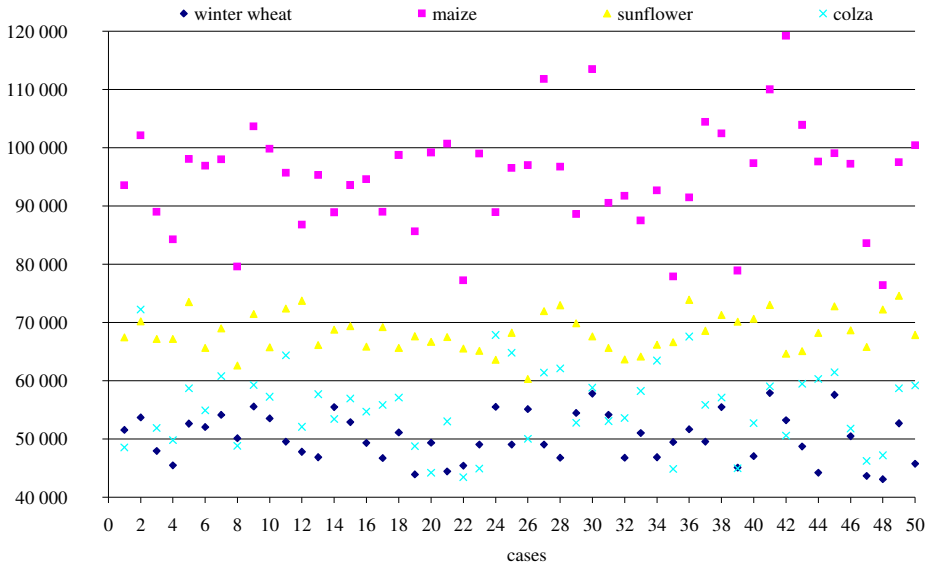
Constrains of the simulation model are typed into the solver (the formerly mentioned provisions and bounds for the model). It is very important that the result must be set to zero before running the solver – it means that the former result must be deleted before re-running the algorithm.

The simulation model was executed 50 times in order to obtain sophisticated results as the way it was previously mentioned. We found that this is enough because the results were very similar to each other.

3. Results and discussion

Applying the framework of the simulation model **maize is the most profitable corp**. The average gross margin was **94,834 HUF/ha** (StDev 9,186 HUF). Sunflower has the second place with averagely 68,223 HUF/ha gross margin (StDev 3,296 HUF). The third crop is Colza that reached 55,426 HUF/ha gross margin (StDev 6,617 HUF). And the least income could be realized by growing winter wheat that has a gross margin 50,201 HUF/ha (StDev 4,039) (Figure 4).

Figure 4. Gross margin (HUF/ha)



Source: own creation

The average **production cost** per hectare of the winter wheat was 59,945 HUF/ha. This number was 70,554 in the case of maize, 59,429 in the case of sunflower and 65,311 HUF/ha in the case of colza.

The average **production values** per hectare of the corps are the follows: (1) winter wheat: 110,146 HUF, (2) maize: 70,554 HUF, (3) sunflower: 127,652 HUF and (4) colza: 120,737 HUF (Table 3).

Table 3. Annual average values of production cost, production value and gross margin

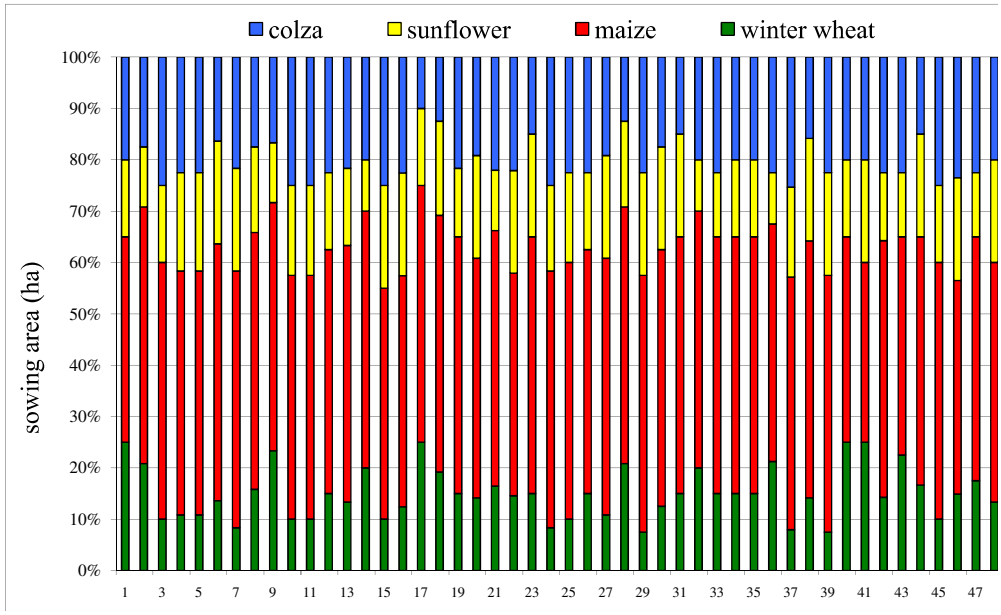
	no. of years	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	Avg
winter wheat	area (ha)	82	25	32	6	32	72	39	28	34	28	**38**
	PC	60 314	59 878	59 668	61 909	60 270	56 150	60 393	59 977	60 187	60 709	**59 945**
	PV	113 908	111 087	107 673	115 441	115 082	95 200	110 524	108 477	112 581	111 494	**110 147**
	GM	53 593	51 209	48 005	53 533	54 812	39 050	50 131	48 500	52 394	50 785	**50 201**
maize	area (ha)	106	127	64	180	168	76	130	106	103	139	**120**
	PC	71 007	69 871	71 089	70 745	71 107	69 629	70 350	70 166	70 683	70 890	**70 554**
	PV	166 578	172 308	157 531	175 360	168 214	141 883	168 065	169 417	167 935	166 583	**165 387**
	GM	95 571	102 437	86 441	104 615	97 108	72 254	97 715	99 251	97 252	95 693	**94 834**
sunflower	area (ha)	24	54	80	21	12	46	14	58	74	34	**42**
	PC	59 338	59 613	59 606	59 199	59 512	59 172	60 027	59 101	59 521	59 204	**59 429**
	PV	129 399	131 323	128 487	129702	127 603	122 284	125 449	127 582	126 985	127 709	**127 652**
	GM	70 062	71 711	68 881	70 502	68 090	63 112	65 423	68 481	67 464	68 504	**68 223**
colza	area (ha)	37	44	74	43	38	57	66	57	39	49	**50**
	PC	65 715	66 332	64 771	65340	65887	64 883	65 032	64 500	65 209	65 446	**65 311**
	PV	121 161	125 271	121 946	126 225	125009	98 831	119 967	122 878	121 333	124 749	**120 737**
	GM	55 446	58 939	57 174	60 885	59 122	33 948	54 935	58 378	56 124	59 303	**55 426**

Source: own creation

Note: PC – Production Cost (HUF/ha), PV – Production Value (HUF/ha), GM – Gross Margin (HUF/ha)

According to simulation results, industrial maize covers 48% (+ 3%), colza 20%, sunflower 17% and winter wheat 15% of the whole area (Figure 5).

Figure 5. Sowing structure per executions

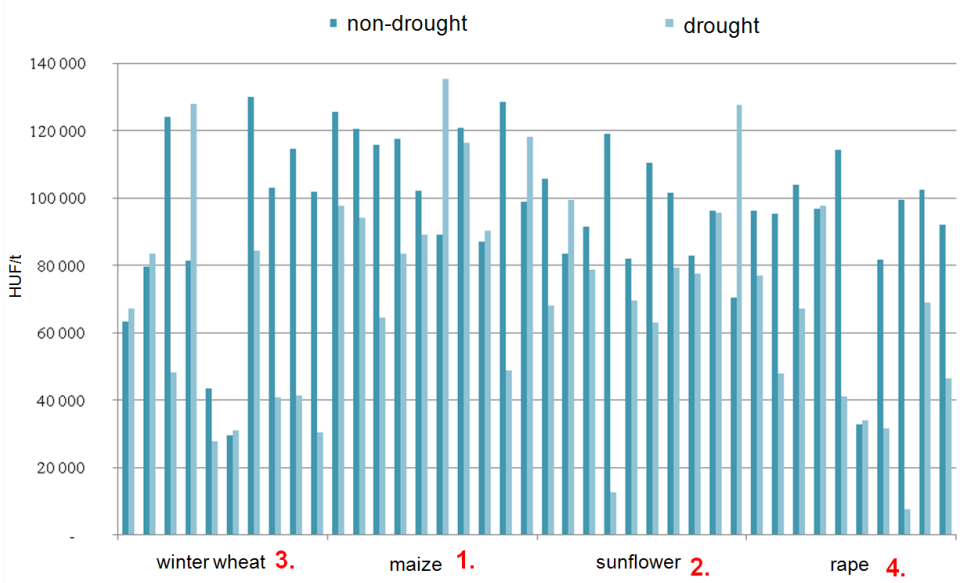


Source: own creation

Figure 5 shows how the sowing structure should be formed in order to obtain the **highest income** within 10 years considering the stipulations of crop rotation and intercropping and the probability of draught.

Finally, a **sensitivity examination** was performed on production values of the examined corps. If the weather conditions are not advantageous and drought is appearing then a **general decreasing** could be observed. Colza and winter wheat shows the highest decline (more than 750 HUF/ton each). The price of the sunflower also shows fall but the value of it is less significant (408 HUF/ton), and the smallest reduction was resulted by maize (140 HUF/ton) (Figure 6).

Figure 6. Effects of the drought on production value of the examined corps



Source: own creation
Note: rape = colza

4. Conclusion

The **optimal sowing structure of a 250 hectare large farm** is the following:

- winter wheat: 38 ha
- maize: 120 ha
- sunflower: 42 ha
- colza: 50 ha

If we would like to deviate from this sowing structure it is expedient to increase the proportion of those corps of which production values are less sensitive to weather changes. These corps are the maize and the sunflower in our case.

The adaptation of PFT could be viable mainly at medium size (250 ha) farms under Hungarian conditions especially when intensive production is used and the rate of the wide row spacing culture is at least 40% of the sowing structure (Lencsés 2009).

The farmers should carry out many technical, technological, informational and economical stipulations in order to be able to adopt PFT. The cost of investment in PFT adaptation is between 17 000 and 34 000 Euros which depends on the farm size. This financial question is the reason why the carefully considered economical

analysis is so important. Besides, ecological aspect should not been forgotten either because PFT is more environmental friendly than the traditional technology of plant cultivation which means a kind of improvement as for sustainability of agribusiness.

Furthermore, the aspect of changes in inputs is also important. Apart from the fact that PFT requires investment in equipments that needs to be maintained, it has a lot of advantages as well for instance more stabile annual yields and reduction of operating expenses (fertilizer, chemicals, pesticides, herbicides, etc.).

References

Barkaszi, L. – Arutyunjan, P. A. - Takácsné-György, K. 2006: Optimisation of the weed sampling system from an economic point of view on wheat (Triticum aestivum) stable with sunflower (Helianthus annuus) forecrop. *Cereal Research Communication,* Vol. 35, Issue 3, Akadémiai Kiadó, Budapest, 1527-1537. p.

Batte, M. T. 1999: Precision Farming – Factors Influencing Profitability. http://www.highbeam.com/doc/1G1-60013061.html

Csathó, P. – Horváth, J. – Mesterházi, P. Á. – Milics, G. – Nagy, L. - Neményi M. - Németh T. - Pecze Zs. - Szabó J. 2007: Hazai gyakorlati tapasztalatok. In: Németh T. - Neményi M. - Harnos Zs. (szerk.): *A precíziós mezőgazdaság módszertana*. JATE Press – MTA TAKI, Szeged, 229-240. p.

Csiba, M. – Milics, G. – Smuk, N. – Neményi, M. 2009: A fenntartható fejlődés kihívása és az erre adható válasz a magyar mezőgazdaságban. Mezőgazdaság és a vidék jövőképe Konferencia kiadvány II. kötet, Mosonmagyaróvár, 264-271. p.

Kalmár, S. – Salamon, L. – Reisinger, P. – Nagy, S. 2004: Possibilities of applying precision weed control in Hungary. *Gazdálkodás*, 48, 8, 88-93. p.

Kis, S. – Takácsné György, K. 2006: Kemikália csökkentésének gazdasági hatásai a mezőgazdasági vállalkozások döntéseiben. *Gazdálkodás*, 50, 3, 52-60. p.

Kovács, A. – Székely, Cs. 2006: A precíziós gazdálkodás hatása a növényvédelem költségeire. In: Takácsné-György K. (szerk): *Növényvédőszer használat csökkentés gazdasági hatásai*. Szent István Egyetemi Kiadó, Gödöllő, 63-70. p.

Leiva, F. R. – Morris, J. – Blackmore, B. S. 1997: Precision farming techniques for sustainable agriculture. 1st European Conference on Precision Agriculture, 8-10 September 1997. Warwick University.

Lencsés, E. 2009: Advantages and disadvantages of precision farming technology from economic aspect. *Annals of The Polish Association of Agricultural and Agribusiness Economists*, Vol. X., No. 6. (XVI. SERiA Congres; Agribusiness in globalization conditions), 83 – 87 p.

Lowenberg-DeBoer, J. – Boehlje, M. 1996: Revolution, Evolution or Dead-end: Economic prespectives on Precision Agriculture; Proceedings of the 3rd International Conference on Precision Agriculture.

Luschei, E. C. – Van Wychen, L. R. – Maxwell, B. D. – Bussan, A. J. – Buschena, D. – Goodman, D. 2001: Implementing and conducting on-farmweed research with the use of GPS. Weed Science 49. July-August 2001.

Pecze, Zs. – Horváth, J. 2004: A precíziós tápanyag-gazdálkodás gyakorlati rendszere az IKR-nél. *Gyakorlati Agrofórum*, 15, 8, 44-46. p.

Pecze, Zs. 2006: personal conversation and interview

Pecze, Zs. 2008: Az IKR Zrt. precíziós gazdálkodási rendszere. In: Takácsné-György K. (szerk.): *Gazdaságilag optimális környezetkímélő herbicid alkalmazást célzó folyamatszervezési, - irányítási és alkalamzási programok kifejlesztése.* Szent István Egyetemi Kiadó, Gödöllő, 103-120. p.

Reisinger, P. 2004: A precíziós gyomszabályozás hazai perspektívái. *Gyakorlati Agrofórum*, 15, 8, 52-54. p.

Sándor, I. – Nagy, J. – Jolánkai, M. 2009: Növébnytermesztésünk helyzete és a szükséges fejlesztés irányai. *Agrofórum*, 20, 4, 5-7. p.

Swinton, S. M. 1997: Precision Farming as Green and Competitive. Paper presented at "Corporate Initiatives in Environmental Management: The Next Generation of Policy?" Learning Workshop, American Agricultural Economics Association Meetings, Toronto, Ontario, Canada, July 26, 1997. http://aec.msu.edu/Smith_Endowment/documents/swinton.htm

Székely, Cs. – Kovács, A. – Györök, B. 2000: The practice of precious farming from an economic point of view. *Gazdálkodás*, English Special Edition 1. 56-65p.

Székelcs, Cs. – Pálinkás, P. 2008: Kockázatkezelés az európai mezőgazdasági vállalkozásokban. *Agrofórum*, 19, 11, 5-8. p.

Takács-György, K. – Györök, B. – Kovács, A. 2002: The effects of precision farmig on the use of chemicals; Xth Congress of European Association of Agricultural Economists (EAAE). Exploring Diversity in the European Agri-Food System. Zaragoza, Congress CD: poster\004_p010_Takacs_Gyorgy.pdf

Takács-György, K. 2007: Economic aspects of chemical reduction on farming – future role of precision farming. Acta Scientiarum Polonorum, Oeconomia 6 (3) 115-120. p.

Takács-György, K. – Barkaszi L. 2006: Economic role of precision weed control in preventing rural areas. Annals of the Polish Association of Agricultural and Agribusiness Economists (SERiA). Warsawa - Poznan. 2006. Tom. VIII. 6. 132-136 p.

Takácsné György, K. 2003: Precíziós növényvédelem, mint alternatív gazdálkodási stratégia? *Gazdálkodás*, 47, 3, 18-24. p.

Weiss, M. D. 1996: Precision Farming and Spatial Economic Analysis: Research Challenges and Opportunities. *American Journal of Agricultural Journal of Agricultural Economics*, Vol. 78, No. 5, 1275-1280. p.

A comparison of simulation softwares in modelling the crop structure management with a stochastic linear programming model

Sándor Kovács[1]

I created a stochastic linear programming model based on crop structure data. As to determine the optimal structure, I perform an MCMC simulation by using WINBUGS and two other Risk Analysis softwares. Best values and the related coefficients of the goal functions, provided by different softwares, were analysed and compared. A deterministic linear programming solution was also compared to all the result of the stochastic simulations. I also determined and compared the optimal solutions of the different softwares according to András Prékopa's study.

Evaluating the results WINBUGS proved to be the most suitable software for establishing management decisions in crop structure modelling. In my study I also presented the way for implementing stochastic linear programming models in WINBUGS.

Keywords: Winbugs, Bayesian Statistics, crop structure, risk analysis, MCMC simulation

1. Introduction

Important decisions influencing the future of the company have to be made under conditions of risk, when reliable information is available only for the most recent time period. Risks must be considered by every economic agent and they should apply methods that are capable of measuring, monitoring and suggesting responses to risks, provided that the information required for decision-making is current and of sufficient amount and quality. The evaluation of this information should enable decision-makers to formulate and analyze multiple decision alternatives. On one hand, the developments in information technology have facilitated the development of applied risk management tools, which have become affordable for even the smallest of enterprises and easy to use. On the other hand new, complex and wide-ranging types of risk have arisen, the measurement of which requires sophisticated mathematical models (Balogh et al. 1999). Simulation models, whose use in

[1] Sándor Kovács, PhD, assistant lecturer, University of Debrecen, Faculty of Applied Economics and Rural Development, Hungary

agriculture has grown rapidly, attempt to mimic the operation of real systems so as to allow accurate measurement of uncertainty and risk.

One of the simplest ways of optimizing the crop structure is to apply linear programing methods (Csipkés et al. 2008). In cultivation LP methods are one of the most favourable methods. However, these methods are hindered by the great measure of uncertainty. The major reasons for that are changes in the prices and yields of crops. That is the reason why stochastic programming has received much attention from 1950 to the present day in many applications (Dantzig 1955, Prékopa et al. 1980, Williams 1966, Prékopa 2003).

The simulation model is a simplified mathematical realization of a real system, aimed at studying the behavior of the original system when changing various conditions and circumstances. The Monte-Carlo method is a generally accepted method of modelling risks which studies the probable outcome of an event characterized by any input parameters and described by well-known functions. The essence of the Monte-Carlo technique is, on the basis of probability distribution assigned to some uncertain factors, to randomly select values, which are used in each experiment of the simulation (Russel–Taylor 1997). Monte-Carlo methods are the statistical evaluations of numerical methods and their characteristics using the modelling of random quantities of mathematical solutions (Szobol 1981). The method is widely used to simulate the likely outcomes of various events and their probability when input parameters are uncertain. In the model to be analyzed the influencing variables and their possible intervals, their likelihood distribution as well as the connections between the variables are fixed. The distribution values of the variables from the given intervals are developed by a random number generator. In the course of my research I examined the possibilities of planning the crop structure taking advantage of the benefits of simulation methods.

2. Material and methods

Bayesian statistic models are often used to model uncertain future events for example stock prices. In effect, this approach handles the unknown parameters in the given model as random variables and samples them from the distribution function based on my preliminary knowledge. Bayesian statistics were unaccomplishable with computer up till the nineties, but then Markov Chain Monte Carlo (MCMC) simulation methods (Metropolis – hasting and Gibbs sampling) received greater emphasis. Translating the problems into Bayesian statistical language based on MCMC is a very difficult task as it demands significant programing and mathematical knowledge as well as a skill in operating random number generators. For the realization of this process, a so-called BUGS (Bayesian interference Using Gibbs sampler) program was developed by the Biostatistics Sub-department of the Cambridge Medical Research Council in 1995 (Spiegelhalter et al. 1996), the

programing language of which is very specifically adjusted to the realization of pure likelihood models based on MCMC. Its version running under Windows operational system is known as WINBUGS.

2.1. MCMC (Markov Chain Monte Carlo Simulation)

Let Y be a mass of facts, $\theta = (\theta_1,...,\theta_k)$ vector of random variables (model parameters), with $\pi(\theta,Y)$ joint distribution function. Let us also suppose that $\pi(\theta,Y)$ is complicated and can hardly be given in an analytical form. In terms of the Bayes statistics, $\pi(\theta \mid Y) \propto \pi(\theta,Y)$ and

$$\pi(\theta \mid Y)\, \alpha \pi(Y \mid \theta)\pi(\theta) \tag{1}$$

is valid (Congdon 2007), so it is easier to work with this function instead of the joint distribution function. On the other hand, my interest is raised by $\pi(\theta \mid Y)$ distribution because the Y data are given, and I want to use those to estimate the parameters. $\pi(Y \mid \theta)$ is the so-called Likelihood function and $\pi(\theta)$ is the apriori distribution function which I know in advance. Let us also suppose that I'm looking for the expected value of a $h(\theta)$ integrable function, this is given by the following integral based on the $\pi(\theta \mid Y)$ distribution function (Jorgensen 2000):

$$E_\pi(h(\theta)) = \int h(\theta)\pi(\theta \mid Y)d\theta \tag{2}$$

It is nearly impossible to calculate this integral in an analytical or numeric way. That is why I use the so-called Monte-Carlo integration, the essence of which is that I take a $\theta^{(0)},...,\theta^{(k)}$ sample from $\pi(\theta \mid Y)$ distribution, and so I can estimate the expected result in the following way (David-Scollnik 2001):

$$E_\pi(h(\theta)) \approx \frac{1}{k}\sum_{i=1}^{k} h(\theta^{(i)}) \tag{3}$$

The combination of the Monte Carlo integration and the Markov Chains is called MCMC simulation, the essence of which is that I simulate possible variations from a Markov chain, the stationary function of which is $\pi(\theta \mid Y)$. It is true that these random samples $\theta^{(0)},...,\theta^{(k)}$ will no longer be independent, but with moderate regularity conditions it is realized that the distribution of $\theta^{(i)}$ converges on $\pi(\theta \mid Y)$ in the case of $i \to \infty$, and as for the expected result it is valid that (David-Scollnik 2001):

$$E_\pi(h(\theta)) \approx \frac{1}{k}\sum_{i=1}^{k} h(\theta^{(i)}) \text{, in the case of } k \to \infty \tag{4}$$

Now the only question remaining is how I can simulate Markov chain variations, the stationary distribution of which is $\pi(\theta \mid Y)$. Researchers have come up with various methods to accomplish this, one of the most simple techniques is the Gibbs sampling process.

At the beginning of the process I start from the original parameter vector $\theta^{(0)} = (\theta_1^{(0)},...,\theta_k^{(0)})$. I take random samples from the so-called full conditional distribution in the following way (Congdon, 2007):

$$\begin{aligned}
&\theta_1^{(1)} \approx \pi(\theta_1 \mid Y, \theta_2^{(0)},...,\theta_k^{(0)}) \\
&\theta_2^{(1)} \approx \pi(\theta_2 \mid Y, \theta_1^{(1)}, \theta_3^{(0)},...,\theta_k^{(0)}) \\
&\theta_j^{(1)} \approx \pi(\theta_j \mid Y, \theta_1^{(1)}, \theta_{j-1}^{(1)}, \theta_{j+1}^{(0)},...,\theta_k^{(0)})
\end{aligned} \tag{5}$$

In fact, this realizes a series of steps by which I go from $\theta^{(0)}$ to $\theta^{(1)}$ parameter vector. After a finite number of iterations I arrive at the Markov chain mentioned above, and the expected result can also be calculated in way (3). Generally speaking I arrive at the desired conditions after a few thousand simulation runs, but often tens of thousands of iterations are needed.

Bayes statistics and simulation methods have a lot in common. Researchers prefer to use these methods in their simulation models primarily because they have an opportunity to build their preliminary knowledge into the model so that this will be able to change during application, influenced by several other factors. David Vose (2006) presents this in his work as a highly efficient analytical tool, based on the Bayes principle and being suitable for estimating parameters on the basis of data more effectively than other methods. The Bayes conclusion theory involves 3 important circumstances (Vose 2006):

- determination of the a priori distribution functions and their parameters,
- determination of suitable likelihood function on the base data,
- determination of posteriori distributions and their parameters

2.2. Stochastic LP model

The general form of the model is as follows:

$$\begin{aligned}
&A\bar{x} \le \bar{b} \text{ (restrictive conditions)} \\
&c\bar{x} -> \max \\
&0 \le \bar{x} \text{ consistent distribution on (0,3) interval}
\end{aligned} \tag{6}$$

In formulae 6 the capacity vector is denoted with „$\bar{b}$" and „A" indicates the technological matrix. The solution vector „$\bar{x}$" consists of variables which are the sowing areas of the different crops in hundred hectares. The goal function coefficients are denoted with „c", they mean the per unit incomes.

The LP task can be considered stochastic, because the goal function coefficients (the values of the per unit incomes) come from a distribution. In the calculation of their values I take into consideration the selling price and quantity of the main product, which also come from a distribution.

In the marking system of MCMC „Y" is the data set, „A" is the technological matrix and „$\bar{b}$" is the capacity vector. c=$h(\theta)$, where θ contains the variables of prices, average yield and area. The $\pi(\theta \mid Y)$ distribution function takes a special form on account of the restrictive conditions:

$$\pi(\theta \mid Y) = \begin{Bmatrix} 1, if\ A\bar{x} \le \bar{b} \\ 0, if\ \bar{b} < A\bar{x} \end{Bmatrix} \tag{7}$$

Because of this, goal function values are formed during the simulation runs if the restrictive conditions are fulfilled. This way the stochastic linear programing task can be rendered into the language of WINBUGS simulation software. WINBUGS saves the θ values (the values of price, average yield and area variables) as well as the goal function values generated during all the simulation runs. On the basis of this, it is possible to find the maximum value for the goal function.

2.3. Optimality Criterion

Since the maximum value does not equal with the optimal value, I had to select an "Optimality Criterion" for evaluating the performance of the applied softwares as well. I formulated the criterion according to Prekopa's work (Prékopa et al. 1980). Based on this study a solution vector x_k is considered optimal if the difference between the values of the goal function at x_{k+1} and x_k does not exceed 1% of the latter and at the same time each individual component of x_{k+1} - x_k does not exceed 2% of the corresponding component of x_k (Prékopa et al. 1980).

3. Result and discussion

In my analysis I have used the data of an agricultural company farming in Löszhát, Hajdúság. The sowing area of the company is 800 hectares, crops grown are winter wheat, maize, winter coleseed, sunflower and green peas. The technology used by the company was built into the model.

When giving the restrictive conditions, I took into consideration the resources available for the company at the time, as well as the professional rules pertaining crop rotation. These per unit data are deterministic in the model.

In the goal function there appears the profit contribution (per unit income). The per unit changing costs in the individual branches, just like the capacity vector and the per unit demands, may be considered fixed. Within the model, i consider the average yields and the selling prices, that is, the return from sales, probability variables.
I determined the distributions of prices and their parameters taking into consideration the time series sales data of the company being analyzed (Table 1):

Table 1. The distributions of prices applied during the simulation

Crop	Distribution	Parameter 1	Parameter 2
Maize	Gamma	17,91	0.83
Wheat	Gamma	39,3	0.33
Coleseed	Gamma	1,54	6.50
Green peas	Normal	50	20.10
Sunflower	Normal	65	33.3

Source: own calculation

I applied the distributions and parameters of average yields based on the farm data of ARI (Agricultural Research Institute), North Plains Region for the years 2001-2005 (Table 2).

Table 2. The distributions applied to the average yields during simulation

Crop	*Distribution*	*Parameter 1*	*Parameter 2*
Maize	Normal	7.55	1.75
Wheat	Normal	4.54	1.15
Coleseed	Gamma	9.53	0.18
Green peas	Normal	5.56	1.25
Sunflower	Normal	2.18	0.61

Source: own calculation, based on data provided by Agricultural Research Institute

After the simulation runs, performed by 3 different programs, I analyzed the goal function with the help of statistical tools (Table 3). According to the measurement of risk by the decision maker, the appropriate decision alternative can be choosen based on the data set. Risk and Crystal Ball programs use the simpler version of the Monte-Carlo simulation, while the WINBUGS applies the Markov Chain Monte Carlo simulation combined with Bayesian Statistics.

Table 3. Characteristics of the profit contribution

Statistic	***CB***	***Risk***	***Winbugs***
Median	42250,81	44025,17	43450,04
Average	43653,99	44398,14	44401,75
Deviation	17089,83	14982,86	18760,57
Relative Deviation	39,15	33,75	42,25
Minimum	3173,57	7807,51	-10350,00
Maximum	102583,12	85510,90	103100,00
Range	99409,55	77703,39	113450,00
Skewness	0,34	0,09	0,29
Kurtosis	-0,27	-0,63	-0,30

Source: own calculation

Table 3 clearly suggests that the widest range, the highest deviation and the highest goal function value can be gained by WINBUGS. An other significant difference between WINBUGS and the two other programs is that WINGUGS can produce negative result for the profit contribution.

Table 4. Comparison of the crop structure and goal function values

Factor	***Applied Softwares***						***Deterministic Linear Programming***
	Best result (highest value)			***Best ten results***			
	Crystal Ball	***Risk***	***Winbugs***	***Crystal Ball***	***Risk***	***Winbugs***	
Goal function*	102583,12	85510,90	103100,05	92481,95	81307,90	97164,02	88460,24
Maize**	2,40	2,43	2,03	2,14	2,36	1,98	2,46
Wheat**	1,92	2,33	2,12	1,88	2,18	1,82	2,36
Coleseed**	0,78	0,78	0,22	0,74	0,44	0,49	0,00
Green peas**	0,16	0,51	0,21	0,33	0,41	0,27	0,53
Sunflower**	1,25	1,77	1,56	1,27	1,81	1,09	2,00

Source: own calculation;*: value in Hungarian Forint; **:sowing area in hundred hectares

Evaluating the results of the simulation runs, it can be observed that by using Risk and Crystal Ball the simulation resulted in higher values for profit contribution than the deterministic linear programming optimum in eight cases, while in the case of WINBUGS in seventeen cases. Table 4 shows the best and the average from the best ten results of the applied softwares, and all the relevant values of the decision

variables. For the sake of the better interpretation, I calculated the standard deviation from the deterministic linear programming solution (Table 5).

Table 5. Standard deviations from the deterministic solution

Factor	Applied Softwares					
	Best result			Best ten results		
	Crystal Ball	Risk	Winbugs	Crystal Ball	Risk	Winbugs
Goal function	0,16	-0,03	0,17	0,05	-0,08	0,10
Maize	-0,06	-0,03	-0,43	-0,32	-0,10	-0,49
Wheat	-0,44	-0,03	-0,24	-0,48	-0,18	-0,54
Coleseed	0,78	0,78	0,22	0,74	0,44	0,49
Green peas	-0,37	-0,02	-0,32	-0,21	-0,13	-0,26
Sunflower	-0,75	-0,23	-0,44	-0,73	-0,19	-0,91

Source: own calculation

Taking the best result in table 5 into consideration, it can be stated that Risk produces the closest result (goal function and sowing areas) to the deterministic linear programming optimum and the solution of WINBUGS and Crystal Ball differs the most from it. Based on the best ten results in table 5, it is obvious that WINBUGS produces the most different result from the deterministic model taking the highest values into consideration and the object function value is the highest in the case of this software.

Table 6. Comparison of the crop structure and goal function values according to Prékopa's optimality criterion***

Factor	Applied Softwares						Deterministic Linear Programming
	Best optimal solution			All optimal solutions			
	Crystal Ball	Risk	Winbugs	Crystal Ball	Risk	Winbugs	
Goal function*	56940.09	57642.19	73870.00	43978.50	52321.20	47212.86	88460,24
Maize**	1.53	1.40	1.69	1.25	1.56	1.38	2,46
Wheat**	1.74	1.77	1.53	1.18	1.32	1.18	2,36
Coleseed**	0.27	0.24	0.30	0.50	0.45	0.40	0,00
Green peas**	0.31	0.34	0.23	0.24	0.30	0.27	0,53
Sunflower**	1.07	1.60	1.58	1.01	1.07	0.97	2,00

Source: own calculation;*: value in Hungarian Forint; **:sowing area in hundred hectares
***: Prékopa et al. (1980)

Table 6 shows the best and all optimal solutions of the applied softwares, and all the relevant values of the decision variables. The optimal solution of the Crystal Ball and Risk Softwares are almost the same, while WINBUGS produce the closest value to the deterministic optimum. Regarding the average of all the optimal solutions, the three softwares produced almost the same solution for the crop structure.

4. Conclusion

Traditional planning is still the most often applied method in cultivation, which provides adequate planning, but also determines an increasing shortfall in economic competition. Due to price and yield fluctuations a methodologically appropriate optimizing planning is necessary. In optimizing planning, linear programing models are most often used, however, because of their deterministic nature, in choosing from among decision alternatives, risks cannot be taken properly into consideration. Applying simulation models may be a solution, in my work I have presented one such application. Out of the three programs, WINBUGS produces the most different result from the deterministic model taking the highest values into consideration and the object function value is the highest in the case of this software, as the software uses Bayesian Statistics. The range of the object function values produced by WINBUGS is wider than in the case of the two other programs. Risk Software produces the closest result (goal function and sowing areas) to the deterministic linear programming optimum. Regarding the optimal solutions, Crystal Ball and Risk Softwares produced almost the same values, while WINBUGS produce the closest value to the deterministic optimum. Regarding the average of all the optimal solutions, the three softwares produced almost the same solution for the crop structure. To sum up, risks and optimal solutions can be modelled and considered in a more widespread and accurate way by using WINBUGS software.

What the crop structure concerns I can state that maize and wheat are of grater importance as these crops constitute a large part in the crop structure. The growing of sunflower might become more important due to the large increase in biodiesel production.

References

Balogh, P. –Bai, A. –Posta, L. –Búzás, F. 1999: Analysis of environmental effects of cereals and energy plantations in case of unfavourable agricultural conditions,

Conference on „The resources of the environment and the sustained development", University of Oradea, Oradea.

Congdon, P. 2007: Bayesian statistical modelling. Wiley, New York.

Csipkés, M. – Etsey, I. – Nagy, L. 2008: Vetésszerkezeti variánsok összehasonlító elemzése a Hajdúságban. XI. Nemzetközi Tudományos Napok, Gyöngyös, 14-21. p.

Dantzig, G. B. 1955: Linear programming under uncertainty. Management Sci. 1, 197–206. p.

David, P. – Scollnik, M. 2001: Actuarial modeling with MCMC and BUGS. North American Actuarial Journal, 5/2, 96-124. p.

Jorgensen, E. 2000: Monte Carlo simulation models: Sampling from the joint distribution of "State of Nature"-parameters. In: Fels, V.D. – Klerx, I. – Mourits, M. (eds): Proceedings of the Symposium on "Economic modelling of Animal Health and Farm Management", Farm Management Group, Dept. of Social Sciences. Wageningen University, Wageningen, p. 73-84. http://gbi.agrsci.dk/~ejo/publications/preprints/specprior.pdf.

Prékopa, A. – Ganczer, S. – Deák, I. – Patyi, K. 1980: The STABIL stochastic programming model and its experimental application to the electricity in Hungary In: Dempster, M. A. H. (ed): Stochastic Programming, Academic Press, London, 369-385. p.

Prékopa, A. 2003: Probabilistic programming. In: Ruszczynski, A. – Shapiro, A. (eds): Stochastic Programming Handbooks Operation Research Management Sci. 10, Elsevier, Amsterdam, 267-351. p.

Russel, R. S. – Taylor, B. W. 1997: Operations Management, Focusing on quality and competitiveness. Prentice Hall, New Jersey, 610-613. p.

Spiegelhalter, D. – Thomas, A. – Best, N. – Gilks, W. 1996: Bayesian inference using Gibbs sampling program Version 0.50. MRC Biostatistics Unit of Cambridge, UK, http://www.mrc-bsu.cam.ac.uk/bugs/documentation/Download/manual05.pdf.

Szobol, M. 1981: A Monte-Carlo módszerek alapjai. Műszaki Könyvkiadó, Budapest.

Vose, D. 2006: Risk analysis, a quantitative guide. Wiley, New York.

Williams, A. 1966: Approximation formulas for stochastic linear programming. SIAM J. Appl. Math. 14, 668–877. p.

Using a multiperiodic linear programming model and a simulation programme for competing field crops and energy orchards

Margit Csipkés[1] *– Lajos Nagy*[2]

Considering the land use of Hungary, there is a need to develop a rationale land use in which, beside the less-favoured areas, the use of set-aside areas are also permitted. There are several opportunities to utilize the less-favoured areas.

We prepared a multiperiodic linear programming model in order to model the crop structure, in which field crops with woody energy orchards were also competed. After having each field and orchard technology compiled, we set the dynamic simulation model, that we prepared in MS Excel. After running the model we analyzed the shadow prices of the constraints and the marginal cost of variables. Considering the result of the analysis and the professional information we made a sensitivity analysis, which gave a basis to create new decision variants. The results of linear programming model were compared with those of Monte Cralo simulation's, where we managed the enterprises' profit contribution as probability variable with normal distribution in the course of modelling.

Keywords: simulation, linear programming, rational land use, biomass

1. Introduction

The aim of this research work is the analysis of the joint applicability of a multiperiodic linear programming model – *LP model* – and the *Monte Carlo* simulation. We used the analysis to plan the medium-term (6 years) crop structure of biomass products for energy purposes.

Reviewing the domestic land use, in the past 80 years there has been significant changes. The rate of forest and set-aside areas has been constantly increasing from 1950, and the rate of arable lands that are registered as non-cropped or waste land year by year are also significant.

[1] Margit Csipkés, PhD student, University of Debrecen CASE Faculty of Applied Economics and Rural Development Department of Economic Analysis Methods and Statistics (Debrecen)

[2] Dr. Lajos Nagy, lecturer, University of Debrecen CASE Faculty of Applied Economics and Rural Development Department of Economic Analysis Methods and Statistics (Debrecen)

Within the utilization of fields we must differentiate the *food production* and the *energy production*. Since food production must be always treated with fundamental significance, therefore this production objective must override the others.

As far as choosing the type of agricultural production is concerned, it is important to analyze that on those fields, where for the producer cannot or can only hard put up the money for the production costs, gainfully sustainable – with less material and energy expenditure - production types have to come to the front. This must be connected with the determination that such a system should not be established and operated that goes with more greenhouse gas emission than the previous systems.

On those agricultural lands, where economic production cannot be executed one possible land use method is *to plant energy orchards*. This is also backed by a study published by the Agricultural Economics Research Institute, that is to say that besides the production of major field crops - mostly cereals and oilseeds – it is important to produce energy crops if we see the bioenergy-production's increased need for raw materials (Udovecz et al. 2007).

In many agricultural studies (Erdős – Klenczner 2000, Gergely 2000a, Gergely 2000b, Erdős 2007) can be read that the agricultural strategy does not deal with sufficient energy orchards as an option for *alternative land use*. One reason for this that producers would not want to adopt orchard planting requiring major employment – compared to crop cultures – on lower quality, less-favoured areas and fields being in the red. The other reason is that the majority of producers see the tasks of agriculture in food production, the result and the income cannot be demonstrated yearly, and because of lack of information farmers are averse from wood production.

The most frequent plants of *energy orchards with short cutting-cycle*, which are grown on fields, are locust, poplar and willow. Under intensive production circumstances all three species can produce high yields; however, they are very sensitive for the endowment of the production site. The average lifespan is about 20 years with harvest in every 1-3 years. They are suitable to reach dependably high yields (8-30 tons/ha) under proper site conditions (Table 1.).

Table 1. Yields and characteristic data of woody energy plants in Hungary

Wood species	Energy content MJ/kg	Average yield kg/ha/year	Moisture content %	Cutting cycle year	Lifespan year
Locust	14,8	7900	15	3	20
Poplar	15,1	20000	15	2	20
Willow	14,8	30000	15	1- 3	25

Source: Marosvölgyi (1998), Führer et al. (2003), Bai (1999), Bai et al. (2002), Defra (a), Gergely (1988)

Beside woody energy orchards, herbaceous plants are of also great significance, from which hemp, Chinese reed and various energy grasses are being dealt with in Hungary. The yield and energy content to be expected are shown in Table 2.

Table 2. Yields and characteristic data of herbaceous energy plants in Hungary

Plant	Energy content MJ/kg	Average yield kg/ha/year	Moisture content %	Cutting cycle year	Lifespan year
Energy grass	15,0	13000	15	1	15
Chinese reed	13,0	17000	15	1	15
Hemp	15,7	12500	10	1	1

Source: Janowszky (2002), Iványi (2001)

Within the energy production cereals, corn, rape and various agricultural by-products will be of greater significance. Yield and energy content data are shown in Table 3.

Table 3. Yield and characteristic data of some field crops in Hungary

Plant	Energy content MJ/kg	Average yield kg/ha/year	Moisture content %	Cutting cycle year	Lifespan year
Cereal grains	14,0	7000	15	1	1
Cereal straw	14,3	6000	15	1	1
Whole cereal	14,1	13000	15	1	1
Rape	14,2	3500	15	1	1

Source: Hartmann-Kaltschmidt (2002)

2. Methodology of modelling

2.1. Short introduction of the applied linear programming model

We made a multiperiodic linear programming model for modelling the production structure, in which we competed the wheat, corn, turnsole and winter colza among field crops, and among woody energy plants locust, poplar and Swedish willow. This is a dynamic simultaneous model that we made by *Microsoft Excel* and the details of technological matrix for each year is in the diagonal of the table. The breakdown of the technology is made by months. The time interval is 6 years that was chosen because of the energy orchards' specific characteristics. Since the linear pro-

gramming model contains *60 variables* and *160 constraints*, thus this model can be only presented in the annex of this study. The variables of the model were given for the planned area under cultivation, planned amount of intermittent labour in shift hours, for the given machine types, and for the number of leased shift hours. The constraints for each year are provided for the area, for machinery work, for leased machinery work, for the available labour force and leased labour force.

In the model we assume that the area of the planted woody energy plants in the beginning of the 6-year-period remains unchanged:

$$x_j^k = x_j^{k+1} \text{ thus}$$

$$x_j^k - x_j^{k+1} = 0$$

k=1,2,…6 : the ordinal number of the actual year

The above variables connect each year together, so these can be named as transfer variables. It was necessary, because in case of orchards the size of the area in the first year has to be run through the period of 6 years.

We look for the *maximum profit contribution* in the model for the whole period – which is now 6 years.

$$\sum_{j=1}^{n} \sum_{k=1}^{6} (T_j^k - C_j^{k\,\text{variable}}) x_j^k - \sum_{k=1}^{6} C_i^{hk} \delta_i^{hk} \Rightarrow \max!$$

T_j^k : Production value of the j^{nth} enterprise in the k^{nth} year;

$C_j^{k\,\text{variable}}$: Variable cost of the j^{nth} enterprise in the k^{nth} year;

x_j^k : Size of the j^{nth} enterprise in the k^{nth} year;

C_i^{hk} : Additional cost of the h^{nth} machine compared to the own resource in the k^{nth} year;

δ_i^{hk} : Number of leased hours for the i^{nth} period in the k^{nth} year from the h^{nth} resource for lease work.

In case of field cultures we built into the model a rate of 4% income change between each year. For orchards the profit contribution of each year changed according to the harvest of the cultures. We calculated with 3 years cutting cycle in case of locust and Swedish willow and with 2 years for poplar.

2.2. Theoretical background of the simulation model

The simulation model is such a simplified mathematical implementation of a real system that studies the behaviour of the original system under changing different conditions, circumstances. In spite of the accurate results provided by analytical models, the simulation process involves the model's run over time and execution to provide representative samples about performance indicators describing the operation of the system (Winston 1997). *Stochastic* and *deterministic models* are differentiated. Randomness is not built into the deterministic models (Kovács et al. 2007a). The gist of stochastic simulation is that we choose values randomly according to the probability distribution assigned to individual uncertain factors that are used in the experiments of the simulation analysis. (Russel-Taylor 1998). In the model to be analysed we set the influence variables and their possible intervals, their probability distributions and the relationship among the variables. The given interval and distribution values of variables are formed by random number generator. The model is run several times in a row, generally *1000-10000 times* and thus we get an expected value and a variance range for the result variable to be determined. By the distribution function the probability can be determined that the value of the given variable will be in the given interval (Kovács et al. 2007b). Most recently income is given as a result variable in the models and the risk is observed at which probability it will be above the value or below. By increasing the number of runs the distribution of the result variable can be given by arbitrary accuracy (Watson 1981, Jorgensen 2000):

$$\psi = E_{\pi}\{U(X)\} = \int U(x)\pi(x)dx \quad \text{, where}$$

$X = \{\theta, \phi\}$ means vectors containing θ decision parameters and ϕ state parameters, and π means x distribution. U(x) is a utility function which means usually the income, Eπ() function gives the expected utility by given distribution.

The advantage of the method is that the model is run for decision variants individually as well, and the risk of different decision variants can be compared. The following formula is applied for the numerical determination of their integral value (Jorgensen 2000):

$$\overline{\psi} = \frac{1}{k}\left\{U(x^{(1)}) + \ldots + U(x^{(k)})\right\}, \text{ where}$$

k means the number of experiments, i.e. the number of runs.

Excellent, easily manageable simulation softwares can be used, for example *Crystal Ball* (Decisioneering, Inc.) and @Risk (Palisade Corporation). These are based on the well-known Excel spreadsheet programme. The model to be applied

will be set up here, which parameters can also be stochastic. The parameters' distribution can be chosen from several distribution types. After the running, the simulation gives the distribution of the result variable, by which it can be stated that at which probability the examined variable will take its value in a given interval.

In the course of our research we applied the Crystal Ball programme package. By this programme we had the opportunity to find the *marginal value of the objective function* and to give the constraints of the resources within the optimal crop structure.

3. Database of the analysis

The database of the analysis is given by Szabolcs-Szatmár-Bereg County, which is located in the north-eastern part of Hungary, amounting 6,4 percent of the national territory. It is the sixth largest county of the country. The area borders on three countries – Romania on the east, Ukraine on the north-east and Slovakia on the north.

Considering the geographic endowments, the following regions belong to this county: the very eastern part of the Great Plain, 78 percent of Nyírség, among the small areas of upper Tisa the whole Rétköz, the flats of Szatmár and Bereg, and the western part of the wetland of Ecsed.

The characteristic climate of this county is continental with the annual mean temperature between 9,0-9,5°C. The annual precipitation was about 400-550 millimetres according to the means of the latest years. The winter is colder and longer usually compared to the Great Plain.

The water requirements of the county are given by mostly artesian water and less by rivers. Its largest river is the Tisa that enters the country at Tiszabecs. Other major watercourses are Szamos, Túr, Kraszna and Lónyai sewer.

In agricultural terms of the county, the acreage is 623.000 hectares, from which 82 percent is production area. Within crop production cereals have a significant role, since compared to the national data 8,6 percent of corn, 25 percent of rye, 8,5 percent of potato, 9,2 percent of turnsole and 80 percent of tobacco are produced in this county. Nationally the production of cabbage, sweet corn, tomato, water melon and cucumber is of definite value.

In the county fruit production is done on 33.000 hectares, which means 34 percent of the country's territory. More than 50 percent of the national produce is yielded in this county. Apple has been of great significance for decades.

Considering forest management, 105.000 hectares from the county's territory are dealt with forest, from which locust and noble poplar are of great significance. These give 47 and 22 percent of all wood types in the county.

The database of our research work was given by a holding of this county with average natural endowments. When we set up the model we used the applied technology of this holding. The holding manages 500 hectares, so we took this also into

consideration. In our analysis we dealt with corn, wheat, winter colza, turnsole, locust, poplar and willow among the crops of alternative crop production.

4. Cost and income calculation

We competed each enterprise according to their profit contribution value in our model. Profit contribution was calculated by the difference of production value and variable costs. The choice of this income category was necessary, because our aim was the *income maximalisation*.

For the calculation of production value we took into account the turnover, the financial assistances and other incomes as well. Raw data for the income was calculated when drafting the enterprise technologies, which is calculated by the multiplication of yield values per hectares and marketing prices. We considered the support levels by the actual laws, and the contractual and estimated prices for the marketing. The financial assistance has two parts: Single Area Payment Scheme (SAPS) and the national envelope Top-up. Among the crop enterprises of the holding we calculated with these two payments for all arable and energy crops.

We took into account the variable costs in case of doing the *enterprises'cost calculation*. We calculated with material costs, labour costs, machinery costs and other direct costs as well. Material costs are the cost of seeds, cuttings' costs, cost of fertilizer and pesticides. Labour costs could be assigned to each enterprise. Special attention was given to the costs of fuel, upkeep and repair as the costs of machinery.

The determination of *profit contribution* was done after the calculation of incomes and costs. Corn had the highest value among arable crops, approximately 143.000 Ft/hectares. Among the examined crops, wheat (128.000 Ft) and winter colza (125.000 Ft) were the second, and after that the turnsole (87.000 Ft) (Table 4.).

Table 4. Profit contribution value of arable crops for one hectare in the examined years

Plant	PC Ft/ha
Year 1	
Corn	147 811
Turnsole	87 906
Winter wheat	128 026
Winter colza	125 147
Year 2	
Corn	149 289
Turnsole	88 785
Winter wheat	129 306
Winter colza	126 398
Year 3	
Corn	150 782
Turnsole	89 673
Winter wheat	130 599
Winter colza	127 662

Plant	PC Ft/ha
Year 4	
Corn	152 290
Turnsole	90 570
Winter wheat	131 905
Winter colza	128 939
Year 5	
Corn	153 813
Turnsole	91 475
Winter wheat	133 224
Winter colza	130 228
Year 6	
Corn	155 351
Turnsole	92 390
Winter wheat	134 557
Winter colza	131 531

Source: own creation

After having determined the value of the profit contribution for the first year, we calculated with a *5 percent of income increase and a 4 percent of cost increase*. The base of the given annual profit contribution is given by the income and cost data of the previous year.

In the course of defining the profit contribution of energy orchards we must take into consideration that *harvesting* does not happen in every year. In case of energy orchards in the years when harvesting is done the profit contribution will be positive, in every other cases we calculate with negative values. Of course, there are exceptions as well. Since we calculated with the amount of assistance also in every year, thus there were some years when we got positive profit contribution. The harvesting cycle for locust and willow is 3 years, in the case of poplar it is 2 years.

The profit contribution values of energy orchards are shown in Table 5. As for the arable crops as well, after having determined the value of the profit contribution for the first year, we calculated with a *5 percent of income increase and a 4 percent of cost increase*. The base of the given annual profit contribution is given by the income and cost data of the previous year.

Table 5. Profit contribution value of energy orchards for one hectare in the examined years

Wood species	PC Ft/ha
Year 1	
Locust	-505 335
Poplar	-423 674
Willow	-531 170
Year 2	
Locust	672
Poplar	134 224
Willow	-2 154
Year 3	
Locust	356 525
Poplar	-182 789
Willow	503 738

Wood species	PC Ft/ha
Year 4	
Locust	-157 188
Poplar	232 401
Willow	-149 911
Year 5	
Locust	24 745
Poplar	-179 151
Willow	24 745
Year 6	
Locust	403 243
Poplar	232 401
Willow	604 233

Source: own creation

5. Evaluation of the results

5.1. The evaluation of the basic model made by linear programming

In the course of our calculations we made a multiperiodic linear programming model, in which we competed among the arable crops the wheat, the corn, the turnsole and winter colza, whereas among woody energy orchards the locust, the poplar and the Swedish willow. The model was made in MS Excel.

We set up the basic model (*LP_BASIC*) according to the initial profit contributions (Table 4. and 5.), than after analysing the shadow prices we ran 3 variants, where we modified the profit contribution of turnsole (*LP_TURN*), the locust (*LP_LOCUST*) and the poplar (*LP_POPLAR*).

The production structure that we got after having solved the basic model can be seen in Table 6.

Table 6. Production structure for 6 years

Name	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Corn	2,50	2,50	2,50	2,02	2,50	2,50
Turnsole				0,62		
Winter wheat	0,92	0,96	0,96	0,78	0,96	0,96
Winter colza	1,00	0,96	0,96	1,00	0,96	0,96
Locust						
Poplar						
Willow	0,58	0,58	0,58	0,58	0,58	0,58

Source: own creation

As we can see in this table, in the production structure corn has an absolutely superiority compared to the other crops for the *period of 6 years*. Corn used up at 100 percent its available maximum field of 250 hectares. The exception was the 4th year, when its area reduced by 48 hectares, and turnsole got into its place by the area of 62 hectares. The winter wheat and the winter colza are on the 2nd and 3rd place with the area of 92-96 hectares. The smallest area has the Swedish willow with the area of 58 hectares.

The maximum *profit contribution* that can be reached by running this model *for 500 hectares and 6 years is approximately 388 million Ft.*

Analysing the basic model's production structure (Table 6.) we can see that the turnsole, the locust and the poplar did not get into. We must analyse the values of profit contribution if we want these crops to be competitive compared to the other field crops. The sensitivity analysis' table for variable cells provides us assistance for this. From this table we can read the shadow prices, marginal costs of the activities. It gives the information about why an activity did not get into the production structure and when it can get into the optimal solution. Besides, it shows with how much the coefficient of an activity must be increased in order to get into the production structure without the decrease of the objective function's value.

We summarized the table of variable cells of the first year's sensitivity analysis in Table 7.

Table 7. The summarized table of variable cells for the first year's sensitivity analysis of the linear programming model (100ha)

Name	Area under cultivation 100ha	Shadow price of the activity eFt/100ha	Objective function value eFt/100ha	Lower limit	Upper limit
Corn Year 1	2,50		14781,1	12556,5	14781,1
Turnsole Year 1		**-3501,94**	**8790,6**	**….**	**12292,6**
Winter wheat Year 1	0,92		12802,6	9300,6	13413,7
Winter colza Year 1	1,00		12514,7	11903,6	12514,7
Locust Year 1			**-50533,5**	**….**	**-18016,6**
Poplar Year 1			**-42367,4**	**….**	**21074,4**
Willow Year 1	0,58		-53117,0	-85633,9	-32361,1

Source: own creation

From this table it can be seen that turnsole (with 87.906 Ft/ha profit contribution) did not get into the production structure. In case if the value of profit contribution would be increased to 122.926 (Figure 1.), than beside the maximum use of the existing resources it would get into the production structure at the expense of wheat and rape (Table 8.).

For the affection of changes, among the woody orchards the willow's role in the production structure would change, since it would occupy the one-third of the total area. However, the locust and the poplar still did not get into the production structure. In accordance with the basic production structure the area of rape has decreased, since it would get into the management only in the fourth year.

Table 8. Production structure for 6 years if the objective function coefficient of turnsole is increased to the minimum threshold price (100ha)

Plant/Year	1st	2nd	3rd	4th	5th	6th
Corn	2,50	2,50	2,50	1,40	2,50	2,50
Turnsole	1,00	1,00	1,00	0,82	1,00	1,00
Winter wheat	0,50	0,50	0,50	0,78	0,50	0,50
Winter colza				1,00		
Locust						
Poplar						
Willow	1,00	1,00	1,00	1,00	1,00	1,00

Source: own creation

5.1.1. The change of the locust's PC value compared to the basic model

From Table 7. we can see that the locust did not get into the basic model's production structure, either. In case if we would like to get a *woody enery orchard* into the production structure instead of the turnsole, than the locust's profit contribution for the first year should be increased from -505.335 Ft to -180.166 Ft. This value was determined by the value of allowable increase in the basic model, which was 32.516 Ft per one hectare. The coefficient of the objective function and the allowable increase must be summed in order to get this value, so in this way we got 180.166 thousand Ft per 100 hectares. The lower profit contribution value was determined like this, and locust would be competitive with the other arable crops under these conditions.

In the present economic state there is not any opportunity for the growth of the locust's profit contribution, however, since the cost of locust cutting for one hectare is approximately 288 thousand forint. Further cost are the handling, cultivation, material and other cost elements. But in case of this reduce would happen in a certain way, than compared to the basic model the locust and the willow would get into the production structure with 34 and 24 hectares (Table 9.).

Table 9. Production structure for 6 years if the objective function coefficient of locust is increased to the minimum threshold price (100ha)

Plant/Year	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**
Corn	2,50	2,50	2,50	2,02	2,50	2,50
Turnsole				0,62		
Winter wheat	0,92	0,96	0,96	0,78	0,96	0,96
Winter colza	1,00	0,96	0,96	1,00	0,96	0,96
Locust	0,34	0,34	0,34	0,34	0,34	0,34
Poplar						
Willow	0,24	0,24	0,24	0,24	0,24	0,24

Source: own creation

5.1.2. The change of the poplar's PC value compared to the basic model

The poplar did not get into the basic model's production structure either (Table 7.). Its profit contribution value should be featured in the model by positive value (210.744 Ft/hectares). This major change is needed, because in the harvesting year this enterprise bears the lowest specific profit contribution.

Table 10. Production structure for 6 years if the objective function coefficient of poplar is increased to the minimum threshold price (100ha)

Plant/Year	1st	2nd	3rd	4th	5th	6th
Corn	2,50	2,50	2,50	2,02	2,50	2,50
Turnsole				0,62		
Winter wheat	0,92	0,96	0,96	0,78	0,96	0,96
Winter colza	1,00	0,96	0,96	1,00	0,96	0,96
Locust						
Poplar	0,34	0,34	0,34	0,34	0,34	0,34
Willow	0,24	0,24	0,24	0,24	0,24	0,24

Source: own creation

If the poplar's profit contribution would reach the minimum marginal cost value, than Swedish willow could also be in the planned production structure (Table 10.). The area sown of the other crops are the same with the values in the basic model.

5.2. Analysis of the simulation model

In the course of our work after structuring and running the multiperiodic linear programming model we made a simulation model consisting the same conditions as we set at the linear programming model. The results of the linear programming model and the Monte Carlo-simulation were compared. We managed the profit contribution of the enterprises as variables with normal distribution. We applied the *Crystal Ball* programme package by which we could find the extreme of the objective function and specify the constraints for the resources within the optimal crop structure.

In the simulation model we set the influence variable, their *probability distributions, the relation between variables*, and the *intervals of the elements' possible change*. We originated the interval and distribution values by random number generation.

In the course of the analysis we calculated with the same economic conditions as at the linear programming model. For the constraints we applied the available resources and observed the rules of rotation cycle. These specific data were *deterministic within the model*.

The profit contribution values of the enterprises are considered as *variates* which was determined by normal distribution in the first year, than from the next year the rate of the annual income and cost increases were considered.

The decision variants were the areas of the enterprises. In the course of making the simulation we choose values according to probability distribution assigned to each uncertainty factor. By using these parameters we made 10.000

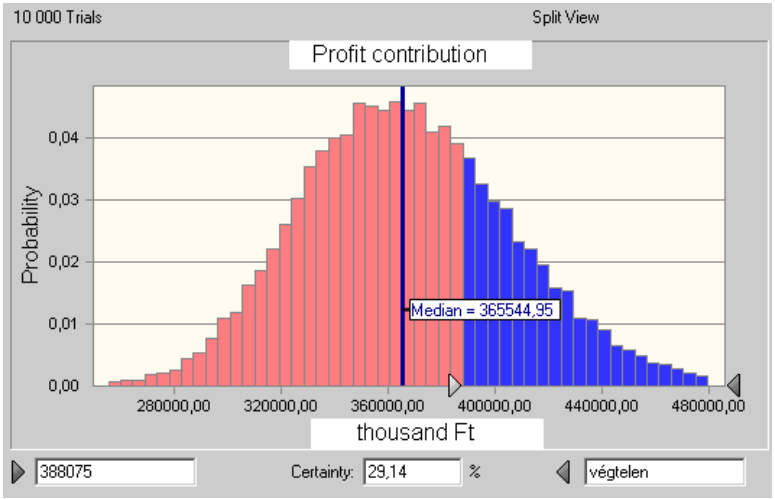
simulationsfor the sake of reaching the maximum profit contribution under the maximum use of the given resources.

Then we analysed the evolution of the objective function by statistical methods, and we created decision variants in accordance with the decision-maker's exposure.

We summarized the results of the linear programming model and the simulation in Table 11. Since we ran the model 10.000 times, we will not report each result in this paper. In the first three coloumns of the summary table the production structure with the highest (PC_max) and the lowest (PS_min) maximum profit contribution is given.

According to the given parameters, the value of the average profit contribution *with simulation* is 427.302 thousand Ft for 6 years and 500 hectares (Table 11., Table 12.), which probability of occurring is 59.9 percent. The value of the average profit contribution with the optimization of linear programming model is 388.075 thousand Ft for 6 years and 500 hectares, which probability of occurring is 80 percent. If we consider the production structure of the linear programming model as permanent when doing the simulation, than the probability of reaching the maximum PC value of the LP is 29.14 percent (Figure 1.).

Figure 1. The profit contribution values by the simulation running



Source: own creation

The median of the simulation model's profit contributions is 365 million Ft, i.e. we experienced at 50 percent of the runs less income than this value (in 6 years on 500 hectares).

The *maximum income* is 466.279 thousand Ft for 6 years and 500 hectares, however, the choice of production structure with higher income entails high risk. This is backed by the high value of relative variance that we got in the course of the simulation model's running.

Table 11. The summerized results of the linear programming model and the simulation

Name		PC_max	PC_mean	PC_min	LP
PC (thousand Ft/6years/500ha)		**466279**	**427302**	**123219**	**388075**
Year 1					
Corn	Area under cultivation (hectare)	250	97,5	121,8	250
Turnsole		-	100	-	-
Winter wheat		201,7	200,4	-	92,2
Winter colza		48,2	100	-	100
Locust		-	-	158,2	-
Poplar		-	-	120	-
Willow		-	2	100	57,8
Year 2					
Corn	Area under cultivation (hectare)	250	218	21,6	250
Turnsole		48,3	81,6	100	-
Winter wheat		201,7	198,4	0,2	96,5
Winter colza		100	100	-	95,7
Locust		-	-	158,2	-
Poplar		-	-	120	-
Willow		-	2	100	57,8
Year 3					
Corn	Area under cultivation (hectare)	250	128,9	21,6	250
Turnsole		100	100	-	-
Winter wheat		50	201,7	100,2	96,5
Winter colza		100	67,4	-	95,7
Locust		-	-	158,2	-
Poplar		-	-	120	-
Willow		-	2	100	57,8

Source: own creation

Table 12. The summerized results of the linear programming model and the simulation

Name		PC_max	PC_mean	PC_min	LP
PC (thousand Ft/6years/500ha)		**466279**	**427302**	**123219**	**388075**
Year 4					
Corn	Area under cultivation (hectare)	247,1	169,6	21,8	202,2
Turnsole		-	65,2	100	61,7
Winter wheat		152,8	163,1	-	78,3
Winter colza		100	100	-	100
Locust		-	-	158,2	-
Poplar		-	-	120	-
Willow		-	2	100	57,8
Year 5					
Corn	Area under cultivation (hectare)	249,8	110,6	21,6	250
Turnsole		-	85,6	-	-
Winter wheat		150,1	201,7	0,2	96,5
Winter colza		100	100	100	95,7
Locust		-	-	158,2	-
Poplar		-	-	120	-
Willow		-	2	100	57,8
Year 6					
Corn	Area under cultivation (hectare)	242,8	199,3	121,8	250
Turnsole		-	37,8	-	-
Winter wheat		201,7	201,7	-	96,5
Winter colza		55,5	59,2	-	95,7
Locust		-	-	158,2	-
Poplar		-	-	120	-
Willow		-	2	100	57,8

Source: own creation

6. Evaluation of results

The aim of our research is to provide assistance for the persons in agriculture in supporting their decisions by methods that are suitable for the description of system relations that requires the simultaneous system-orientated consideration of many variables.

In the course of our work we made a multiperiodic linear programming model and a simulation model for modelling the crop structure, in which we competed ar-

able crops and woody energy orchards as well. After having run the linear programming model we analysed the shadow prices of constraints and the marginal cost of the variables. Considering the results of the analysis and professional information we made a sensitivity analysis, which gave a basis to create new decision variants.

In our research we analysed that by how much profit contribution can be considered the energy orchards as competetive beside arable crops.

References

Bai, A. 2002: A biomassza felhasználás. Szaktudás Kiadó Ház. ISBN 963 942 46 0 Budapest

Erdős, L. – Klenczner, A. 2000: A fatermelés szerepe és lehetőségei a távlati földhasználatban. Gazdálkodás XLIV, 4, Budapest

Erdős, L. 2007: Ültetvényerdők szerepe a távlati földhasználatban. Előadás és tanulmány a Debreceni Agrártudományi Egyetem Nemzetközi Agrár- és Vidékfejlesztési Konferencián; Gazdálkodás 51, 4, Debrecen

Führer, E. – Rédei, K. – Tóth, B. 2003: Ültetvényszerű fatermesztés. Mezőgazda Kiadó és az ERTE közös kiadása, Budapest

Gergely, S. 2000a: Az Észak-magyarországi régió energiaerdő programjának lehetőségei és korlátai I. Gazdálkodás 2000, 6, Budapest

Gergely, S. 2000b: Az Észak-magyarországi régió energiaerdő programjának lehetőségei és korlátai II. Gazdálkodás 2000, 7, Budapest

Golberg, D. E. 1989: Genetic Algorithms in Search. Optimization and Machine Learning. Addision-Wesley

Iványi, I. 2001: A szántóföldi rostnövény-termesztés lehetséges szerepe az energiatermelésben. Tessedik Sámuel Főiskola Tudományos Közlemények. Tom1. No. 2. 165-174.

Janowszky, J. – Janowszky, Zs. 2002: „Szarvasi-1" energiafű. Mezőgazdasági Kutató Fejlesztő Kht.

Jorgensen, E. 2000: Monte Carlo simulation models: Sampling from the joint distribution of "State of Nature"-parameters. In: Van der Fels-Klerx, I.; Mourits, M. (eds). Proceedings of the Symposium on "Economic modelling of Animal Health and Farm Management", Farm Management Group, Dept. of Social Sciences, Wageningen University, p. 73-84.

Kovács, S. – Ertsey, I. – Balogh, P. 2007a: Technological and economic risk analysis of laying hen breeding applying simulation. Joint IAAE-EAAE Seminar, Agricultural Economics and Transition: „What was expected, what we observed, the lessens learnd". Corvinus University, Budapest

Kovács, S. – Ertsey, I. – Balogh, P. 2007b: Bayesi statisztikán alapuló Monte Carlo szimuláció alkalmazása a tojótyúk nevelés technológiai kockázatának vizsgá-

lata során, Agrárgazdaság, Vidékfejlesztés, Agrárinformatika AVA3 nemzetközi konferencia CD-kiadvány, Debrecen

Kovács, S. – Nagy, L. 2008: Szimulációs modellek alkalmazása a növénytermesztés vetésszerkezetének optimalizálásában.

KSH 2009: http://portal.ksh.hu/pls/ksh/docs/hun/agrar/html/tabl1_3_1.html Downloaded: 2009-07-13

Marosvölgyi, B. 1998: Faültetvények és azok energetikai hasznosítása. XL. Georgikon Napok, Keszthely

Russel, R. S. – Taylor, B. W. 1998: Operations Management, Focusing on quality and competitiveness, New Jersey: Prentice Hall, p. 610-613.

Udovecz, G. – Popp, J. – Potori, N. 2007: Alkalmazkodási kényszerben a magyar mezőgazdaság. Agrárgazdasági Kutató Intézet 19. p., Budapest

Watson, H. 1981: Computer Simulation in Business. New York: Whiley

Winston, W. L. 1997: Operations Research Applications and Algorithms, Wadswoth Publishing Company, 863-870.

Analysis of dairy farms' efficiency by using Data Envelopment Analysis (DEA)

Tímea Gál[1]

In Hungary the dairy sector is in a long-term critical period, the stock has been in constant decline. The consumption of milk and dairy products in Hungary is slightly rising compared to the world tendency, and it is fallen behind the level in 1990. The milk consumption per capita in 2006 was 75 litres less than the EU-15 average.

Dairy enterprise is a very risky activity: the profitability of the enterprise is affected by the fluctuation of feed and animal health products prices from the side of inputs, and by the fluctuation of end-product prices. Under these circumstances it is vital for the cattle breeders, in order to survive, to harness the reserves in the breeding as effectively as possible.

In our research I made a multi-faceted efficiency analysis of an agricultural holding's three dairy farms. The chosen method for the analysis was Data Envelopment Analysis (DEA). The selection of the method is justified by the fact that there is not such a reliable database by which I could define production functions, and that DEA makes possible to manage several inputs and outputs, i.e. multiple decision problems, simultaneously. By using DEA the sources that causes shortfalls can be identified, analyzed and quantified on farms that does not operate efficiently, thus it can help the corporate decision support successfully.

In the model inputs are the cost data per one litre milk – feed, medicinal product use, logistic cost -, and the main parameters concerning the keeping and rearing. Outputs are indicators concerning milk production, milk quality and others.

I prepared the model in MS Excel, the linear programming model series were programmed by Visual Basic. After solving the model, in light of the shadow prices we can determine why either of the farms is not efficient.

Keywords: DEA analysis, efficiency, bovine sector

[1] Tímea Gál, PhD student, University of Debrecen, Centre of Agricultural Engineering and Economics, Coordination Centre for Logistics (Debrecen)

1. About DEA in general

The idea of *Data Envelopment Analysis* (hereinafter DEA) method is originated by Farrel (1951), who wanted to develop a method that is more suitable for measuring productivity. However, in 1978 Charnes et al. reformed this as a mathematical programming problem. This technique is a relatively new *"data-oriented" process*, which can be applied for measuring the performances of decision making units (DMU's) producing from several inputs several outputs (Cooper et al. 2004). Recent years the method of DEA has been used in many applications for performance measurement. It has been used for measuring the efficiency of a service's internal quality (Soteriou-Stavrinides 2000; Becser 2008), efficiency measurement of banks (Sherman-Ladino 1995; Tóth 1999), of educational (Tibenszkyné 2007) and other public bodies, and also for measuring the efficiency of business parks (Fülöp-Temesi 2000). However, its application in agricultural practice was not significant. The efficiency analysis of animal farms and agricultural production processes can be carried out by *simulation methods* (Szőke et al. 2009; Kovács-Nagy 2009), however, the quality of available database does not always allow the full mapping of technological processes. In these cases DEA is a more efficient tool.

DEA process has two known approaches: *input-oriented* (cost-oriented) and *output-oriented* (result-oriented). In case of the input-oriented approach we examine how much and which proportion the inputs should be used to minimize the cost at the same emission level. In the output-oriented approach we determine the partial increase of outputs without changing the quantity of inputs (Farrel 1957; Charnes et al. 1978).

This is complicated by the fact that we must take into consideration at efficiency measurement that not every input benefits in the same way: if we calculate with the intake on the same level we count with *Constant Return to Scale* (CRS), if not, with *Variable Return to Scale* (VRS) (Cooper et al. 2004).

2. Description of DEA model

It is an often arising question on a farm that in the course the operation of enterprise how efficient its units are working. The investment analysts are interested in the efficiency of competing participants within an industrial enterprise. DEA is a *linear programming application* by which the above-mentioned problems can be solved. In the course of DEA analysis we get the result that at what efficient level the inputs are transformed into outputs, so it is suitable to find the unit (plant, university, restaurant, etc.) which has the "best-practice" (Albright-Winston 2007). I apply the

method of DEA to determine the frontier efficiency by the efficiently operating units (Tofallis 2001; Bunkóczi-Pitlik 1999).

Efficiency can be measured by output/input indices, thus:

$$E_i = \frac{\sum_{j=1}^{n_o} O_{ij} w_j}{\sum_{j=1}^{n_i} I_{ij} v_j} \quad \text{where} \tag{1.1.}$$

E_i : the efficiency of the i^{nth} unit

O_{ij} : the value of the i^{nth} unit's j^{nth} output factor

n_o : number of outputs

w_j : the evaluation of one unit of the j^{nth} output

I_{ij} : the value of the i^{nth} unit's j^{nth} input factor

n_j : number of inputs

v_j : the evaluation of one unit of the j^{nth} input

Objective function of the model:

$$\sum_{j=1}^{n_o} O_{ij} w_j \Rightarrow MAX\,! \tag{1.2.}$$

For every examined unit we solve a separate LP exercise by which the economic content of the objective function is the same, namely my aim is to maximize the value of the units' weighted outputs. After having solved all LP models, we get the best evaluation (input and output weights) as a result (Ragsdale 2007).

Constraints:

1. The efficiency of any unit cannot be greater than 100%.

$$\sum_{j=1}^{n_o} O_{kj} w_j \le \sum_{j=1}^{n_i} I_{kj} v_j \text{ (k=1,2,..., the number of units to be taken under the analysis)}$$

that is

$$\sum_{j=1}^{n_o} O_{kj} w_j - \sum_{j=1}^{n_i} I_{kj} v_j \le 0$$

2. For the sake of the calculations input prices should be scaled in a way that the input cost of the inth economic unit shall be 1 (Ragsdale 2007).

$$\sum_{j=1}^{n_I} I_{ij} v_j = 1 \qquad (1.5.)$$

3. Case study on dairy farms and the results of the analysis

In this DEA model I examined the efficiency of *three dairy farms*. These farms belong to one agricultural holding in the county of Hajdú-Bihar. For setting up and solving the model I used the holding's production and financial data for 2008. I took into consideration the variable costs and the average number of dairy cows as input factors, and the milk production for 305 days, from the milk quality data the average milk fat and milk protein content and the turnover as outputs. The aim of the analysis was *to examine the competitiveness* of the three farms within the holding, and in case of the inefficient farms to explore the critical factors, and to determine the direction of the occurring further analyses.

For the sake of the easier clarity in the linear programming model I marked the constraints with red colour, the variables with green colour and the objective function with blue colour. The maximum output values that I got by solving the LP model are in the DEA efficiency column (Figure 1.). On this basis it can be stated that considering the given input and output constraints the first and the third farms are operating efficiently, while the third one is not (DEA efficiency value is less than 1).

Figure 1. The DEA model that is suitable for doing efficiency analysis on dairy farms

	Outputs				Inputs					
Farm code	Milk for 305 days	Milk protein	Milk fat	Turnover	Variable costs	Number of dairy cows	Weighted output	Weighted input	Difference	DEA efficiency
	litre	%	%	M Ft	M Ft	heads				
1	8900	3,2	3,4	719	665	1050	1,6935	1,6935	0,0000	1,0000
2	8480	3,1	3,41	352	314	525	0,8468	0,8468	0,0000	1,0000
3	8630	3,25	3,39	389	379	620	0,9341	1,0000	-0,0659	0,9163
Weights	0,00000	0,01198	0,00000	0,00230	0,00000	0,00161				

	3	
Output	0,9341	objective function
Input	1,0000	

Source: Own model by using farm data

The analysis of shadow prices provides a further help to answer why the third farm is not operating efficiently compared to the first two farms. The LP model's shadow prices for the third farm can be seen in Table 1.

Table 1. Shadow prices after solving the LP model of Farm 3

Cell	Name	Final value	Shadow price
D16	Input 1	1,0000	0,9341
L7	Farm 1 Difference	0,0000	0,0565
L8	Farm 2 Difference	0,0000	0,9900
L9	Farm 3 Difference	-0,0659	0,0000

Source: Own calculations

Using the *shadow prices* of Farm 1 and Farm 2, which are considered efficient according to DEA, as weights we can 'make' such an input and output average vector that regards to a complex, hypothetic farm. The input and output characteristics of this farm can be compared with the current data of Farm 3, so the shortcomings and the factors that worse the efficiency can be explored (Table 2.).

Table 2 The characteristics of the composite farm 'made' by using shadow prices as weights

	Outputs				Inputs		
Farm code	Milk for 305 days	Milk protein	Milk fat	Turnover	Variable costs	Number of dairy cows	Weights
	litre	%	%	M Ft	M Ft	heads	
1	8900	3,20	3,40	719	665	1050	5,65%
2	8480	3,10	3,41	352	314	525	99,00%
3	8630	3,25	3,39	389	379	620	0,00%
Composit farm	8899	3,25	3,57	389	349	579	

Source: Own calculations

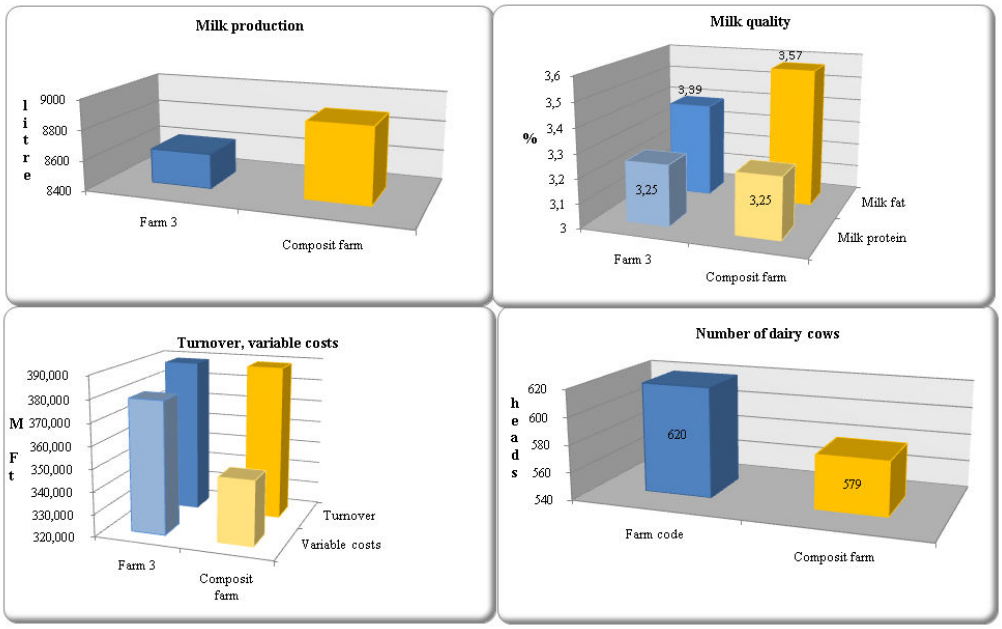
For the input factors, both the number of dairy cows and the variable costs, a lack of balance can be seen between the composite farm and Farm 3. The number of dairy cows and the variable costs shall be also cut down. It can be observed that this should be done without reducing turnover, which presupposes the betterment of milk production and the quantity of milk fat from the quality parameters as well (Figure 2.).

The above data suggest that there are several significant technological, management or animal health problems, since the specific turnover for one cow is

the lowest, 627 thousand Ft/year (687 thousand Ft on Farm 1, 667 thousand Ft on Farm 2). At a constant level of turnover we shall reach cost-reduction, yield-growth and quality improvement.

The results of the analysis suggest *the need for further examinations*. For the sake of production results and the improvement of quality considering the animal health and technological factors the proposal solution can be further fined, shaded.

Figure 2. Input and output factors on the inefficient Farm 3 and on the composit farm



Source: own calculations

References

Albright, S.C. – Winston, W.L. 2007: *Management Science Modeling*. Revised Third Edition, Thomson, 18-192. p.

Becser, N. 2008: Szolgáltatásminőség modellek. BCE Vállalatgazdaságtani Intézet, 89. sz. *Műhelytanulmányok*, 2008. január, 19.

Borsi, B. 2005: *Tudás, technológia és a magyar versenyképesség*. PhD. értekezés, 26-30. p.

Bunkóczi, L. – Pitlik, L. 1999: A DEA (Data Envelopment Analysis) módszer alkalmazási lehetőségei üzemhatékonyságok mérésére. *IA'99*, Debrecen,

1999. augusztus 26. http://miau.gau.hu/miau/08/dea.doc Letöltve: 2009. szeptember 5.

Charnes, A. - Cooper, W. - Rhodes, E. 1978: "Measuring the efficiency of decision-making units." *European Journal of Operational Research,* Vol. 2, 429–444. p.

Cooper, W. W. – Seiford, L. M. – Zhu, J. 2004: *Data Envelopment Analysis. International Series in Operations Research & Management Science,* Vol. 71. Chapter 1. 1-39. p.

Farrell, M. J. 1957: The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society,* Vol. 120, 253–281. p.

Fülöp, J. – Temesi, J. 2000: A Data Envelopment Analysis (DEA) alkalmazása ipari parkok hatékonyságának vizsgálatára. *Szigma* 32, 3-4, 85-110. p.

Kopasz, M. 2007: Útban a versenyképesség felé. *AVA3 konferencia*, Debrecen, 2007. március 20-21., www.avacongress.net/ava2007/Program_Booklet.pdf, Letöltve: 2009. szeptember 5.

Kovács, S. – Nagy, L. 2009: An application of Marcov chain Monte Carlo simulation. *Internal Congress on the aspects and visions of applied economics and informatics* (AVA4), Debrecen, 1333-1338. p.

Ragsdale, C. T. 2007: *Spreadsheet Modeling & Decision Analysis: A Prectical Introduction to Management Science*. Fifth Edition, Thomson, 102-113. p.

Sherman, H. D. – Ladino, G. 1995: Managing Bank Productivity Using Data Envelopment Analysis (DEA), *Interfaces* 25, No. 2, 60-80. p.

Szőke, Sz. – Nagy, L. – Kovács, S. – Balogh, P. 2009: Examination of pig farm technology by computer simulation. *Internal Congress on the aspects and visions of applied economics and informatics* (AVA4), Debrecen 1317-1325. p.

Tibenszkyné, F. K. 2007: *Az oktatás hatékonyságának mérése a ZMNE 2006-ban végzett hallgatóin Data Envelopment Analysis (DEA) módszer használatával*. Doktori PhD. értekezés, ZMNE, 149-165. p.

http://193.224.76.4/download/.../phd/.../tibenszkyne_forika_krisztina.pdf Letöltve: 2009. szeptember 5.

Tofallis, C. 2001: Combining two approaches to efficiency assessment. *Journal of the Operational Research Society* 52, 11, 1225–1231. p.

Tóth, Á. 1999: Kísérlet a hatékonyság empirikus elemzésére. *MNB füzetek* 1999/2, 11-33. p.

Estimating agricultural demand for electricity in Iran (1975-2007)

Vahid Ghorbani[1] *- Nader Mehregan*[2] *- Mahmoud Haghani*[3]

High electricity consumption in agricultural sector is an important issue in Iran economy. The main reason is the low price of electricity. In this study we estimated demand function of electricity in agriculture sector (1975-2007) using ARDL method. Results indicated that (1) the short-term & long-term price elasticity is -0.1 and -0.49, respectively, (2) short-term & long-term income elasticity is 0.43 and 2.07, respectively, and (3) the lagged error correction term was significant with expected negative sign (-0.21).The CUSUM and CUSUMSQ tests have been done and the results showed a stabilization of coefficient in confidence level 5%.

Keywords: electricity demand, ARDL, elasticity

1. Introduction

High electricity consumption in agricultural sector in Iran is an important issue that Price liberalization and omitting electricity subsidy is one of the solutions to reduce electricity consumption in agriculture sector. Electricity price in Iranian agriculture sector is 21 rials per kilowatt /h while cost Price is 834 rials per kilowatt /h in 2007 and agriculture sector use 8% of total electricity use in Iranian economy. In this study we applied CPI index to calculate real price. The investigation of nominal and real price of electricity for the period 1975-2007 shows that in spite of increasing nominal price of electricity between these years, the real price decreased. Also, results showed that the average growth of electricity consumption in agricultural sector was 14%. Because of some facts such as low price of electricity in Iran, none existence of a suitable substitution, high consumption and its effects on macroeconomic variables it would be necessary to investigate the main factors which affects the electricity consumption in agricultural sector. Therefore this study

[1] *Vahid Ghorbani,* Msc student in energy economics , Abbaspour university of Technology ,Tehran-Iran
[2] *Nader Mehregan*, Associate professor in economic , university of Bo Ali ,Hamedan-Iran
[3] *Mahmoud Haghani*, Assistant professor in management , Abbaspour university of Technology, Tehran-Iran

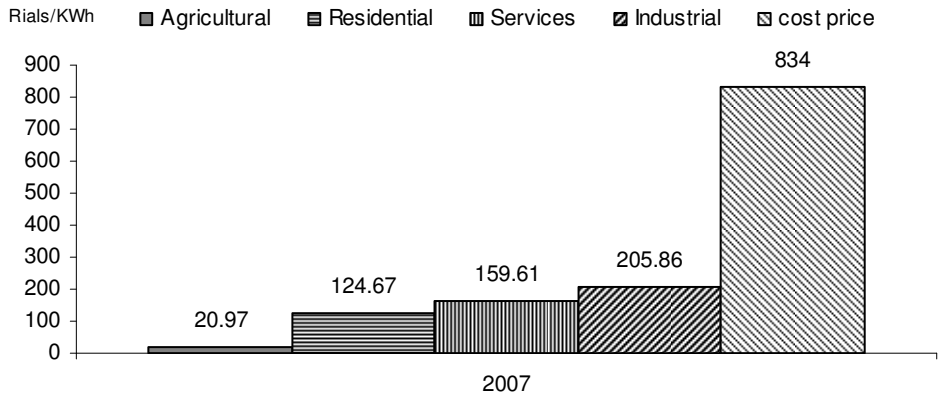
aimed to estimate long and short- term electricity consumption function in agricultural sector in Iran.

A few recent examples for these studies are noted here: (Filippini - Pachauri 2004) for urban Indian households; (Dergiades - Tsoulfidis 2008) for residential in theUnited States; (Narayan - Smyth 2005) for residential in Australia; (Amusa et al. 2009) for Aggregate electricity demand in South Africa; (Zachariadis - Pashourtidou 2007) for Cyprus; (Razak - Al-Faris 2002) for GCC countries; (Kumar et al. 1999) for India show that electricity consumption in agriculture sector are income inelastic (<1). The short – run price elasticity is -1.35 in agriculture sector; (Atakhanova - Howie 2007) in Kazakhstan. This paper estimates Kazakhstan's aggregate demand for electricity as well as electricity demand in the all sectors . Results show that firstly, price elasticity of demand in all sectors is low. Secondly, income elasticity of demand in the aggregate and all sectors is less than unity.

The paper is organized as follows: After the introduction the second section is provides the econometric specification of the model for agricultural demand for electricity and discusses the ARDL cointegration technique. The third part presents the data and evaluates the results of the econometric analysis. Finally there is a conclusion and recommendation regarding to the results of the study.

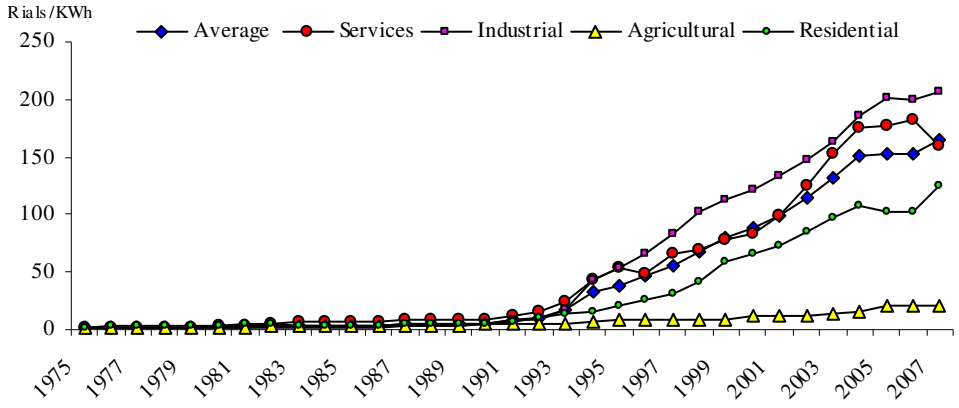
2. Electricity Price in Iran

Figure 1.



*Source:*own creation

Figure 2.



Source: own creation

3. Model and methodolpogy

A modified agricultural electricity demand model in logarithmic form is adopted based on :

$$\ln Co_t = \alpha_0 + \alpha_1 \ln Y_t + \alpha_2 \ln P_t + \alpha_3 D_t + \varepsilon_t \qquad (1)$$

Where Co_t is the agricultural electricity consumption (KWh) , Y_t is the real added value in agricultural sector, P_t is the real agricultural electricity price (rials/KWh) and D is the dummy variable for drought in Iran that occur in 1988 and 1989 ln is the natural logarithm transformation.

As for the expected signs in Eq. (1), one expects that $\alpha_1>0$ because higher added value in agricultural sector should result in greater economic activity and accelerated purchases of electrical technology. The coefficient of price level is expected to be less than zero for usual economic reasons, therefore, $\alpha_2<0$. The coefficient of dummy variable is expected to be more than one because drought in 1988-1989 cause to decrease the agricultural product.

In the last two decades, several econometric procedures were employed to investigate the electricity demand functions. With regards to cointegration approaches, there are several example including (Engle - Granger 1987, Narayan - Smyth 2005, Razak - Al-Faris 2002, Halicioglu 2007) and (Amusa et al. 2009). A

recent single cointegration approach, known as autoregressive distributed lag (ARDL) of' Pesaran et al. (2001).An ARDL representation of Eq. (1) is formulated as follows:

$$\Delta lnCo_t = \propto_0 + \sum_{i=1}^{m} \propto_{1i} \Delta lnCo_{t-i} + \sum_{i=0}^{m} \propto_{2i} \Delta lnY_{t-i} + \sum_{i=0}^{m} \propto_{3i} \Delta lnP_{t-i} + \sum_{i=0}^{m} \propto_{4i} \Delta D_{t-i} + \propto_5 lnCo_{t-1} + \propto_6 lnY_{t-1} + \propto_7 lnP_{t-1} + \propto_8 D_{t-1} + v_t \quad (2)$$

Once a long-run relationship has been established, Eq. (2) is estimated using an appropriate lag selection criterion. At the second stage of the ARDL cointegration procedure, it is also possible to perform a parameter stability test for the selected ARDL representation of the error-correction model Halicioglu (2007).

A general error – correction model (ECM) of Eq. (2) is formulated as follows:

$$\Delta lnCo_t = \propto_0 + \sum_{i=1}^{m} \propto_{1i} \Delta lnCo_{t-i} + \sum_{i=0}^{m} \propto_{2i} \Delta lnY_{t-i} + \sum_{i=0}^{m} \propto_{3i} \Delta lnP_{t-i} + \sum_{i=0}^{m} \propto_{4i} \Delta D_{t-i} + \lambda \mathrm{EC}_{t-1} + \mu_t \quad (3)$$

Where λ is the speed of adjustment parameter and EC_{t-1} is the residuals that are obtained from the estimated cointegration model of Eq. (1).

The Granger representation theorem suggests that there will be Granger causality in at least one direction if there exists a cointegration relationship among the variables in Eq. (1), providing that they are integrated order of one. Engle and Granger (1987) cautions that the Granger causality test, which is conducted in the first-differences variables by means of a vector autoregression (VAR), will be misleading in the presence of cointegration. Therefore, an inclusion of an additional variable to the VAR system, such as the error-correction term would help us to capture the long-run relationship. To this end, an augmented form of the Granger causality test involving the error-correction term is formulated in a multivariate *p*th order vector error-correction model (Halicioglu 2007).

$$\begin{bmatrix}\Delta lnCo_t\\ \Delta lnY_t\\ \Delta lnP_t\\ \Delta D_t\end{bmatrix}=\begin{bmatrix}C_1\\ C_2\\ C_3\\ C_4\end{bmatrix}+\begin{bmatrix}d_{11i}d_{12i}d_{13i}d_{14i}\\ d_{21i}d_{22i}d_{23i}d_{24i}\\ d_{31i}d_{32i}d_{33i}d_{34i}\\ d_{41i}d_{42i}d_{43i}d_{44i}\end{bmatrix}\begin{bmatrix}\Delta lnCo_{t-i}\\ \Delta lnY_{t-i}\\ \Delta lnP_{t-i}\\ \Delta D_{t-i}\end{bmatrix}+ \qquad (4)$$
$$\begin{bmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\\ \lambda_4\end{bmatrix}[EC_{t-1}]+\begin{bmatrix}\omega_{1t}\\ \omega_{2t}\\ \omega_{3t}\\ \omega_{4t}\end{bmatrix}$$

EC_{t-1} is the error–correction term, which is obtained from the long-run relationship described on Eq. (1), and it is not include in Eq. (4) if one finds no cointegration amongst the vector in question.

The existence of a cointegration derived from Eq. (2) does not necessarily imply that the estimated coefficients are stable. Stability tests of Brown et al. (1975), which are also known as cumulative sum (CUSUM) and cumulative sum of squares (CUSUMSQ) tests based on the recursive regression residuals, may be employed to that end. These tests also incorporate the short-run dynamics to the long-run through residuals. The CUSUM and CUSUMSQ statistics are updated recursively and plotted against the break points of the model. Provided that the plots of these statistics fall inside the critical bounds of 5% significance, one assumes that the coefficients of a given regression are stable. These tests are usually implemented by means of graphical representation.

4. Data and empirical results

The data used in this paper are annual time series spanning the period 1975-2007. The sources of our data on price of electricity in rials per KWh for agricultural sector were obtained from the TAVANIR information bureau (annual publication)[4]. agricultural electricity consumption in millions of KWh and agricultural added value in milliard rials were obtained from the Iranian central bank[5].

We performed the Augmented Dickey – Fuller (ADF) test and Phillips-Perron (PP) test to verify the exact order of integration of the variables. Table 1 below displays the results of ADF tests and Table 2 below displays the results of PP tests. Results show that all variables used in our study are an I(1) in 1% level of critical values.

4 . TAVANIR is depended on Iranian power ministry. For details see http://www.tavanir.org.ir
5 . For details see http://www.cbi.ir

Table 1. ADF tests

Level			1 st Differences			Order of Integration
Variable	**ADF stat.**	***p*- value**	**Variable**	**ADF stat.**	***p*- value**	
Co	-1.99	0.28	ΔCo	-4.63	0.00	I(1)
Pr	-0.24	0.92	ΔPr	-5.43	0.00	I(1)
Ad	-2.88	0.18	ΔAd	-7.77	0.00	I(1)

Source: own creation

Note: ADF stands for the Augmented Dickey – Fuller test. All level variables are in logs. Δ is the first difference operator.

Table 2. PP tests

Level			1 st Differences			Order of Integration
Variable	**ADF stat.**	***p*- value**	**Variable**	**ADF stat.**	***p*- value**	
Co	-1.98	0.29	ΔCo	-4.62	0.00	I(1)
Pr	-2.43	0.35	ΔPr	-5.43	0.00	I(1)
Ad	-2.9	0.17	ΔAd	-7.75	0.00	I(1)

Source: own creation

Note: PP stands for the Phillips-Perron test. All level variables are in logs. Δ is the first difference operator.

Having estimated Eq. (2) by means of OLS, the ARDL approach to cointegration requires the testing of the following null hypothesis : α_5 to $\alpha_8 = 0$ against the alternative that at least one of these coefficients is different from zero. We performed the two-step method of Engle and Granger to verify the cointegration of the regression. Engle - Granger (1987) proposed a two-step method of testing for cointegration which looks for a unit root in the residuals of a first-stage regression. Results indicated that if appoint optimum lag based on Akaike Information Criterion (AIC), optimum lag equal zero and this lag test statistics is -5.61 that greater than 95% critical value for the Dickey-Fuller statistic (-4.93) therefore cointegration relationship matter and existence the long-run equilibrium. Moreover We could appoint optimum lag based on Schwarz Bayesian Criterion (SBC) and Hannan-Quinn Criterion (HQC) that same result will occur (see Table 3).

Table 3. Bounds testing for cointegration

	Test statistic	**LL**	**AIC**	**SBC**	**HQC**
DF	-5.61	27.97	26.97	26.3	26.76
ADF(1)	-3.79	27.97	25.97	24.63	25.56
ADF(2)	-3.02	27.97	24.97	22.97	24.36

Source: own creation
95 % Critical value for the Dickey-Fuller statistic = -4.9327
LL = Maximized log-likelihood AIC = Akaike Information Criterion
SBC = Schwarz Bayesian Criterion HQC = Hannan-Quinn Criterion

Taking for granted the existence of a long-run equilibrium, we estimated by setting the maximum lag-length to two and using the Schwarz Bayesian Criterion (SBC) for the selection of model's lag order. The specification finally selected was the ARDL (1,0,0,0). The derived long-run elasticities resulting shown in Table 4 and diagnostic tests for the ARDL model are shown in Table 5. The estimated elasticities display the expected signs which are negative for the price of electricity and positive for the added value and drought variable and moreover all long-run elasticities are significant. Result show that long-run price elasticity is -0.49 that mean if agricultural electricity price increase ten percent decrease the agricultural electricity consumption 4.9 percent and long-run income elasticity is 2.07 that mean if added value in agricultural sector increase ten percent, increase electricity consumption 20.7 percent. Moreover result shows that drought in Iran (1988-1989) cause to increase the electricity consumption. Finally, diagnostic tests for the underlying ARDL model verify that the residuals are non-serially correlated, correct functional form, normal, and non-heteroscedastic

Table 4. Long-run coefficient for the ARDL (1, 0, 0, 0) model

Variable	**C**	**Pr**	**Ad**	**D**
Coefficient	-13.49	-0.49	2.07	0.43
t-statistic	-1.87	-2.93***	3.64***	1.76*
p-value	0.07	0.07	0.00	0.09

Source: own creation
Note: The ARDL (1, 0, 0, 0) specification was selected based on the Schwarz Bayesian criterion.
The maximum lag length was set to 2.
D is dummy variable for Drought in Iran (1988-1989)
***and * indicate 1% and 10% levels of significance, respectively.

Table 5. Diagnostic tests of the ARDL model

Variable	t-statistic	p-value
Serial Correlation	0.18	0.67
Functional Form	0.04	0.85
Normality	2	0.35
Heteroscedasticity	0.00	0.95

Source: own creation

The short-run model results are presented in Table 6 below. As expected all short-run elasticities are lower in absolute value than those in the long-run. The short-run price elasticity -0.1 that mean if agricultural electricity price increase hundred percent decrease consumption only ten percent and long-run income elasticity is 0.43 that mean if added value in agricultural sector increase hundred percent electricity consumption will increase 43 percent. The lagged error correction term is statistically significant with the expected negative sign. Diagnostic tests for the short-run ARDL model are shown in Table 7.

Table 6. Error-correction representation results. ΔCo, is the dependent variable

Variable	ΔC	ΔPr	ΔAd	ΔD	EC_{t-1}
Coefficient	-2.82	-0.1	0.43	0.09	-0.21
t-statistic	-1.43	-1.69^{*}	1.96^{**}	1.39^{NS}	-2.48^{**}
p-value	0.16	0.1	0.05	0.17	0.02

Source: own creation
Note: The error-correction term is given by:
EC = Co + 0.49670*Pr - 2.0746* Ad - 0.43639*D + 13.4918*C.
** and * indicate 5% and 10% levels of significance, respectively. NS indicate not significance.

Table 7. Diagnostic statistics

R^2-adjusted	0.2548	**Schwarz criterion**	23.2064
***F*-statistic**	5.223	**Akaike criterion**	26.7913
***DW*-statistic**	2.0875	**RSS**	0.2334

Source: own creation
Note: DW is the Durbin-Watson statistic and RSS is the residual sum of squares

In order to ensure the stability of the long-run parameters of our econometric specification, we applied the CUSUM and the CUSUMQ tests in the residuals of the error-correction. Figs. 1 and 2 below, display the results of CUSUM and CUSUMQ

tests, respectively. In both figures the dotted lines represent the critical upper and lower bounds at the 0.05 level of significance. The visual inspection of Figs. 1 and 2 reveals that there is no evidence of parameter instability, since the cumulative sum of the residuals and the cumulative sum of the squared residuals move within the critical bounds.

Figure 3. CUSUM test

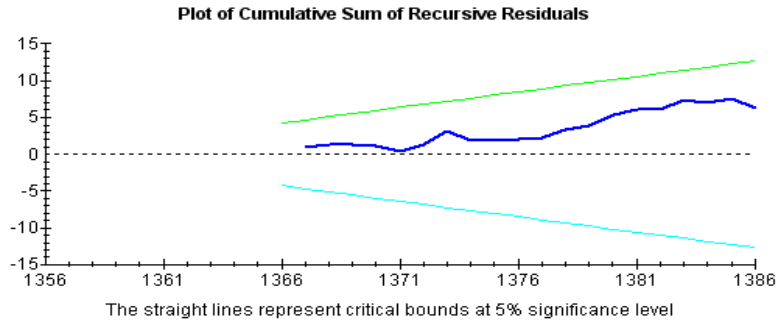
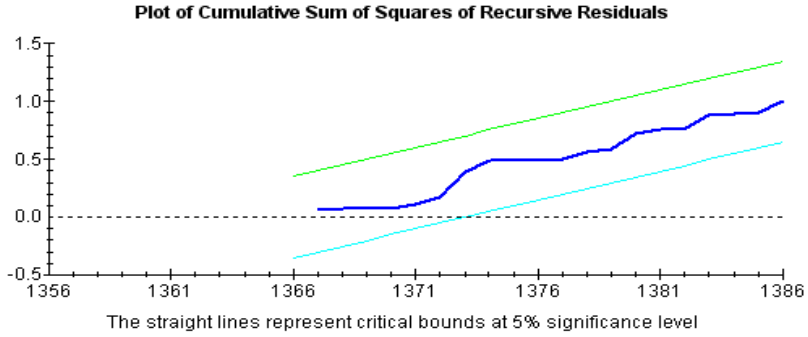


Figure 4. CUSUMQ test



5. Summary and conclusions

This paper has examined the determinants of the agricultural demand for electricity in Iranian economy. The econometric specification assume that the agricultural demand for electricity depends on the price of electricity, added value of agricultural sector and extensive drought that occur in 1988-1989 in Iran. For estimate the model we used advanced ARDL cointegration technique. The error correction model was consistent with the expectations about the signs of the short-run parameters and their magnitude which was found lower than their long-run counterparts. Results show that (1) price and income elasticity in long-run almost 4.9 times price and income

elasticity in short-run. (2) low price elasticity lead to low efficiently of price policy then Price liberalization and omitting electricity subsidy is one of the solutions to both increase efficiently of price policy and reduce electricity consumption.

References

Amusa, H. -Amusa, K. - Mabugu, R. 2009: *Aggregate demand for electricity in South Africa: An analysis using the bounds testing approach to cointegration*. Energy Policy, ARTICLE IN PRESS.

Atakhanova, Z. - Howie, P. 2007: *Electricity demand in Kazakhstan*. Energy Policy, 35, 3729–3743. p.

Dergiades, T. - Tsoulfidis, L. 2008: *Estimating residential demand for electricity in theUnited States*. 1965–2006. Energy Economics, 30, 2722–2730. p.

Engle, R. -Granger, C. 1987: *Cointegration and error correction representation: estimation and testing. Econometrica*. 55, 251–276. p.

Filippini, M. - Pachauri, S. 2004: *Elasticities of electricity demand in urban Indian households*. Energy Policy, 32, 429–436. p.

Halicioglu, F. 2007: *Residential electricity demand dynamics in Turkey. Energy Economics*, 29, 199–210. p.

Hondroyiannis, G. 2004: *Estimating residential demand for electricity in Greece*. Energy Economics, 26, 319– 334. p.

Holtedahl, P. - Joutz, F. L. 2004: *Residential electricity demand in Taiwan.* Energy Economics, 26, 201–224. p.

Iranian central bank http://tsd.cbi.ir/IntTSD/Display/Display.aspx

Iranian energy ministry. Annual report, 2007. Available at: http://iranenergy.org.ir/statistic%20info/-energy%20balance/86/chapter04-Electricity.pdf

Iranian energy ministry, TAVANIR. Available at: http://www.tavanir.org.ir

Kumar, B. – Shukla, R. M.1999: *Elasticities of electricity demand in India*. Energy Policy, 27, 137-146. p.

Narayan, P.K. - Smyth, S. 2005: *The residential demand for electricity in Australia: an application of the bounds testing approach to cointegration*. Energy

Policy, 33, 467–474. p.

Pesaran, M.H. - Shin, Y. - Smith, R.J. 2001: *Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics* 16/3, 289–326. p.

Razak, F. - Al-Faris 2002: *The demand for electricity in the GCC countries*. Energy Policy, 30, 117–124. p.

Zachariadis, T. - Pashourtidou, N. 2007: *An empirical analysis of electricity consumption in Cyprus*. Energy Economics, 29, 183–198. p.

The transformation and the ownership structure in Hungary

Katalin Botos[1]

Was there a transition at all in Hungary? Did the system changed at all in 1990? The answer of the author is yes. Not in the sense that from the society's top became bottom- as during the transformation after 1945- but in the sense that private property became again to be the basic feature of the economic system. We can follow the evolution of the economy in the mirror of the financial accounts. The players of the market, households, state and financial and non financial enterprises have changed their financial account positions during the period 1990-2006 markedly. The study based on HNB data follows carefully the changes in the asset and liability structures of households, state and enterprises and the financing capacity of each sector. The state's asset position has diminished, the households' has grown. But the great winner is the foreign owner's sector. It has an influence on the per capita GDP and GNI creating a marked difference between them. It is very important to have internal financing capacity because state budget has a deficit since decades. We don't have enough in the household sector therefore the country needs external financing. The study examines the roots of the international indebtedness of the country, the role of the economic policy and the banking sector's strategy (selling foreign-financed mortgage loans to the households). Today the state is in a much worse situation than before transition: State debt (and foreign debt of the country) is even higher than in 1990 and state's ownership (covering the national debt) is now on a minimum level.

Keywords: *asset and liability structure of firms, state and households, financing capacity, economic policy, foreign financing, international indebtedness*

1. Introduction

"Did political transformation took place in Hungary at all?"- sounds the passionate question one often hears everyday talks. .

If we look at the changes of the ownership structure in the country we may say that very basic transformation happened., actually twice. The first after the II. World War, the second after 1990.

[1] Dr. Katalin Botos, University of Szeged Hungary, Faculty of Economics and Business Administration (Szeged)

As a result of nationalisation the private ownership had been abolished by the communist forces in 1949, but regained its previous economic function by the privatisation after 1990. Legal experts claim that amendments to the Hungarian constitution in 1989 solidified the function and significance of private property to the highest degree. More, than in any country.

The newly emerged socio-economic system in Hungary could be presented in several ways. I chose to describe these changes on the basis of financial accounts data. (I rely on a recent publication of the Hungarian National Bank: Financial Accounts of Hungary. Data, analysis, methodology. National Bank of Hungary. 2008.) The sources of statistical figures in this article are referred to by page numbers of the quoted issue.

The present study has historical perspective and does not aim to analyse the effects of the financial-economic crisis of 2008 but the long term trends since 1990.

2. The citizen as proprietor

The economic model of the "existing socialism" after 1956 households were able to save some money. Their savings were generally disposed at mutual saving associations or branches of the National Savings Bank. People had holiday houses around the cities. The housing , too, became more and more financed by private credits from the Savings Bank (OTP).

In the years preceding the fall of communism, when the IMF surveyed public expenditure and did not allow money creation for financing the budget deficits, money was raised by issuing so called "dwelling fund bonds". Indebtedness was growing and so did subsidies from the state budget for housing. As a consequence of the political transformation in 1990, *the composition of household monetary assets has changed quite a lot*

In order to cut budgetary expenditures the new government of 1990 offered to abolish the other half of credits if debtors paid back their credits before the deadline. A number of advance repayments occurred, a lot of people, who could afford, took. Massive repayments happened. The savings of the population diminished .Later it started growing again. Between 1993 and 1998 financing capacity of the households stabilised around 10 % of the GDP but again gradually decreased in the late 90ies. In 2003 it amounted only to half a percent. (p. 75) The reduction was partly the result of a housing boom due to increased state subsidisation of housing. Private investments in housing inevitably used up household savings and as a result these savings could no longer cover budget deficits. Without household savings there is a heavy pressure for external financing of budget deficits, which became a cardinal problem by now in Hungary.

Originally most of the savings was in form of bank deposits or cash. By the beginning of the 1990's the rate of cash or deposits within household's monetary assets decreased by half. Stock ratios and business shares added up to 25%, non-

stock based securities and insurance reached only 7 %. By 2007 only 36% of household monetary assets were found in cash and bank deposits, business share ratios stayed at 25-26% (p. 46). Business share assets were about fourfold higher than stock share assets. These assets arose in the early years of political transformation and were due mainly to privatisation. On the other hand a great proportion of citizens were forced to launch their own enterprises in order to escape from collapsing workplaces. So they became owners of business shares. Although little profit could be expected from these enterprises they offered self-employment at least.

Following a wave of privatisation in 1995, investment in stock exchange shares increased to 5% but fell back to 1,5 % after the Russian crisis. In the portfolio composition of the households we find investment coupons, public securities and mortgage bonds, too. The most spectacular growth can be seen in insurance and pension insurance savings. The rate of pension insurance savings rose from 4% to 18%. It must be admitted though that the above transition did not begin spontaneously. It involved a certain legal compulsion since entrant employees were obliged to enter pension insurance funds in 1998.

As a consequence of amendments to taxation laws in 2006, investment yields became exempt from the 20% interest rate tax. Additionally, the taxation of the gain on the exchange was abolished. This measure drove remarkable amounts into unit trusts. Nevertheless, significant capital market boom should not be expected from that, since unit trusts put the money in bank deposits. This means that *people keep their money in banks indirectly rather than directly.*

Financing capacity *of the private sector* is greatly influenced by its liabilities. Let us take a look at the credit portfolio of the population. Prior to the fall of communism the population was given significant credits connected to housing. In addition to housing loans, collateral loans, car purchase loans, free credits on mortgage and student loans were introduced later on. According to data from 1990 the ratio of personal credits to the GDP had scarcely exceeded 10%, whereas this ratio trebled by 2007 reaching 29% (p. 47). A sudden increase in currency credits can be seen after 2002. By 2007 and 2008 60% of the total personal credit portfolio was credit in foreign exchange- mainly in Swiss francs. Interest rates of those currency credits were lower; therefore they were more easily financed, unless exchange rates tottered. In the case of devaluation of the forint currency credits become very expensive for Hungarians

Our credit portfolio is still behind the average of the European Union, it amounts to 29% of the GDP (in 2007), while in the European Union this figure is 65% (p. 48). Even in two decades we did not manage to catch up with European countries whose citizens were more successful in achieving middle-class status. No wonder, since Hungarian salaries are one-fifth of the average European salaries.

To sum up financial situation of the households: Notions have different meanings in Hungary: shares in investment mean (indirect) bank deposits rather than

real investments in securities, and business shares are rather for self-employment than for real competitive ventures, so how could we possibly reach a strong, wealthy middle class? *If GDP per capita is one half of the European average but salaries reach only one-fifth of the European level, relative numbers will hardly resemble that of the European Union.* In other words, Hungary has only quasi middle-class with quasi property structure. Not to mention the dispersion of incomes, the analysis of which, however, is not within the scope of this study based on financial accounts.

3. The state as (the) proprietor (of monetary assets)

The question of state property is of primary importance when one intends to assess the economic processes of the past twenty years. *As real balance sheets were not available at the time of transformation and are not at our disposal even to this day,* I rely in this respect on the information available in the publications of the National Bank of Hungary. From the publication of the HNB the *financial ratios* can be used for the evolution of the gross and net capital of the state. The majority of corporations were public corporations before the fall of communism. *The value of public finance assets, that is the value of gross capital, was one and a half times as great as the contemporary GDP. By 2007 this value decreased to 20 percent of the GDP.* What concerns liabilities: At the end of 1989 the nominal value of the public debt of Hungary mounted to 1264 billion forints which equalled to 73% of the GDP.

A substantial portion of public debt derived from foreign currency credits, yet the state paid these debts to the National Bank of Hungary in Hungarian forints until 1997 when the so called " debt exchange schemes" were introduced.

Preceding the democratic transformation the National Bank of Hungary was considered by foreign creditors as the debtor. Naturally, it was seen as souverain debt. A number of analyses in the field of Hungarian history of recent past (Ignác Romsics, András Vígvári, János Honvári, Csaba Nagy, Katalin Botos) pointed out that facilities for the repayment of currency credits were not available. These credits did not result in exportable national production which would have yielded convertible foreign exchange. After 1979 public foreign debt increased remarkably because of a dramatic increase in international interest rates and the devaluation of the dollar. To make things worse the Hungarian forint was revaluated, although the deterioration of the balance of payments would have demanded the opposite. The reason for such a measure was to avoid inflation yet, suppressed inflation caused an even greater deficit in the balance of payments.

In 1982 Hungary had joined the IMF The foreign exchange policy of the 1970's, based on voluntarism could not be continued. As the balance of payments deteriorated, the forint was devaluated. This process continued for after the fall of communism, though the forint in the first half of the decade was still over -valuated. Under the era of György Surányi, President of the NBH a sliding devaluation of the forint was introduced.

The index for debt and GDP ratio rose above 90% by 1994, the end of the first democratic governmental period, since gross domestic product had decreased but debts had to be met at the same time. The balance of foreign trade closed with a deficit in the first two years of the second democratic government (1994-98), which meant that the deterioration of the balance of payments was due to the debt of the public sector. The public debt and GDP ratio was gradually decreasing from 1996 until 2001. Several factors, such as budget constrains, the introduction of the Bokros-package, and an increase in GDP during the second half of the 1990's – especially in the third governmental period – played a role in the decrease of the above ratio. Public debts were partly refinanced, partly repaid through selling national assets.

Unfortunately, budget deficit was continuously produced; by 2002 the public debt/GDP ratio stopped decreasing, and at the end of 2007 public debt amounted to 67% of the GDP. (Even this proportion has been substantially exceeded by 2009 November).

We have seen that household – sector was not able to finance budget deficit, which means a straight way to external indebtedness. The over-valuated forint contributed to it, too.

Taking into consideration the fact that current balance of payments was continuously showing a deficit one inevitably ponders over the phenomena of continues revaluation of the forint. Two processes took place that both influenced exchange rates. On the one hand an intense inflow of foreign currency began as a result of privatisation; massive privatisation process could be observed during the second governmental era (1994-1998). On the other hand a special foreign exchange policy introduced band based fixing of exchange rates. This policy opened the way to speculative attacks which, making use of higher interest rate levels, pushed exchange rates upwards. We might arrive at a strange conclusion from the above: *prior to the fall of communism it was the economic policy based on voluntarism which revaluated foreign currency, after the fall of communism currency was revaluated by the market, although real economic situation would not demand for that just the contrary.* Does it seem that we get the same results no matter what political system we have?

In *the first years of the democratic transformation monetary assets of the Hungarian state amounted to 250% of its liabilities, it decreased to 30% by 2007.* This fact clearly reveals that our present economic position is much more unfavourable than it used to be at the fall of communism since the *property coverage of our public debt decreased* in the highest degree.

4. Public finance deficit

This fact is rather shocking No one reckoned with the fact that the majority of the Parliament would not dare and *would not even want to protest* against governmental proposals creating deficit 15-20 years after the transformation.. The concern of MP-s was that whichever governmental party secured a place for them in the Parliament that party should stay in power. This dependence oriented them to accept the yearly budgets with huge deficits.

Debts had to be paid, so selling national assets became the a way of the financing the repayment of the country's debts. By selling state property for cash the ownership portfolio of the state was exchanged for funds; the most of which was used for debt reimbursement and a smaller portion covered current budget deficits. If a portion of those funds had been used for creating economic development funds, a faster growth of the GDP might have been facilitated, even though net public debt wouldn't have diminished. Debt service in that case would have remained a considerable item of the expenditure- side of the state budget. But if there is considerable economic growth, the relative value of debt might have even decreased! It is true, the operation of economic development funds requires civil servants of impeccable character. It must be admitted that temptation was (and still is) rather great for those in power to use governmental means for political purposes, although this is a rather sad reason for rejecting a rational alternative of economic policy.

Budget deficit and public debt are significant markers of economic policy. Low budget deficit (3 %to GDP) and lower than 60% public debt to GDP are preconditions for joining the euro zone, as well as diminishing the rate of inflation.[2]

If we take a look at financial data of member states from 2007 (based on a press release of the European Union, October 2008) it appears that most member states (15 members) closed the financial year of 2007 with a deficit. Only Greece (3, 5%) and Hungary (5%) closed with a deficit greater than 3%. The above data again represent the unique failure of recent Hungarian economic policy. If we take a close look at the dynamics of the Hungarian deficit index, we see that it slummed in 2002 and stayed at a low level, 8-9% of the GDP in 2006.

Let us not console ourselves with the fact that the leading economy of the European Union, that is Germany, also struggles with budget deficit — the shouldering of the German Democratic Republic by the German Federal Republic meant a serious burden on the German budget. Let us compare ourselves to Finland

[2] It is worth mentioning that the Maastricht concept of debt differs from the one used in financial accounting, and Hungarian authorities are to create a special notion for it deducted from the financial account data. The concept of debt accepted in Maastricht does not include portions of property, only cash, deposits, securities (without derivatives), and credits. It is a gross concept; claims are not deducted from debts. Particular debt items are calculated at face value rather than at marketable value.

who realised 5, 3%budget surplus, and we must wait the coming 20th anniversary of the democratic transformation of Hungary with discontent.

5. Finances of the corporate sector

To use a precise technical term in international statistics, we look at the financial accounts of non-financial corporations.[3]

It is a technicality of financial corporations that their assets and commitments change in the same rate, thus the amount of their net monetary assets, their balance is almost zero. Net monetary assets of non-financial corporations are, however, usually negative, since these corporations do not invest their financial resources into assets. (Exceptions to this are those *special purpose enterprises* whose function is to intermediate money between their foreign partners instead of financing production. For this reason they are not included in corporate account statistics.)

Finance requirements of corporations reached their peak during the second half of the 1990's, due to a boom of investments. The year of 2002 was a confine – enterprises became savers. Most obviously, this is an unfavourable phenomenon since it indicates that in the economy has less realisable income than in the financial sector, that is it is better to save income then reinvest… Where does the financial sector transfer the savings of the enterprises? The market shows that, because of the high yields on state securities, financial intermediaries concentrated on financing the state. This means that the corporate sector provided credit for the state occasionally...For instance in 2002, when the population's financing capacity decreased almost to zero, corporations became net financier. The corporate sector became net borrower again in the early years of the 2000.

In the resource structure of the corporate sector owners' shares and credits dominate: the owners' shares constitute about fifty percent. (The EU average is 55%, that is, somewhat higher.) Credits reach one-third of resources. (This rate is 29% in the European Union). Interestingly, credit ratio is much lower in Poland or Slovakia: it is around 20%. Commercial credits are more significant in these countries. *Foreign currency debts* of the corporate sector amount to more than 40% of the GDP (p. 61). Non-stock based securities are not common in former socialist countries such as Hungary, and they are rarely seen in the European Union either. Large-scale presence of such securities can be detected only in the USA, where credit ratio is very low, only 9% (p. 62).

[3] Financial corporations are too, enterprises as all the others, even if they operate with specifically greater external financial resources. For exactly this reason they form a different category. Undeniably, if we emphasise the entrepreneurial nature of banks their public service function is underplayed

Corporate resources are to provide for operation costs and the corporation's assets. The importance of monetary assets out of all assets is ascending, which is due to ownership and credit relations of corporations. Out of all assets cash, deposits, credits and ownership claims have a significant ratio. The greatest proportion is formed by other active debts, which amount to one third and include outstanding liabilities. Regrettably enough, the enforcement of lending through delayed settlements of outstanding liabilities has become fashionable recently. This process is not among the positive parameters of market relations; a process which we cannot be proud of in relation to democratic transition in Hungary.

As regards the *proprietary structure* of the corporate sector, we can establish that state ownership –as we have seen studying the state's asset structure- was characteristic of Hungary in 1989 as most corporations were owned by the state. In the past 18 years the ratio of state-ownership has lowered from 85% to 7 %. Private ownership rate reached 20% by 2007, while the percentage of foreign ownership has increased even more spectacularly. *Foreign* ownership rate reached 25% by the middle of 1990's and doubled by 2007. Property value of foreign investments increased from 39 billion forints (in 1989) to 15.200 billion forints. *Nowadays, foreign investors own fifty percent of Hungarian enterprises.*

The ratio of capital inflow and outflow is what counts. If national investors have at least as much invested capital abroad as the inflow of foreign capital –no problem should arise. At least, if their yields transferred back and forth are – by and large –balanced. In this case only the advantages of the division of labour are experienced and all parties make a profit. When the flow of capital is one-sided and economic development is based on external resources, a fundamental difference between the GDP and the GNI should be expected.

In an OECD report (OECD 2004) one can find significant and relevant data concerning the issues above. There is no significant difference between GDP and GNI in most of the OECD countries , there were no major differences between the development of the two indexes, *except in the cases of two or three countries – including Hungary.*

In Hungary the outflow of incomes was far higher than inflow; thus *GNI remained much lower than GDP*. This indicates indirectly that the ratio of foreign-owned corporations, the profit of which is mostly repatriated, is significantly high.

The welfare of a country's population depends on the income remaining inside the country. National consumption is (or might be) increased in the long run, if national income is spent or invested at home. Incomes transferred abroad work the same way – only in a foreign country. IF an FDI, it creates in the other country incomes in form of wages. Besides these issues, the dynamics of GDP and GNI are also strongly influenced by the *terms of trade*. Unfavourable terms of trade mean

that through the exchange of home made goods for imported goods income is drawn out of the country. According to certain calculations (see Botos J., 2009), the decrease in the terms of trade in the last ten years were almost as much as the GDP of a base year. Because of the price formations of external trade, that amount, one base year's GDP, flowed out of the country in 7 years. This could happen since one or two years of improvement was counterweighted by the deterioration of these terms in the remaining years. Analysts may come to think that the explanation lies in foreign-owned corporations. These corporations may influence the income remaining in the country by means of import, which in many cases comes from their subsidiaries... The formulation of transfer prices can be especially interesting from the point of view of taxation – besides, it can also influence GNI indexes. Keeping the setting of prices under control is a complicated task, the importance of which, in our view, has so far been overlooked by Hungarian economic leaders.

6. Summary

After the fall of communism financial accounts of market participants changed profoundly. On the one hand, the ratio of business share assets increased in the housholds' portfolio; on the other hand, it has decreased to the minimum in the state portfolio. The winners of all this are foreign investors– who became the owner of a significant proportion of Hungarian national wealth. Both the 'inherited' and the continuously 'produced' public finance deficit - financed from external resources- played a part in the above process. We had to sell out our wealth for foreign owner to finance our debts.

Economic policy after the fall of communism was mainly based on export-led development. We can conclude that such an economic policy, which was guided by external markets, is a cause and an effect at the same time – and did not serve social welfare in Hungary properly in the past decades. It is an effect of the significant outstanding total debt inherited from the socialist controlled economy, which made the economy become export-directed for the sake of foreign currency acquisition. It is also a cause because it invited too much foreign capital into Hungary as FDI-s. These foreign investments took the majority of their profit out of the country, blocking the way in certain sense for both present and future welfare development. It seems that Hungarian standards of living have little chance of reaching the Western European level in the near future.

We seem to have manoeuvred ourselves into an economic policy that proves to be dead end-if the final aim of any economic policy is to increase welfare in the society.

References

Antalóczi, K. – Botos, K. 1990: Lesz-e konvertibilis a magyar forint? A klíringtől a konvertibilitásig. Közgazdasági és Jogi Kiadó, Budapest.

Botos, J. 2009: Mi a teendő? Georgikon konferencia előadás, Keszthely.

Botos, K. 1987: Világméretű pénzügyi egyensúlyhiány. Közgazdasági és Jogi Kiadó, Budapest.

Botos, J. – Botos, K. 2005: *What can we expect from joining the EU?* In: Essays in Economics and Economic History. Heller Farkas Füzetek 1, Budapest.

Boriani, R - Johanson, A- d'Ercole, M. M. 2006: *Alternativ measures of well–being.* OECD, Social employment and working papers, 33.

Honvári, J. 2006: XX. századi magyar gazdaságtörténet. AULA, Budapest.

Magyarország pénzügyi számlái 2008. Magyar Nemzeti Bank, Budapest.

Nagy, Cs. 2007: Az IMF-csatlakozás elvetélt kísérletei. Heller Farkas Füzetek, Budapest.

Romsics, I. 2005: Magyarország története a XX. században, Osiris, Budapest.

Vígvári, A. 1990: *On the History of Indebtedness in Hungary*. Paper to the third International Karl Polanyi Conference 1990. November.

Dynamic Estimation of the Hungarian Term Structure

Szabolcs Kopányi [1]

This paper focuses on dynamic properties of the Hungarian term structure. As Hungary is a key European emerging market empirical findings might offer value for both researchers and practitioners. The yield curve and its dynamics are first characterized by descriptive statistical analysis that is followed by a Principal Component Analysis (PCA). A semi non-parametric (SNP) study investigates structural dynamics of the yield curve without making parametric assumptions, then a stochastic mean reverting affine model (3-factor Vasicek model) is calibrated to the sample which is shown to work relatively more accurately in the Hungarian bond market than in the American one. The last section is devoted to forecasting future yield curves, where empirical results are somewhat less convincing.

Keywords: term structure of interest rates, affine model, Kalman filter, emerging markets

1. Introduction and motivation

Bonds represent claims for future cash flows, show the time value of money. The term structure of interest rates (yield curve) summarizes market expectations of a given time regarding the aforementioned time value of money; in our case shows how bond yields depend on time-to-maturity at each moment of time. It produces discount rates for risk-adjusted cash-flows in numerous financial applications.

Despite its key importance, the term structure is not directly observable. Therefore we need to estimate it. Term structure estimation evolved into two distinct though still related problems of finance. The first tries to produce a continuous yield curve on the back of some traded prices: this is the *static approach*. The curve is a snapshot of a given market, just as shown on Figure 1 depicting the Hungarian government bond market. The data source is the Hungarian Government Debt Management Agency (GDMA).

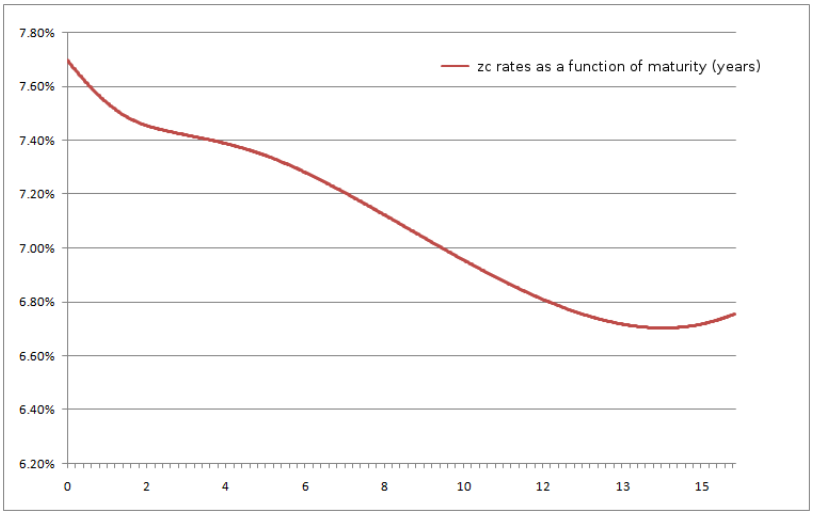
Static curve estimation is feasible via bootstrapping, Ordinary Least Squares (OLS) or Generalized Least Squares (GLS) and yield curve fitting techniques (e.g. cubic splines).

The second problem in finance, which this paper is devoted to, is related to the panel study approach and *focuses on dynamics of interest rates and the term structure*. The question is: how can we describe the evolution of interest rates over

[1] Szabolcs Kopányi, PhD student, Corvinus University Budapest, Investments and Corporate Finance Department (Hungary)

time? The concept is similar to how the evolution of a share price or a foreign exchange rate is estimated over time. Only similar, because the term structure – unlike share prices and FX rates – is not a scalar quantity. Different points (i.e. maturities) of the term structure cannot relate to each other in arbitrary ways, one must ensure that no-arbitrage rules apply.

Figure 1. GDMA zero coupon yield curve on 2 Jan 2008



Source: Government Debt Management Agency (GDMA)

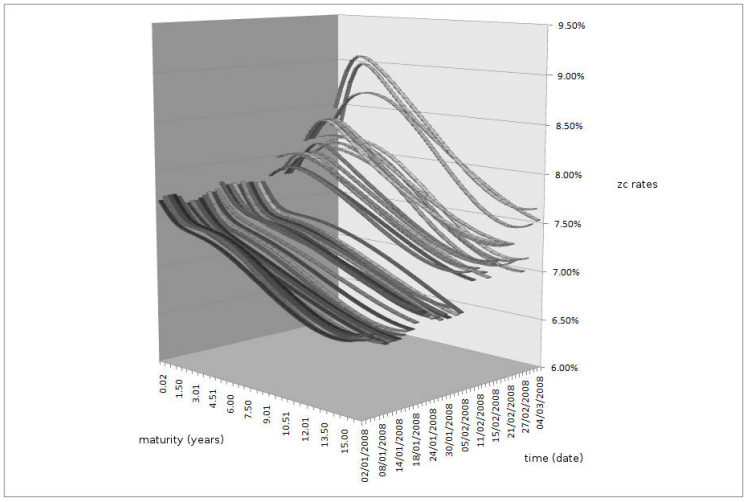
In the case of dynamic structured modelling estimation begins with selection of an interest rate model. To find an appropriate interest rate model is in itself a highly complicated issue, since there are dozens of frequently quoted models in literature. There is no universal interest rate model, therefore researchers often find their models as part of the estimation procedure (non-parametric estimation).

Estimation with structured models focuses on obtaining the distribution of the underlying stochastic variable(s) in the interest rate model. Shall this be infeasible (the pricing stochastic differential equation has no analytic solution) various moments of these distributions are estimated. The underlying stochastic variable is often not observable (e.g. volatility in models with several factors), therefore it has to be estimated as part of the estimation procedure. Only the sky and the lack of creativity limit the scope for empirical models.

The viewpoint of the dynamic approach is shown on Figure 2, displaying the evolution of the zero coupon yield curve as shown in Figure 1 over the period 2 January 2008 to 3 March 2008.

The *goal of this paper is to investigate Hungarian term structure dynamics via econometric methods*. My empirical models might help market actors *understand evolution of the yield curve* and *generate out-of-sample forecasts*. The *structure of this paper* is as follows. The first section presents descriptive statistical measures for the yield curve and a PCA model is run to guide stochastic modelling of the term structure, in terms of how many factors are truly needed for a proper fit. Then a semi non-parametric model is estimated to investigate yield curve dynamics without making any parametric assumptions. This is followed by stochastic modelling, where parametric assumptions are made and results of the PCA studies are facilitated. Here belongs the main conclusion of this paper: the 3-factor Vasicek model works particularly well for the Hungarian term structure. In the last section out-of-sample forecasting potential of the 3-factor Vasicek model is investigated.

Figure 2. GDMA zero coupon term structure between 2 Jan and 3 Mar 2008

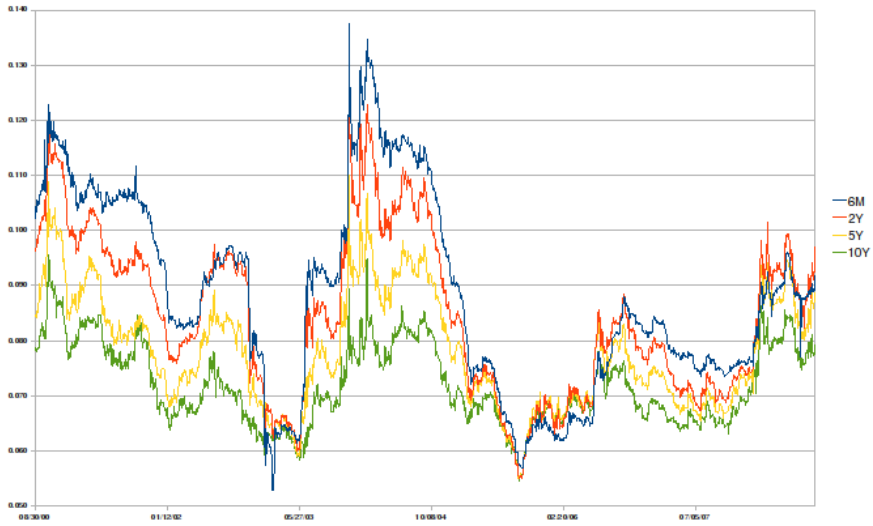


Source: Government Debt Management Agency (GDMA)

2. Descriptive statistics and PCA

For my empirical research I used a zero coupon sample of government bond data, collected on a daily basis between 1998 and 2008 by the GDMA. Recorded maturities are 2 week, 1 month, 3 month, 6 month, 9 month, 1 year, 2 year, 3 year, 4 year, 5 year, 6 year, 7 year, 8 year, 9 year and 10 year. Selected maturities are displayed on Figure 3.

Figure 3. Time series of sample data (N=2007)

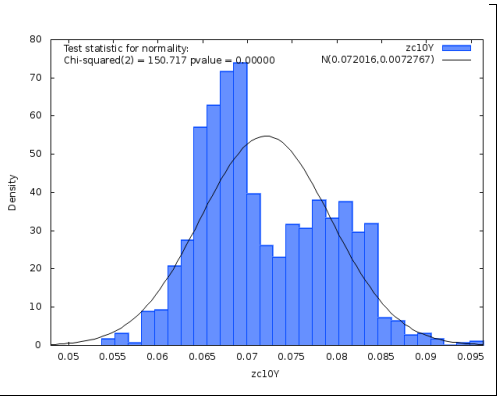


Source: Government Debt Management Agency (GDMA)

As the chart on Figure 3 shows, Hungarian interest rates have been volatile over the observation period. Also, the histogram of the 10 year tenor (shown on Figure 4) perfectly highlights the dual mode nature of the Hungarian market. It underscores a fact in a statistical way that is widely known by market participants from their own experience: things are either "very good" or "very bad" in the Hungarian bond market. "The pendulum often moves excessively into both extremes and only reverts with cumbersome slowness to standstill". The question whether that particular *standstill* can be regarded as fair value remains open to be answered by a macroeconomic study. If yes, we deal with excessive market reactions, if no, the bipolar nature of Hungarian economic processes warrant bimodal yield levels.

Figure 4. Descriptive statistics of the 10 year tenor

Average	0.07201
Median	0.06997
Minimum	0.05445
Maximum	0.09561
Standard deviation	0.00728
C. V.	0.10104
Skewness	0.35442
Ex. kurtosis	-0.64330



Source: own calculations

To judge how many factors are needed to describe Hungarian term structure dynamics, it is worth to carry out a Principal Component Analysis (PCA). After the famous paper of Litterman and Scheinkman (1991), it is common to assume that three factors, namely *level, steepness* and *curvature* drive the whole yield curve. Running a PCA on the Hungarian data set revealed that the first three factors cumulatively explain 99.81% of yield levels' covariance; results for daily yield changes show 95.71% cumulative explained covariance for the first three principal components.

Table 1. PCA of yield levels and their first differences

Principal components[2]	**Cumulated explained covariance (yield levels)**	**Cumulated explained covariance (daily yield changes)**
Factor 1	0.9470	0.6692
Factor 2	0.9933	0.8994
Factor 3	0.9981	0.9571
Factor 4	0.9994	0.9848
Factor 5	0.9998	0.9970

Source: own calculations

With the results shown in Table 1, we can conclude that with appropriate calibration, 3-factor stochastic models should produce a good fit for the Hungarian term structure.

[2] Principal components are different in the two PCA studies.

3. Semi non-parametric model calibration

Over the course of semi non-parametric (SNP) analysis of the data sample, I started with plain VAR models (with lags 1, 2, 3 and 4), first these have been calibrated to the sample. Going further I continuously extended parameterisation of the auxiliary model (e.g. with ARCH and GARCH processes) and judged their significance by using two information criteria (AIC and BIC). Optimisations have been carried out with a method by Gallant and Tauchen (1996), using control runs to ensure robust results (i.e. not falling in the trap of a local minimum).

Results showed that Hungarian term structure dynamics are governed by a semi non-parametric GARCH process. The conditional variance of the auxiliary model is a VAR(1), GARCH(1,1) process, innovations are given by a 6th order polynomial with a time lag of 1. I investigated if using polynomials as coefficients of the innovation polynomial or introducing asymmetric volatility (leverage effect) into the model improve auxiliary model fit; but I did not get a confirmation in any of the two cases. SNP model calibration has been carried out for the tenors 6 month, 2 year, 5 year and 10 year in a pure time series approach, and for all these maturities combined in a panel approach. Different estimations confirmed each other.

My results for the Hungarian market are easily comparable with a study by Dai and Singleton (2000) referring to US markets. Authors there found the best score for a VAR(1), GARCH(1,2) auxiliary model, with innovations represented by a 4th order polynomial with a time lag of 1. All these let us conclude that *structural dynamics of the American and Hungarian yield curves are quite similar*, despite the fact that the Hungarian term structure had been an inverted one throughout the entire observation period. A partial explanation for this similarity might be, that *emerging market investors follow developed markets closely*, and *core market developments usually lead to important repercussions in emerging markets*.

4. Kalman filter calibration of the 3-factor Vasicek model

In this section, parametric assumptions are made, and a 3-factor stochastic model is calibrated to the sample. I chose the affine (i.e. constant plus linear) model family, and a mean reverting stochastic model by Vasicek (1977) for its simplicity and favourable applicability. In the Vasicek model, the vector of state variables, $X_{i,t}$ is driven by a mean reverting affine diffusion:

$$dX_{i,t} = \kappa_i(\theta_i - X_{i,t})dt + \sigma_i dW_{i,t},$$

for $i = 1, 2$ and 3. κ_i denote the strength of mean reversion, θ_i stand for long term factor means, σ_i mark factor volatilities and W_i are independent Wiener processes.

Model calibration has been carried out via the Kalman filter[3]: the likelihood function of the stochastic process has been reproduced with the Kalman filter, and this reproduced likelihood function has been estimated with maximum likelihood (ML).

In order to make an international comparison, I calibrated the 3-factor Vasicek model for the US bond market with a similar sample (daily observations of 15 maturities between 2001 and 2009), too. The results are shown in Table 2[4].

Table 2. Estimated parameters of the 3-factor Vasicek model

Parameter	Vasicek (HUF)	Vasicek (USD)
θ_1	0.000	0.020
θ_2	0.000	0.000
θ_3	0.000	0.039
κ_1	0.170	0.004
κ_2	0.675	0.246
κ_3	1.000	0.581
σ_1	0.022	0.007
σ_2	0.099	0.045
σ_3	0.033	0.012
λ_1	-0.217	-0.054
λ_2	-0.271	-0.383
λ_3	-0.330	-0.007
Average in-sample fit (bp)	8	5
Average sample yield level (%)	8.17	4.64
Relative model fit (bp per 100 bp yield level)	0.98	1.08

Source: own calculations

How do we interpret results? Comparing 3-factor model results in the Hungarian and American samples we find that: 1) long run factor means in the Hungarian sample are zero (θ_i parameters are bound to zero value due to admissibility issues), whereas in the US case we get numbers different from zero, these are the values where factors converge to. 2) The mean reversion process is stronger (κ_i values are higher) in the Hungarian sample than in the American case. 3) Estimated volatility parameters can be interpreted quite intuitively: factor volatilities in the emerging Hungarian

[3] The Kalman filter, which is named after the Hungarian Rudolf Kalman, is particularly well suited to deal with problems where state variables (or a part of them) are latent. It involves an *a priori* estimate via system dynamics, then this prediction is combined with information gained from observations to refine the state estimate; the improved estimate is termed the *a posteriori* state estimate.

[4] λ_i parameters characterize market prices of risk.

market are several-fold of those in the US treasury market which is the most important and also the most liquid in the world. 4) Average model mismatch (i.e. in-sample forecast residuals) seems to be slightly worse in the Hungarian sample, but if we normalize results with average yield levels, the *Hungarian model calibration turns out to be relatively more accurate than that for the American market*. Relative model fit (average in-sample residual per 1%p yield level) is 0.98 bp in the Hungarian market versus 1.08 bp in the American sample.

To demonstrate how well the 3-factor Vasicek model fits the sample, Figure 5 plots the 2 year maturity and its in-sample model estimate.

Figure 5. In-sample model fit of the 3-factor model (2 year tenor)

Source: own calculations

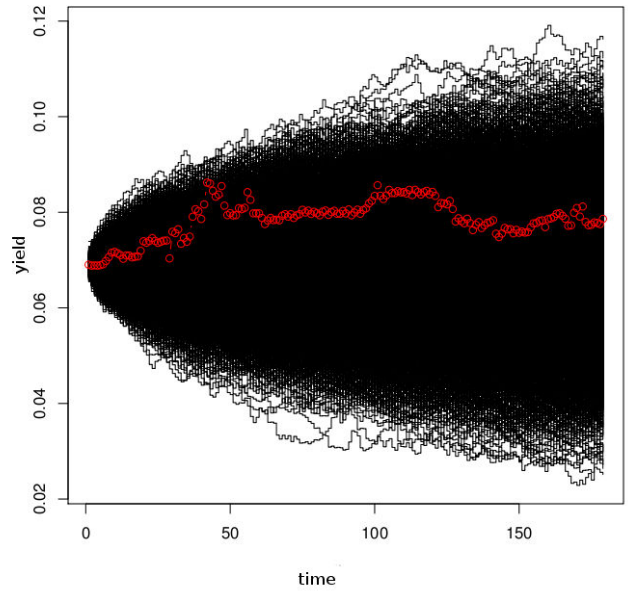
5. Forecasting future yield curves

With calibration of the 3-factor Vasicek model, I investigated its in-sample forecasting capabilities. Given the fact that the 3-factor model fit the Hungarian sample quite well, Hungarian term structure dynamics became easily understandable in a *retrospective way*. It remains an open question, though, whether the investigated model is capable to produce valuable out-of-sample forecasts. That constitutes a further step ahead, *"if we understand so much about the drivers of the yield curve, let us tell what the future holds"*.

Using a sample shortened by 180 days (i.e. $N' = 2007 - 180 = 1827$) I recalibrated the 3-factor Vasicek model. The results, including average residuals, dif-

fered only marginally from original-sample-results, therefore I do not disclose them separately. Still, the use of a shortened sample makes the calibration process realistic and correct. With the resulting parameters, I ran simulations for the whole panel consisting of 15 maturities, producing N''=180 observations (i.e. the forecast horizon is 6 months). I "manufactured" missing term structure observations this way. I then repeated the simulation process 10 thousand times, to evaluate the model not on a single trajectory but rather on a fanchart-like, but empirically plotted graph. Comparing simulated trajectories with the last 180 original daily observations, we find how accurately the model would have forecasted future interest rates (*backtesting*). Average daily forecast errors have been saved for all trajectories, in order to generate an average measure for all 10 thousand trajectories. Figure 6 shows the probabilistic nature of realizations: simulated trajectories of the 10 year maturity are plotted with thin black lines (the darker trajectories are, the more likely they are to realize), last 180 original daily observations are dotted with red marks.

Figure 6. Simulated trajectories of the 3-factor Vasicek model (10 year tenor)



Source: own calculations

The average forecast error of the 10 year tenor on a 6 month forecast horizon is 108 bps (as shown in Table 3), however the mode of the distribution (i.e. the 10 thousand

trajectories) is around 50 bps. That is an acceptable level for the volatile Hungarian market.

Still, Figure 6 underlines that out-of-sample forecasts have to be dealt with healthy cautiousness: the presented simulation is a valuable tool to forecast expected ranges of future interest rates, but it is definitely not an oracle to tell the winning market bet.

Table 3. Backtesting results based on 10 thousand trajectories

Maturity	**Average out-of-sample forecast errors (bp)**
2 week	282
1 month	279
3 month	270
6 month	258
9 month	247
1 year	241
2 year	215
3 year	190
4 year	173
5 year	159
6 year	144
7 year	134
8 year	123
9 year	116
10 year	108
Average of all tenors	196

Source: own calculations

Average backtesting errors are shown in Table 3, referring to 10 thousand simulated trajectories of all 15 maturities. It is clearly visible that *out-of-sample forecast errors are significantly higher* (25-fold more) *than their in-sample counterparts*. Also of note, is that *longer maturities* (i.e. less volatile parts of the yield curve) *are easier to predict*. In this example, I had the most accurate results with the 10 year tenor in the Hungarian sample. *Out-of-sample forecasting accuracy can be naturally improved by applying shorter forecast horizons*. With a forecast horizon of 1 week I got 13 bp average forecast error for the 10 year tenor, based on 10 thousand trajectories. This amounts approximately to 1.5-fold market bid-ask spread, i.e. it is a relatively acceptable result. Given that, the combined 2-hour runtime for the estimation and simulation algorithms is quite luring.

6. Concluding thoughts

This paper reaches the main conclusion that the 3-factor Vasicek model is a favourable choice for dynamic models on the Hungarian term structure. This statement is supported by empirical evidence regarding the model's in-sample forecasting potential. The 8 basis point average estimation error is first negligible with relation to the Hungarian market (practically it amounts to one unit bid-ask spread) and second reveals better relative (as corrected with average yield level) in-sample fit in the Hungarian market than in the US one.

When a researcher has to deal with structural modelling of the yield curve, they might consider the following points. It is sensible to choose an interest rate model for structural modelling which we have an efficient tool for to estimate. They shall avoid having a too complicated model which has to be calibrated by a non-linear estimation method which in turn cannot even recognize simple functional dependencies? On these grounds, I chose the affine model family and the Vasicek model. Decisions regarding the number of modelling factors are best guided by PCA studies. For empirical research on the Hungarian term structure I recommend the use of 3-factor models. Regarding estimation methods I had positive experience with the Kalman filter.

Considering out-of-sample forecasting potential of the 3-factor Vasicek model, one has to note that the model has somewhat limited potential for true forecasting purposes and results have to be interpreted by healthy cautiousness. The model is not an oracle to "tell the winning lottery draw", but a tool to show a range of expected future interest rates. As a practical note, less volatile maturities of the term structure should be used for forecasting purposes.

Concluding from the results detailed above the target audience of the presented methodology is rather the National Bank of Hungary, the Government Debt Management Agency and the Hungarian Financial Supervisory Authority. Commercial banks might find the methodologies useful, too; though their benefits are more likely to show up as more efficient risk management than hard profits of proprietary trading desks.

References

Ahn, D. H. 1995: A generalized squared autoregressive intertemporal term structure model. *University of North Carolina working paper*, 1995.

Ahn, D. H.-Dittmar, R. F.-Gallant, A. R. 2002: Quadratic term structure models: Theory and evidence. *The Review of Financial Studies*, 2002/1, 243-288. p.

Aït - Sahalia, Y. 1996a: Nonparametric pricing of interest rate derivative securities. *Econometrica*, 1996/5, 527–560. p.

Aït - Sahalia, Y. 1996b: Testing continuous-time models of the spot interest rate. *Review of Financial Studies*, 1996/2, 385–426. p.

Balduzzi, P. - Das, S. R. - Foresi, S. - Sundaram, R. 1996: A simple approach to three factor affine models of the term structure. *Journal of Fixed Income*, 1996/12, 43–53. p.

Bansal, R.- Zhou, H. 2002: Term structure of interest rates with regime shifts. *Journal of Finance*, 2002/10, 1997–2043. p.

Beaglehole, D. R. - Tenney, M. S. 1991: General solutions of some interest rate-contingent claim pricing equations. *Journal of Fixed Income*, 1991/9, 69–83. p.

Beaglehole, D. R. - Tenney, M. S. 1992: A nonlinear equilibrium model of the term structure of interest rates: Corrections and additions. *Journal of Financial Economics*, 1992/12, 345–353. p.

Brennan, M. J. - Schwartz, E. S. 1979: A continuous time approach to the pricing of bonds. *Journal of Banking and Finance*, 1979/7, 133–155. p.

Brigo, D. - Mercurio, F. 2006: *Interest Rate Models – Theory and Practice*. Springer Finance.

Chen, L. 1996: *Stochastic mean and stochastic volatility – a three-factor model of the term structure of interest rates and its application to the pricing of interest rate derivatives*, Blackwell Publishers.

Chen, R. R. - Scott, L. 1993: Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. *Journal of Fixed Income*, 1993/12, 14–31. p.

Constantinides, G. 1992: A theory of the nominal structure of interest rates. *Review of Financial Studies*, 1992/4, 531–552. p.

Cox, J. C. - Ingersoll, J. E. - Ross, S. A. 1985: A theory of the term structure of interest rates. *Econometrica*, 1985/3, 385-408. p.

Dai, Q. - Singleton, K. J. 2000: Specification analysis of affine term structure models. *Journal of Finance*, 2000.10., 1943-1978p.

Duffee, G. R. 2002: Term premia and interest rate forecasts in affine models. *Journal of Finance*, 2002/2, 405–443. p.

Duffie, D. - Kan, R. 1996: A yield-factor model of interest rates. *Mathematical Finance*, 1996/10, 379–406. p.

Duffie, D. - Singleton, K. 1997: An econometric model of the term structure of interest rate swap yields. *Journal of Finance*, 1997/12, 1287–1323. p.

Gallant, A. R. - Tauchen, G. 1996: Which moments to match? *Economic Theory*, 1996/12, 657-681. p.

Heath, D. - Jarrow, R. A. - Morton, A. 1992: Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 1992/1, 77–105. p.

Ho, T. S. Y. - Lee, S. B. 1986: Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 1986/12, 1011–29. p.

Leippold, M. - Wu, L. 2002: Asset pricing under the quadratic class. *Journal of Financial and Quantitative Analysis*, 2002/6, 271-295. p.

Litterman, R.-Scheinkman, J. A. 1991: Common factors affecting bond returns. *Journal of Fixed Income*, 1991/6, 54-61. p.

Longstaff, F. A. 1989: A nonlinear general equilibrium model of the term structure of interest rates, *Journal of Financial Economics*, 1989/2, 195-224. p.

Sargent, T. J. 1979: A note on maximum likelihood estimation of the rational expectations model of the term structure. *Journal of Monetary Economics*, 1979/1, 133-143. p.

Tuckman, B. 1995: *Fixed Income Securities: Tools for Today's Markets*. Wiley.

Vasicek, O. 1977: An equilibrium characterization of the term structure. *Journal of Financial Economics*, 1977/11, 177–188. p.

Simple Methods to Assess the Credibility of the Hungarian Inflation Targeting Regime

Miklós Erdélyi[1] *–Szabolcs Szikszai*[2]

In our paper, we introduce simple approaches to assess the credibility of the Hungarian inflation targeting (IT) monetary policy regime in the period between July 2001 and April 2009. First, we present the findings of a non-representative survey we conducted amongst the major stakeholders of monetary policy. The results of the survey suggest that, according to financial market players, the National Bank of Hungary (MNB) lost some of its credibility between 2002 and 2007.We then apply simple methods to assess the success of the IT regime and to measure the predictability and gradualism of the decisions of the MNB's Monetary Council (MC) and the consistency of central bank communication. The results of this analysis show that the time series of the inflation measure using unchanged tax rates became stationary after the introduction of IT in 2001, while the predictability and gradualism of MNB MC's policy rate decisions and the consistency of central bank communication each displayed low levels in the observed period.

Keywords: communication, credibility, inflation targeting, predictability

1. Introduction

1.1. Credibility of the inflation targeting monetary policy regime

Since price stability became the primary objective of monetary policy authorities around the world, inflation targeting has been the most widely acclaimed central banking technique. No country that has ever introduced an IT regime abandoned it later. Blinder (2006) points out that countries which desire disinflation are more

[1] Miklós Erdélyi, Msc (IT), Bsc (Applied Economics), PhD student, University of Pannonia, Faculty of Information Technology (Hungary)
[2] Szabolcs Szikszai, Msc (Economics), senior lecturer, University of Pannonia, Faculty of Economics (Hungary)

Acknowledgement: We are grateful for exchange rate transaction data kindly made available to us by Reuters Hungary Ltd. and fixed income data provided by the National Bank of Hungary. We also thank László Andor, Attila Csajbók, Júlia Király, Csaba László, Péter Mihályi, Gábor Oblath, Klára Pintér and Balázs Vonnák for useful comments. All remaining errors are entirely ours.

likely to introduce inflation targeting but its advantages for any type of economy are clear. In a wide cross-country panel study, Mishkin and Schmidt-Hebbel (2007) show that inflation targeting helps keep inflation low in the long run, improves the efficiency of monetary policy and strengthens central bank independence in various types of economies ranging from large industrial economies such as Germany, Canada or the United Kingdom to emerging economies such as Brazil, Turkey or Korea. However, their analysis also concludes that the monetary policy performance of industrial-country inflation targeters dominates that of emerging countries. That is, inflation targeting has so far been more successfully adopted by industrial economies than by emerging ones.

For an inflation targeting regime to be successful in an emerging economy – i.e. to achieve its inflation target –, it needs to be more than the simple announcement of a medium-term inflation target. According to Mishkin (2004), the strategy of an inflation targeting central bank also needs to be information inclusive, transparent and supported by good communication, with policy-makers held accountable for the success of their policy actions. All this is what, Mishkin (2004) says, it takes to make inflation targeting succeed in reducing inflation in emerging economies.

1.1.1. The role of expectations in monetary policymaking

Using Michael Woodford's phrasing, central banking is an art of "management of expectations" (Woodford, 2005, p. 3), and central banks can consider themselves successful when they manage market expectations in a way that leads to lower inflation in the long run. A part of these expectations is formed of the evolution of market prices of financial assets and another part is formed of the policy moves of the central bank itself as reacting to changing economic and market conditions. Based on certain signals, financial investors might expect a central bank to act in a way that is in line with its strategy (e.g. reducing inflation) or they might expect otherwise. To put it differently, investors might think of a central bank as a more or a less credible one. Either way, the result of their expectations is fed back into the prices of financial assets, continuously changing the market setting in which the central bank operates. Thus, monetary policy success hinges to a great extent on the ability of the central bank to make itself look credible in the eyes of financial market investors.

1.1.2. Concepts of credibility of monetary policy

A general perception of a central bank's credibility is a key component of success not only in itself but also because it can contribute the most to reaching the policy target. In fact, all of the points mentioned by Mishkin (2004), above, have to do with the drive on the central bank's part to establish credibility of its actions among the investors in financial markets. This is because, if a monetary policy regime is thought to be credible, expectations of the target variable become well-anchored to the target and fewer and more gradual policy moves are required to reach this target

(Amisano-Tronzano 2005). In this context, market perceptions of the inflation aversion of an inflation targeting central bank play a crucial role in making inflation targeting operate more effectively. As Blinder (1999) points out, a more credible central bank can engineer disinflation with a lower level of employment sacrifice as changes in the key monetary policy instrument will cause smaller GDP/employment gap volatility.

In his study based on a survey among central bankers and economists, Blinder (1999) shows that central bankers mostly identify credibility with "dedication to price stability" (Blinder 1999). An important, and difficult, issue is to measure this dedication. Obviously, when an inflation targeting regime is successful in keeping inflation low, it becomes credible. In other words, a central bank becomes credible when it matches its words with its deeds, i.e. if it says that it will and, indeed, does bring inflation down. But this is a simple, *ex post* assessment of credibility, based on the track record of the inflation-targeting central bank. A more dynamic assessment of credibility, suited to real-life circumstances, comes from the measure of perceived, *ex ante* credibility as seen by investors in financial markets. In this context, an inflation targeting central bank, dedicated to price stability, is said to be credible when the perception of its pre-commitment to disinflation or, in other words, its aversion to inflation as perceived by and implied by market prices of financial assets is high.

1.1.3. Predictability of monetary policy

After Mishkin (2004), transparency is an important element of a credible inflation targeting regime, but this can generally be said of any kind of monetary policy regime. Perez-Quiros and Sicilia (2002) opine that transparency "facilitates the understanding of what the central bank does and by doing so, it helps central banks to foster their credibility" (Perez-Quiros-Sicilia 2002). The practical manifestation of the transparency of a central bank is the predictability of its policy moves. Assessing the degree of predictability of policy moves, therefore, has been given great consideration in the academic literature, lately. The degree of predictability as implied by financial market prices, according to Perez-Quiros and Sicilia (2002), shows how transparent a central bank is in the eyes of financial market participants. In other words, predictability is a measure of how well market participants understand monetary policy strategy and how precisely they anticipate future policy actions under various market conditions. In this vein, higher predictability of policy moves helps build central bank credibility. While measuring credibility itself can prove to be a difficult task, measuring predictability of both market interest rates in general and base rate decisions in specific is more straightforward. Evaluating credibility as implied by market prices gives a good proxy for a central bank's credibility as seen by market participants.

1.1.4. Central bank communication

An issue of high importance is how credibility can be established. Still after Mishkin (2004), a key component of building credibility by making it more predictable is good communication, which has lately become a central goal in monetary policy-making (see Woodford 2005). The most important task of a central bank striving for good market perception of its credibility is to communicate well. This is because communication gives a firsthand feedback to financial investors as to whether the central bank consistently follows the monetary strategy laid out for itself. First of all, investors can test whether communication is aligned with the monetary policy strategy. In case of an inflation targeter, each act of communication has to reflect the logic of disinflation. Financial investors will then match central bank communication with policy actions and see for themselves whether communication is useful or useless in signaling policy moves. If financial investors see a central bank that "will do what it says" (Blinder 1999), they will conclude that the signals of this particular central bank are indicative and will act accordingly. This will be reflected in the prices of major financial assets, i.e. they will move in the direction in which the central bank ushers them with communication. This is the process of building credibility. In case communication is inconsistent with policy actions, however, financial investors will conclude that central bank communication is confusing and will ignore it and focus instead on the very policy actions. Investors will reach the same conclusion when communication is inconsistent within itself. It might happen e.g. when policy-makers communicate the direct opposite of what they communicated a day before or, when policy-makers communicate ambiguously. These are the processes of destroying credibility.

In short, besides being a dynamic endogenous variable, credibility is also an important exogenous factor in policy-making. After economic agents obtain a certain perception of the inflation aversion of a central bank – partly through its communication –, this perception of credibility then feeds back into monetary policy actions and increases (when credibility is high) or decreases (when credibility is low) their effectiveness.

1.2. Evolving perceptions of MNB's credibility in the IT regime

The issue of the market perception of credibility is especially delicate in Hungary, where, in June 2001, the Monetary Council decided to switch to another monetary policy regime by introducing a converging inflation target. As it has been shown in earlier examples[3], the need to establish a strong feeling of an inflation-averse author-

[3] These examples include any central bank switching from one monetary policy regime to another – e.g. inflation targeting –, or, as it happened to the European Central Bank (ECB), any central bank being established from the scratch. The emphasis in the case of MNB is on the switch to another regime coupled with new legislation warranting full independence for MNB – a constellation of conditions never seen before.

ity amongst financial investors is of key importance for the central bank to operate a successful inflation targeting regime (see the case of the European Central Bank in Goldberg-Klein 2005). In other words, for a central bank switching to another regime, like the National Bank of Hungary (MNB) did in 2001, the most important, but also the most tedious and, sometimes, painful task is the establishment of a strong market perception of credibility.

Seemingly, the success of the new inflation targeting regime in Hungary was warranted by new legislation that brought major changes into the operational background of monetary policymaking[4]. But, as it eventually turned out, the MNB had to grapple with the still dual nature of the Hungarian monetary policy framework. There emerged an inherent contradiction between the IT regime and MNB's legal obligation to keep the euro-forint exchange rate within a ±15% floatation band around the parity rate. Raising the policy rate to bring inflation down strengthened the forint against the euro and, thus, threatened to push the exchange rate out of the floatation band. Still, the MNB had to establish the credibility of its inflation targeting regime and maintain the credibility of the floatation band of the euro-forint exchange rate at the same time. The eventual clash of these monetary policy tasks had an adverse impact on the market perception of credibility of MNB's IT regime[5]. Sadly enough, the market perception of the central bank's credibility was exacerbated by inconsistent central bank communication. Apart from that, the conspicuous lack of coordination between fiscal and monetary policy-makers – an example of which is presented in Madár (2003) – contributed to MNB's missing the inflation target in the majority of the cases and further damaged the credibility of the IT regime in the eyes of the investors.

[4] The new legislative background of monetary policy was created in three key steps. First, on May 3, 2001, MNB MC and the government jointly decided to widen the floatation band of the euro-forint exchange rate from ±2.25% to ±15% as of May 4. The crawling devaluation of the floatation band was later abolished as of October 1, 2001. Second, MNB MC decided to introduce the IT regime on June 12, 2001. Third, the new Act LVIII on the National Bank of Hungary came into effect on July 13, 2001, providing MNB full operational and personal independence in its efforts to achieve and maintain price stability. All this information along with the latest version of the central bank act is available at www.mnb.hu.

[5] Press evidence (see Nyíri 2003 or The Economist 2004, amongst others) supports the view that the market perception of credibility of MNB's IT regime suffered a decisive blow in the turbulent period between October 2002 and November 2003, when the central bank was continually faced with the dilemma to decide between acting in the logic of the IT regime and acting to keep the euro-forint exchange rate within the floatation band. When, due to speculative pressure, the exchange rate reached the strong limit of the band, the MNB decided that maintaining the band was more important. But, later, it decided with the government to shift the parity of the band in the weaker direction, which, then, forced it to put aside the IT logic again to defend the exchange rate of the forint. This chain of events was especially detrimental to the perception of credibility of the IT regime and, apparently, MNB later could by no means recover the substantial loss of credibility suffered in this sub period (Rozkrut et al. 2007, p. 194).

The purpose of this paper is to provide insight into the evolution of the financial market perception of the credibility of MNB's IT regime in the period between June 2001 and April 2009. We seek answers to the following basic questions related to the credibility of the IT regime, which might serve as springboards to more complex issues of central bank credibility[6]:

1. What does narrative evidence tell us about the evolution of the perceived credibility of the Hungarian monetary policy in the observed period?
2. Was the introduction of inflation targeting in Hungary successful in anchoring inflation?
3. Were MNB MC's interest rate decisions predictable in the observed period?
4. Did central bank communication signal future policy rate decisions properly?

2. Narrative consensus on the concept of credibility in Hungarian monetary policy

We conducted a short survey with the purpose of investigating the market perception of credibility of monetary policy in general and that of the monetary policy conducted by the MNB between 2002 and 2007. Between May and August 2008, we asked chief financial officers of domestic corporations, analysts of leading commercial banks and consulting firms and other economists following monetary policy events about what they think about credibility in general and about MNB's credibility in specific.

2.1. Participants of the survey

The questionnaire and the list of responding institutions are available on request, but we insert here, in table 1, the composition of the respondents according to sectors.

[6] This paper is part of an extensive research by the authors into the evolving market perception of credibility of the Hungarian monetary policy. In this paper, we concentrate – without going into historical details – on the fundamental issues of the IT regime alone and disregard credibility issues connected with the exchange rate band and fiscal policy.

Table 1. Number of respondents by sectors

Energy, public utility	9	22%
Finance	9	22%
Manufacturing	7	17%
Commerce	5	12%
Transport, postal, info communications	5	12%
Consulting	3	7%
Construction	1	2%
Regulation	1	2%
Education	1	2%
Total	**41**	**100%**

Source: own creation

Our methods of distributing the questionnaire included mailing the questionnaire to 85 companies listed in the 2007 edition of TOP 200 – an annually printed special issue of weekly Figyelő, introducing the biggest Hungarian companies (Figyelő 2008) –, plus sending direct e-mails containing the questionnaires to some TOP 200 CFOs and 25 analysts and economists working in domestic banks, consulting firms and research institutes. We approached 7 other, middle-sized, Hungary-based companies not listed in TOP 200 by way of personal acquaintance with their managers. We also created an on-line version of the questionnaire[7], where CFOs and analysts could answer the questions anonymously. We have e-mailed the link to this on-line version of the questionnaire to only a few CFOs but to all the 25 analysts mentioned above.

The number of respondents was 41, which gives a response ratio of 35%. (Surprisingly, the addition of the on-line version of the questionnaire failed to improve this ratio significantly.) Companies operating in the energy sector and in finance both gave 22% of all respondents, and are closely followed by respondents from manufacturing (17%). These 3 groups made up 61% of the total number of respondents, while the share of these sectors in the Hungarian GDP[8] was 29.4% in 2007. We must add, however, that distortions in the representation of individual sectors vis-à-vis their actual weights in GDP are caused partly by the fact that the biggest of our respondent companies operates in more than one sectors. Finance is over-represented in our survey – 22% as opposed to GDP-weight of 4.6% - as analysts follow monetary policy decisions the most closely and their opinions are the most relevant. Manufacturing is adequately represented, somewhat below its GDP-weight of 21.9%. The weight of respondents from commerce is 12%, which equals the sector's weight in GDP, while transport-communications is over-represented – 12% instead of 8.2%, its GDP-weight. The representation of the rest of the sectors is be-

7 The questionnaire is available at http://5mp.eu/fajlok/szixai/kerdoiv_www.5mp.eu_.html.

8 The GDP-weights of sectors are available at www.ksh.hu.

low their GDP-weights. As for the weight of the respondent companies in the corporate sector, 49% of all respondents are listed in the TOP 200. They give 24% of the total net revenue of the first 200 Hungarian companies and can, thus, be considered a sizable and representative group.

Whatever the weights, however, based on the above mentioned number of respondents one can hardly consider this survey a representative one. Its findings are, nevertheless, very informative since we get a clearer picture on the opinion of financial experts about the meaning of credibility of monetary policy and the credibility of the Hungarian central bank.

2.2. Findings of the survey

In the questionnaire, we asked how closely respondents followed monetary policy events between 2002 and 2007 and weighted answers accordingly. Answers of respondents who kept close track of all events (67% of total) - including both decisions and communication - were given a weight of 1. Answers of those respondents who only followed either monetary policy communication or decisions (14%) were given a weight of 2/3. And, finally, answers of those who only followed monetary policy events occasionally (19%) were given a weight of 1/3.

Two-thirds of all respondents agreed that the credibility of monetary policy is determined by the strategy of the central bank ("a coherent central bank strategy aiming to reach a set of announced monetary policy goals"). 59% of all respondents also agreed that credibility means good communication ("plausible central bank communication, capable of orienting markets"). 41% said that credibility is warranted by a good track record ("reaching the announced monetary policy goals") and 37% said that credibility means transparency ("transparent monetary policy rules"). Only 17% agreed that credibility means accountability ("accountability of central bank decision-makers"). These results suggest that strategy and communication seem to be the core components of monetary policy credibility, according to our respondents, while accountability is hardly an issue. The track record and transparency of monetary policy are somewhere in the middle: they are of some, but not decisive importance.

As for the evaluation of monetary policy-making in Hungary between 2002 and 2007, our respondents gave the MNB the highest point for its transparency: an average of 3.12 out of 5. Strategy came second with an average of 3.07 and track record followed with an average of 2.86. Falling behind was communication and accountability with averages of 2.47 and 2.28, respectively. This suggests that respondents valued the MNB's transparency and strategy the most, while they were more critical of its communication and accountability over the observed period. It is fair to say that they were neutral on its track record.

We also wanted to find out more about the gradual evolution of opinions on these credibility issues over the investigated period. According to the majority who answered questions related to this issue, we can say that the MNB was worse off in

all but one aspect at the end of the observed period compared to the beginning. The only exception was strategy, which they thought was more coherent at the end of the entire period than it was in 2002. The other four aspects – communication, transparency, track record and accountability – all received less points at the end of the period that in the beginning. Our respondents also indicated in their answers that their overall assessment of the MNB's credibility worsened considerably after October 19, 2002 – the beginning of the sub period of market turbulence and high volatility of the exchange rate. The setback was the most severe (-23%) in the assessment of the MNB's track record and transparency. We think that the formal was probably a direct consequence of the central bank's having missed its 2003 and 2004 inflation targets, while the latter is a manifestation of the confusion – caused by sharp course reversals of monetary policy – over the relative importance of reaching the inflation target and maintaining the floatation band of the exchange rate. Although the overall assessment of the MNB recovered somewhat after the volatility of domestic financial markets subsided, this recovery was, nevertheless, insufficient to improve the perception of the MNB's credibility back to its initial levels. This latter suggests that the MNB could not fully repair the damage done to the assessment of its credibility during the times of market turbulence.

2.3. Divergent opinions about credibility outside the survey

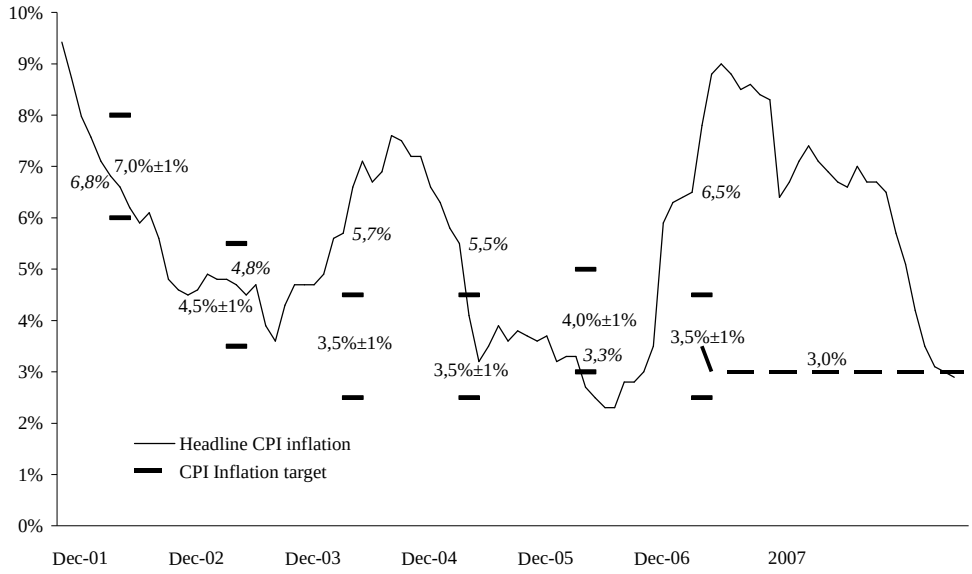
Concluding from answers to the questions of our questionnaire, economists working in the business environment identify the concept of credibility of monetary policy with different, although interrelated concepts. But, besides the results of the survey, it is also worth looking at the remarks that our questionnaire provoked our respondents to make. One analyst who declined to participate in the survey said that the questionnaire was too complex and that he did not have an opinion about the evolution of credibility in the individual sub periods. He pointed out that the only measure of credibility was that investors who are wrong in their assessment of credibility of monetary policy in a turbulent period will have to pay the price for their wrong assessment.

Furthermore, three respondents called our attention to the fact that we should also deal with the contribution of fiscal policy to the deterioration of perceived credibility of monetary policy and it was partial on our side to only examine the credibility of monetary policy. They said that it was probably fiscal rather than monetary policy mistakes that caused most of the turbulence in financial markets and monetary policy alone could not help avoid it. Another respondent opined that changes in the volatility of the euro-forint exchange rate should not be considered as the basis of differentiation between the perceptions of credibility within the individual sub periods. We understood this as saying that changes in exchange rate volatility had little or nothing to do with changes in the perception of credibility.

3. Statistical analysis of the success of MNB's IT regime

The first, *ex post*, approach to the concept of credibility, as we put forward in the introduction, is judging the track record, that is, the success of the central bank in meeting its target. The findings of our survey also suggest that this analysis is an important element of assessing credibility. In case of the Hungarian IT regime, it is plain to see from chart 1 that the inflation target – which is set together by MNB and the Ministry of Finance – was missed 5 out of 9 years (2003, 2004, 2006, 2007 and 2008).

Chart 1. The inflation target and actual inflation rates

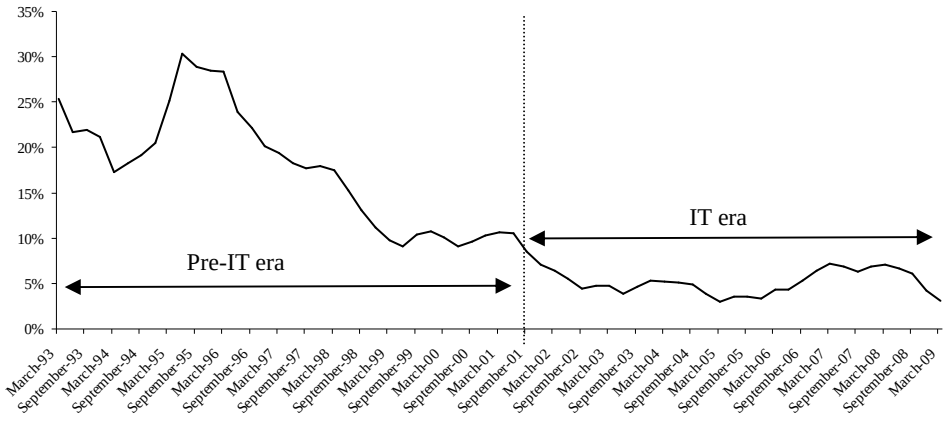


Source: own collection from www.ksh.hu and www.mnb.hu

Since the Hungarian headline CPI inflation target is a common nominal target of the monetary and fiscal policy, and the target can only be attained through the coordination of the monetary and fiscal sides, the simple fact that the target was missed in a majority of the cases so far tells little about the credibility of the IT regime. To investigate MNB's responsibility in this issue, it is necessary to narrow our analysis. We will first look at a simple comparative statistical overview of the period before and after the introduction of IT in June 2001 using a seasonally adjusted inflation measure, which filters out the impact of tax rate changes (VAICPI computed by the Central Statistical Office (KSH)). Clearly, it was most often the exogenous changes in VAT, excise and other tax rates by fiscal policy that deterred the headline CPI

inflation rate from the target. Filtering out tax rate changes from the headline CPI inflation measure better shows just how much MNB contributed to lowering and/or stabilizing the inflation rate. It can be seen from a look at chart 2 that VAICPI became lower and less volatile in the IT era compared with the pre-IT era.

Chart 2. Seasonally adjusted quarterly average change of year-on-year VAICPI



Source: own collection from www.ksh.hu, www.mnb.hu and MNB staff

For a detailed comparison of the two periods of similar length, table 2 presents the fundamental statistics.

Table 2. VAICPI dynamics in the pre-IT and the IT eras

	Q1 1993-Q2 2001	**Q3 2001-Q1 2009**
Number of Observations	34	31
Average	17.75%	5.24%
Standard Deviation	6.46%	1.40%
Coefficient of Variation	36.37%	26.64%
Persistence	84.64%	85.53%
Range	21.23%	5.55%
Minimum	9.10%	2.98%
Maximum	30.33%	8.54%

Source: own calculations based on data from www.ksh.hu and www.mnb.hu

Apparently, all relevant statistics point to an amelioration of inflation dynamics after the introduction of IT, as attested by the VAICPI measure. The volatility and average of inflation subsided, along with their ratio, the coefficient of variation. The range, maximum and minimum all decreased substantially. The only exception is the

persistence of inflation, which is the extent to which the average inflation rate of the previous quarter determines the average inflation rate of the following quarter. The fact is that slightly higher inflation persistence is even desirable in a generally lower inflation environment, which characterizes the IT period vis-à-vis the pre-IT era. Therefore, it does not alter the overall perception that, from a static point of view, inflation was harnessed in following the introduction of IT.

3.1. Inflation stationarity before and after the introduction of IT

It should be noted, however, that inflation is a dynamic phenomenon, so this simple statistical comparison should be complemented by a more advanced test that accounts for the dynamic nature of inflation. We apply the augmented Dickey-Fuller test with drift to see if the above time series of the seasonally adjusted quarterly average change of year-on-year VAICPI has a unit root. The equation we use in this test is as follows:

$$\Delta y_t = a_0 + \sigma y_{t-1} + \sum_{j=1}^{k} \theta_j \Delta y_{t-j} + u_t ,$$

where y_t is the logarithm of the inflation measure and k, the lag order of the autoregressive process, is initially set to equal 5.

The results of the test are unsurprising given the precedents. In the pre-IT period, we cannot conclude to reject the null hypothesis of $\sigma=0$, meaning that VAICPI inflation was a non-stationary process until June 2001. Regardless of the lag length, the test statistics show that, in the pre-IT era, the inflation rate of previous periods failed to significantly determine the inflation rate of the following period. On the other hand, the test statistic for the IT-era was significant at 5% level, with a p-value of 0.0223 and a Durbin-Watson statistic of 2.0683. The test statistic was significant using shorter or longer lag operators, as well. Based on this result, one can conclude that VAICPI inflation became mean-reverting or, in other words, anchored, after the introduction of IT. Although the cause-effect link is unclear, one can reasonably assume that this was at least partly owing to the anti-inflation measures taken by MNB according to the logic of the IT regime.

4. Comparative analysis of MNB MC's predictability

We will now go on to discuss the simple *ex ante* aspects of credibility. We begin with the concept of predictability, which, as we stated in the introduction, is the manifestation of a central bank's transparency and, as such, is a prerequisite for building credibility. Here, we follow the methodology first introduced by Gaspar, Perez-Quiros and Sicilia (2001) and later used by Ross (2002) and Perez-Quiros and Sicilia (2002) to assess the predictability of the European Central Bank. Pintér and

Wenhardt (2004) applied this methodology first on the Hungarian experience. This simple methodology can be used in a relative context to compare the predictability of the decisions of MNB MC with those of the decision-making bodies of other central banks.

This model compares the actual short-term market interest rate after the interest rate decision with expectations as reflected in short-term market interest rates before the decision. The model assumes that interest rate decisions can have two possible outcomes: one is when the MC decide to change the base rate (the interest rate paid on the two-week MNB bond, which is the key policy-making tool of the central bank) by either 25 or 50 basis points and another is when the base rate is left unchanged. That is

$$E_t(i_{t+1}) = 25\beta + i_t, \text{ or} \quad [1]$$

$$E_t(i_{t+1}) = 50\beta + i_t, \quad [2]$$

where $E_t(i_{t+1})$ is the base rate expected for time t+1 at time t as implied by the short-term market interest rate (2-week Bubor in the Hungarian case) ahead of the rate decision, and i_t is the base rate before the decision, both measured in basis points. β is the probability of an interest rate change – positive when a rate hike is expected and negative otherwise. When β is more than or equal to 50% (or 0.5), we judge that the market expects a rate change ahead of the rate decision. Models [1] and [2] only differ in the presumption that interest rate changes usually amount to either 25 or 50 basis points in the monetary policy routine. In MNB's case, model [2] seems a more suitable choice.

We have created three plus one different categories to evaluate the predictability of base rate decisions based on how well these decisions were anticipated by the short-term market interest rate. Based on this, a decision is precisely anticipated, imprecisely anticipated, or unanticipated and there is also the possibility that the market expects a change in vain. We consider a decision precisely anticipated when the difference between $E_t(i_{t+1})$ and i_{t+1} is less than 12.5 basis points in model [1] or less than 25 basis points in model [2]. A decision is considered imprecisely anticipated when the market expects the direction of the change well but misses its extent by 12.5 basis points or more in model [1] or by 25 basis points or more in model [2]. A decision is deemed unanticipated if the market misses the direction of the change or if it expects a change when there is none or if it expects no-change when there is change. The fourth (plus one) category includes instances when interest rate changes are anticipated in vain, that is, when either a rate cut or a rate hike is expected but does not happen.

The predictability of MNB MC's base rate decisions is worse in most aspects than that of the central banks of 4 developed and 4 emerging countries – or region in

the case of the Euro zone – in our sample[9]. According to model [1], MNB ranks 5th (2nd in the emerging peer group) in the overall hit rate that quantifies the share of all – precisely or imprecisely – anticipated decisions. It has the worst record (9th) in the anticipation of interest rate changes, and has the 3rd highest record in the anticipation of interest rate no-changes. As for the reliability of the indication of interest rate changes expected by the market, MNB's record is only better than that of the NBP (8th).

According to model [2], respective ranks are 6th, 8th, 5th and 7th. Using this model, the relative forecasting ability of market interest rates is slightly lower because model [2] shows deterioration in the predictability of interest rate no-changes compared to model [1]. On the other hand, the relative reliability of anticipations of interest rate change is higher than in model [1][10]. Model [2] predictability statistics are presented in table 3.

Table 3. Predictability of central bank rate decisions by model [2]

	MNB	**ECB**	**Fed**	**Riks-bank**	**BoE**	**SARB**	**CBRT**	**CNB**	**NBP**	**Average**
Overall Hit Rate	*67%*	89%	75%	82%	76%	61%	63%	70%	56%	**73%**
Rate Changes	*22%*	63%	70%	34%	31%	62%	68%	19%	38%	**43%**
No-changes	*91%*	94%	78%	95%	96%	60%	58%	97%	67%	**86%**
Reliability of Changes	*58%*	70%	63%	73%	70%	58%	54%	75%	38%	**64%**
Number of Meetings	*139*	157	130	72	144	54	43	138	100	-

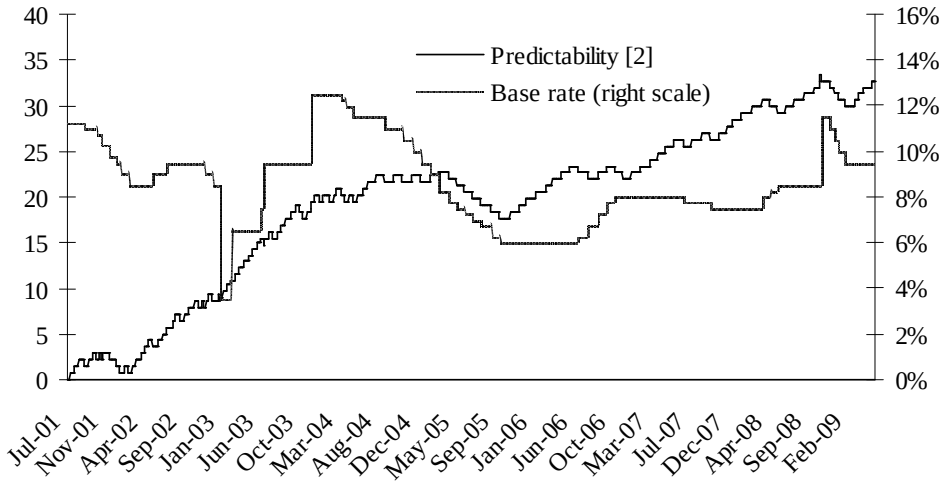
Source: own calculations based on data available on central bank websites

In assessing the predictability of MNB MC's decisions in itself, it should be noted that the predictability of their interest rate changes is made low by the poor predictability of their decisions to lower the policy rate. To illustrate this fact, we sketched in chart 3 the cumulated predictability as of model [2] on a scale of ±100 points together with the policy rate in the period stretching from July 2001 to April 2009.

[9] Developed peers in the sample are the ones with the world's oldest, most established monetary systems: the United States (Fed), the Euro zone (ECB), the United Kingdom (BoE) and Sweden (Riksbank). Emerging peers are taken from the so called Europe, Middle East, Africa (EMEA) region: the Czech Republic (CNB), Poland (NBP), Turkey (CBRT) and South Africa (SARB). In the entire sample, the Fed is the only central bank that does not operate a formal IT regime, but is also responsible for maintaining price stability as a central task.

[10] This has to do with the specifications of the model, since surprises caused by rate decisions according to model [1] are not necessarily surprises in model [2]. Therefore, model [2] tends to improve the anticipation ratio of policy rate no-changes, while it reduces the anticipation ratio of rate changes.

Chart 3. Cumulated market predictability of MNB MC's base rate decisions



Source: own calculations based on data from www.mnb.hu

Chart 3 shows that MNB MC's cumulated predictability as of model [2] is +33 points (+13 points in model [1]) and there were both ups and downs in the path of the curve. Most of the downward spirals were triggered by serial rate cut decisions (e.g. the series of rate cuts in early 2002, 2005 or late 2008). According to model [2], financial investors anticipated only 12% of MNB MC's rate cuts (24% in model [1]), while they anticipated 44% of the rate hikes (88% in model [1]) in the observed IT period. It seems that, for some reason, financial investors find it harder to predict MNB MC's rate cut decisions, which weighs considerably on the relative assessment of MNB's credibility.

On a whole, there is no significant difference for MNB between the results of model [1] or [2]. In both models, MNB's average rank is between 6th and 7th in the overall sample, and MNB's scores exceed the sample average only in the case of anticipation of interest rate no-changes. It means that the relative predictability of MNB MC's interest rate changes is low. Table 4 summarizes the rankings of central banks by the aspects of the two predictability models.

Table 4. Relative predictability of MNB MC decisions

	Model [1] – 25-basis-point Changes				Model [2] – 50-basis-point Changes			
Rankings	**Overall Hit Rate**	**Rate Changes**	**No-changes**	**Reliability of Changes**	**Overall Hit Rate**	**Rate Changes**	**No-changes**	**Reliability of Changes**
1.	ECB	ECB	CNB	CNB	ECB	Fed	CNB	CNB
2.	CNB	Fed	ECB	Fed	Riksbank	CBRT	BoE	Riksbank
3.	Fed	SARB	*MNB*	CBRT	BoE	ECB	Riksbank	ECB
4.	BoE	CBRT	Fed	SARB	Fed	SARB	ECB	BoE
5.	*MNB*	BoE	BoE	BoE	CNB	NBP	*MNB*	Fed
6.	CBRT	Riksbank	Riksbank	ECB	*MNB*	Riksbank	Fed	SARB
7.	Riksbank	CNB	CBRT	Riksbank	CBRT	BoE	NBP	*MNB*
8.	SARB	NBP	NBP	*MNB*	SARB	*MNB*	SARB	CBRT
9.	NBP	*MNB*	SARB	NBP	NBP	CNB	CBRT	NBP

Source: own calculations based on data available on central bank websites

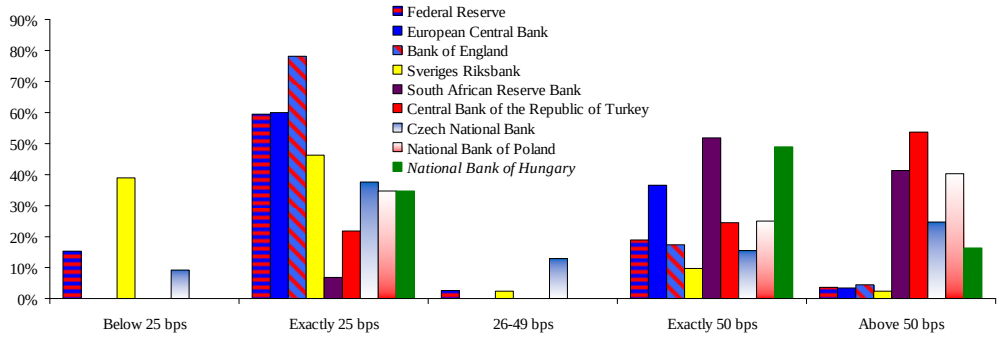
4.1. Gradualism of MNB MC's interest rate decisions

As suggested by Blinder (2006), an important element in the predictability of central bank decisions is the frequency distribution of interest rate changes, which shows how gradual these changes are. In MNB's case, a comparison with the average of the rest of the central banks in the sample seems useful[11].

The frequency distribution chart, presented in chart 4, reveals that MNB MC displayed a tendency to make less gradual interest rate decisions than the average of the sample. Comparing the frequency distribution of MNB MC's rate decisions to that of the average of all 8 other central bank in the sample, we can conclude that MNB has operated with changes of 25 basis points or lower less often than the average (35% vs. 56%), while it has more 50-basis-point changes on record than the average of the sample (49% vs. 24%).

[11] When computing the averages we weight by the actual number of interest rate decisions.

Chart 4. Gradualism of central bank rate decisions



Source: own calculations based on data available on central bank websites

Comparing MNB routine to that of the central banks of developed countries in the sample, the contrast is even starker as 77% of all interest rate changes in the developed countries' peer group have been changes of 25 basis points or lower, and only 18% have been changes of 50 basis points. In comparison with the average of the emerging peer group – which excludes Hungary, of course –, the ratio of small changes (0.25% or lower) is almost equal (35% for MNB MC vs. 33% for the peer group). On the other hand, the ratio of MNB MC's 50-basis-point changes has been almost double that of the emerging peer group (49% vs. 25%), while the ratio of changes of more than 50 basis points is lower (16% vs. 37%). Clearly, MNB MC relied significantly more on 50-basis-point rate changes and less on changes of more than 50 basis points in the observed period than the average of the emerging peer group[12].

When we restrict the timeframe of observation to a shorter period, in which all the central banks in the sample were given independence in changing the policy rate according to the logic of the IT regime – expect, of course, for the Fed, which does not operate such a regime –, MNB MC's gradualism appears to have fallen behind even the average of the emerging peer group. In conclusion, MNB MC have made less gradual policy rate decisions than the average of the peer group in both observed periods, and their decisions become less gradual when we only take into account the more recent years.

[12] Especially, the SARB and the CBRT tend to have frequently changed policy rates by more than 0,5%. As we show in the next subsection, it has to do with the fact that the level of interest rates in these two countries was mostly above 10% throughout the period. This later argument is bolstered by the fact that the Polish Monetary Policy Council abandoned the practice of changing the policy rate by more than 0.5% after its policy rate fell below 10%.

4.2. The correlation of ranks by predictability, gradualism and the average level of policy rates

This can, of course, be explained partly by the lower level of interest rates in developed economies: the lower the level of interest rates, the smaller the extent of interest rate changes necessary to reach a certain monetary policy goal. It can be said that 50-basis-point changes become scarce under interest rate levels of 5-6%. For instance, the Hungarian policy rate was only below 6% temporarily – for 5 weeks – in the entire examined period[13], which is one explanation for MNB's lower rank in gradualism.

Table 5 shows the rankings of central banks according to predictability, gradualism, maximum and minimum levels of policy rates in the respective IT periods[14], all ending in April 2009. Apparently, there is a clear gap between the groups of central banks of developed and emerging countries. In the developed group and in the Czech Republic, the average level of the policy rate was below 5%, and these central banks proved to be more gradual and predictable in their decisions than the rest of the sample.

Table 5. Central bank rankings in the IT period

Central Bank	Average ranking	Predictability[15] [1]	Predictability [2]	Gradualism[16]	Average Level of Policy Rate	Average Level of Policy Rate	Observed Period
ECB	2.25	3.	1.	4.	1.	3.06%	01.1999-
Fed	3.25	2.	5.	3.	3.	3.94%	01.1995-
CNB	3.25	1.	3.	5.	4.	4.44%	01.1998-
BoE	3.5	4.	4.	1.	5.	4.99%	06.1997-
Riksbank	3.5	8.	2.	2.	2.	3.12%	01.2000-
MNB	*6.75*	*7.*	*6.*	*7.*	*7.*	*8.62%*	*07.2001-*
NBP	7.5	9.	9.	6.	6.	6.75%	01.2001-
CBRT	7.5	5.	8.	8.	9.	15.84%	01.2006-
SARB	7.5	6.	7.	9.	8.	9.81%	01.2001-

Source: own calculations based on data available on central bank websites

[13] From January 17 until February 24, the policy rate was 3.5%, the rate of interest MNB paid on overnight deposits placed by commercial banks with the central bank. The availability of the 2-week deposit was restricted throughout this period.

[14] This applies for all except for the cases of Poland and Sweden. For Poland, we could only find market interest rates from January 2001 to calculate predictability, although the NBP has been operating IT since 1999. For Sweden, we applied January 2000 instead of January 1993 – the formal start of the IT regime – as the beginning of the period of observation since the Riksbank were granted full independence in their interest rate decisions at that point of time.

[15] Ranks are calculated as the average of ranks by the overall hit rate and the reliability of changes.

[16] Ranks are set in order of the share of interest rate changes of 25 basis points or lower.

Calculating the Spearman rank correlation[17] for the correlation of the different rankings in the IT periods, we conclude the followings. The link is very strong (0.85) between the ranking by predictability as of model [2] and the ranking by the average level of policy rates throughout the period. The ranking by gradualism and by the average level of policy rates also show strong correlation with each other (0.77). (The ranking by the first predictability model does not show significant correlation with any other rankings.)

Table 6. Spearman rank correlations in the IT period

	Predictability [1]	**Predictability [2]**	**Gradual-ism**	**Average Level of Policy Rate**
Predictability [1]	-	0.47	0.25	0.33
Predictability [2]	-	-	0.67	***0.85***
Gradualism	-	-	-	*0.77*

Source: own calculations based on the rankings in Table 5

In table 7, we present rankings based on a broader period of observation, which is the longest period that we found relevant for measuring the gradualism of policy rate decisions and the average level of policy rates for the respective central banks. MNB's rank improves one notch as it takes 6th place in gradualism and the average level of interest rate from NBP. However, this improvement is, in fact, deterioration as it indicates that NBP started off in a less favorable interest rate environment and could still lower its policy rate and improve the gradualism of its decisions more successfully than MNB.

[17] The Spearman rank correlation is $\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$, where d_i is the difference between the individual country ranks by two different aspects and n is the total number of ranks in the sample. *Italic* figures in the tables show significance at the 5% level of significance (ρ>0,683), **bold** at 2% (ρ>0.783) and ***bold and italic*** at the same time at 1% (ρ>0.833).

Table 7. Central bank rankings in the broader period

Central Bank	Gradualism[18]	Average Level of Policy Rate	Average Level of Policy Rate	Maximum Level of Policy Rate	Minimum Level of Policy Rate	Observed Period
Riks-bank	1.	2.	3.98%	8.91%	0.5%	06.1994-
BoE	2.	4.	4.99%	7.5%	0.5%	06.1997-
Fed	3.	3.	4.60%	9.625%	0 - 0.25%	08.1987-
ECB	4.	1.	3.06%	4.75%	1.25%	01.1999-
CNB	5.	5.	5.80%	39%	1.75%	12.1995-
MNB	*6.*	*6.*	*8.62%*	*12.5%*	*3.5%*[13]	*07.2001-*
NBP	7.	7.	9.55%	24%	3.75%	01.1998-
CBRT	8.	9.	23.42%	57%	9.75%	02.2002-
SARB	9.	8.	10.05%	13.5%	7%	11.1999-

Source: own calculations based on data available on central bank websites

The Spearman rank correlations based on the broader period show an even closer correlation between the ranking by predictability as of model [2] based on data from the previously observed IT period, and the ranking by gradualism and by the average level of policy rates, as seen in table 8.

Table 8. Spearman rank correlations in the broader period

	Predictability [2]*	Gradualism	Average Level of Policy Rate
Predictability [2]*	-	**0.78**	**0.78**
Gradualism	-	-	***0.87***

*Based on the shorter, IT period.
Source: own calculations based on the rankings in Table 7

These results suggest that the rankings of central banks by gradualism and predictability are closely correlated with their ranking by the average levels of interest rates, which is a way of saying that central banks appear to be more predictable in their decisions in an international comparison if the average level of their policy interest rate is lower and they take a more gradual approach in their decisions. The reverse of this argument partly explains why MNB ranks behind central banks of developed countries and the Czech Republic in predictability. MNB could neither lower interest rates to the critical level nor take a gradual enough approach in its decision-making to make its decisions become significantly more predictable by financial investors.

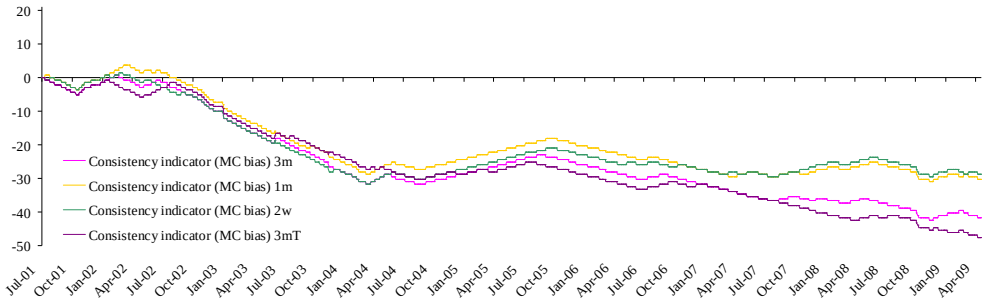
[18] Ranks are set in order of the ratio of interest rate changes of 25 basis points or lower.

5. Consistency of communication

Communication, as shown by the literature review as well as the results of the survey, is a seminal component of successful monetary policy-making. Much has been said about MNB's communication with financial markets but little actual analysis has been carried out to see whether the communication of MNB' governor and the Monetary Council helped build the credibility of the IT regime or not. An exception is Rozkrut et al. (2007), who, amongst others, introduce a new measure, the so called consistency indicator, to appraise the consistency of central bank communication. The consistency indicator can be applied to measure the consistency of the communication of the Monetary Council or that of the central bank governor ahead of each interest rate decision. Similarly to cumulated predictability, it is calculated as a cumulative measure building in the values given for each rate decision. We give a value of 1 if the sign of the difference between the geometric mean of the base rate in the three months following the rate decision and a market (e.g. 2-week, 1-month, 3-month Bubor, 3-month T-bill) interest rate on the day before the decision coincides with the sign of communication[19] by the MC or the governor in the period leading up to the rate decision. This is because if the two signs coincide then it can be said that the central banks oriented investors in the right direction. Taking the cumulated value of this indicator and plotting it on a ±100-point scale, we get chart 5, which can be viewed as a track record of communication.

[19] Naturally, a positive value in communication is an early indication of tightening and a negative is an early indication of easing. In case of the MC, we discern the monetary policy bias from the communiqué published after the most recent rate decision and consider it as indication of future monetary policy ahead of the upcoming rate decision. MC communication is given a value of 1 when it is deemed to be hawkish – hinting a rate hike in the future –, -1 when it is thought to be dovish – hinting a rate cut in the future – and 0 when it is considered to be neutral in its predictions of the path future monetary policy will take. In the case of the governor we take the average of his comments in the period between two consecutive rate decisions and use it as indication of the direction of the next rate decision. Similarly to the MC case, values of the individual comments made by the governor are given 1, -1 or 0 depending on their monetary policy bias. The only difference is that, in the case of the governor, there can be more than one comment in the interim periods. Thus, the sum of the values attached to the individual comments is divided by the total number of comments made in the interim period.

Chart 5. Consistency indicator of MNB MC's monetary policy bias

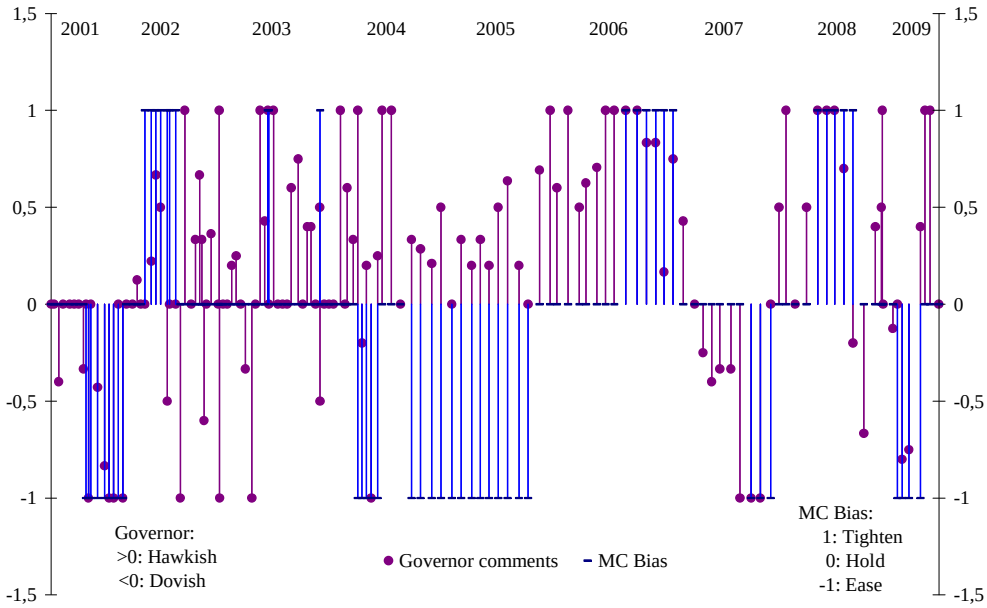


Source: own calculations based on data from www.akk.hu and www.mnb.hu

The values of this consistency indicator for MNB MC range from -29 to -47 depending on which market interest rate we use in our calculation. We get similar figures (-32 to -49) when we plot the consistency indicator of the governor's communication ahead of rate decisions[20]. These figures send a strong message to policy-makers: they suggest that neither the bias statements included in the MC's post-decision communiqués nor the statements made by the central bank governor in the interim period served as good predictors of subsequent policy changes. To put it in a starker form, the central bank's MC and its governor more often misled financial investors speculating over the future path of policy rates than guided them in the right direction. And not only did they fail to indicate the future path of policy rates, the messages conveyed by the two sources of communication were more often contradictory than unequivocal. In chart 6, we plotted the direction of indications of the future path of policy rates as implied by MNB MC communiqués and comments by the governor in between rate decisions in the period from July 2001 until April 2009. Times when the indications point to different directions – e.g. in 2004 and 2005, the governor continuously communicated tightening, while the MC communiqués signaled easing or maintaining monetary conditions – are when central bank talk is simply counterproductive and destroying credibility.

[20] To calculate these latter figures we collected all comments made by the governor in the observed period. Our sources included www.budacash.hu, www.ebroker.hu, www.fn.hu, www.hirtv.net, www.index.hu, www.mno.hu, www.napi.hu, www.nol.hu, www.otpbank.hu, www.portfolio.hu, www.raiffeisen.hu, www.reuters.hu, www.tozsdeforum.hu, www.vg.hu. Our database is available on request.

Chart 6. Communication on future policy by MNB MC and the governor



Source: own calculations based on data from www.mnb.hu and own press collection

6. Conclusions

The above analysis is only a preliminary investigation into the complex concept of evolving credibility of the Hungarian IT regime. It avoids a number of delicate issues – such as the inherent conflict between the IT regime and the exchange rate band or the (lack of) coordination with fiscal policy – in order to simplify the analysis. It presents the results of the observations mostly without explaining the facts. This analysis should be complemented by a thorough account of the historical conditions in which MNB had to operate in the observed period, as some of these conditions played a decisive role in central bank decisions. We plan to elaborate on these aspects in another paper.

Nevertheless, we can draw a few conclusions from the results by answering the questions posed in the introduction:

1. Narrative evidence tells us that the credibility of the Hungarian monetary policy as perceived by the stakeholders of monetary policy deteriorated between 2002 and 2007. The respondents of our short survey confirmed that

central bank communication, which most of them deemed an integral component of credibility, became worse by the end of the period.

2. The introduction of inflation targeting in Hungary brought about the anchoring of inflation dynamics. Although the headline CPI inflation targets were missed in the majority of years between 2001 and 2009, the analysis of the VAICPI time series shows that this measure of inflation – which ignores the price shocks caused by tax changes – became stationary after the introduction of IT in June 2001.
3. The predictability and gradualism of MNB MC's interest rate decisions were relatively low in the observed period compared to other central banks in the developed peer group and mediocre compared to other central banks in the emerging peer group.
4. Central bank communication failed to signal future policy rate decisions properly in the observed period. The consistency of communication by both MNB MC and the central bank governor were low and their indications of future policy moves were sometimes contradictory.

References

Amisano, G. – Tronzano, M. 2005: Assessing ECB's Credibility During the First Years of the Eurosystem: A Bayesian Empirical Investigation. *University of Brescia, Department of Economics Working Papers*, 0512, 1-25. p.

Blinder, A. S. 1999: Central Bank Credibility: Why Do We Care? How Do We Build It? *NBER Working Paper Series*, 7161, 1-23. p.

Blinder, A. S. 2006: Monetary policy today: Sixteen questions and about twelve answers. *Center for Economic Policy Studies Working Papers*, 129, 1-65. p.

Figyelő 2008: *Figyelő TOP 200 '2007*. Figyelő, Budapest.

Gaspar, V. – Perez-Quiros, G. – Sicilia, J. 2001: The ECB Monetary Policy Strategy and the Money Market. *International Journal of Finance and Economics*, 6, 4, 325-342. p.

Goldberg, L. S. – Klein, W. M. 2005: Establishing Credibility: Evolving Perceptions of the European Central Bank. *NBER Working Paper Series*, 11792, 1-28. p.

Madár, I. 2003: Az árfolyam nem javíthat versenyképességet. *Világgazdaság*, June 26, retrieved December 30, 2009 from http://vg.hu/velemeny/interju/az-arfolyam-nem-javithat-versenykepesseget-36094.

Mishkin, F. S. 2004: Can Inflation Targeting Work in Emerging Market Countries? *NBER Working Paper Series*, 10646, 1-34. p.

Mishkin, F. S. – Schmidt-Hebbel, K. 2007: Does Inflation Targeting Make a Difference? *NBER Working Paper Series*, 12876, 1-64. p.

Nyíri M. 2003: Fokozódó bizalomhiány. *Magyar Hírlap Online*, July 28, retrieved May 12, 2009, from

http://www.magyarhirlap.hu/Archivum_cikk.php?cikk=70030&archiv=1&next=0

Perez-Quiros, G. – Sicilia, J. 2002: Is the European Central Bank (and the United States Federal Reserve) Predictable? *ECB Working Papers*, 192, 1-61. p.

Pintér, K. – Wenhardt, T. 2004: A jegybanki kamatdöntések előrejelezhetősége és hatása a hozamokra. *MNB Műhelytanulmányok*, 31, 1-48. p.

Ross, K. 2002: Market Predictability of ECB Monetary Policy Decisions: A Comparative Examination. *IMF Working Papers*, 233, 1-53. p.

Rozkrut, M. – Rybinski, K. – Sztaba, L. – Szwaja, R. 2007: Quest for central bank communication: Does it pay to be "talkative"? *European Journal of Political Economy*, 23, 1, 176-206. p.

The Economist 2004: The Balcerowicz effect. *The Economist*, January 29, retrieved December 28, 2009, from http://www.economist.com/businessfinance/displaystory.cfm?story_id=E1_NPRRSPN&source=login_payBarrier

Woodford, M. 2005: Central Bank Communication and Policy Effectiveness. *NBER Working Paper Series*, 11898, 1-65. p.

The list of websites mentioned or referred to in the text:

www.akk.hu
www.bankofengland.co.uk
www.budacash.hu
www.cnb.cz/en/
www.ebroker.hu
www.ecb.int
www.federalreserve.gov
www.fn.hu
www.hirtv.net
www.index.hu
www.ksh.hu
www.mnb.hu
www.mno.hu
www.napi.hu
www.nbp.pl/Homen.aspx?f=/srodeken.htm
www.nol.hu
www.otpbank.hu
www.portfolio.hu
www.raiffeisen.hu
www.reservebank.co.za/
www.reuters.hu
www.riksbank.com

www.tcmb.gov.tr/yeni/eng/
www.tozsdeforum.hu
www.vg.hu

Liquidity and Short Term Asset-Liability Structure: Evidence from a Panel of Hungarian Firms

Eva Palinko[1] *–Marta Szabo*[2]

Financial crisis has made a hard effect on the solvency and liquidity position of companies. The increasing number of bankruptcy signals that less liquid companies come easily to unsafe situation. We analyze the features of short-term liquidity conditions of the Hungarian manufacturing companies. The liquidity condition is interpreted broadly, connecting with the asset-liability structure. We investigate whether the structure of short-term assets is determined by particularity of sectors of industry and liabilities by general financial conditions. The features of asset-side analysis are confronted by the liability-side, and we make conclusion on trends of liquidity position of Hungarian manufacturing sectors. We use AMADEUS company level database. The balance sheet and income statement data from 2004 to 2007 provide information for panel data analysis.

Keywords: *Company liquidity, working capital, financial constraints*

1. Introduction, theoretical background

In corporate finance the focus is traditionally on long term financial decisions. The financial analysis of investment decisions and the financing of investments are in the focus of inquires.

In every project decisions we analyze the contribution to corporate values. Investment decisions define corporate value creation in the long term and reversing their impact claims significant cost and time. Meanwhile we cannot leave out of consideration, that in the balance sheet of companies the level of current assets has a relevant measure. The level of current assets is shaped partly by the impact of corporate operational/sectoral peculiarities and partly because of project decisions made by the management. In the case of firms operating in the manufacturing sector the ratio of current assets is more than 50% of the total assets. Contrary to the invested assets the effects of decisions which influence the current assets can be reversible in the short-term and on lower costs. Nevertheless, the weight of current assets in the

[1] Eva Palinko, PhD, associated professor, Budapest University of Technology and Economics, Faculty of Economics and Social Sciences, Hungary
[2] Marta Szabo, PhD, associated professor, Budapest University of Technology and Economics, Faculty of Economics and Social Sciences, Hungary

balance sheet and the frequency of decisions concerning current assets make is reasonable to give more attention to this issue.

In literature of corporate finance, several pieces of empirical research confirm the importance of working capital. According to Gentry et al. (1979) there is a contradiction between that managers spend a significant part of their time with managing working capital and the fact that the literature of corporate finance handles it almost as a peripheral question and there is limited research in this field. The most common explanation of not giving significance to this question is that working capital management is often a daily routine for companies, their impact is not irreversible. Another explanation is that the right level of working capital and restructuring is principally a question of production, prices and market forecast, that is not especially financial field. Surveys carried out in four countries conclude that there is found out that the most active function of working capital management is the support of revenues from cash side by the active management of, inventories, account receivables and account payables. Its secondary role is the assurance of financing buffer to unanticipated financial situations. Its tertiary role is to minimize the elements of working capital and decrease costs.

Howorth et al. (2003) examined working capital routines in the UK in the scope of small companies and found that small companies cannot be considered homogeneous at all regarding working capital management. Based on their research they found that those companies gave little attention to working capital the profitability of which was high and which were not growth oriented, used less external finance, had shorter production cycles and had fewer cash flow problems.

Working capital management plays an important role in the shaping of corporate profitability and liquidity risk. The level of working capital can be held in a relatively high level, but this is problematic because of the alternative cost of capital, decreases profitability but in the same time it mitigates the risk related to production and consumer relations. The relation of risk and reward is different in the case of short term liabilities. In normal economic situations when short term interest rates are smaller than the long term ones, the high level of short term liabilities has a positive impact on corporate profitability, at the same time it raises the level of liquidity risk. Deloof (2003), Lazaridis et al. (2005) proved based on empirical data that there is a significant connection between the size of working capital, its inner components and corporate profitability.

Working capital is an important input factor and an easily variable tool of providing liquidity (Fazzari et al. 1993). Net working capital which is the difference between current assets and short term liabilities shows the net position of the company in respect of liquid assets. Among current assets inventories serve the company's production directly, while the accounts receivables help the realization of revenues. Among short term liabilities the accounts payable are not just the input of production, but an important tool of managing liquidity. The change of working capital, the variation of its inner components the sectoral determination/operational

specialties and beside the impacts coming from the risk-reward positioning of companies it is strongly determined by the 'behavior' of the company's management.

The management and the owners of the company, its partners do not possess the same information about the inner processes of the company, which causes additional costs (adverse selection costs) to the external partners of the firm. The agency theory, the corporate governance is a relevant issue not only regarding the whole company but concerning working capital management as well. Asymmetric information affects inventories, accounts receivable, accounts payable and the size of credit extension. According to Jensen (1986) the company's management is interested in holding relatively high cash pile and working capital. Higher level of working capital increases the management's flexibility in making decisions, it means higher collateral for credit extension and decreases the chance and cost of possible financial difficulties. It makes the performance and effort of management weaker in continuously maintaining liquidity in a way that is not negligible. In case of larger companies because of greater information asymmetry greater amount of cash and through this a bigger pile of working capital can be observed (Ozkan et al. 2004, Dyck et al. 2004).

From the elements of working capital cash holding attracts great attention, because in this case the questions of agency theory and corporate governance appear in a concentrated way. (Guney et al. 2007, Ozkan et al. 2004, Drobetz et al. 2007).

From the above mentioned surveys we can conclude that one of the main stream of the studies related to working capital management trends towards those parameters (profit, risk, liquidity) which are in the centre of decisions of corporate finance and which parameters can direct decisions. The other part of the surveys is based on the testing of parameters which form the level of working capital.

Earlier studies revealed many company-specific factors in connection with working capital management. In most cases we have to take it as a general feature that working capital management depends on the size of the company, the asset intensity of its activity, the possibilities of company growth, the leverage and the volatility of the cash flow (Guney et al. 2007, Delannay et al. 2004, Ozkan et al. 2004). Cash holding decreases the probability of financial distresses and through this the probable costs of financial difficulties. Cash holding shows the attendance of agency problem as well, the greater amount of cash pile can refer to the management's freedom of decision making. The management can be interested in keeping larger cash pile which can help the management to achieve their goals. Relatively few studies deal with the specialties of working capital management in developing countries. The study of Delannay et al. (2004) deals with nine Central and Eastern European countries, he examined some elements of working capital (accounts receivable, accounts payable) and he found that there is a relatively big difference among certain countries, which can be explained by the differences of their level of development, but he couldn't find unambiguous effects that can be generalized.

In the operation of companies the long-term and short-term asset groups need different handling, managing. In the centre of our research there is the short-term asset-liability management. On the one hand we need both the analysis of assets and liabilities, to explore those factors, which affects the level and inner components of working capital and current liabilities. On the other hand with the connection of assets and liabilities, and with the examination of working capital, we can find explanation for the other relations of the company's operation and thus the set of short-term liquidity.

The aim of the study is to find out the connection in reference to the level of current assets and short-term liabilities and their inner structure based the of balance sheet and profit and loss statement of the Hungarian companies in manufacturing sector. The other goal of the study is to make establishments to the liquidity status of companies, and to those factors that shape liquidity by jointly handling current assets and short term liabilities. We examine the specialties of working capital management in reference to two company groups, large-and medium sized companies. Our goal is to find out if there is any difference in the level of working capital and its components among Hungarian companies in the manufacturing sector.

2. Data, variables

We obtained data form the AMADEUS database[3]. The sample comprises medium and large-sized firms from Hungarian manufacturing sector. Companies on AMADEUS for Central European countries (for Hungary) are considered to be large when they have operating revenue greater than 10 million euros, or total assets greater than 20 million euros, or employees greater than 150. The medium sized firms meeting at least one of the following criteria: operating revenue greater than 1 million euros, or total assets greater than 2 million euros, or number of employees greater than 15[4].

In addition we applied series of filters. We eliminated the observations of firms with anomalies in their account, such as negative values in the current assets, fix assets, current liabilities, and tested the outliers.

In the first round, in the course of choosing variables we constituted three groups of variables. The first group shows the level of current assets and their inner

[3] The AMADEUS database was created and distributed by Bureau van Dijk, and contains company level financial and economic data, information on European companies.

[4] The selection criteria is different for UK, Germany, France, Italy, Spain, Ukraine and Russian Federation. Companies that are located in these countries are considered to be large when they have operating revenue greater 15 million euros, total assets greater than 30 million euros, employees greater than 200, considered medium sized when they have operating revenue greater 1,5 million euros, total assets greater than 3 million euros, employees greater than 20.

components, the second group shows the level of debts and their inner components, the third group shows the connection between the assets and liabilities and the size of liquidity.

Current asset's ratios:

CA_TA	Current assets to total assets
Stock	Stocks (inventory) to total assets
Debtor	Account payable to total assets
Cash	Cash and cash equivalent to total assets

Debt ratios:

CL_TF	Current liabilities to total financing (total assets)
TL_TF	Current liabilities and long term debt to total financing
Loan	Short term debt to total financing
Creditor	Account payable to total financing

Liquidity ratios:

NCA_TA	(current assets – current liabilities) / total assets
CA_CL	Liquidity ratio (current assets / current liabilities)
ACID	Quick ratio (current assets – stock) / current liabilities

In the second round of choosing variables we aimed at taking that kind of variables which expectedly affect the size of working capital and the set of liquidity. These variables can be seen among the independent variables of linear regression. However we cannot leave out of consideration that the factors of the asset and liability side can be independent variables of each other. The size of current assets can be influenced by whether reduced rate credits or patient suppliers are available, and conversely the size of credit pile is influenced by the size of current assets.

Operating ratios:

OPC	Operating cycle. Number of days account receivable + inventory
CCC	Cash cycle. Number of days account receivable+inventory–account payable
FINS	Financial stability. Current assets to turnover (operating revenue, sales)
CF_S	Cash flow to turnover

Profitability ratios:

Oprof	Operating profit (EBIT) to total assets
Nprof	Net profit (profit after tax) to total assets
Omarge	Operating profit to turnover
Nmarge	Net profit to turnover

There are two types of questions in connection with the empirical test attempts to find connections for the peculiarities of working capital management of Hungarian companies in manufacturing sector and confirm them. On the one hand we wait for the test to confirm the sectoral determination by identifying the factors which define working capital. Our hypothesizes are the followings:

1. From the parameters that form working capital the level of current assets is strongly sector-determined, so there is no significant difference in the case of the two company groups.
2. In the level of short-term liabilities/or in their volatility we expect strong difference referring to the examined company groups, assuming that there is a greater sectoral independence in deciding the size of this level.
3. The problems of paying discipline and queuing in case of Hungarian firms can be seen in negative working capital that affects significant range of companies.
4. In case of higher working capital we count with stronger profitability worsening effect.

On the other hand from the empirical tests we expect to confirm that the set of working capital strongly affects the profitability of the firm and its level of liquidity and risk.

Based on preliminary examinations we expect that the peculiarities of operation significantly affect the company's level of working capital. Between the length of the operating cycle and the size of working capital we suppose a strong positive relationship, larger amount of working capital belongs to longer operating cycle. The length of cash cycle is defined by two factors, the relative size of inventories and account receivables to the revenue (operation cycle) and the relative size of accounts payable to the revenue. The cash cycle can equally be linked to the asset and liability side, it expectedly shows larger variance than the operating cycle. Given the operating cycle is strongly determined by the sectoral/business activity, and the cash cycle carries the management/owners greater freedom of decision making in company financing.

The level of current assets forms the level of profitability and liquidity risk of the company. The risk grows together with the rise of the expected return and because of the higher amount of working capital the decreasing return reduces the liquidity risk of the company. From the value of the financial stability indicator we can typically draw conclusions about the set of liquidity, smaller indicator value comes together with higher level of company liquidity. From the indicators that ex-

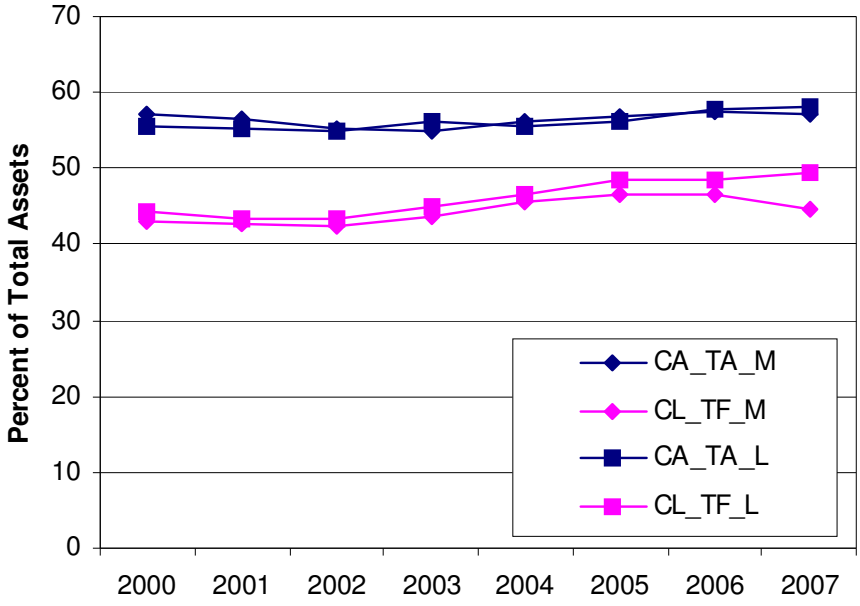
press liquidity, we expect that higher cash flow indicator comes together with higher level of liquidity, possibly with lower debt level. The value of cash flow is the sum of net profit and depreciation.

Based on the relationship of different profitability indicators and the size of current assets we expect that for higher profitability belongs lower level of current assets and normally (in case of normal yield curve) higher short-term liability/debt level.

3. Descriptive statistic

Fig.1. shows the average working capital level of Hungarian firms in the manufacturing sector. According to international tendencies the level of current assets approaches 60% in the total assets. The gap between the current assets formation and the level of short-term liabilities refers to healthy short term asset financing, so that the current assets in line with the size of the gap are financed by long-maturity or infinite maturity shareholder capital. This relationship confirms the realization of the general principle, which says that the maturity structure of liabilities which is adjusted to the assets' maturity structure is the necessary condition of providing liquidity on the long run. No significant difference can be observed between large and medium sized companies, the same tendency can be seen in this front, current assets and short-term liabilities are nearly moving together. One possible answer for this can be that while the dominance of large companies which can be observed against business partners delegates lower amount of working capital to the large sized companies, the principal-agent problem appeals more intensively in large companies than in the case of medium sized companies, because of this there is an ambition to plus liquidity which increases the current assets. These two impacts together can result the joint move of the elements of working capital in large and medium sized companies.

Figure 1. Current assets, current liabilities to total assets ratios of Hungarian large and medium sized manufacturing firms (2000-2007)



Source: own creation
Information extracted from the Amadeus database.
CA_TA_M current assets to total assets, medium sized firms,
CL_TF_M current liabilities to total finance, medium sized firms,
CA_TA_L current assets to total assets, large firms,
CL_TF_L current liabilities to total finance, large firms.

Descriptive statistics of working capital management in 2007 can be found in Table 1. The means are considered primary the aspect of comparison, but the minimum and maximum values of the indicators and the standard deviation carry important information as well.

The level of current assets and their inner structure are almost the same within the sphere of large and medium sized companies. Some difference can be observed in the level of cash pile. Large sized companies averagely hold 6.98% of their assets in cash. Medium sized companies hold higher cash pile, 9.99% of their assets. Cash holding cannot be considered high based on international comparison. Publicly listed firms in Switzerland held an average cash pile between 10 and 15% between 1994 and 2004 (Drobetz et al. 2007). As for other researches the highest cash pile can be observed in the case of Japanese companies (17%), in France (12%), in England (11%), while in the USA cash holding was relatively low, 7% (Guney et al. 2007).

International data apply not only to companies in the manufacturing sector, but can be considered as informative.

The lower Hungarian cash level can be explained theoretically by the level of lower information asymmetry between the owner and the management, but knowing the Hungarian company conditions, we think that the often mentioned under financing doesn't make possible to hold higher cash level. And in connection with this the acceptance of the lack of generally appearing paying discipline 'doesn't require' the higher cash level and doesn't increase the Hungarian level of cash holding to the European average. In the sphere of Hungarian companies the higher cash holding of medium sized companies is in connection with the fact that medium sized companies can get instant loans more difficultly in the case of financial hardness and there is less likely they can reach shorter days sales outstanding by using their power. As a result of all the above mentioned medium sized companies cannot keep down the cash cycle to the level of large sized companies. Standard deviance of the cash pile is the greatest in both company types.

In the valuation of the liability side it is conspicuous that medium sized companies have less credit pile. All the rate of short-term liabilities (44.57%), all the size of loan (9.64%), all the sum of current liabilities and long term debts (54.36%) are lower, than in the case of large sized companies. In the sphere of medium sized companies it occasionally can refer to companies with 15-30 employees, which can get loans more difficultly, primarily because of the lack of credit collateral. (In the sample of 4229 medium sized companies the average number of employees was 45.) Based on preliminary expectations regarding suppliers we expected greater value in the sphere of large sized companies with the explanation that their dominant role can be realized in the payment deadlines, but contrary to our expectations the value of this indicator is nearly the same in the case of large and medium sized companies. The general selling pressure delegates longer duration to both company groups or moreover a large sized company is not forced to strict liquidity follow-up which is typical to the medium sized companies, this can be seen in the lower cash cycle for example. At the same time this can be seen in the lower profitability of large sized companies.

The lower credit pile and higher cash pile of medium sized companies strongly influence the liquidity indicator. The average of liquidity indicators is significantly better in the sphere of medium sized companies than in the case of large sized companies. At the same time the maximum values and the standard deviation of liquidity indicators are significantly higher in the case of medium sized companies. The value of liquidity indicator in the case of medium sized companies is 1.8653 and in the case of large sized companies this value is 1.567. These can refer to a relatively good average liquidity situation, but other bias effects can appear in these values (for example few easily available external sources). Frequent liquidity difficulties which are typical for Hungarian companies cannot be felt with the help

of this indicator. Liquidity is reported to the balance sheet preparation date but the generally good liquidity level is reinforced by the previously confirmed positive net working capital which testifies long term liquidity.

From the peculiarities of operation there is no significant difference in the length of the operating cycle, in the case of large sized companies this is averagely 100 days, in the sphere of medium sized companies it is 108 days, the value of standard deviation is relatively high. The value of cash cycle in the case of large sized companies is 62 days, this value is 75 days in the case of medium sized companies, which difference is already higher. The deviation of cash cycle, its minimum and maximum values also show significant difference compared to the average value.

The indicator of financial stability measures the level of short term liabilities to the revenues. The current liabilities take 38.42% to revenues in case of large sized companies and 32.59% to revenues in case of middle sized companies. The difference can be explained by the higher liability level of large sized companies. The maximum value of this indicator and its standard deviation is significantly higher in case of large sized companies. The value of operating cash flow is related to the revenues, from which we can partly make conclusions to the stability of the company and its responsible operation. The value of this indicator shows a significant difference in case of the two company groups. It is lower in case of large sized companies (5.65%) than in the case of medium sized companies (7.66%), this difference is in accordance with the profitability indicators.

We examined profitability with four indicator groups, as for every indicator the profitability of large sized companies is half of the medium sized companies. Explanation can be found on the cost side with high probability, the lower level of current assets and lower credit pile can make partial explanation to the difference. (Of course except for current assets, our field of examination, other stronger reasons can cause the difference which can be seen in profitability, as for example in the sphere of medium sized companies because of their flexibility there is a higher chance to gain competitive edge, and to preserve it on the longer run.) At the same time if we reverse the question we can make a better approach, the stronger motivation of middle sized companies make them possible to reach higher profitability and lower amount of current assets. We take the nature of owner's control as an important factor. In the sphere of middle sized companies the combination of owner-manager role is more frequent, the stronger owner control can cause higher profitability.

Table 1. Descriptive statistics for the main variables

	Large companies					Medium sized companies				
	N	Min	Max	Mean	Std. Dev.	N	Min	Max	Mean	Std. Dev.
Asset's side										
CA_TA	840	,00	1,00	**,5817**	,21072	4229	,00	1,00	**,5701**	,24241
Stock	825	,00	,71	**,1944**	,13805	3971	,00	,88	**,1786**	,16402
Debtor	689	,00	,84	**,1942**	,16479	1530	,00	,87	**,2058**	,16623
Cash	839	,00	,79	**,0698**	,10858	4211	,00	,90	**,0999**	,13096
Liability's side										
CL_TF	795	,00	2,12	**,4936**	,23979	4183	,00	1,62	**,4457**	,23340
TL_TF	795	,00	2,12	**,58,43**	,25137	4183	,0	3,43	**,5435**	,25487
Loan	637	,00	1,1	**,1277**	,15547	1484	,00	,82	**,0964**	,12732
Creditor	691	,00	,8	**,154**	,13866	933	,00	2,01	**,1579**	,16974
Liquidity										
NCA_TA	795	-1,2	0,86	**,0848**	,26133	4182	-1,08	,94	**,1235**	,27645
CA_CL	795	,01	23,09	**1,567**	1,58472	4178	,03	61,89	**1,8653**	2,49460
ACID	781	,04	23,09	**1,0593**	1,3008	3923	,01	44,34	**1,2976**	2,05429
Operation analysis										
OPC	686	,00	852,09	**100,27**	80,8514	1524	,00	758,73	**108,45**	82,8924
CCC	685	-277,8	624,94	**62,255**	72,7074	1478	-979,40	648,37	**75,53**	85,6593
FINS	789	,02	8,43	**,3842**	,52020	4177	,00	5,61	**,3259**	,34556
CF_S	788	-1,58	,55	**,0565**	,13820	4147	-1,47	1,64	**,0766**	,11225
Profitability										
Oprof	835	-1,19	,82	**,0544**	,14014	4209	-1,82	3,27	**,1025**	,21363
Nprof	833	-1,27	,66	**,0384**	,13454	4208	-1,60	1,31	**,0644**	,13238
Omarge	832	-2,4	,60	**,0280**	,15295	4223	-1,76	1,23	**,0453**	,11829
Nmarge	829	-2,43	,54	**,0182**	,15125	4203	-1,72	1,29	**,0356**	,10946

Information extracted from the Amadeus database. CA_TA is the ratio of current assets to total assets, Stock is the ratio of stock to total assets, Debtor is the ratio of account receivable to total assets, Cash is the ratio of cash to total assets, CL_TF is the ratio of current liabilities to total financing, TL_TF is the ratio of long term debt and current liabilities to total financing, Loan is the ratio of short term debt to total financing, Creditor is the ratio of account payable to total financing, NCA_TA is the ratio of net current assets to total assets, CA_CL is the liquidity ratio, ACID is the liquidity quick ratio, OPC is the number of days operating cycle, CCC is the number of days cash cycle, FINS is the ratio of current liability to sales, CF_S is the ratio of operating cash flow to sales, Oprof is the ratio of return on assets, compare EBIT to total assets, Nprof is the ratio of net profit to total assets, Omarge, operating margin is the ratio of, EBIT to sales, Nmarge is the ratio of net profit to sales.

Source: own creation

4. Correlation and regression analysis

With the help of correlation analysis we intend to give explanation to what kind of factors can affect the size of current assets and short term liabilities, and what kind of factors play important role in defining liquidity.

Instead of presenting the results of correlation matrixes one by one, we summarized the results in tables contain large and medium sized company data for 2007 (Table 2-4).

Table 2. Correlation coefficients

		Large	Medium
CA_TA_2007	Pearson Correlation	1	1
	N	840	4229
DEBT_2007	Pearson Correlation	,494**	,501**
	N	689	1529
STOCK_2007	Pearson Correlation	,386**	,413**
	N	825	3971
CASH_2007	Pearson Correlation	,334**	,361**
	N	839	4211
CL_TA _2007	Pearson Correlation	,328**	,326**
	N	795	4182
FINS_2007	Pearson Correlation	-,172**	,106**
	N	789	4182
NMARG_2007	Pearson Correlation	,124**	-,142**
	N	829	4176

Source: own creation

**. Correlation is significant at the 0.01 level (2-tailed)

In accordance with the expectations, *current assets* are primarily determined by the sectoral peculiarities of operation, production and services. From the factors that determine the level of current assets in the first place there are those effects which form the specialties of operation (inventories, accounts receivable, cash). According to our preliminary expectations the size of inventories affects the level of current assets the most. However, based on our data, accounts receivable affect current assets more significantly in case of large sized companies and in the sphere of medium sized companies as well. In the case of medium sized companies the average size of accounts receivable are high as well (0.2058), this refers to the fact that the payment deadline for customers, the allowances for customers, the customer-

focused managing of revenues play a significantly more important role in shaping the rate of current assets, than managing inventories (0.186).

The liability side significantly affects the current assets as well. The strength of the relationship is 0.33 between short term liabilities and current assets in both company types. The financial stability and the net margin don't show unambiguous and substantial effects. The importance of profitability indicators was negligible in every regression analysis as well. The values of ROA type indicators can be regarded as low (net profit: large companies 3.84%, medium sized companies 6.44%). The relationship between the working capital and the level of profitability which was showed in international researches (Lazadiris et al. 2005) cannot be pointed in companies in the Hungarian manufacturing sector. This can be traced back to the fact that profitability is not a dependent variable. Meanwhile the level of working capital can be a dominant factor in shaping profitability, and reversely, in shaping working capital, the parameters coming from the peculiarities of sectoral/business activity affect the level of working capital more strongly than profitability. We ran the variables separately to the large sized and medium sized companies, the order of the first six companies was the same in both company types, after this there where smaller differences regarding the strength of the relationship.

Table 3. Correlation coefficients

		Large	Medium
CL_TF_2007	Pearson Correlation	1	1
	N	795	4183
TL_TF_2007	Pearson Correlation	,852**	,796**
	N	795	4183
CA_TA-2007	Pearson Correlation	,328**	,326**
	N	795	4182
LOAN_2007	Pearson Correlation	,467**	,424**
	N	633	1480
CCC_2007	Pearson Correlation	-,139**	-,082**
	N	660	1475
FINS_2007	Pearson Correlation	,215**	,336**
	N	789	4177

Source: own creation

**. Correlation is significant at the 0.01 level (2-tailed)

In the course of examining those factors that explain *short term liabilities* we analyzed if there is any common factor among those factors that shape current assets

and the level of short term liabilities, to what extent those factors that shape the asset side and the liability side can be linked. Short term liabilities show close correlation primarily with the credit pile (short-term and long-term credits). However suppliers are not among the significant independent variables. The impact of suppliers can be demonstrated through cash cycle indicator, but the impact of cash cycle is not significant. Short term liabilities correlate with the average level of current assets (0.32) from the factors of the asset side, so we can state that the size of current assets affects the liabilities. The effect of the amount of cash is not among the independent variables it is not provable that firms handle cash pile as a buffer, reserve source. From profitability indicators none is among independent variables, it couldn't be pointed that more profitable companies have lower liabilities.

Table 4. Correlation coefficients

		Large	Medium
ACID_2007	Pearson Correlation	1	1
	N	781	3923
TL_TA_2007	Pearson Correlation	-,464**	-,567**
	N	781	3923
CASH_2007	Pearson Correlation	,306**	,524**
	N	780	3912
STOCK_2007	Pearson Correlation	-,226**	-,300**
	N	781	3923
CL_TA_2007	Pearson Correlation	-,450**	-,532**
	N	781	3923
NPROF_2007	Pearson Correlation	,131**	,135**
	N	781	3923

Source: own creation

**. Correlation is significant at the 0.01 level (2-tailed)

From the connection of the asset and liability side we can conclude the *liquidity* situation of the company. Dealing with variables that explain liquidity we intend to find the answer to the question whether the peculiarities of operation or the financing motivations determine the liquidity situation of a company. From liquidity indicators we show the values of independent variables which were used to explain the quick ratio, which is the most narrowly interpreted liquidity indicator. The liability side affects liquidity the most in both company types. The strongest relationship can be pointed in middle sized companies, the strength of the relationship between the total liabilities and the quick ratio is 0.567, the correlation coefficient of short term liabilities refers to strong relationship with a value of 0.532. The direction of

the relationship is negative, the higher the level of liabilities the smaller the quick ratio. On the second place we emphasize the relationship of cash and quick ratio. The relationship is positive among middle sized companies, the strength of the relationship is 0.524 between the level of cash and quick ratio, in the case of large sized companies the strength of the relationship is 0.306. In connection with liquidity its relationship with inventories can be emphasized from the factors of the asset side. The direction of the relationship is negative, in case of middle sized companies it is stronger (0.30) than in large companies (0.226). Higher the level of stocks less the liquidity position of companies.

The correlation analyse helped to identify the factors determine the liquidity and factors effect the size of current assets and short term liabilities. The regression model offers more specific information on factors form firm's liquidity.

The basic regression model for liquidity:

$$LIQ_{i,t} = (\alpha_i + \alpha_t) + \sum_{j=1}^{n} \beta_j X_{i,j,t} + \varepsilon_{i,t.}$$

Where $LIQ_{i,t}$ is the dependent variable of firm i at the end of year t. α_i and α_t represent firm-specific and time-effects. $X_{i,j,t}$ dependent variables for i firm, t time and j variables, β_j coefficient of independent variables. The null hypothesis is the correlation between the independent and dependent variable do not differ from zero.

The regression results, concerning quick liquidity ratios of firms, are reported by firm size in Table 6. The regression analysis concerns business years of 2004-2007 and were run separately to the large and mediums sized sample. We use SPSS statistic (stepwise diagnosis, forward methods) to built up linear regression model variables and we use simple panel model. The time effects were built in using SPSS statistic restructure data wizard method.

Table 6. Panel model. Dependent variable: Liquidity quick ratio

Large			Medium		
	Unstandardized Coefficients	t		Unstandardized Coefficients	t
(Constant)	0,984**	29,392	(Constant)	1,0771**	36,404
CASH	0,482**	3,667	CASH	0,459**	3,676
CA_TA	2,382**	33,270	CA_TA	2,909**	34,755
STOCK	-2,240**	25,994	STOCK	-2,736**	-0,543
CL_TA	-1,854**	-24,127	CL_TA	-2,440**	-31,320
DEBT	-,239**	-3,194	DEBT	-0,585**	-6,474
TL_TF	-,150**	-2,190	TL_TF	,001**	0,013
R Square	,736		R Square	,647	
Adjusted R Square	,735		Adjusted R Square	,646	
Std. Error of the Estimate	,27705		Std. Error of the Estimate	0,44798	

Source: own creation
**:significant at the 0.01 level

The value of R Square is 73.6% in case of large companies and it falls to 64.7% in medium sized companies. The R Square can be considered high. In the literature of corporate finance can be often found a value around 30% in the relevant topic. The other estimation attributes of the model are on an acceptable level.

The regression results support our main hypothesis that both asset and liability side makes effect on company liquidity position. Current assets and current liabilities to total assets ratios have the highest beta value. In case of the medium sized companies both current assets and current liabilities have a strong influence on the company liquidity position with a value of 2.909 and -2,440 respectively. The positive beta value of current assets signals that company liquidity ratio increases when the level of short lived assets, as "collateral" increases, and the negative value of current liabilities indicates that firm's liquidity is depressed when the liabilities increases. The beta value of current liabilities indicates less stronger effect in the group of large firms. The negative relationship between stock, debtor and liquidity explains that the high level of stock and account receivable has a negative effect on the available cash stocks, and the firm's liquidity. The cash holding also shows a strong effect on liquidity especially in large sized firms. The description and correlation analyse marks the distinguished role of cash holding as well. No significant difference can be observed between the two company groups except the effect of cur-

rent liability and debtor, which stress the importance of the managing the account payable and receivable.

5. Conclusions

Theoretically the size of working capital is defined by two important factors: current assets and short term liabilities. The results of correlation matrixes and linear regressions show that however the impact of the asset side is strong, by managing short term liabilities more significant effects can be reached regarding the size of working capital and company liquidity.

From the results of the tests and referring to our hypothesizes we state the followings:

1. We consider the 1st and 2nd hypothesizes confirmed. Regarding both current assets and short term liabilities, and regarding the level of working capital as well we discovered strong correlation which reinforces the strong sectoral/business determination. Positive net working capital confirms the validity of the matching the term structure of assets and liabilities in the case of small and medium sized companies.
2. We didn't find our hypothesis valid that deviation is higher regarding CL/TF indicators than in the case on CA/TA.
3. In the sphere of medium sized companies the indicators coming from the liability side don't show higher deviation. (Only the higher volatility of the cash cycle refers to the fact that the sectoral determination of the asset side is stronger and the independence of decision making is greater on the liability side.)
4. Lower profitability level which is expected along with higher level of working capital is not confirmed by the data of medium sized companies. However this doesn't mean that this effect of higher net working capital is not valid, we just say that the level of profitability in the case of this parameter is overwritten by other factors. (We examine those parameters that shape working capital, not those, that affects profitability.)

Regarding our research the followings can mean further possibilities for additional studies: Examining the relationship of cash-cash equivalent and corporate governance. Besides the examination of trade credit, as an examination of the source of corporate financing could make it possible to clarify the role and significance of accounts payable and accounts receivable, and to explore the relationship between trade credit and bank credit.

References

Delannay, A. – Wiell, L. 2004: The Determinants of Trade Credit in Transition Countries. *Economics of Planning*, 37, pp. 173-193.

Deloof, M. 1998: Corporate Groups, Liquidity, and Overinvestment by Belgian Firms Quoted on the Brussels Stock Exchange. *Managerial and Decision Economics,* 19/1, pp. 31-41.

Deloof, M. 2003: Does Working Capital Management Affect Profitability of Belgian Firms? *Journal of Business Finance and Accounting,* 30/3-4, pp. 585-596.

Drobetz, W. – Grüninger, M.C. 2007: Corporate Cash Holdings: Evidence from Switzerland. *Financial Markets and Portfolio Management,* 21, pp. 293-324.

Dyck, A. – Zingales, L. 2004: Private Benefits of Control: an International Comparison. *Journal of Finance,* 59/2, pp.537-600.

Fazzari, S.M. – Petersen, B.C. 1993: Working Capital and Fixed Investment: New Evidence on Financing Constraints. *The RAND Journal of Economics,* 24/3, pp. 328-342.

Gentry, J.A. – Mehta, D.R. – Bhattacharyya, S.K. – Cobbaut, R. – Scaringella, J. 1979: An International Study of Management Perceptions of the Working Capital Process. *Journal of International Business Studies,* 10/1, pp.28-38.

Guney, Y. – Ozkan, A. – Ozkan N. 2007: International Evidence on the Non-linear Impact of Leverage on Corporate Cash Holdings. *Journal of Multinational Financial Management,* 17, pp. 45-60.

Howorth, C. – Westhead, P. 2003: The Focus of Working Capital Management in UK Small Firms. *Management Accounting Research,* 14, pp. 94-111.

Jensen, M. 1986: Agency Cost of Free Cash Flow, Corporate Finance, and Takeovers. American Economic Review, 76/2, pp. 3-43.

Lazadiris, I. – Tryfonidis, D. 2005: The Relationship Between Working Capital Management and Profitability of Listed Companies in the Athens Stock Exchange. *http:// ssrn.com/abstract=931591*

Ozkan, A. – Ozkan, N. 2004: Corporate Cash Holdings: An Empirical Investigation of UK Companies. *Journal of Banking and Finance*, 28, pp. 2103-2134.

Analyses of Extreme Events on Emerging Capital Markets

Gábor Dávid Kiss[1] *–László Dudás*[2]

This study deals with the statistical methods of contagion-effects on emerging capital markets. After fitting probability distribution on the empirical data, and cross-market correlation sensibility test were used on time series (2002-2009) of Hungarian, Polish Russian and US government bond, stock and currency markets to study their behavior under extreme and normal circumstances. The aim of this analysis is to identify the possible differences between emerging and developed capital markets to investigate the validity of economic axioms according to the relation of bond, stock and currency markets on the emerging markets.

Keywords: power-law test, cross-market correlation

1. Introduction

This study analizes global capital market as a network of national economies, which are networks of markets themselves. There are market actors such as investment banks, state treasuries, national banks, etc. The global network operates under the following rules: quasi free movement of production factors (labor, capital, goods, and intellectual properties), technical progress, international markets, deregulation, liberalization, lack of coordination of economic policies, and liberalized capital accounts. (Wang – Chen 2003) .

To examine singularity and historicity of market distortions, this study models economic networks as graphs, where nodes are individuals or organizations and the edges are the social interactions between them (Barabási – Albert 1999). It is necessary to examine the human factor in case of nodes on the market to understand the

[1] Gábor Dávid Kiss, PhD student, University of Szeged, Faculty of Economics and Business Administration, Hungary.
[2] László Dudás, PhD student, University of Szeged, Faculty of Science and Informatics, Hungary.

Acknowledgement: The authors would like to acknowledge and extend their gratitude to the following persons who have made the completion of this article possible: Péter Kovács, Beáta Farkas, Ling Xiao, Daniel Stavárek, Péter Erdős, Anita Pelle, Dávid Tátrai and Zoltán Násztor.

occurrence of bubbles[3] and decoupling effects[4] in a non-linear world. The conditions of the bounded rationality model, in which decisions are made under uncertainty resulting in current prices and expected future prices co-evolving over time with mutual feedback, are closer to reality by giving a deeper description of the mispricing problems than the rational homo oeconomicus model. During operations, market actors, who learn by using heuristics, are possibly biased[5], while their reactions are non linear[6]. (Barabási – Albert 1999, Hommes – Wagenne 2008, di Mauro et al. 2008)

Many real-life complex networks are neither completely regular[7] nor completely random. There are complex systems, in which conditions are constantly changing giving a rise to endogenously engendered novelty. Simple complexity models are characterized by fat tails in returns distribution, long memory and interacting agents (Hommes – Wagenne 2008). The extended characteristics are the followings: (1) particular states of the system are singular, (2) processes are non-linear and frequency-dependent, (3) strength and direction of causal relations are highly divergent in terms of magnitude and power, (4) exogenous events are influencing the system but events in the system are not completely dependent on the environment, (5) there is a hierarchical order between elements and particular emergent properties (Herrmann-Pillath 2000).

Scale-free[8] networks are special cases of complex networks. They are inhomogeneous in nature, which means that nodes have very few link connections and yet a few nodes have many connections. In comparison with a random network, complex networks have the same size and an average degree, but the average path length is somewhat smaller. The clustering coefficient is much higher as well, while

[3] Occurrence of bubbles were simulated in a rational and well informed market environment too when markets deviate from full rationality in asset pricing. Therefore they do not request uncertain circumstances (Hommes – Wagenne 2008).

[4] A divergence between development in a financial market benchmark and its effect on real economy (di Mauro et al. 2008).

[5] Their sentiment varies over time, according to prevailing market mood (Hommes – Wagenne 2008).

[6] Economic agents do not respond strongly to relatively small changes in prices, but larger price movements may trigger a disproportionately larger response with a strong effect on other economic variables (di Mauro et al. 2008).

[7] There are two generic aspects to understand real networks: the nearest-neighbor coupled network (a lattice) and the randomly connected network (Erdősi-Rényi model). In a lattice, every node is joined only by a few of its neighbors creating a homogenous network with low level of dynamism, which is clustered, without small-world effect. Random networks can appear quite suddenly. They are homogenous without showing clustering in general, but have small-world effect. The connectivity approximately follows a Poisson distribution. (Wang – Chen 2003)

[8] scale-free: "The shape of the degree distribution does not change over time, namely, does not change due to further increase of the network scale." (Wang – Chen 2003)

there are a few "big" nodes (hubs) with very large degrees (very large number of connections to bring the other nodes of the network close to each other). A quantity x obeys a power law if it is drawn from a probability distribution, $P(x) \sim x^{-\alpha}$, where α is a constant parameter of the distribution known as the exponent or scaling parameter[9]. The probability P(x) that a vertex in the scale-free network interacts with x other vertices decays as a power law, following therefore majority of cases has very low probability presenting a dissonance between expected value and mode which is the opposite to harmony at random networks. (Wang – Chen 2003, Csermely 2008)

To understand the singularities of financial crises, we have to study how scale-free complex networks can describe synchronization, transition processes and failures. Singularities are the results of non-linear dynamics in the economy, when small changes in systemic characteristics cause large-scale implications at the macro level (Herrmann-Pillath 2000). Synchronizability of a scale-free network is about the same as a star-shaped coupled network driven by tiny fraction of distant links (small world effect); so hubs are playing a similar role as single star-center. In these networks, that is the explanation for the error tolerance and the attack vulnerability phenomenon. The network is robust during the random removal of a fraction of nodes; but after the preferential removal of key nodes, the system performance decreased. Phase transition was described between scale-free and random networks. The connectivity number is emerging after the collapse of the "stable" scale-free state. The system is random in its "chaotic" phase with high evolutionary performance until a new "stable", scale-free combination is established (Grubestic et al. 2008, Wang – Chen 2003, Yuan – Wang 2007).

By using the complex scale-free network model, a lot of unusual events on the financial market could be described. The "singularities", which are resulted by the scale-freeness of real networks, are summarized in three levels of complexity. The first level of time series means, that the autocorrelation function of returns is a monotonically decreasing function holding at least for approximately 20 trading days. Price returns and volatility are locally nonstationary, but asymptotically stationary. The second level, the so called event-based trade allows synchronous interaction in the same economic sector of each time series with all the others. The third level of complexity means a collective behavior during extreme market events. (Bonano et al. 2001)

2. The aim of this paper

This assumption about a hierarchical market structure could be obvious, if we are thinking on the US capital market's 45% share from global capitalization. Therefore

[9] The scaling parameter typically lies in the range $2 < \alpha < 3$, but there could be occasional exceptions. (Clauset et al. 2009)

it is crucial to take a closer look on emerging market's deviations from the probable, "normal" stage. This study presents several theoretical concepts of identifying extreme market movements and spillover of correlation.

"Perfect storms" of capital markets are characterized by big falls in one equity price, which is accompanied by simultaneous big falls in other equity prices – multivariate normal distributions are unfeasible tools to describe heavy tail's "garden of improbable events".

This paper structured as follows to study the properties of contagion: after the overview of the mythological issues from the theoretical literature, the existences of fat tailness on the sample markets are presented by tail distribution tests. The systemic cause of these extreme events is explained in the third part of this work – dealing with a static and semi-dynamic modeling of return-correlation-time triangle.

3. Methodology

The introduction of methodology was divided on two parts: after tests of fat tailness of each participant market and instrument come the tests of network behavior.

3.1. Tail distribution tests of each market or instrument

The proof of the existence of heavy tailness (the "garden") structured as follows:

- Possible asymmetries and tailness were tested at first with the indispensable skewness and kurtosis tests. If the positive and negative sides behave different according these tests, it is necessary to analyze only the separated tails of the empirical distributions – after the rejection of normal distribution hypothesis with the usage of Kolmogorov-Smirmov test of normality[10].

- The rejection of normality suggests that tails could be bigger or smaller than the case of normal distribution. Therefore it is necessary to run basic R^2 based fitting-test[11] with general models of exponential and power-law distributions to separate the tailed and quasi non tailed cases from each other.

- Than estimated power-law properties were studied deeper by Clauset, Shalizi and Newman's (2007) improved quantile-based maximum likelihood estimation (MLE) method[12,] to estimate the scale parameter α. Size of the tails is determined by the scale parameter α – as smaller the α, as fatter is the tail. P-values are given by Monte Carlo procedures: the power-law model is fitted for generated syn-

[10] Jacque-Bera test of normality is also a common procedure.

[11] R-square – measures how successful the fit is in explaining of the variation of the data. A value closer to 1 indicates a better fit. General model for exponential: f(x) = a*exp(b*x) and for power-law: f(x) = a*x^b+c.

[12] Scripted in MATLAB, see http://www.santafe.edu/~aaronc/powerlaws/.

thetic data sets, and the number of times is counted when the Kolmogorov-Smirnov is larger than observed goodness-of-fit (maximum distance between the tail probability or cumulative distribution function of the empirical data and the fitted power-law model). (Clauset et al. 2009, Quismorio 2009)

3.2. Test of common developments under normal and extraordinary events

Power law properties on one or both sides of the empirical distributions signs the existence of the "garden of improbable events", but it is necessary to model the interconnections between the main markets and the emerging markets under these circumstances. Several issues have to be considered during the modeling process:

- Constantly high level of correlation could be interpreted as a sign of financial convergence (Stavárek 2009). But some strange developments could be find, when we investigate dynamics of the correlation over time. There are several ways to model cross correlations on network of markets. Each market or instrument has the same role in equal models, but this holistic view has some disadvantages: definition of correlation period, weekend distortions and too much smoothing or black out effect during aggregation (Kiss 2009). If there is one "stable points" of exchange for the selected instruments as main or leader markets with dominant level of liquidity and capitalization, hierarchic, or top-bottom structures are able to build. Copula-based approach is the most progressive idea, which deals with each market or instrument as an univariate margin of a "global" multivariate distribution – where the multivariate dependence structure is represented by a copula as Sklar's theorem suggests (Embrechts et al. 2001).
- There are several approaches to identify market interdependencies, we can employ both static and moving window approaches or we can calculate time-varying correlation from conditional covariances and variances given by BEKK-GARCH methods (Babatskaja et al. 2008).
- Linear correlation is a natural scalar measure of dependence in elliptical distributions (as multivariate normal and t-distribution), but it can be misleading under heavy-tailed distributions as power-law or t_2-distributions. Using of Kendall't tau[13] or Spearman's rho[14] could be better alternative to calculate dependence (concordance[15]) for nonelliptical distributions as Embrechts et al. (2001) suggest[16].

[13] Kendall's tau for the random vector $(X;Y)T$ is defined as
$\tau(X;Y) = P\{(X - X')(Y - Y') > 0\} - P\{(X - \sim X)(Y - \sim Y) < 0\}$;where $(X'; Y')^T$ is an independent copy of $(X;Y)^T$. Hence Kendall's tau for $(X;Y)T$ is simply the probability of concordance minus the probability of discordance (Embrechts et al. 2001).

[14] Spearman's rho for the random vector $(X;Y)T$ is defined as $\rho_S(X;Y) = 3(P\{(X - X')(Y - Y'') > 0\} - P\{(X - X')(Y - Y'') < 0\})$; where $(X;Y)^T$,$(X';Y')^T$ and $(X'';Y'')^T$ are independent copies (Embrechts et al. 2001).

[15] Let $(x; y)^T$ and $(x'; y')^T$ be two observations from a vector $(X;Y)^T$ of continuous random variables. Then $(x; y)^T$ and $(x'; y')T$ are said to be concordant if $(x - x')(y - y') > 0$, and discordant if $(x - x')(y - y') < 0$ (Embrechts et al. 2001).

- To compare more than two correlation coefficients, a χ^2 test is required. After the Z score standardization of the empirical correlation coefficients (r_1, r_2, … r_k) which were calculated from a sample with n_1, n_2, … n_k components, theoretical correlation coefficients (R_1, R_2, … R_k) have to be tested with a H_0, that they are equal (H_0: $R_1=R_2= \ldots =R_k$) (Lukács 1999).

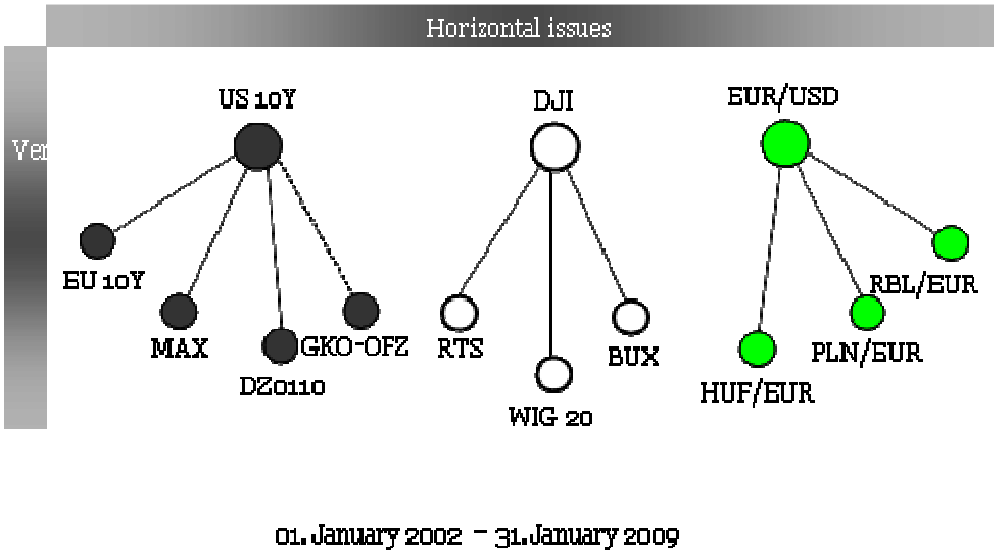
Due to the central role of US capital markets, usage of hierarchical models is more discursiveness. In this case we have to assume that the contagion developments are top-bottom processes. To measure correlations, moving window approach is the most feasible for our research, because the common developments were analyzed as the follows:

On the top of the vertical (stock, bond and currency markets) dimensions there are indicators form US, due to the central role of their capital markets. Hungarian, Polish and Russian markets[17] are containing the "emerging" part of the sample. Inflation targeting monetary policy of the European Union underlined the necessity of a long term bond market benchmark from the continent (as EU 10Y bond).

[16] If X and Y are continuous random variables whose copula is C, then Kendall's tau and Spearman's rho satisfy the following properties for a measure of concordance:
1. κ is defined for every pair X;Y of continuous random variables.
2. $-1 \leq \kappa_{X;Y} \leq 1$, $\kappa_{X;X} = 1$and $\kappa_{X;-X} = -1$.
3. $\kappa_{X;Y} = \kappa_{Y;X}$.
4. If X and Y are independent, then $\kappa_{X;Y} = \kappa_{\Pi} = 0$.
5. $\kappa_{-X;Y} = \kappa_{X;-Y} = -\kappa_{X;Y}$.
6. If C and C' are copulas such that C < C',then $\kappa_C \leq \kappa_{C'}$.
7. If $\{(X_n;Y_n)\}$ is a sequence of continuous random variables with copulas C_n,and if $\{C_n\}$ converges pointwise to C, then $\lim_{n\to\infty} \kappa_{Cn} = \kappa_C$. (Embrechts et al. 2001)

[17] There are some distortions of validity in the case of bond markets due to the data collection, because in the case of Russia, there is only data available for federal bond market (GKO-OFZ) after 9. 1. 2006.; while there are only single bond prices in Poland (so DZ0110 was taken into consideration thanks to it's fitting to the examined interval), which is opposite of Hungarian bond market index (MAX).

Figure 1 Observed markets



Source: own edition

Logarithmic return series were used (1) with daily frequencies and are constructed from the price series (S_t) between 2. 1. 2002 and 31. 1. 2009 to study an entire conjuncture cycle in the world economy.

$$Y_t=[\log(S_t)-\log(S_{t-1})]*100 \quad (1)$$

At hierarchical view we have to study the return fluctuations on the main market (Ym_t). Therefore it is necessary to split its development (2) on monotonic stages.

$$D_{Ym} = D_{Ym}(t_n) = (Y_m(t_n) - Y_m(t_{n-1})) \quad (2)$$

Turning points were set by $D_{Ym}(t_n) \times D_{Ym}(t_{n+1}) < 0$ cases, therefore monotone emerging and declining periods $\{ l_1, l_2, \ldots , l_k. \}$ could be defined. Than we have to study, how the return developments on the n emerging markets (Yen_t) were determined in these intervals $\{ l_1, l_2, \ldots , l_k. \}$.

$$D_{Yen} = D_{Yen}(t_n) = (Y_{en}(t_n) - Y_{en}(t_{n-1})) \quad (3)$$

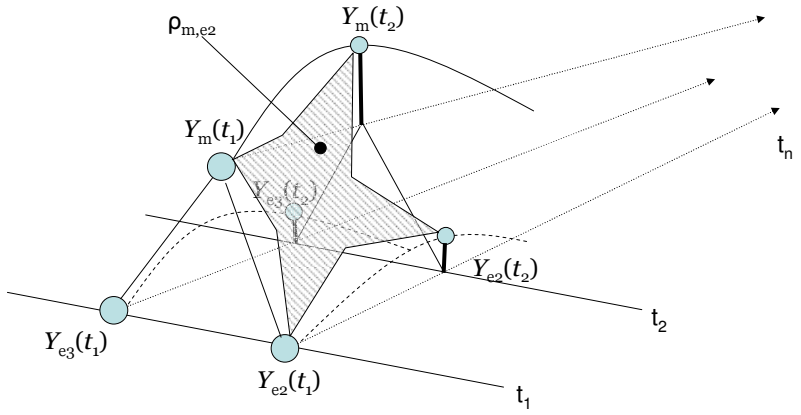
The resulted row vectors have to divide on emerging and declining group (determined by $D_{Ym}(t_n)$), than the extraordinary and normal stages were separated – according to their outlying properties in the tail distributions. Synchronizations between $D_{Ym}(t_n)$ and $D_{Ye1}(t_n)$, $D_{Ye2}(t_n)$, …, $D_{Yen}(t_n)$ raw vectors were signed by Spearman's rho correlation coefficients (4).

$$\rho_{DYm(li),\ DYek(li)} = \mathrm{cov}(D_{Ym}(l_i), D_{Yek}(l_i)) \,/\, \sigma_{DYm(li)}\, \sigma_{DYek(li)} \quad (4)$$

To separate normal and extreme log return developments on the main market, normal distribution was fitted on its empirical distribution. Heavy tails were identified,

where the theoretical and empirical distribution diverged. Than the significance of difference between normal and extreme average returns, standard deviations and correlation coefficients were compared with the usage of chi-square test.

Figure 2 Semi-dynamic hierarchical windowed correlation



Source: own edition

4. Results

4.1. Descriptive statistics

Standard picture of efficient markets are showed by the higher level of standard deviation at higher mean level of logarithmic means. But there is a lack of normal distribution as Kolmogorov-Smirnov test for normality suggest. Negative skewness means fatter left (dropping) side, against a robust right (emerging) side. Therefore extreme events had higher occurrence during negative developments, while positive movements were similar mostly in the case of all emerging stock markets opposing to Dow Jones Industrial. Bond markets behave quietly similar – except GKO-OFZ, which is the most fragile instrument in the sample. We can find the opposite in the case of currency markets due to the opposite logic of exchange rates. While the main currencies are determined by the devaluation trend of USD against EUR with rare test backs, emerging actors behave extremely during their devaluation. Kurtosis is the indicator of the representation of mode inside the distribution. High kurtosis

means a robust mode with dangerous backyard filled with extreme events – so the market usually operates normally, but it collapses when something weird happen.

Table 1 Data overview and descriptive statistics for the return series

						Kolmogorov-Smirnov	
	N	Skewness	Kurtosis	Mean	SD	Z	Asymp. Sig. (2-tailed)
DJI	1817	0.1241	10.2356	-0.0001	0.0128	4.1869	0.0000
RTS	1817	-0.6625	14.1959	0.0004	0.0225	4.8399	0.0000
BUX	1817	-0.3164	9.2518	0.0002	0.0159	2.6750	0.0000
WIG 20	1817	-0.2585	2.4823	0.0001	0.0161	2.0375	0.0005
US 10Y	1817	-0.1076	2.0352	0.0000	0.0007	1.8553	0.0020
EU 10Y	1817	-0.1124	1.3083	0.0000	0.0004	1.7123	0.0057
MAX	1817	-0.1237	18.6461	0.0002	0.0044	7.0802	0.0000
DZ0110	1537	-0.1507	5.6904	0.0000	0.0022	5.9033	0.0000
GKO-OFZ	780	-3.6813	42.6656	0.0000	0.0014	9.0965	0.0000
USD/EUR	1817	-0.2348	3.7052	0.0002	0.0063	2.0230	0.0006
HUF/EUR	1817	1.3987	12.8395	0.0001	0.0057	4.2770	0.0000
RBL/EUR	1817	0.7882	8.5373	0.0003	0.0053	3.5858	0.0000
PLN/EUR	1790	0.6517	5.2166	0.0001	0.0064	3.1172	0.0000

Source: own calculations

4.2. Test for power-law and exponential logarithmic yield distributions

According to literature suggestions, tail asymptotic distribution was examined on two ways. At first exponential and power-law tail distributions had to be separated by a simple R^2 based fitting-test[18]. Than estimated power-law properties were studied by Clauset, Shalizi and Newman's (2007) improved quantile-based maximum likelihood estimation (MLE) method[19], to estimate the scale parameter α. Size of the tails is determined by the scale parameter α – as smaller the α, as fatter is the tail. P-values are given by Monte Carlo procedures: the power-law model is fitted for generated synthetic data sets, and the number of times is counted when the Kolmogorov-Smirnov is larger than observed goodness-of-fit (maximum distance between the tail probability or cumulative distribution function of the empirical data and the fitted power-law model). (Clauset et al. 2009, Quismorio 2009)

[18] R-square – measures how successful the fit is in explaining of the variation of the data. A value closer to 1 indicates a better fit. General model for exponential: f(x) = a*exp(b*x) and for power-law: f(x) = a*x^b+c.

[19] Scripted in MATLAB, see http://www.santafe.edu/~aaronc/powerlaws/.

Table 2 Tail distributions

	Power-law distribution				Exponential distribution	
	Negative tail		Positive tail		Negative tail	Positive tail
	α	p	α	p	R^2	R^2
WIG 20	**1.9738**	**18.80%**	**2.1418**	**75.20%**	91.08%	94.30%
DJI	**1.8995**	**80.00%**	**1.8557**	**80.80%**	86.59%	81.05%
BUX	**1.7643**	**90.40%**	**2.0204**	**7.00%**	79.25%	80.93%
RTS	**1.8351**	**7.70%**	**1.9407**	**7.80%**	84.86%	72.34%
US 10Y	**2.0591**	**98.80%**	**2.1089**	**80.80%**	93.94%	91.90%
DZ0110	**2.2477**	**96.40%**	1.5072	0.00%	90.46%	92.95%
MAX	**1.7265**	**21.90%**	1.7709	***6.80%***	88.72%	77.89%
EU 10Y	2.5059	***5.90%***	1.7372	0.00%	**95.78%**	92.87%
GKO OFZ	1.2255	0.00%	1.2343	0.00%	**95.79%**	**97.25%**
USD/EUR	**2.0549**	**70.70%**	**2.0764**	**18.60%**	**98.00%**	91.58%
RBL/EUR	**1.9410**	**53.10%**	**1.8987**	**60.80%**	91.04%	82.85%
PLN/EUR	**1.9573**	**36.30%**	**1.8881**	**83.10%**	88.97%	87.27%
HUF/EUR	**1.9826**	**46.30%**	**1.8470**	**20.60%**	**98.66%**	78.38%

Source: own calculations

As literature said, there are significant differences in the thickness, which means that emerging markets have fatter negative tails than developed markets. During periods of boom longer and thicker tails were detected with a power-law exponent α close to 3, while periods of stagnation characterized by shorter and thinner tails with an exponential decay close to 1. (Quismorio 2009)

Such asymmetries are identified at the negative tails of the sample. Stock and currency markets are behaving as power-law in both cases of market developments, and their α exponent remains under the literally presented 2.5 to 3.5 interval – which signs a lot of opportunities for unpredictable events in each conjuncture phase. Bond markets have only negative power-law tails, but the fitting process could be too difficult due to the observed enormous kurtosis. EU 10Y indicator is smoother than US 10Y, which could be the result of strict inflation targeting monetary policy of the ECB.

According to the forms of distribution, there are no strict differences between developed and emerging markets, the difference in α is marginal.

4.3. Test of common developments under normal and extraordinary events

4.3.1. Static correlation on the entire observed period

In the case of stock markets on the entire period between 2002 and 2009, weak but significant (Sig.≈0) connection were signed by Spearmen's 2 tailed correlation between DJI and the emerging capital markets ($\rho_{RTS, DJI}$= 0.236; $\rho_{WIG\ 20, DJI}$= 0.239), while significant relation was only flagged between WIG20 and RTS ($\rho_{WIG\ 20, RTS}$= 0.046). Hungarian stock market had no significant correlation.

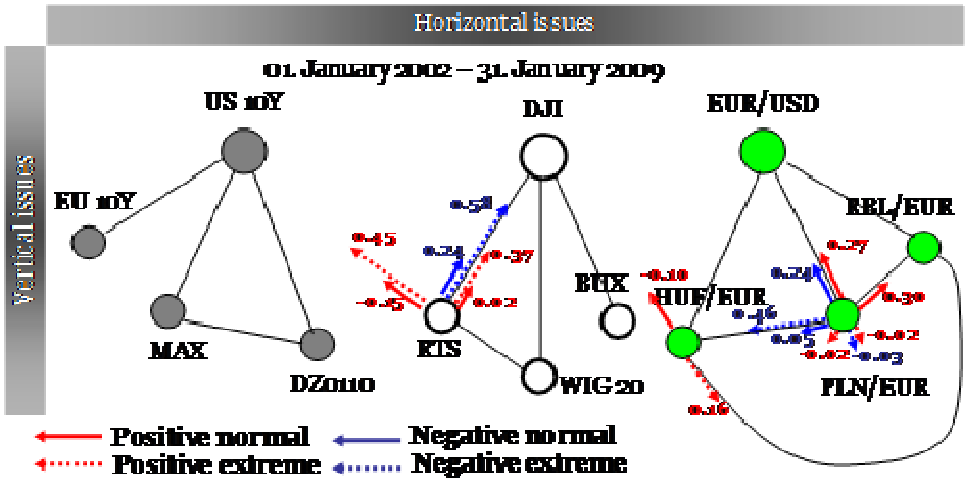
Developed bond markets had only significant but strong correlation ($\rho_{US\ 10Y, EU\ 10Y}$= 0.4726) during this period. Currency markets[20] had significant but varied correlations ($\rho_{EUR\ /USD, RBL/EUR}$= 0.915; $\rho_{EUR\ /USD, HUF/EUR}$= -0.064; $\rho_{EUR\ /USD, PLN/EUR}$= 0.11). PLN and HUF had also a strong significant connection ($\rho_{HUF\ /EUR, PLN/EUR}$= 0.44), while PLN and EUR relation remained weaker ($\rho_{HUF\ /EUR, PLN/EUR}$= 0.164). In Hungary, there was lack of significant correlation between bond, stock and currency markets, while WIG 20 was sensible on bond and currency market developments ($\rho_{WIG\ 20, US\ 10Y}$ = -0.111; $\rho_{WIG\ 20, EU\ 10Y}$= -0.106; $\rho_{WIG\ 20, MAX}$= 0.126; $\rho_{WIG\ 20, DZ0110}$= 0.06; $\rho_{EUR\ /USD, WIG\ 20}$= -0.055; $\rho_{WIG\ 20, HUF/EUR}$= -0.24; $\rho_{WIG\ 20, PLN/EUR}$= -0.21).Hungarian MAX index had converse relation with HUF and PLN exchange rates ($\rho_{MAX, HUF/EUR}$= -0.357; $\rho_{WIG\ 20, PLN/EUR}$= -0.249).

4.3.2. Dynamic correlation phenomena and tail distribution

The outputs of static spearmen's 2 tailed correlation signed poor inter connections on horizontal and vertical scale, but more detailed picture was provided after the establishment of quasi dynamic intervals ("windows"). Synchronisation of markets could depend on the probability of return developments.

[20] In the case of currency markets, it is necessary to study the exchange rate arrangements in analyzed countries as Stavárek (2009) suggest.

Figure 3 Calculated semi-dynamic hierarchical windowed correlation



Source: own calculations

The hypothesis, that the US (as "main") markets playing a dominant role with on influence on the emerging markets, was mostly rejected. There were only several cases, when significant difference occurred between market common movements according to the main market's normal or extreme stage. As Stavarek (2009) suggested, results were the inverse on currency market than on the others. USD-EUR relations had an indirect impact on Euro-zone candidate countries – PLN and HUF had a mediocre correlation under extreme strong periods of the euro, while PLN had a weak correlation with the EUR under normal circumstances. But this test was useful to sign the special role of Russian stock and currency markets – RTS moved only together with the main market (DJI) under extreme boosts and falls, and had a contrary connection to the WIG 20 under positive developments; while HUF, PLN and RBL tended to move together differently under stronger US dollar.

5. Discussion

Market risk management had to face with several problems in the last decade: market data returns are likely uncorrelated, but dependent, their probability distributions are heavy tailed, and extreme events appear in clusters while volatility is random (Embrechts et al. 2001). The conceptual model, which was presented in this study, had the same problems. More progressive solutions are given by the copulas as we mentioned in the methodological chapter, because probabilities and common devel-

opments are unified by them. But there are two causes to be careful during their usage: it is not clear from the basic models that the events on two or *n* margin's tails are happening in the same time, and correlation remained static in several applications.

To pay attention on the separation of monotonic or quasi monotonic developments of main market yields, ARIMA models could be more feasible – because short run fluctuations could be controlled better, than the presented multiplications of first derivates (or their rolling mean smoothed version). (Ramathan 2003)

6. Conclusion

Financial engineering facilitates the transformation and reshaping of risk, but some authors emphasize the disadvantage of network operations, in which the high efficiency on liquidity allocation capability and the ability of fast feedback through current account cause immanent instability of the current financial system with short-term orientation and unrelenting concentration of wealth (Brunnhuber et al. 2005, Magas 2005).

Imbalances of transitional countries were indicated by financial markets as models of bounded rationality and complex scale-free networks suggested. Developed countries have to face with extreme bullish and bearish circumstances, in which volatility of liquidity diffuse through the market. Emerging markets behave slightly altered, but it is not necessary to deeply revalidate the basic axioms on them as our results suggests. Differences could be found in their relationships to the occurrence of improbable events and their weak synchronisations under fat tailness. Capital flows are non-linear, therefore, the bottleneck-effects are caused by the withdrawal of the capital that is faster than its inflow. Occurrence of asymmetries means that negative shocks hit much harder these markets than positive news – so their collapse is much more sudden, while the recovery is slower. To handle emerging coutries as homogenous groups by foreign investors is the main reason of contagion effects – so if anything werid happens in these countries, funds are being extracted from other markets too (Égert – Koubaa 2004).

References

Babetskaia-Kukharchuk O., Babetskii I., Podpiera J. 2008: Convergence in exchange rates: market's view on CE-4 joining EMU. *Applied Economics Letters*, 15, pp. 385–390

Barabási, A-L., Albert, R. 1999: Emergence of Scaling in Random Networks. *Science.* 286, 509,

Bonanno, G., Lillo, F., Mantegna, R. N. 2001: Levels of Complexity in Financial Markets. *Physica A.*, 299, pp. 16-27

Brunnhuber, S. - Finkc, A. - Kuhlec J. 2005: The financial system matters: future perspectives and scenarios for a sustainable future. *Futures*, 37, 4, pp 317-332

Clauset, A., Shalizi, C. R., Newman, M. E. J. 2009. Power-law distributions in empirical data. *SIAM Review, to appear.*

Csermely, P. 2008: Creative elements: network-based predictions of active centres in proteins and cellular and social networks. *Trends in Biochemical Sciences*, 33, 12, pp. 569-576.

di Mauro, F., Rüffer, R., Bunda, I. 2008: The Changing Role of The Exchange Rate in a Globalised Economy. *ECB Occasional Paper Series.* NO 94 / September 2008

Embrechts P., Lindskog f., McNeil A. 2001: *Modeling Dependence with Copulas and Applications to Risk Management*. Department of Mathematics ETHZ, Zürich

Grubesic, T. H., Matisziw, T. C., Murray, A. T., Snediker, D. 2008: Comparative Approaches for Assessing Network Vulnerability. *International Regional Science Review*, 31, 88

Herrmann-Pillath, C. 2000: How to Research Complex Systems: A Methodological Comparison of "ORDO-Liberalism" and "Regulation Theory". In: Labrousse, A., Weisz, J.-D., (eds.), *Institutional Economics in France and Germany*, Reihe Economics and Philosophy, Heidelberg: Springer, pp. 272-301.

Hommes, C., Wagene, F. 2008: Complex Evolutionary Systems in Behavioral Finance. In Hens, T., Schenk-Hoppé, K.R. (eds.) *Handbook of Financial Markets: Dynamics and Evolution*, Tinbergen Institute TI 2008-054/1 Tinbergen Institute Discussion Paper, ACADEMIC PRESS CeNDEF, School of Economics, University of Amsterdam

Kiss, G. D. 2009: Analysis of Diversification Problems in the Case of Hungarian and Russian Pension Investments. In Mgr. Cervinek, P. (ed.) New Economic Challenges - 1st International PhD Students Conference, Masaryk University - Faculty of Economics and Administration, Brno, pp. 25-30.

Lukács, O. 1999: *Matematikai Statisztika*. Műszaki Könyvkiadó, Budapest

Magas, I. 2005: A pénzügyi integráció hozadékai a világgazdaságban - empirikus tapasztalatok 1970-2002. In: Botos Katalin (szerk): *Pénzügyek és globalizáció*, SZTE GTK Közleményei, JATEPress, Szeged pp. 139-161

Quismorio, B. A. 2009: The Tail Distribution of the Philippine Stock Price Index. *Working Paper, University of Philippines-Diliman.*

Ramanthan R. 2003: *Bevezetés az önonometriába*. Panem, Budapest

Stavárek, D. 2009: Assessment of the Exchange Rate Convergence in Euro-Candidate Countries. *Amfiteatru Economic Journal*, 11, 25, pp. 159-180.

Wang, X. F., Chen, G. 2003: Complex networks: small-world, scale-free and beyond. *Circuits and Systems Magazine, IEEE*, 3, 1, pp. 6-20.

Yuan, Baosheng – Wang, Bing-Hong – Li, Baowen 2007: Evolutionary Dynamics in Complex Networks of Adaptive and Competing Agents. *Working Paper, Great Eastern Life Assurance Co.*

Testing the UK stock market overreaction

Mohammad Abdul Washad Emambocus[1] *- Gurjeet Dhesi*[2]

The overreaction hypothesis asserts that investors tend to violate the Bayes' rule where they overreact to unanticipated and dramatic news or events in the financial world. This is consistent with the Representativeness Heuristic (Grether 1980) which is used as a basis for testing overreaction in the market. This paper scrutinises the behaviour of the UK stock prices in relation to the financial crisis (credit crunch). An event study is carried out for the period of 2004-2009 for the portfolios of Winners and Losers which have been assigned using quintiles of the best and worst performances of the sampled securities. Monthly data over the period shows clear evidence of stock overreaction behaviour in the UK market. The data is then split into sub-periods for further investigation which analyses the development of the overreaction of the market. The results of the overreaction have also been used as a proxy for the confidence level of investors in the market following global financial health. This provided evidence that in this case as the overreaction in the market decreases, the confidence in the market increased providing an indication of the mean starting to revert to its position.

Keywords: Overreaction Hypothesis, mean reversion, market instability

1. Introduction

Traditional theories assume that agents are able to process the relevant information at their disposal and form biased probability judgment on the basis of the Bayes theorem where investors tend to overweight recent information and underweight base rate (prior) data. Thus, overreaction of the market and investors is inconsistent with the Bayes rule which states that the market and rational investors update their

[1] Mohammad Abdul Washad Emambocus, PhD student, teaching assistant, London South Bank University, Faculty of Business (London)
[2] Dr Gurjeet Dhesi, PhD student, senior lecturer, London South Bank University, Faculty of Business (London)

Acknowledgment: We thank Dr Lee Rose, Professor Warwick, Tony Kilmister and Dr D'Silva for their valuable comments and suggestions. We express our deep gratitude to Jim Snaith, Head of Business Department at London South Bank University, for all the help provided in the research and attending conferences.

beliefs correctly. Therefore, this discrepancy has been at the core of modern financial economics and of particular interest to investors as it allows the opportunity for arbitrage leading to the irrational behaviours. There is now considerable evidence supporting the fact that returns of assets are predictable over long and short period of time for both the market and individual stocks. In this paper emphasis on a potentially crucial aspect of long-run mean reversion in stock returns known as the Overreaction Hypothesis has been scrutinised. This hypothesis, as put forward by De Bondt and Thaler (1985) analysing the US market showed that stocks which have witnessed poor performance (losers) over the past three to five years tend to surpass previous high performance stocks (winners) for the equivalent period length. This study looks at measuring the overreaction of UK stock market during 2007-2009 taking into consideration the financial crisis.

2. Literature Review

Overreaction Hypothesis states that *"extreme movements in stock prices will be followed by subsequent price movements in the opposite direction"* (De Bondt-Thaler 1985). This means that assets which have experienced considerable downfall in relation to the market in their returns over a period of time will surpass the market over the equivalent period of time (normally one to five years). Before deeper analysis of the phenomenon, a clear understanding of the aspect of overreaction in relation to an "appropriate reaction" is required. This distinction is provided by the two important theories in behavioural finance which are Bayes' Theorem and Representativeness Heuristic.

Bayes' theorem is concerns with the conditional and marginal probability distributions of random variables, especially a *priori* with a *posteriori* probabilities:

$$\boldsymbol{P(A \setminus B) = \frac{P(B \setminus A)P(A)}{P(B)}}$$

Bayes' Theorem assumes that investors are rational and therefore update their beliefs correctly with the application of conditional probability in order to weight both current and prior information.

On the other hand, Representativeness Heuristic asserts that investors are insensitive to prior probability of results and do not consider base rate frequencies. The subjective probability of an event, or a sample, is determined by the degree to which it is: firstly similar in essential characteristics to its parent population; and secondly reflects the salient features of the process by which it is generated (Kahneman-Tversky 1979). Therefore, investors tend to underweight prior data and overweight current data. This shows that is assigning probabilities, investors

normally violates the Bayes Theorem which might then leads to an overreaction in the market with the advent of a herding behaviour.

De Bondt and Thaler (1985, 1987) were the first put forward in their paper significant evidence to prove the phenomena of long-run overreaction. They ascertain that the return or prices of the stocks under consideration would adopt temporary swings away from their original fundamental values because of waves of confidence and pessimism in the market. Their findings demonstrated that losers/winners, measured and classified by their performance relative to the aggregate stock market over the past 3 to 5 year period, consistently surpassed or underperformed the market in the following 3 to 5 year period.

There are two main explanations that have been proposed for the overreaction hypothesis which are subject to debates. The first reason is related to the size effect, whereby the portfolio of the losers tends to be smaller and usually the smaller companies outperform large corporations (Zarowin, 1990). The second explanation is that the reversal in the share prices reflects changes in equilibrium required returns (Chan 1988). That is the significant movements in the leverage will lead to high changes in a companies' CAPM beta. Therefore, betas of the extreme losers go beyond the betas of the top winners. Ball and Kothari (1989) found that these changes in betas exceeded by 0.76. These differences in the betas clearly explain large differences in realised returns therefore providing evidence for overreaction.

3. Methodology and Data

Definitions of Winners and Losers

For the purpose of this study, winners and losers are being defined as being the best and worst performing stocks over the specific period of time respectively which in this case is from 2007-2009. In this perspective, quintiles method has been used in relation to the performance of the stock returns where the assets were classified into two portfolios. Note that this study considers only the top and bottom quintiles of stocks by their performances. Therefore, portfolios of stocks are formed on the criteria of prior period performance.

In order to test the Overreaction Hypothesis in the UK stock market, event study techniques were applied similar to De Bondt and Thaler (1985). The end of the month dividend adjusted returns of the sampled stocks on the FTSE All Shares were used for the classification of the portfolios. Moreover, the FTSE All Shares was used to represent the market to which the shares' performances were marked to form three different periods which involved 1, 12, 24 and 36 months. Therefore, the market adjusted returns for the end of the month were calculated as follows:

$$U_{it} = R_{it} - R_{mt} \qquad t = 1...n. \qquad (1)$$

where:

n = 1 .. 12, or n = 1 .. 24, or n = 1 .. 36.

$\boldsymbol{U}_{it}$ = the difference between the log return on stock *i* at period *t*

$\boldsymbol{R}_{it}$ = the log return of the individual companies at period *t*

$\boldsymbol{R}_{mt}$ = the log return of the FTSE All Shares at period *t*

Then, for the every stock for the period starting on the 1st of January 2007, we calculate the cumulative excess return which is denoted as $\sum_{t=35}^{t=0} \boldsymbol{U}_{it}$ for the overall 36 months. Hence, the $\boldsymbol{CU}_i$'s are taken into consideration in order to ranked them from low to high performing stocks where the quintile method is applied to form the portfolios. The top 17 companies are therefore classified as being the equally weighted portfolio of winners (*W*) and the companies at the bottom of the quintiles are grouped together to form the loser portfolio (*L*). Moreover, the Cumulative average returns of the portfolio ($\overline{\boldsymbol{ACUR}_p}$) are computed over the post portfolio formation period, this is determined through an average of the $\boldsymbol{CU}_i$ of the stocks from each portfolios. Therefore, there will be the results for the individual cumulative average return of individual returns ($\overline{\boldsymbol{ACUR}_i}$) and also the cumulative average return of the portfolios. The cumulative averages of the portfolio returns for the winner quintile and loser quintile are denoted as $\overline{\boldsymbol{ACUR}_P^W}$ and $\overline{\boldsymbol{ACUR}_P^L}$ respectively. Consequently, by subtracting both cumulative average return we get the $\overline{\boldsymbol{ACUR}_{Dt}}$. Therefore, if the return on the later is insignificantly different from zero then we cannot accept the simple overreaction hypothesis bearing the assumption that differences in transaction costs between both portfolios do not influence the results of $\overline{\boldsymbol{ACUR}_{Dt}}$. On the other hand, a significant positive value for $\overline{\boldsymbol{ACUR}_{Dt}}$ will eventually confirms overreaction for the UK stock market for the period under consideration.

Moreover, the differences in the average cumulative returns also points out one of the most important aspect of overreaction, which is the contrarian strategy. Here the contrarian investor attempts to profit by investing in a manner that differs from conventional wisdom, which is related to the herding behaviour, where the investors try to exploit the mispricing in the security market.

3.1 Test 1

The first test in order to investigate the existence of overreaction in the UK stock market is to compare the means of both portfolios returns (that is, the winner and loser) by regressing $\overline{CUR_{Dt}}$ against a constant:

$$\overline{CUR_{Dt}} = \overline{CUR_P^L} - \overline{CUR_P^W} = \alpha_1 + n_t \qquad \text{t = 1 … n} \qquad (2)$$

where

α_1 is a constant

$\boldsymbol{n_t}$ = white noise error term

t = the period after portfolio formation

n = 1, 12, 24 and 36 months

A *t-test* has been carried out in order to test the significance of the constant α_1 which will provide evidence whether there is a difference in the means of the winner and loser portfolios. As such, a significantly positive result for the constant can be interpreted as being a confirmation for the overreaction hypothesis. This indicates that the average performance of the losers is greater than the average performance of the winners. Consequently, an insignificant result indicates no overreaction hypothesis.

Moreover, $\overline{CUR_{Dt}}$ is also used to control for possible risk differences which is carried out by regressing it against the market risk premium. This test enables the identify the exposures of the portfolios to systematic risks thus providing explanation for the differences in returns. This is provided in the equation (3):

$$\overline{CUR_{Dt}} = \alpha_2 + \beta(RM_t - RF_t) + \varepsilon_t \qquad \text{t = 1 … n} \qquad (3)$$

where:

α_2 = the Jensen performance index

β = the difference between market beta for returns of the two portfolios

R_f = the risk free rate (UK government month T-Bill rate)

ε = a white noise error term

t = the period after portfolio formation

n = 1, 12, 24 and 36 months

Therefore, if the α_2 has a significantly positive value, this will confirm an overreaction in the UK market. Moreover, if β is significantly different from zero then it can said that systematic risk is responsible for the differences in returns whereby the losers may contain more systematic risk than the winners.

3.2 Test 2

The second test which is used to confirm or reject the conclusions of the first analysis uses the cumulative average returns of the stocks in the portfolios for the 36 months. As such, $\overline{\boldsymbol{CUR}_{Pt}^{W}}$ and $\overline{\boldsymbol{CUR}_{Pt}^{L}}$ (where *t = 1* through *t = 36*) are taken in order to determine the average cumulative returns of both portfolios denoted as $\boldsymbol{ACUR}_{W,t}$ and $\boldsymbol{ACUR}_{L,t}$. For *t > 0*, if $\boldsymbol{ACUR}_{W,t} < 0$ and $\boldsymbol{ACUR}_{L,t} > 0$ indicates overreaction in the market, therefore, $\{\boldsymbol{ACUR}_{L,t} - \boldsymbol{ACUR}_{W,t}\} > 0$. Further analysis to test whether average cumulative return provides indication that there is a statistically significant difference in the investment performance; the pooled population of the cumulative average return are estimated as follows:

$$\boldsymbol{S}_t^2 = [\sum_{n-1}^{N}(\boldsymbol{CUR}_{W,n,t} - \boldsymbol{ACUR}_{W,t})^2 + \sum_{n=1}^{N}(\boldsymbol{CUR}_{L,n,t} - \boldsymbol{ACUR}_{L,t})^2] / 2(\boldsymbol{N} - 1). \quad (4)$$

The variances of the two portfolios mean having with two equal sample size *N* is $2\boldsymbol{S}_t^2 / \boldsymbol{N}$ and the t-statistic is as follows:

$$\boldsymbol{T}_t = [\boldsymbol{ACUR}_{L,t} - \boldsymbol{ACUR}_{W,t}] / \sqrt{2\boldsymbol{S}_t^2 / \boldsymbol{N}} \quad (5)$$

The *t*-statistics can be computed for each of the 36 months however they do not represent independent evidence.

Moreover, tests have been conducted in order to determine whether at any month *t,* the average return makes a contribution to $\boldsymbol{ACUR}_{W,t}$ and $\boldsymbol{ACUR}_{L,t}$. The tests comprise the verification whether the contributions are statistically significant from zero. Therefore, the sample standard deviation of the winner is:

$$\boldsymbol{s}_t = \sqrt{\sum_{n=1}^{N} (\boldsymbol{AR}_{W,n,t} - \boldsymbol{AR}_{W,t})^2 / \boldsymbol{N} - 1} \qquad (6)$$

$\boldsymbol{s}_t / \sqrt{\boldsymbol{N}}$ is the sample estimate of the standard error of $\boldsymbol{AR}_{W,t}$, the *t*-statistics is:

$$\boldsymbol{T}_t = \boldsymbol{AR}_{W,t} / (\boldsymbol{s}_t / \sqrt{\boldsymbol{N}}) \qquad (7)$$

The same procedure is repeated for the analysis of the loser portfolio.

4. Discussion

The results of the tests are illustrated in tables 1 and 2. They provide evidence to support the assumption of Overreaction Hypothesis. Over the period analysed for the UK stock market, the results show that prior losers subsequently earn positive risk-adjusted excess returns and prior winners subsequently earn negative risk-adjusted excess return. Referring to the $\boldsymbol{ACUR}_{C,t}$ which provides an indication of stock market overreaction and contrarian strategies adopted by investors, clearly points out that the market have been overreacting during the whole three years. Moreover, as a measure of contrarian, the $\boldsymbol{ACUR}_{C,t}$ suggests a herding behaviour on behalf of investors to take advantage of the mispricing or chaos in the financial market. For the individual period of months (1, 12, 24 & 36), $\boldsymbol{ACUR}_{W,t} < 0$ and $\boldsymbol{ACUR}_{L,t} > 0$, further indicates overreaction in the UK securities in relation to the market, thus $\{\boldsymbol{ACUR}_{L,t} - \boldsymbol{ACUR}_{W,t}\} > 0$ and the accompanying *t*-statistics statistically support the hypothesis of overreaction in the UK market. Consequently, in line with the work of De Bondt and Thaler (1985), the results shows that the portfolio of prior losers are found to outperform the portfolio of prior winners by about 60% for the first month. These facts prove that investors violate the Bayes' theorem where they tend to overweight current information and underweight prior

information. These have been described as the Representativeness Heuristics where investors move away from the use of conditional probabilities in their strategies for portfolio assignment.

Table 1. Overreaction Hypothesis Test

Nature of the portfolios (Test 2)	**ACAR (t-statistics) into the test periods**			
	Number of months into the test periods			
	1	**12**	**24**	**36**
Loser	**0.14487**	**0.51236**	**1.37967**	**1.41962**
Winner	**-0.09771**	**-1.27111**	**-3.34595**	**-2.38167**
$ACUR_{C,t}$	**0.24259**	**1.78347**	**4.72562**	**3.80129**
S_t^2	**0.24431**	**15.0258**	**104.791**	**61.5013**
$2S_t^2 / N$	**0.02874**	**1.76775**	**12.3283**	**7.23545**
$t-statistics$	**1.43095****	**1.34139****	**1.34588****	**1.34139****
$S_{t,W}$	**0.13058**	**0.15585**	**0.14063**	**N/A**
$S_{t,L}$	**0.31954**	**0.03653**	**0.01816**	**N/A**
$s_{t,W} / \sqrt{N}$	**0.03167**	**0.03780**	**0.03410**	**N/A**
$s_{t,L} / \sqrt{N}$	**0.07750**	**0.00886**	**0.00440**	**N/A**
$T_{t,W}$	**-2.12800**	**-1.44864**	**-2.0452**	**N/A**
$T_{t,L}$	**1.25091**	**1.48558**	**5.90915**	**N/A**

** Significant at 95% level

Source: own creation

The *t*-statistics from table 1 shows that the performances of the different portfolios are statistically significant at 95% confidence level. That is, investors focus on securities that are involved with extreme return experiences. However, for the mean-reversion, it is not to be in line with the works of De Bondt and Thaler (1985), where they showed that greater the mispricing in the market, quicker will be the reversion of the mean. In this case, due to the nature of the crisis, slow reversion has been observed as the overreaction will take longer to revert back to its original position. The crisis is of a global phenomenon where there is a high interdependence of stock markets. This element makes the recovery of the market to be more painful which huge consequences to the economy.

The next section of the analysis is related to measure the degree of the overreaction in the UK market. It has already been specify that there has been evidence of overreaction in the market; however it is crucial to measure the extent or gravity of the overreaction. These provide an indication on the investors' sentiment in relation to the financial instability. It has also been used as a proxy for measuring the riskiness of the market or the confidence of investors to the reliability in the market.

Table 2. Market Overreaction

	N=12	N=24	N=36
Return on Loser	**0.013163**	**0.094011**	**0.02176**
Return on Winner	**-0.05476**	**-0.01961**	**-0.03667**
TEST 1			
$\overline{CUR_{Dt}}$ =	**0.067921**	**0.113623**	**0.058431**
Annualised Return Difference	**120%**	**200%**	**64%**
Systematic Risk			
α_2	**-0.001**	**0.022**	**0.042**
Significant F	**Sig F <0.05**	**Sig F <0.05**	**Sig F <0.05**
β	**0.71**	**1.01**	**1.00**

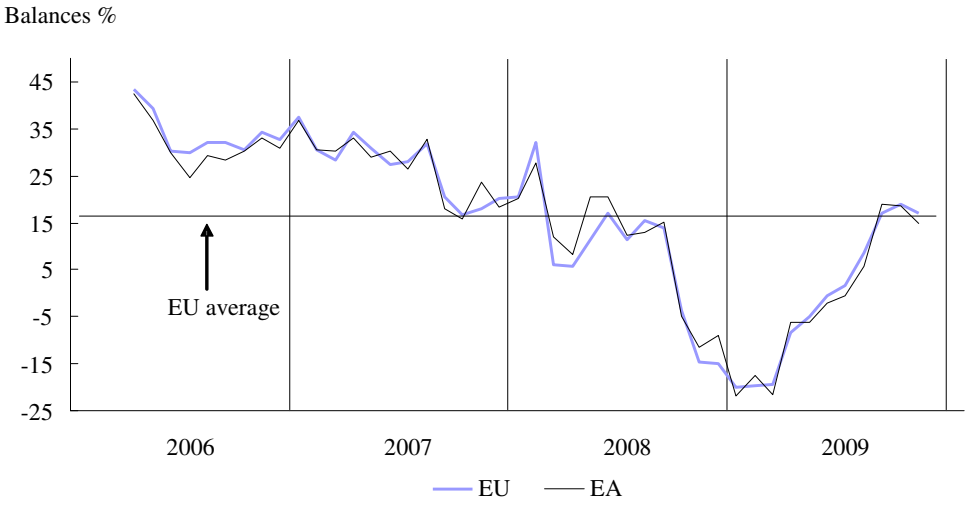
Source: own creation

The systematic risk analysis shows that as all the significant F < 0.05, so the null hypothesis that there is no evidence of overreaction in the UK market is rejected. Moreover, the β clearly points out that much of the overreaction in the market is mainly due to the volatility in the market which is a result of the subprime crisis. Table 2 also shows that during the first 12 months (i.e. in 2007), the overreaction amounted to 120% indicating a herding behaviour of investors due to the dramatic news events in the world and local market. In 2008, the overreaction escalated to a 200%. This can be explained by the several events on a global level leading investors to be pessimistic about the market. In September 7th 2008, the world witnessed a historic US mortgage bailout where two large American mortgage lending companies, Fannie Mae and Freddie Mac were rescued by the US government. Another blow came from Lehman Brothers on September the 15th of 2008 where the company filed for bankruptcy. On the same day, Merrill Lynch was saved by another huge American bank. This was the worst period for the Wall Street and the risk premium on the bank's debt soared, as the solvency concerns increased resulting with alarming funding issues for vulnerable banks. Moreover, in

October 8th of 2008, the UK government in turn announced a £500 billion rescue package for the banking sector. However, even all the measures taken by the US and UK governments to save the financial system, these did not reassured or restored the confidence of investors resulting into an overreaction in the market both in UK and the US.

On the other hand, in 2009 the UK financial sector witnessed a surge of the confidence level in the market leading to a fall of the overreaction to 64%. Even though the world financial health was still in jeopardy with the advent of the UK economy falling into a recession on the 23rd of January 2009 and with a fall in the GDP by 0.8% in July 2009; investors were better prepared which did not led into frantic mania in the financial world. The Economic Sentiment Indicator (ESI) showed that the European businesses and consumers demonstrated an increased in their level of optimism about the economy. In November 2009, the ESI rose to 87.9 (+1.9) in the EU and 88.8 (+2.7) in the European area (European Commission 2009). Even though the European economies' indicator suggest that they are still below the long-term average, nonetheless there have been an improvement in the sentiment level to the upward direction which is a sign of the investors regaining back their confidence over the financial system. Refer to figure 1:

Figure 1. Financial Services Confidence Indicator

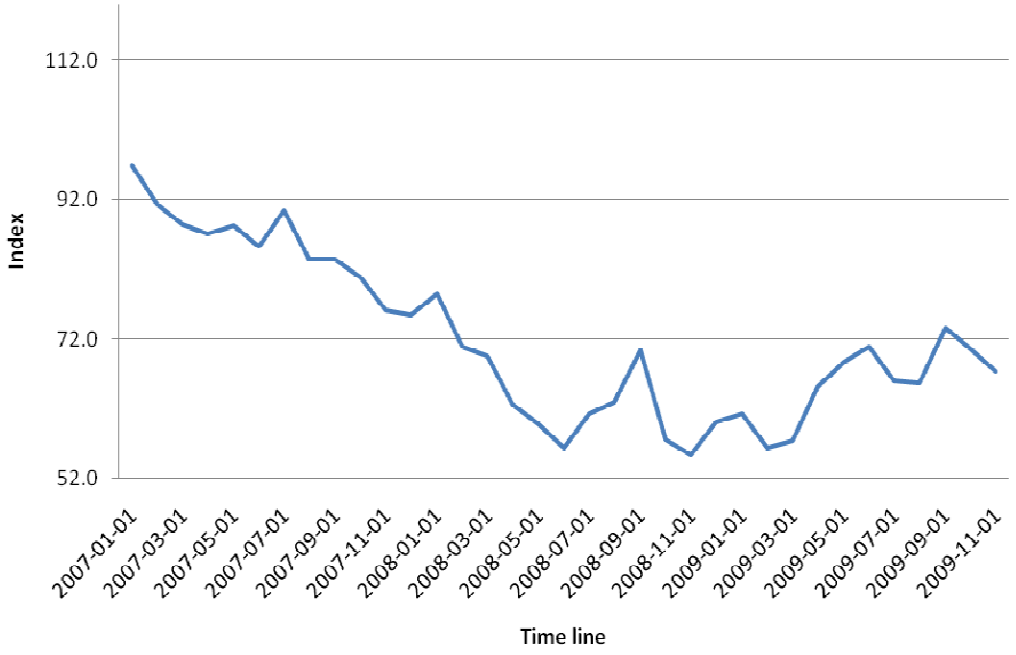


Source: European Commission 2009

Figure 1 clearly illustrates that the confidence over the financial sector is improving which confirms the fact that the degree of overreaction went down to 64% in 2009. This change in the confidence level can also be attributed to the fact that various member states in the European Union reported an increase in their economic sentiment. In Netherlands the sentiment rose sharply to a +6.3, France (+2.2) and Germany (+1.7). Even though UK did not witnessed a rise in their sentiment, the general improvement in the EU sentiment resulted into a fall in the overreaction in the UK market.

On a world level, the Michigan Consumer Sentiment Index can be used to reflect the level of sentiment gained in the financial system. Refer to figure 2 for an indication of the confidence level in the world market:

Figure 2. Consumer Sentiment Index for 2007-2009



Source: Derived from data collected from Michigan Consumer Sentiment Index

Figure 2 is in line with the findings in table 2 where the US confidence level can be taken as a proxy for the UK sentiment degree over the financial system as both markets are highly correlated. It can observed that during the first 12 months the confidence level fell by significant level of 22% which led to a huge overreaction

level of 160% in 2007. However, for the 24 months the sentiment index has been fluctuating nonetheless the herding behaviour of investors resulted into a figure of 200% due to the reasons already mentioned above. On a positive note, in 2009 the sentiment index started to take an upward movement which led to a significant improvement in the confidence level in the UK financial system.

5. Conclusions

The research conducted provided evidence that investors tend to overreact to unexpected and dramatic news events which in this case is the financial crisis initiated from a housing bubble. This kind of investors' behaviour is crucial to investigate as it constitutes a main determinant to the financial health of the market.

The results is in line with the Overreaction Hypothesis and also with the Representativeness Heuristics relating mainly to the behaviours of investors at the advent of an extreme event. The paper showed that the losers outperformed prior winners which mean that investors have been underweighting prior information and overweighting recent information or data applying subjective probability distribution.

The evidence of the evolution of the overreaction in the UK market has also been used in this paper as a proxy to analyse the confidence or sentiment level of the market. The research showed that as the herding behaviour in the market went down, the confidence level went up even through in is below the long-term average level of confidence. This has been further justified through the analysis of Financial Services Confidence Indicator and the Michigan Consumer Sentiment Index.

Consequently, the behaviour of investors constitute a crucial role in identifying the predictability of the market. However, the research showed that the identification of this arbitrage opportunity or looking for the period where the mean will revert back to its position is not evident.

References

Ball, R. - Kothari, S. 1989: Nonstationary Expected Returns: Implications for Tests of Market Efficiency and Serial Correlation in Returns. *Journal of Financial Economics,* 25, 51-74. p.

Business Wire 2003: Greenspan Taking "Preemptive Move" Against Market Overreaction. [Online] Available from http://www.allbusiness.com/economy-economic-indicators/economic-conditions-deflation/5719904-1.html [Accessed 12/12/09]

Chan, K.C. 1988: On the Contrarian Investment Strategy. *Journal of Business,* 61, 147. p.

Chang, T. – McLeavey, D. – Rhee, S. 1995: Short-term abnormal returns of the contrarian strategy in the Japanese stock market. *Journal of Business Finance and Accounting* 22, 1035-1048. p.

De Bondt, W. - Thaler, R. 1985: Does the Stock Market Overreact? *Journal of Finance,* 40, 793-805. p.

DeBondt, W. - Thaler, R. 1987: Further evidence on investor overreaction and stock market seasonality. *Journal of Finance*, 42, 557-581. p.

Economic Research 2009: Michigan Consumer Index [Online] Available from http://research.stlouisfed.org/fred2/graph/?s[1][id]=UMCSENT# [Accessed 23/11/09]

European Commission 2009: Business and Consumer Survey Results [Online] Available from: http://ec.europa.eu/economy_finance/db_indicators/db_indicators8650_en.htm [Accessed 16/12/09]

Grether, D. 1980: Bayes Rule as a Descriptive Model: The Representativeness Heuristic. *Quarterly Journal of Economics,* 95, 537-557. p..

Kahneman, D. – Tversky, A. 1979: Prospect theory: an analysis of decision under risk. *Econometrica,* 47, 263–291. p.

Kahneman, D. – Tversky, A. 1982: Intuitive Prediction: Biases and Corrective Procedures. Cambridge University Press, London.

Zarowin, P. 1990: Size, seasonality and stock market overreaction. *Journal of Financial and Quantitative Analysis* 25, 113-125. p.

Changes in the Monthly Effects from the Romanian Foreign Exchange Market

Ramona Dumitriu[1] *– Costel Nistor*[2] *- Razvan Stefanescu*[3]

This paper investigates the systematic patterns displayed by the Romanian Foreign Exchange Market in some months of the year. In our analysis we employ monthly values of the Romanian national currency rates against the United States dollar and the euro. We find that since the Foreign Exchange Market settlement in Romania until present significant changes occurred in the monthly effects. We associate these changes to the new tend in the Romanian foreign trade and to the new monetary strategy of the central bank.

Keywords: Seasonality, Romanian Foreign Exchange Market, Foreign Trade, Monetary Policy

1. Introduction

The financial markets seasonality is among the subjects largely approached in the specialized literature. Especially for the stock markets there were studied different forms of seasonality: quarterly, monthly, weekly, daily a.s.o. (Wachtel 1942, Officer 1975, Rozeff and Kiney 1976, Gultken and Gultken 1983, Agrawal and Tandor 1994, Schwert 2001). From the financial decisions perspective it is important not only the seasonality of price levels but also the seasonality of prices volatility (Tang 1998).

In many aspects the exchange rates behavior is similar to the prices of other financial assets, their evolutions reflecting the market expectations (Frenkel 1981). However, there are some circumstances that differentiate the exchange rates seasonality. Along with expectations other factors, such as the central bank interventions or the external transactions could influence decisively the prices of the foreign currency.

[1] Ramona Dumitriu, PhD student, lecturer, „Dunarea de Jos" University Galati, Faculty of Economics, Romania
[2] Costel Nistor, PhD, assoc. prof., „Dunarea de Jos" University Galati, Faculty of Economics, Romania
[3] Razvan Stefanescu, PhD, lecturer, „Dunarea de Jos" University Galati, Faculty of Economics, Romania

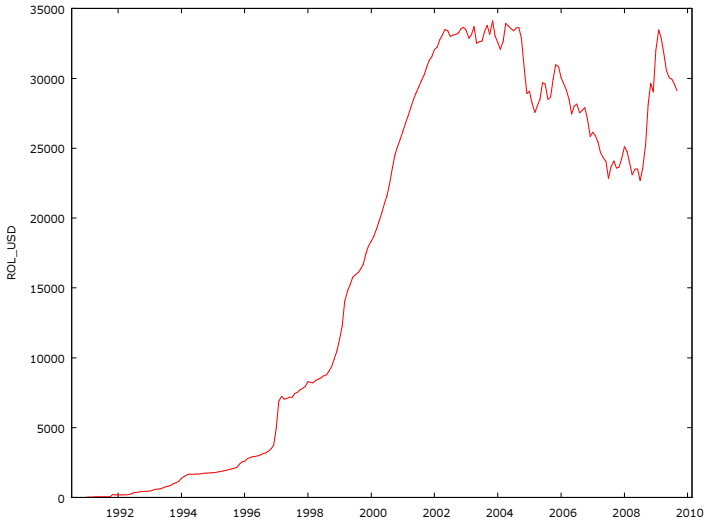
The Romanian foreign exchange market evolution in the last two decades could be split in two stages:

- a transition stage, between 1991 and 1998;
- a consolidation stage, between 1999 and 2009.

During the first stage, in the context of transition to the market economy, the Romanian foreign exchange market experienced some substantial transformation. After the communist regime collapse in December 1989 the Romanian authorities implemented transition reforms, which included the foreign exchange market liberalization. However, this process was quite slow and many restrictions, especially the ones for the foreign capitals, lasted for a long time.

In most of the 1990s in Romania there was a serious macroeconomic instability with a severe decline of the industrial production, significant deficits of the trade balance and a high inflation. Financing the trade balance deficits was very difficult in the absence of the foreign investment and since borrowing in foreign currencies was not easy. In these circumstances, after the significant decrease of the foreign reserve, the national currency (ROL) devaluation was unavoidable (Figure 1). However, in order not to aggravate the monetary instability, the National Bank of Romania (NBR) tried to obtain d a slow and controlled depreciation. In these years the trust in the national currency was quite low, so in many transactions the dollar and other foreign currencies were preferred. Such currencies were bought on the black market at higher prices than the official exchange rates.

Figure 1. Evolution of ROL/USD exchange rate from January 1991 to September 2009



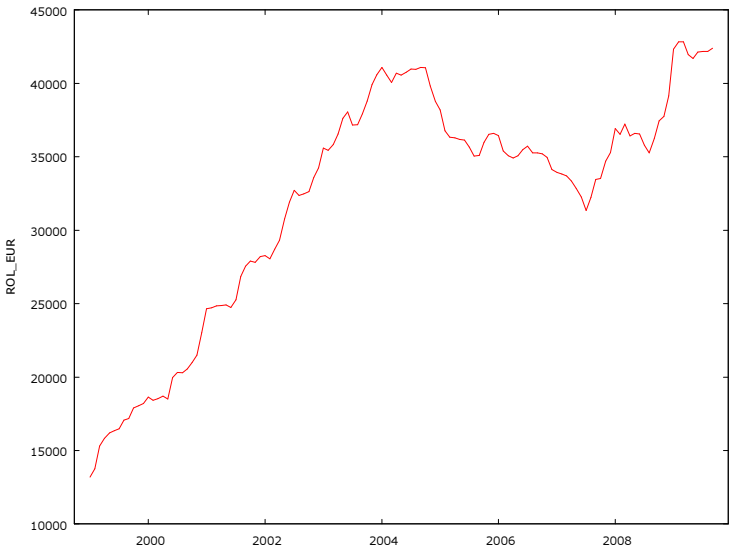
Source: own creation

The intervention in the exchange rates evolution was subordinated to the price control applied by the Government. Most of the prices with major impact on population were related to the exchanges rates, so the authorities tried to avoid abrupt depreciations. However, especially in some winter months, when the imports of oil and gas substantially led to the foreign reserves decrease, NBR had to allow significant devaluation. These events led to some expectations about significant depreciations during the cold seasons. The high inflation contributed, by deteriorating of national currency purchasing power, to the devaluation from this period of time.

The consolidation stage of the Romanian foreign exchange market coincided with the recovery of the national economy. The new European currency, euro, replaced, step by step, the US dollar in many transactions and NBR added the evolution of ROL / EUR exchange rate to its main targets. In the first years some restrictions for the foreign capitals were eliminated and the other could be easily avoided. In these circumstances, the speculative operations with foreign capitals began to play a major role. The privatization process was accelerated and most of the prices were liberalized. Since the inflationary pressure was attenuated, NBR allowed a substantial devaluation in the first years. In 2005 NBR officially adopted the inflation targeting regime for its monetary policy and it liberalized the foreign capital inflows. It was announced the intention of a lower intervention of NBR on the foreign exchange market, a measure with potential significant impact on the volatility of the exchange rates. In the next years the foreign direct investment, the remittances of the Romanian workers from abroad and the real interest rates high level contributed to the appreciation of the national currency. However, between 2008 and 2009, in the financial crisis context, the national currency depreciated again (Figure 2).

In this article we investigate the monthly effects exhibited by the foreign exchange market during the two stages. We also try to identify the presence of the exchange rates volatility monthly seasonality before and after the adoption of the inflation targeting regime and the liberalization of the foreign capital inflows. For that purpose we use classical regression techniques with dummy variables for each month.

Figure 2. Evolution of ROL/EUR exchange rate from January 1999 to September 2009



Source: own creation

The rest of the paper is organized as follows. In the second part we describe the data and the methodology used in our analysis. In the third part we present the empirical results and in the fourth part we conclude.

2. Data and Methodology

We employ average monthly values of ROL / USD and ROL / EUR provided by NBR. For the ROL / USD exchange rate we used data from January 1991 to September 2009. For the ROL / EUR exchange rate the data are from January 1999.

For both series of time we calculated the monthly returns as it follows:

$$R_t = 100 * [\ln (S_t) - \ln (S_{t-1})] \quad (1)$$

where S_t *and* S_{t-1} are average exchange rates in the months t and t-1, respectively.

We use two variables to express the returns of the two time series:

- *RUSD* as the returns for ROL / USD monthly exchange rates;
- *REUR* as the returns for ROL / EUR monthly exchange rates.

We estimate the volatilities of the two exchange rates by monthly coefficients of variation:

- *CVUSD* – monthly coefficient of variation for ROL / USD exchange rates from January 1998 to September 2009;

- *CVEUR* - monthly coefficient of variation for ROL / EUR exchange rates from January 1999 to September 2009.

We also define for every month of a year a dummy variable which takes the value of one in the respective month and value of zero otherwise.

In order to capture the changes in the seasonality of the returns induced by the economic recovery and the foreign capital inflows we split the sample of ROL/USD monthly values in two periods of time: one from January 1991 to December 1997 and the other from January 1998 to September 2009. Also, we analyze the seasonality of the exchange rates volatilities for two periods of time: before and after liberalization of foreign capital inflows.

The stationarity of the time series will be investigated by the Augmented Dickey – Fuller tests. In case of a variable is proved to be not stationary in level we will use in the further analysis its first differences.

We estimate the seasonalities of the time series by performing regressions in which the dummy variables are included. We use two types of models: a simple one and an autoregressive one.

The simple model has the equation:

$$R_t = \sum_{i=1}^{12} a_i * d_{it} + u_t \qquad (2)$$

where d_{it} is a monthly dummy variable taking the value one for the month I and zero otherwise. An a_i coefficient could be interpreted as the average returns in the month i.

The autoregressive model has the form:

$$R_t = \sum_{i=1}^{12} b_i * d_{it} + c * R_{t-1} + u_t \qquad (3)$$

We determine the coefficients of the two equations by the OLS technique.

3. Empirical Results

3.1. Stationarity of the variables

In the Table 1 there are presented the results of ADF Tests for RUSD and REUR. It indicates that both variables are stationary in level.

Table 1. Augmented Dickey-Fuller Test for RUSD and REUR

Variable	Deterministic terms	Lagged differences	Test statistics	Asymptotic p-value
RUSD	No constant and no trend	13	-2.76197	0.005584
	Constant and no trend	13	-2.83042	0.05402
REUR	No constant and no trend	6	-2.9024	0.003605
	Constant and no trend	6	-3.22484	0.01862

Note: The number of the lagged differences was chosen based on Akaike Information Criteria.
Source: own creation

The results of ADF tests for CVUSD and CVEUR are presented in the Table 2. According to them both variables are not stationary in levels but stationary in their first differences.

Table 2. Augmented Dickey-Fuller Test for CVUSD and CVEUR in levels and in their first differences

Variable	Deterministic terms	Lagged differences	Test statistics	Asymptotic p-value
CVUSD	Constant and no trend	13	-1.11572	0.7119
	Constant and trend	13	-1.85226	0.6791
d_CVUSD	No constant and no trend	12	-5.60457	0.00001
	Constant and no trend	12	-5.57005	0.00001
CVEUR	Constant and no trend	19	-1.54392	0.5114
	Constant and trend	19	-1.84073	0.6849
d_CVEUR	No constant and no trend	18	-2.67083	0.007342
	Constant and no trend	18	-2.67273	0.0788

Note: The number of the lagged differences was chosen based on the Akaike Information Criteria.
Source: own creation

3.2. Monthly effects for the simple model

In the Table 3 there are presented the monthly effects of the ROL / USD exchange rates for the simple model from January 1991 to December 1997. It indicates significant effects in seven months of the year.

Table 3. Monthly effects for ROL/USD from January 1991 to December 1997 in a simple model

$$RUSD_t = \sum_{i=1}^{12} a_i * d_{i_t} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	10.4569	4.98841	2.0962	0.03963**
dfeb	7.98733	4.70278	1.6984	0.09381*
dmar	4.98565	1.79982	2.7701	0.00715***
dapr	8.36844	7.628	1.0971	0.27632
dmay	2.69409	1.73302	1.5546	0.12450
djun	4.77079	2.35424	2.0265	0.04647**
djul	6.71823	4.26268	1.5761	0.11946
daug	2.78312	1.21943	2.2823	0.02547**
dsep	3.1629	1.2339	2.5633	0.01249**
doct	3.87256	1.76445	2.1948	0.03145**
dnov	20.9283	18.0418	1.1600	0.24994
ddec	2.23515	2.18932	1.0209	0.31075

Mean dependent var	6.533585	S.D. dependent var	15.21644
Sum squared resid	16906.26	S.E. of regression	15.43102
R-squared	0.109554	Adjusted R-squared	-0.028402
F(11, 71)	3.678957	P-value(F)	0.000369
Log-likelihood	-338.4108	Akaike criterion	700.8215
Schwarz criterion	729.8476	Hannan-Quinn	712.4826
rho	-0.018846	Durbin-Watson	2.032077

Source: own creation

The monthly effects of ROL / USD exchange rates for the simple model from January 1998 to September 2009 are presented in the Table 4 from which monthly effects in January, June and September result.

Table 4. Monthly effects for ROL/USD from January 1998 to September 2009 in a simple model

$$RUSD_t = \sum_{i=1}^{12} a_i * d_{i_t} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	2.18401	1.15592	1.8894	0.06110*
dfeb	0.515866	0.930067	0.5547	0.58010
dmar	0.922604	1.33115	0.6931	0.48951
dapr	0.66254	0.900559	0.7357	0.46326
dmay	0.0663186	0.794724	0.0834	0.93363
djun	0.989395	0.558894	1.7703	0.07906*
djul	-0.307903	0.657667	-0.4682	0.64046
daug	0.871036	0.713578	1.2207	0.22446
dsep	1.71822	0.696274	2.4677	0.01492**
doct	1.94194	1.20064	1.6174	0.10825
dnov	1.5259	1.15677	1.3191	0.18949
ddec	-0.169547	1.14274	-0.1484	0.88229

Mean dependent var	0.896873	S.D. dependent var	3.103229
Sum squared resid	1255.164	S.E. of regression	3.131448
R-squared	0.062313	Adjusted R-squared	-0.018270
F(11, 128)	1.684748	P-value(F)	0.083831
Log-likelihood	-352.1879	Akaike criterion	728.3758
Schwarz criterion	763.6755	Hannan-Quinn	742.7206
rho	0.588832	Durbin-Watson	0.822593

Source: own creation

The monthly effects for the ROL / EUR exchange rates from January 1999 to September 2009 are presented in the Table 5 from which significant monthly effects for September and October result.

Table 5. Monthly effects for ROL/EUR from January 1999 to September 2009 in a simple model

$$REUR_t = \sum_{i=1}^{12} a_i * d_{i_t} + u_t$$

Variable	Coefficient	Std. Error	t-ratio	p-value
djan	2.43866	1.07312	2.2725	0.02490
dfeb	-0.547504	0.658745	-0.8311	0.40761
dmar	1.19852	1.05675	1.1342	0.25907
dapr	0.403064	0.565048	0.7133	0.47708
dmay	0.649171	0.604497	1.0739	0.28509
djun	1.21239	0.794743	1.5255	0.12985
djul	-0.0293818	0.59449	-0.0494	0.96067
daug	0.624084	0.778824	0.8013	0.42459
dsep	1.29251	0.401408	3.2199	0.00166***
doct	1.58612	0.491951	3.2241	0.00164***
dnov	1.05854	0.666747	1.5876	0.11509
ddec	1.31103	0.907528	1.4446	0.15126

Mean dependent var	0.912304	S.D. dependent var	2.331058
Sum squared resid	619.2492	S.E. of regression	2.310488
R-squared	0.102663	Adjusted R-squared	0.017571
F(11, 116)	3.095475	P-value(F)	0.001110
Log-likelihood	-282.5187	Akaike criterion	589.0374
Schwarz criterion	623.2618	Hannan-Quinn	602.9429
rho	0.499233	Durbin-Watson	0.964994

Source: own creation

3.3. Monthly effects for the autoregressive model

In the Table 6 there are presented the monthly effects of ROL / USD exchange rates obtained in an autoregressive model from January 1991 to December 1997. Significant monthly effects for six months of the year result.

Table 6. Monthly effects for ROL/USD from January 1991 to December 1997 in an autoregressive model

$$RUSD_t = \sum_{i=1}^{12} b_i * d_{i_t} + c * RUSD_{t-1} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	10.5	5.03075	2.0872	0.04057**
dfeb	9.80749	5.57727	1.7585	0.08310*
dmar	5.13618	2.4354	2.1090	0.03858**
dapr	8.4624	7.71053	1.0975	0.27623
dmay	2.8518	2.43568	1.1708	0.24569
djun	4.82156	2.43999	1.9761	0.05215*
djul	6.80814	4.37335	1.5567	0.12411
daug	2.90973	1.85554	1.5681	0.12143
dsep	3.21535	1.37137	2.3446	0.02193**
doct	3.93217	1.89206	2.0782	0.04141**
dnov	21.0013	18.1944	1.1543	0.25237
ddec	2.62957	1.95894	1.3423	0.18389
RUSD_1	-0.018846	0.200671	-0.0939	0.92545

Mean dependent var	6.634619	S.D. dependent var	15.28205
Sum squared resid	16789.61	S.E. of regression	15.59897
R-squared	0.112450	Adjusted R-squared	-0.041906
F(12, 69)	3.598469	P-value(F)	0.000345
Log-likelihood	-334.5466	Akaike criterion	695.0932
Schwarz criterion	726.3806	Hannan-Quinn	707.6546
rho	-0.001991	Durbin-Watson	2.003836

Source: own creation

The monthly effects of ROL / USD exchange rates derived from an autoregressive model from January 1998 to September 2009 are presented in the Table 7, indicating significant monthly effects in January, June and September.

Table 7. Monthly effects for ROL/USD from January 1998 to September 2009 in an autoregressive model

$$RUSD_t = \sum_{i=1}^{12} b_i * d_{it} + c * RUSD_{t-1} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	2.28393	1.1173	2.0441	0.04302**
dfeb	-0.655514	0.62206	-1.0538	0.29400
dmar	0.618603	0.999993	0.6186	0.53729
dapr	0.118846	0.656184	0.1811	0.85657
dmay	-0.324118	0.693645	-0.4673	0.64112
djun	0.950313	0.497551	1.9100	0.05841*
djul	-0.890956	0.550654	-1.6180	0.10816
daug	1.05248	0.894413	1.1767	0.24152
dsep	1.20491	0.513756	2.3453	0.02057**
doct	0.752258	0.930629	0.8083	0.42042
dnov	0.381507	0.969847	0.3934	0.69471
ddec	-1.06876	0.851091	-1.2558	0.21153
RUSD_1	0.589303	0.0997313	5.9089	0.00001***

Mean dependent var	0.908767	S.D. dependent var	3.111248
Sum squared resid	821.6483	S.E. of regression	2.553628
R-squared	0.384911	Adjusted R-squared	0.326331
F(12, 126)	6.327789	P-value(F)	1.12e-08
Log-likelihood	-320.7227	Akaike criterion	667.4455
Schwarz criterion	705.5936	Hannan-Quinn	682.9479
rho	0.034527	Durbin-Watson	1.925820

Source: own creation

In the Table 8 there are presented the monthly effects of ROL / EUR exchange rate from January 1999 to September 2009 obtained from an autoregressive model, indicating significant monthly effects in five months of the year.

Table 8. Monthly effects for ROL/EUR from January 1999 to September 2009 in an autoregressive model

$$REUR_t = \sum_{i=1}^{12} b_i * d_{i_t} + c * REUR_{t-1} + u_t$$

Variable	Coefficient	Std. Error	t-ratio	p-value
djan	1.78408	0.731899	2.4376	0.01633**
dfeb	-2.24076	0.443378	-5.0538	0.00001***
dmar	1.47188	0.875622	1.6810	0.09551*
dapr	-0.195339	0.528239	-0.3698	0.71222
dmay	0.447927	0.484974	0.9236	0.35764
djun	0.888268	0.82032	1.0828	0.28117
djul	-0.634712	0.529254	-1.1993	0.23291
daug	0.638754	0.78363	0.8151	0.41670
dsep	0.980909	0.415042	2.3634	0.01980**
doct	0.900345	0.539799	1.6679	0.09808*
dnov	0.266616	0.690838	0.3859	0.70027
ddec	0.782514	0.755417	1.0359	0.30245
REUR_1	0.499286	0.100323	4.9768	0.00001***

Mean dependent var	0.886344	S.D. dependent var	2.321640
Sum squared resid	440.1699	S.E. of regression	1.964978
R-squared	0.351873	Adjusted R-squared	0.283649
F(12, 114)	5.619611	P-value(F)	1.68e-07
Log-likelihood	-259.1340	Akaike criterion	544.2680
Schwarz criterion	581.2425	Hannan-Quinn	559.2903
rho	-0.062543	Durbin-Watson	2.009944

Source: own creation

3.4. Monthly effects in the exchange rates volatility

In the Table 9 there are presented the monthly effects if the first differences of CVUSD, obtained in an autoregressive model from January 1998 to December 2004. The F test indicates the model has a poor adequacy, ignoring that we may identify a significant June effect.

Table 9. Monthly effects for ROL/USD volatility from January 1998 to December 2004 in an autoregressive model

$$d\_CVUSD_t = \sum_{i=1}^{12} \alpha_i * d_{it} + \beta * d\_CVUSD_{t-1} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	-0.149994	0.218675	-0.6859	0.49506
dfeb	0.106202	0.0908348	1.1692	0.24636
dmar	0.793027	0.57892	1.3698	0.17518
dapr	-0.59916	0.492206	-1.2173	0.22764
dmay	-0.305797	0.310497	-0.9849	0.32813
djun	-0.304035	0.156067	-1.9481	0.05547*
djul	-0.0821137	0.191549	-0.4287	0.66949
daug	0.224472	0.145234	1.5456	0.12678
dsep	-0.0191179	0.142585	-0.1341	0.89373
doct	0.0674666	0.184382	0.3659	0.71555
dnov	0.376559	0.329408	1.1431	0.25693
ddec	-0.0397356	0.282373	-0.1407	0.88850
d_CVUSD_{t-1}	-0.468677	0.317654	-1.4754	0.14465

Mean dependent var	-0.000506	S.D. dependent var	0.842076
Sum squared resid	35.26282	S.E. of regression	0.714881
R-squared	0.386055	Adjusted R-squared	0.279282
F(13, 69)	1.512937	P-value(F)	0.134956
Log-likelihood	-81.75346	Akaike criterion	189.5069
Schwarz criterion	220.7943	Hannan-Quinn	202.0683
rho	-0.125205	Durbin-Watson	2.232658

Source: own creation

The monthly effects of d_CVUSD derived from an autoregressive model from January 2005 to September 2009 are presented in the Table 10, indicating a significant May effect.

Table 10. Monthly effects for ROL/USD volatility from January 2004 to September 2009 in an autoregressive model

$$d\_CVUSD_t = \sum_{i=1}^{12} \alpha_i * d_{it} + \beta * d\_CVUSD_{t-1} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	0.05159	0.383964	0.1344	0.89373
dfeb	-0.463897	0.67236	-0.6900	0.49385
dmar	0.162811	0.196511	0.8285	0.41186
dapr	0.0144892	0.458628	0.0316	0.97494
dmay	-0.519684	0.172434	-3.0138	0.00427***
djun	0.00491891	0.35362	0.0139	0.98896
djul	-0.200716	0.158579	-1.2657	0.21228
daug	0.466686	0.521224	0.8954	0.37546
dsep	0.126091	0.225849	0.5583	0.57947
doct	-0.272655	0.461987	-0.5902	0.55809
dnov	0.326619	0.482784	0.6765	0.50224
ddec	0.159904	0.364137	0.4391	0.66272
d_CVUSD_{t-1}	-0.52803	0.126308	-4.1805	0.00014***

Mean dependent var	0.000151	S.D. dependent var	0.877525
Sum squared resid	25.77338	S.E. of regression	0.765349
R-squared	0.402325	Adjusted R-squared	0.239323
F(13, 44)	3.239090	P-value(F)	0.001742
Log-likelihood	-58.25879	Akaike criterion	142.5176
Schwarz criterion	169.0772	Hannan-Quinn	152.8396
rho	-0.192724	Durbin's h	-4.417340

In the Table 11 there are presented the monthly effects of the d_CVEUR in an autoregressive model from January 1999 to December 2004, indicating a lack of monthly seasonality.

Table 11. Monthly effects for ROL/EUR volatility from January 1999 to December 2004 in an autoregressive model

$$CVEUR_t = \sum_{i=1}^{12} \alpha_i * d_{it} + \beta * CVEUR_{t-1} + u_t$$

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
djan	-0.042574	0.382163	-0.1114	0.91169
dfeb	-0.264436	0.243017	-1.0881	0.28111
dmar	0.738598	0.836401	0.8831	0.38091
dapr	-0.429418	0.565968	-0.7587	0.45114
dmay	-0.144954	0.459065	-0.3158	0.75334
djun	-0.0933063	0.178064	-0.5240	0.60231
djul	0.202301	0.382872	0.5284	0.59929
daug	-0.201588	0.226216	-0.8911	0.37660
dsep	0.0517125	0.270542	0.1911	0.84909
doct	-0.239831	0.186164	-1.2883	0.20286
dnov	-0.00152802	0.323706	-0.0047	0.99625
ddec	0.450125	0.3021	1.4900	0.14174
d_CVEUR_{t-1}	-0.470678	0.314984	-1.4943	0.14061

Mean dependent var	0.002922	S.D. dependent var	0.993445
Sum squared resid	44.08408	S.E. of regression	0.879434
R-squared	0.352647	Adjusted R-squared	0.216363
F(13, 57)	1.205778	P-value(F)	0.299282
Log-likelihood	-83.14182	Akaike criterion	192.2836
Schwarz criterion	221.5141	Hannan-Quinn	203.8943
rho	-0.184427	Durbin-Watson	2.056205

Source: own creation

The monthly effects results for d_CVEUR obtained from an autoregressive model in the period January 2005 – September 2009 are presented in Table 12. They indicate a significant March effect but the model wasn't validated by F test.

Table 12. Monthly effects for ROL/EUR volatility from January 2005 to September 2009
in an autoregressive model

$$CVEUR_t = \sum_{i=1}^{12} \alpha_i * d_{it} + \beta * CVEUR_{t-1} + u_t$$

Variable	Coefficient	Std. Error	t-ratio	p-value
djan	0.118404	0.179887	0.6582	0.51383
dfeb	-0.228863	0.398898	-0.5737	0.56907
dmar	-0.508421	0.23073	-2.2035	0.03284**
dapr	0.035796	0.338221	0.1058	0.91619
dmay	-0.116115	0.217752	-0.5332	0.59655
djun	0.0685377	0.286699	0.2391	0.81217
djul	-0.050834	0.260621	-0.1950	0.84625
daug	0.163609	0.329078	0.4972	0.62154
dsep	0.226031	0.207657	1.0885	0.28231
doct	0.0958658	0.592274	0.1619	0.87216
dnov	0.245661	0.630942	0.3894	0.69889
ddec	-0.347009	0.299709	-1.1578	0.25318
d_CVEUR$_{t-1}$	-0.460536	0.160002	-2.8783	0.00615***

Mean dependent var	-0.015417	S.D. dependent var	0.688209
Sum squared resid	18.10583	S.E. of regression	0.641480
R-squared	0.317711	Adjusted R-squared	0.131632
F(13, 44)	1.522416	P-value(F)	0.147504
Log-likelihood	-48.19520	Akaike criterion	122.3904
Schwarz criterion	148.9501	Hannan-Quinn	132.7124
rho	-0.166113	Durbin-Watson	2.329403

Source: own creation

4. Conclusions

In this paper we approached the monthly seasonality exhibited in the last decades by the Romanian foreign exchange market. We began with the investigation of the monthly effects on the exchange rates returns in the transition and the consolidation stages. For the transition stages we found, by simple and autoregressive models, seasonal effect of the ROL/USD exchange rates returns for seven months: January, February, March, June, August, September and October. There are different explana-

tions for these seasonal effects. Quite often, during the winter months January and February, NBR had to allow, because of the increased imports of oil and gas, sharp devaluation of the national currency. In general, in the cold season the justified increase of the prices with major impact for the population was avoided. This measure was postponed for the beginning of spring and it could be considered as responsible for the March effect. In June and August there were also operated quite often increases of some administrated prices. Because July and August are holiday months for the Parliament and for many members of the Government, many important economic and political decisions were postponed for the next months, generating significant changes in the exchange rates returns in September and October.

In the consolidation stage we found some changes in the seasonality of the foreign exchange market. There are significant differences between ROL/USD and ROL/EUR and between the simple and the autoregressive model. In this period of time the US dollar was replaced by euro as the main instrument for savings and transactions. For all the models we obtained a significant September effect explained by the July and August holidays of the decision factors. For some models we found monthly effect in January, February and March, which could be connected with the circumstances of the cold season.

We did not identify many changes in the exchange rates monthly volatilities after the foreign capital inflows liberalization and the inflation targeting adoption, except a May effect for the ROL / EUR exchange rate. This situation could be explained by the facts that massive foreign capital yet occurred before the liberalization and NBR still has significant interventions on the foreign exchange market.

References

Agrawal, A. – Tandon, K. 1994: Anomalies or illusions? Evidence from stock markets in eighteen countries. *Journal of International Money and Finance*, 13, 1, 83–106. p.

Ariel, R. A. 1987: A monthly effect in stock returns. *Journal of Financial Economics*, 17, 161–174. p.

Bekaert, G. - Campbell R. H. 2002: Research in Emerging Markets Finance: Looking to the Future. *Emerging Markets Review* 3, 429–448. p.

Chinn, M. D. – Ito, H. 2002: *Capital Account Liberalization, Institutions and Financial Development: Cross Country Evidence.* NBER working paper no. 8967.

Dornbusch, R. 1976: Expectations and Exchange Rate Dynamics. *Journal of Political Economy*, Vol. 84 Issue 6, 1161-1176. p.

Elysiani, E. - Perera, P. – Puri, T. 1996: Market efficiency and calendar anomalies in emerging capital markets: Evidence from the Colombo stock exchange. *Journal of International Financial Markets*, Institutions and Money, 6, 4, 59–77. p.

Frenkel, J. A. 1981: *Flexible* Exchange Rates, Prices and the Role of 'News': Lessons from the 1970s. *Journal of Political Economy*, Vol. 89, 665-705. p.

Froot, K.A. – Ito, T. 1989: On the Consistency of Short-Run and Long-Run Exchange Rate Expectations. *Journal of International Money and Finance*, Vol. 8 Issue 4, 487-510. p.

Froot, K.A. – Thaler, R.H. 1990: Anomalies: Foreign Exchange. *Journal of Economic Perspectives*, 4, 179-192. p.

Froot, K.A. – Thaler, R.H. 1990: Foreign Exchange. *Journal of Economic Perspectives*, Vol. 4 Issue 3, Summer, 179-192. p.

Gultekin, M. N. - Gultekin N. B. 1983: Stock Market Seasonality: International Evidence. *Journal of Financial Economics*, 12, 469-481. p.

Hansen, P. R. – Lunde, A. 2003: *Testing the Significance of Calendar Effects*, Brown University, Department of Economics. Working Paper No. 2003-03, January.

Isard, P. 1995: *Exchange Rate Economics*, Cambridge University Press, Cambridge.

MacDonald, R. – Taylor, M.P. 1992: *Exchange Rate Economics: A Survey.* IMF Staff Papers, Vol. 39 Issue 1, March, 1-57 p.

Mark, N. 1995: Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. *American Economic Review* 85, 201-218. p.

Mussa, M. 1979: *Empirical Regularities in the Behavior of Exchange Rates and Theories of the Foreign Exchange Market*, In Karl Brunner and Allan H. Meltzer, (eds.): Carnegie-Rochester Conference Series on Public Policy, Vol11, (Policies for Employment, Prices, and Exchange Rates), North Holland.

Officer, R. R. 1975: Seasonality in Australian Capital Markets: Market Efficiency and Empirical Issues. *Journal of Financial Economics*, Vol. 2, 29-51. p.

Rozeff, M.S. - Kinney, W.R. 1976: Capital Market Seasonality: The Case of Stock Returns. *Journal of Financial Economics*, Vol. 3, 379-402. p.

Schwert, G. W. 2001: *Anomalies and market efficiency*, in Forthcoming in Handbook of the Economics of Finance, ed. by M. H. George Constantinides, and R. Stulz. Elsevier Science B.V.

Takagi, S. 1991: *Exchange Rate Expectations: A Survey of Survey Studies.* IMF Staff Papers. 38, 1, March: 156-183. p.

Tang, G.Y.N. 1998: *Monthly Pattern and Portfolio Effect on Higher Moments of Stock Returns: Empirical Evidence from Hong Kong*, Asia-Pacific Financial Markets, Vol. 5, 275-307. p.

Wachtel, S. B. 1942: Certain observations in seasonal movements in stock prices. *Journal of Business*, 15, 184–193. p.

Analysis of the Dynamic Relation between the Currency Rates and the Interest Rates from Romania and Euro Area before and after the Financial Crisis

Razvan Stefanescu[1] *– Ramona Dumitriu*[2] *- Costel Nistor*[3]

This paper examines the changes induced by the actual financial crisis in the dynamic relation between the currency rates and the differentials of the interest rates from Romania and euro area. In the framework of the Uncovered Interest Rate Parity hypothesis we apply the Vector Autoregressive methodology for daily values of the currency rates and the interest rates during the crisis. We compare the results obtained with a similar analysis for a period of time before the crisis began and we find significant differences.

Keywords: Uncovered Interest Rates Parity, Vector Autoregressive Model, Financial Crisis, Romanian Foreign Exchange Market

1. Introduction

The relation between the exchange rates and the interest rates differentials is among the topics of the international finance. It was approached in many theoretical models, most known of them being the uncovered interest rates parity (UIRP), which stipulates that interest rates differentials between two countries could be considered as an estimator for the expected change in the exchange rate.

The basic equation of UIRP is:

$$(1 + i_t^h) / (1 + i_t^f) = E_t (S_{t+T}) / S_t \quad (1)$$

where:

[1] *Razvan Stefanescu*, PhD, lecturer, *„Dunarea de Jos"* University *Galati*, Faculty of Economics, Romania
[2] *Ramona Dumitriu*, PhD student, lecturer, *„Dunarea de Jos"* University *Galati*, Faculty of Economics, Romania
[3] *Costel Nistor,* PhD, assoc. prof., *„Dunarea de Jos"* University *Galati*, Faculty of Economics, Romania

- i_t^h is the return at time t on a domestic asset which arrived to maturity at the moment t + T;
- i_t^f is the return on a comparable foreign asset;
- S is the nominal exchange rate expressed as the price, in the domestic currency, of a unit of the foreign currency;
- $E_t (S_{t+T})$ is the expected value of the exchange rate at time t+T, based on the information available at time t.

On the assumption of rational expectations we may consider that on average:

$$E_t (S_{t+T}) = S_{t+T} \quad (2)$$

Taking natural logarithms we obtain:

$$\ln (S_{t+T}) - \ln (St) \approx \ln i_t^h - \ln i_t^f \quad (3)$$

The fundamental equation of UIRP can be expressed as:

$$\ln (S_{T+t}) - \ln (S_t) = \alpha + \beta [\ln (i_t^h) - \ln (i_t^f)] + \varepsilon_t \quad (4)$$

Meredith and Chinn (1998) formulated the unbiasedness hypothesis of UIRP imposing the conditions $\alpha = 0$ and $\beta = 1$.

According to UIRP, the national currency of a country with high interest rates tends to depreciate. However, empirical studies revealed numerous situations in which UIRP wasn't validated. Froot and Thaler (1990) surveyed 75 studies on UIRP and they found the unbiasedness hypothesis of UIRP was confirmed only in few cases. Empirical researches identified some particularities of UIRP. Meredith and Chinn (1998) found that in general UIRP was validated on long term but invalidated on short term. McCallum (1994) explained such a situation by the role of short term interest rates as monetary policy instruments. When an economy is affected by a negative shock the national currency depreciates. The monetary authorities react to this shock by raising the interest rates. After the shock is dissipated the national currency appreciates and the monetary authorities lower the interest rates. In the specialised literature there were revealed some particularities of the relation between the exchange rates and the interest rates in the emerging markets or during the financial crisis. Francis et al. (2002) found the emerging market liberalization provoked mixed effects on the uncovered interest rate parity for the countries from Latin America and Asia. Flood and Rose (2001) proved that for countries in crisis, when the volatility of the exchange rates and the interest rates increase and the monetary authorities have to deal with speculative pressure, UIRP may work differently.

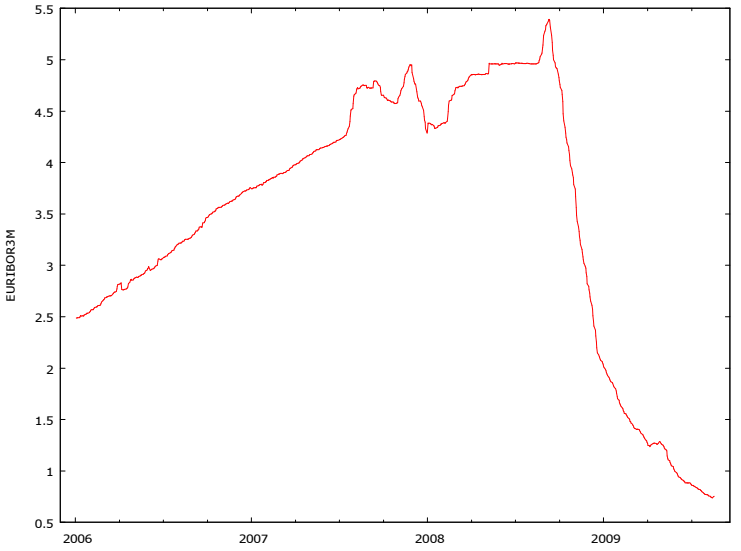
In this paper we approach the relation between the currency rates and the interest rates from Romania and from the Euro Area before and after the financial crisis. In the last years these variables were significantly affected by the global crisis. The nominal exchange rates RON / EUR which reflect the price of a unit of euro in the Romanian national currency decreased from 2006 to the first half of 2007. During the second half of 2007 they experienced a significant depreciation followed by a relative stabilization in the first half of 2008. In the second half, in the global crisis context a sharp depreciation occurred and it was followed by a more calm evolution during the first three quarters of 2009 (Figure 1).

Figure 1. Evolution of nominal exchange rate RON/EUR from January 2006 to September 2009

Source: own creation

The interest rates in the Euro Area experienced an almost regular growth at the beginning of 2006 to the end of 2007. After a fluctuant evolution, in the first three quarters of 2008 the European Central Bank cut the interest rates in order to stimulate the economic activity, affected by the global crisis (Figure 2).

Figure 2. Evolution of EURIBOR 3M from January 2006 to September 2009



Source: own creation

In Romania from the beginning of 2006 to the end of 2007 the interest rates were relatively stable and much higher than in the Euro Area. During the first half of 2008 the National Bank of Romania (NBR) slowly increased the interest rates. In the third quarter of 2008, in the context of the global crisis, speculative attacks were directed against the national currency. NBR reacted by increasing for a short period of time the interbank interest rates to almost 50 percent per annum. After that, in order to stimulate the economy, the interest rates were reduced (Figure 3).

In our analysis we use sub-samples of data from two periods of time: before and after the global crisis affected the exchange rates and the interest rates. We test the unbiasedness hypothesis of UIRP using simple regressions. Then we analyze, in a Vector Autoregressive (VAR) framework, the interactions between the interest rates differentials and the exchange rates.

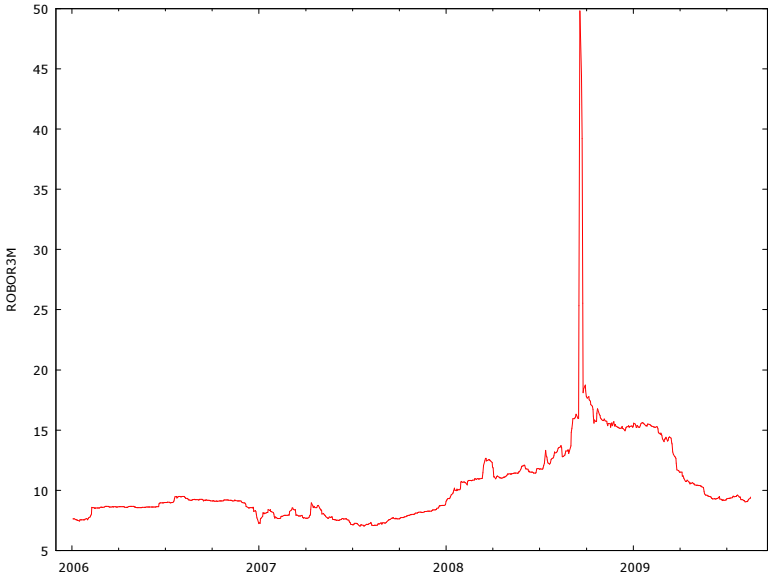
The rest of the paper is organized as follows. In the second part we describe the data and the methodology used in our investigation. In the third part we present the empirical results and in the fourth part we conclude.

2. Data and Methodology

In our investigation we employ daily values of the exchange rate RON / EUR and the interbank offered rates from the Euro Area and from Romania. We use two variables:

- neer, as natural logarithm of the nominal exchange rate RON / EUR;
- dif, as differences between the natural logarithms of ROBOR 3M and natural logarithms of EURIBOR 3 M;
- d_neer, as first differences of neer;
- d_dif, as first differences of dif.

Figure 3. Evolution of ROBOR 3M from January 2006 to September 2009



Source: own creation

The data are from January 2006 to September 2009. We divide this sample in two sub-samples:

- first sub-sample, from the 3rd of January 2006 to the 23rd of July 2008, corresponding to a relative tranquil period;
- second, from the 24th of July 2008 to the 30th of September 2009, when the impact of the global crisis on the exchange rates and on the interest rates was significant.

In the Table 1 there are presented the descriptive statistics of the two variables for both sub-samples. We may notice that there are significant differences between the two periods of time.

The stationarity of the time series used in our investigation will be analysed by the Augmented Dickey – Fuller Test and by the test proposed by Saikkonen and Lütkepohl (2002) and Lanne et al. (2001), which allow us to take into account the eventual structural breaks. The deterministic terms of the equations will be established based on the graphical representation of time series, while the number of the lagged differences will be chosen based on Akaike Information Criteria.

The unbiasedness hypothesis of UIRP will be tested by simple regressions. We shall study the cointegration between the two variables using Johansen Test and a test proposed by Saikkonen and Lütkepohl (2000). Then we shall analyze the interactions between the interest rates differentials and the exchange rates in a VAR framework in which the number of lagged differences will be chosen based on the Schwartz Bayesian Criteria. Finally we shall test the Granger causality between the two variables.

Table 1. Descriptive statistics of neer and dif for the two sub-samples

Indicator	**neer**		**dif**	
	Sub-sample 1	**Sub-sample 2**	**Sub-sample 1**	**Sub-sample 2**
Mean	1.24616	1.39792	0.827150	1.88476
Median	1.25623	1.43561	0.838918	2.12955
Minimum	1.13501	1.25857	0.399489	0.896964
Maximum	1.32694	1.46156	1.20204	2.52759
Std. Dev.	0.0435187	0.0649417	0.219542	0.514667
C.V.	0.0349223	0.0464558	0.265419	0.273068
Skewness	-0.526704	-0.908079	-0.175695	-0.704678
Ex. kurtosis	-0.303264	-0.747136	-1.07854	-0.970115
Jarque - Bera test for normality	32.8447	46.2797	35.1703	35.1289
p-value for Jarque - Bera test	0.0001	0.0001	0.0001	0.0001

Source: own creation

3. Empirical Results

3.1. Analysis of the first sub – sample

Based on the graphical representation of the four variables we chose the constant and the trend as deterministic terms for neer and dif and only the constant for their first differences. The results of Augmented Dickey – Fuller Tests are presented in the Table 2. They indicate that both variables are not stationary in level but stationary in their first differences.

Table 2. Augmented Dickey-Fuller Test for the observations from the first sub-sample

Variable	**Deterministic terms**	**Lagged differences**	**Test statistics**
neer	Constant and trend	3	-1.6734
d_neer	Constant and no trend	2	-15.7165***
dif	Constant and trend	2	-0.5540
d_ dif	Constant and no trend	1	-13.2585***

Note: The number of the lagged differences was chosen based on Akaike Information Criteria.
Source: own creation

Table 3 presents the results of the unit root tests with structural breaks. Again the both variables proved not to be stationary in level but stationary in their first differences.

We test the unbiasedness hypothesis of UIRP by a simple regression between d_neer and d_dif. The results indicate the rejection of this hypothesis.

Table 3. Unit root tests with structural breaks for the observations from the first sub-sample

Variable	Deterministic terms	Shift Function	Break Date	Lagged differences	Test statistics
neer	Constant and trend	Impulse dummy	503	3	-1.1877
		Shift dummy	526	3	-1.0668
d_neer	Constant and no trend	Impulse dummy	503	2	-15.5430***
		Shift dummy	502	2	-3.5980***
dif	Constant and trend	Impulse dummy	332	2	-1.1327
		Shift dummy	332	2	-1.0282
d_ dif	Constant and no trend	Impulse dummy	332	1	-13.0031***
		Shift dummy	324	1	-13.4044***

Note: The number of the lagged differences was chosen based on the Akaike Information Criteria.
Source: own creation

In the Table 4 there are presented the results of Johansen tests that indicate the lack of cointegration between neer and dif.

Table 4. Johansen cointegration tests for the first sub sample (Case 3: Unrestricted constant)

Rank	Eigenvalue	Trace test	p-value	Lmax test	p-value
0	0.038899	31.320	0.0001	25.988	0.0003
1	0.0081071	5.3318	0.0209	5.3318	0.0209

Source: own creation

The results of the Saikkonen and Lütkepohl tests, presented in the Table 5, indicate again the lack of cointegration between neer and dif.

Table 5. Saikkonen and Lütkepohl cointegration tests for the first sub sample (intercept included)

Rank	LR	p-value
0	18.92	0.0028
1	0.65	0.4745

Source: own creation

Since dif and neer are integrated at order 1 but not cointegrated we study their interactions of their first differences in a VAR framework. The results, presented in the Table 6, indicate a low interaction between the two variables.

Table 6. VAR system for the first sub-sample

Equation 1: d_neer

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	-8.01895e-05	0.000176544	-0.4542	0.64982
d_neer_1	0.147616	0.0580516	2.5428	0.01123**
d_neer_2	-0.0496225	0.0501503	-0.9895	0.32280
d_dif_1	-0.0303478	0.0162896	-1.8630	0.06291*
d_dif_2	-0.000568234	0.0154784	-0.0367	0.97073

F(4, 649) = 2.522382; P-value(F) = 0.039970

Equation 2: d_dif

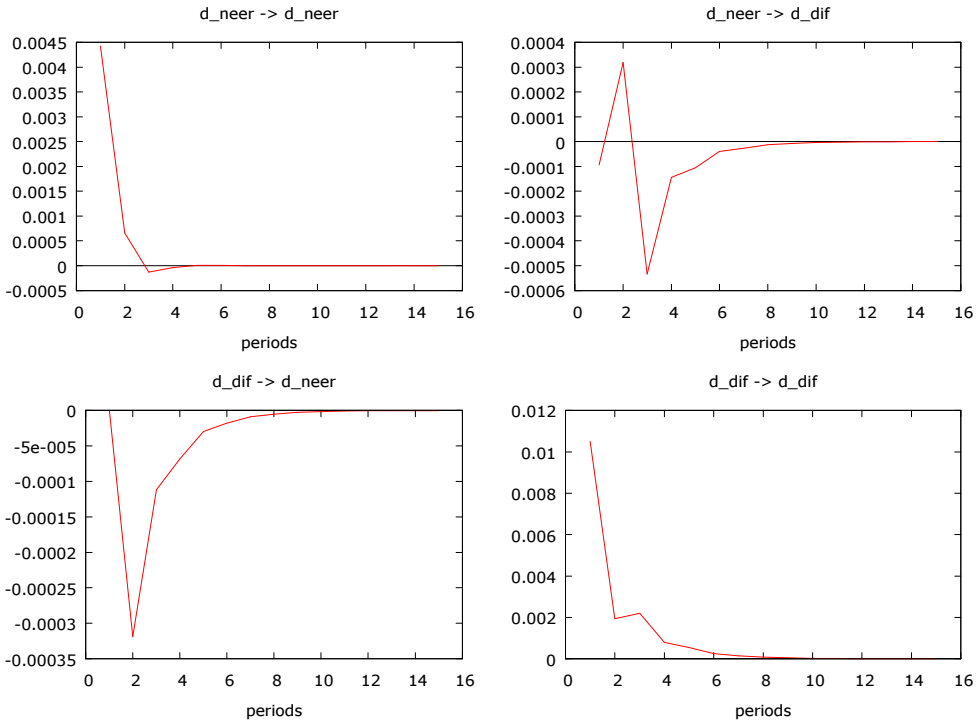
Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	-0.000217414	0.000424883	-0.5117	0.60903
d_neer_1	0.076577	0.0841299	0.9102	0.36304
d_neer_2	-0.142957	0.103215	-1.3850	0.16651
d_dif_1	0.184186	0.075134	2.4514	0.01449**
d_dif_2	0.177109	0.0603779	2.9333	0.00347***

F(4, 649) =3.531997; p-value(F) = 0.007302

Source: own creation

The impulse – response analysis indicate that a shock from d_dif leads to a fall of d_neer for a short period of time, but finally it is back to the initial level. Instead, a shock of d_neer provokes a raise of d_dif which, after a fluctuant evolution, arrives to the initial level (Fig. 4).

Figure 4. VAR impulse-response analysis for the first sub-sample



Source: own creation

The Granger causality tests indicate no causality between d_neer and d_dif (Table 7).

Table 7. Tests of Granger causality between the variables for the first sub-sample

Null hypothesis	**F-statistic**	**P-value**	**Causal inference**
H_0: "d_neer" do not Granger - cause "d_dif"	0.2548	0.2548	"d_neer" do not Granger-cause "d_dif"
Null hypothesis	**F-statistic**	**P-value**	**Causal inference**
H_0: "d_dif" do not Granger-cause "d_neer"	2.0006	0.1357	"d_dif" do not Granger-cause "d_neer"

Source: own creation

3.2. Analysis for the second sub-sample

The graphical representation of the four variables suggests that we may use the constant and the trend as deterministic terms for neer and dif and only constant for their first differences. The results of Augmented Dickey – Fuller Tests, presented in the Table 8, indicate that both variables are not stationary in level but stationary in their first differences.

Table 8. Augmented Dickey-Fuller Test for the observations from the second sub-sample

Variable	**Deterministic terms**	**Lagged differences**	**Test statistics**
neer	Constant and trend	14	-1.4043
d_neer	Constant and no trend	13	-4.8567***
dif	Constant and trend	12	-1.9369
d_ dif	Constant and no trend	11	-6.0376***

Note: The number of the lagged differences was chosen based on Akaike Information Criteria.
Source: own creation

In the Table 9 there are presented the results of the unit root tests with structural breaks. Again we found neer and dif as non stationary while their first differences are stationary.

Table 9. Unit root tests with structural breaks for the observations from the second sub-sample

Variable	Deterministic terms	Shift Function	Break Date	Lagged differences	Test statistics
neer	Constant and trend	Impulse dummy	48	8	-2.0441
		Shift dummy	49	8	-2.0695
d_neer	Constant and no trend	Impulse dummy	80	9	-5.5298***
		Shift dummy	44	9	-3.0325***
dif	Constant and trend	Impulse dummy	57	7	-2.1142
		Shift dummy	80	7	-1.8860
d_ dif	Constant and no trend	Impulse dummy	49	12	-6.1406***
		Shift dummy	52	11	-2.6276*

Note: The number of the lagged differences was chosen based on the Akaike Information Criteria.
Source: own creation

The results of the Johansen cointegration tests, presented in the Table 10, suggest that dif and neer are not cointegrated.

Table 10. Johansen cointegration tests for the second sub sample (Case 3: Unrestricted constant)

Rank	Eigenvalue	Trace test	p-value	Lmax test	p-value
0	0.088261	30.198	0.0001	26.519	0.0002
1	0.012737	3.6791	0.0551	3.6791	0.0551

Source: own creation

In the Table 11 there are presented the results of the Saikkonen and Lütkepohl cointegration tests which confirm the lack of cointegration between dif and neer.

Table 11. Saikkonen and Lütkepohl cointegration tests for the second sub-sample (intercept included)

Rank	LR	p-value
0	6.61	0.3707
1	0.17	0.7380

Source: own creation

The two equations of the VAR model for the second sub-sample are presented in the Table 12. They suggest an interaction much significant than for the first sub sample.

Table 12. VAR system for the second sub-sample

Equation 1: d_neer

Variable	Coefficient	Std. Error	t-ratio	p-value
const	0.000694996	0.000339933	2.0445	0.04184**
d_neer_1	0.224282	0.0963506	2.3278	0.02064**
d_neer_2	-0.124962	0.0763382	-1.6370	0.10276
d_dif_1	-0.0145743	0.0090842	-1.6044	0.10976
d_dif_2	-0.0134358	0.0111667	-1.2032	0.22991

F(4, 280) = 4.444529; P-value(F) = 0.001692

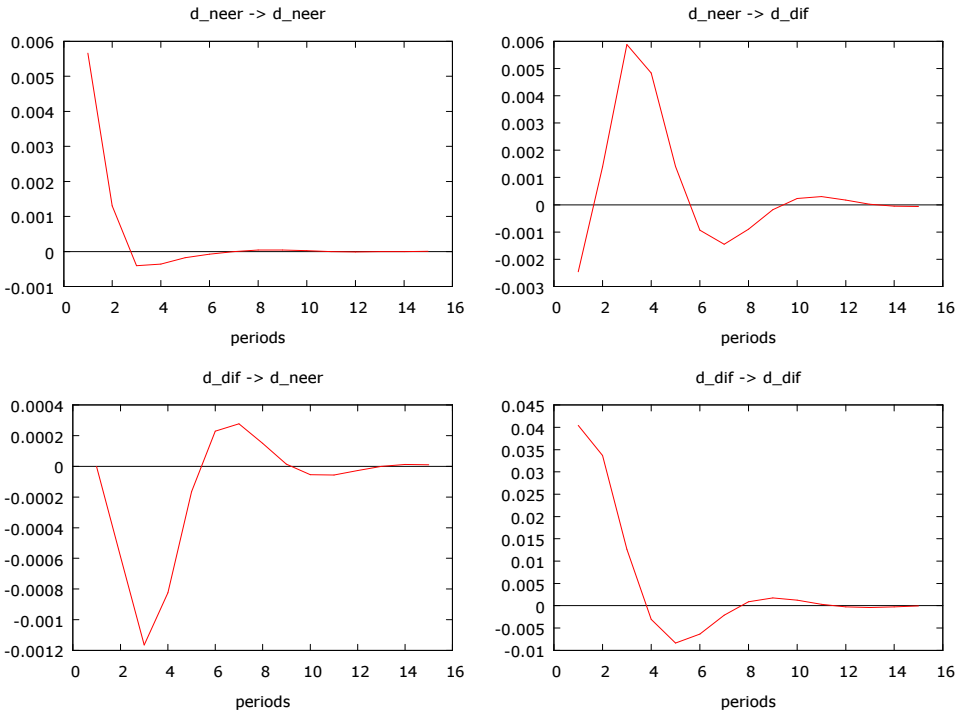
Equation 2: d_dif2

Variable	Coefficient	Std. Error	t-ratio	p-value
const	0.00252239	0.00298876	0.8440	0.39941
d_neer_1	0.608058	0.509638	1.1931	0.23383
d_neer_2	0.53222	0.531244	1.0018	0.31729
d_dif_1	0.833495	0.123193	6.7658	0.00001***
d_dif_2	-0.372968	0.131799	-2.8298	0.00499***

F(4, 280)= 12.10619; P-value(F) = 0.00001
Source: own creation

The impulse – response analysis indicates that a unit shock of d_dif provokes a fall of d_neer which, after a fluctuant evolution, is back to the initial level. A unit shock of d_ner leads to a fall of d_dif, followed by a fluctuant evolution, until the stabilization to the initial level (Figure 5).

Figure 5. VAR impulse-response analysis for the second sub-sample



Source: own creation

The Granger causality tests for the second sub-sample indicate a unidirectional relation: d_dif Granger causes d_neer but d_neer does not Granger cause d_dif (Table 13).

Table 13. Tests of Granger causality between the variables for the second sub-sample

Null hypothesis	**F-statistic**	**P-value**	**Causal inference**
H_0: "d_neer" do not Granger-cause "d_dif"	2.2272	0.1088	"d_neer" do not Granger-cause "d_dif"
H_0: "d_dif" do not Granger-cause "d_neer"	8.7425	0.0002	"d_dif" Granger-cause "d_neer"

Source: own creation

4. Conclusions

In this paper we analysed the impact of the financial crisis on the relation between the exchange rates and the interest rates differential. We used daily data of RON / EUR exchange rates and three months interest rates from Romania and the Euro area.

We found no evidence in favor of UIRP. By contrary, the VAR analysis indicated that a raise of the interest rates differential led to an appreciation of the national currency.

From the VAR impulse – response analysis it resulted the financial crisis affected significantly the relation between the interest rates differentials and the exchange rates. This conclusion was confirmed by the Granger causality tests. For the first sub-sample we found no causality among the variables but for the second sub – sample it resulted the interest rates differential Granger caused the exchange rates.

The lack of causality for the first sub – sample may be explained by the significant control of NBR of the exchange rates and the interest rates in this period of time. Instead, in the financial crisis context, NBR preferred to relax the control of the exchange rates to make easier the defense of the national currency in the case of speculators attacks.

References

Amisano, G. - Giannini, C. 1997: *Topics in Structural VAR Econometrics*, 2nd ed., Springer, Berlin.

Bekaert, G. - Harvey, C. R. 2000: Foreign Speculators in Emerging Equity Markets. *Journal of Finance,* 55, 565-613. p.

Bekaert, G. – Wei, M. – Xing, Y. 2002: Uncovered Interest Rate Parity and the Term Structure. *NBER Working Paper* No. 8795.

D'Amato, M. – Pistoresi, B. 2001: Interest Rate Spreads between Italy and Germany, 1995-1997. *Applied Financial Economics* 11 (December): 603-12. p.

Dornbusch, R. 1976: Expectations and Exchange Rate Dynamics. *Journal of Political Economy.* Vol. 84 Issue 6, December, 1161-1176. p.

Flood, R. P. – Rose, A. K. 2001: *Uncovered Interest Parity in Crisis: The Interest Rate Defence in the 1990s*. Working Paper No. 01/207, International Monetary Fund, Washington.

Francis, B. – Hasan, I. – Hunter, D. 2002: Emerging Market Liberalization and the Impact on Uncovered Interest Rate Parity. *Federal Reserve Bank of Atlanta.* Working Paper 2002/16 August.

Froot, K. A. – Ito, T. 1989: On the Consistency of Short-Run and Long-Run Exchange Rate Expectations. *Journal of International Money and Finance*, Vol. 8, Issue 4, 487-510. p.

Froot, K. A. 1990: Short Rates and Expected Asset Returns. Working Paper No. 3247. *National Bureau of Economic Research*, Cambridge, Massachusetts.

Froot, K. A. – Thaler, R.H. 1990: Foreign Exchange. *Journal of Economic Perspectives*, Vol. 4, Issue 3, Summer, 179-192. p.

Guy, M. – Chinn, M. 1998: Long Horizon Uncovered Interest Parity, Working Paper No. 6797. *National Bureau of Economic Research*, Cambridge, Massachusetts.

Isard, P. 1995: *Exchange Rate Economics*. Cambridge University Press, Cambridge.

Lanne, M. - Lütkepohl, H. - Saikkonen, P. 2001: *Test procedures for unit roots in time series with level shifts at unknown time*. Discussion paper, Humboldt - Universität Berlin.

Loopeseko, B.E. 1984: Relationships Among Exchange Rates, Intervention, and Interest Rates. *Journal of International Money and Finance*, 3, 257-277. p.

MacDonald, R. – Taylor, M.P. 1992: *Exchange Rate Economics: A Survey*. IMF Staff Papers, Vol. 39, Issue 1, March, 1-57. p.

McCallum, B. T. 1994: A Reconsideration of the Uncovered Interest Parity Relationship. *Journal of Monetary Economics*, Vol. 33, Issue 1, February, 105-132. p.

Meredith, G. - Chinn, M. 1988: Long-horizon uncovered interest rate parity, *NBER Working Paper* 6797.

Saikkonen, P. - Lütkepohl, H. 2000: Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process. *Journal of Time Series Analysis*, 21, 435-456. p.

Saikkonen, P. - Lütkepohl, H. 2002: *Testing for a unit root in a time series with a level shift at unknown time.* Econometric Theory, 18, 313-348. p.

Takagi, S. 1991: *Exchange Rate Expectations: A Survey of Survey Studies.* IMF Staff Papers. 38, 1, March, 156-183. p.

Wu, J. 1999: A Re-Examination of the Exchange Rate-Interest Differential Relationship: Evidence from Germany and Japan. *Journal of Money and Finance*, Vol. 18, Issue 2, April, 319-336. p.

Model for sequential dynamic competition between random investment portfolios and portfolios selected by collective expert opinions

Angel Marchev, Jr.[1]

"The future is something which everyone reaches at the rate of sixty minutes an hour, whatever he does, whoever he is."
--C. S. Lewis

Abstract: This paper discusses the issue of market efficiency and proposes an approach for its empirical testing. The essence of the methodology is comparing expert opinions with randomly selected portfolios and implementing all the good practices of Delphi method for conducting an expert survey. The proposed approach also shares some similarities with the idea of prediction markets. The approach is yet to be validated empirically.

Keywords: Efficient market hypothesis, Online expert opinion survey, Delphi approach, Collective intelligence
JEL: G11, G14, C83

1. Research formulation

There is a prolonged international and interdisciplinary dispute on the topic of financial market efficiency. Ever since the financial markets emerged as the favorite destination for trading risk capital, scholars begun to speculate on market predictability. Starting the twentieth century was the remarkable but non-honored thesis of Louis Bachelier (Bernstein 1998) which raised the question whether the "past, present or even the discounted future events" reflected in the market price, show relation to the price change (Bachelier 1900). After him there was hardly a single decade in which the concept of efficient market hasn't been elaborated upon (Dimson et al. 1998).

In one of the founding theories of modern investment science Sir Maurice Kendall has stated that "in series of prices … the random changes … are so large as

[1] Angel Marchev, Jr., assistant professor, University of National and World Economy, Department of Management (Sofia)

to swamp any systematic effect which may be present. The data [behaves] almost like wandering series" (Kendall 1953). Kendall shows that the probability of a given stock (or commodity) price to rise is equal to the probability to fall.

The biggest names after him to build on his ground work are Eugene Fama and Burton Malkiel – pioneers of the Random Walk Hypothesis, and the closely related Efficient-market hypothesis.

The free market is a social phenomenon which means it is subjective in nature, depending on the subjective decisions of all market participants. The random walk hypothesis states that because of the complexity of the market, the prices follow a random walk trajectory where the changes in future and the past states are independent from change in the current state so much that the price movement of a single stock is unpredictable (Malkiel 1973).

The efficient market hypothesis states that due to the information rich environment in which every investor makes decisions the current market stock price already reflects all known information defined as a set of past, present and (expectations for) future events and thus no one could utilize unique knowledge to profit from it (Fama 1965). The hypothesis was proposed in the 1960's and even if it has been partially correct then it surely must be much more so now with the modern communication and IT capabilities.

On one hand famous researchers such as Maurice Kendall, Burton Malkiel and Eugene Fama have studied the properties of financial time series with the general idea that the phenomenon financial market is principally unpredictable. The best suggestion for the investor in such information-rich environment is to invest in randomly selected wide portfolio and not to follow any analysis and forecasting. "Taken to its logical extreme," says Malkiel "[the theory] means that a blindfolded monkey throwing darts at a newspaper's financial pages could select a portfolio that would do just as well as one carefully selected by the experts" (Malkiel 1973).

Of course on the other hand the notion that there could be no methodology for predicting financial markets objects most of the theoreticians and practitioners in the field of financial investments. There are enough opposing texts (see for example Lo et al. 1999, Dorsey 2003, Lo 2004) to justify the ever-springing theoretical and empirical tests of market efficiency. Looking through an alchemist eyes the topic has become the Philosopher's stone of investment theory.

In this paper the issue of market efficiency is treated unprejudiced and neutral following the epistemological principals of empiricism. The proposed approach for market efficiency testing is as much a methodology as it is an experiment proposition. Its validation with real life data is yet to be conducted and so the author has not yet formed conclusive opinion on the topic.

2. "Dartboard contest" of Wall Street Journal

For the last several decades numerous tests of market efficiency have been conducted – both scientifically sophisticated and more wide-public oriented with the Wall-Street Journal Dartboard Contest being the most outreaching and well commented. Starting in October 1988 it had run for fourteen years while the rules underwent only slight changes. Every month each of four "professionals" selected one long or short position for the next six months. The professional portfolio competed against a portfolio of four positions selected randomly by throwing a dart. The selected security had to comply with limitations on market capitalization, average daily volume, minimal price and market listing. Dow Jones Industrial Average was used as a benchmark measuring the market return. After six months the active returns[2] of the professional and the random portfolios are compared.

By the end of the competition Dow Jones Industrial Average had an average rise of 5.6% over the time period. The professional portfolios had an average of 10.2% investment gain (4.6% active return) while the random portfolios had 3.5% average gain (-2.1% active return) (Jasen 2002).

The most critical question to the competition (which is also topical for the current research) is whether the success of the professionals was self-inclined (Rasp et al. 2003)? Have their professional publicly stated opinion inclined the investors to trade along the professional selections[3] and thus drifting the market in gaining way?

Other major drawbacks of the procedure include:

- Very limited number instances (realizations) of expert opinions.
- Very limited number (only four) of instances (realizations) of randomness against which the expert opinion are confronted for a given time period.
- There were no weights in the random selections nor there were in the expert predictions. The portfolios were not intra-structured and optimized.
- The respective contests are analyzed separately, so there is no way to approximate any reasonable conclusion within the given time period.
- The pace of the experiment was very slow – only 48 seeds per year, six months for each contest.
- Relatively small number of securities complied with the limitations

Any sensible analysis within the above-mentioned drawbacks would have to be made post-factum after decades of experimentation and still would not be soundly proven. This is the reason why they had to analyze the results with some significance only after fourteen years of competition (142 six-month contests) have past.

[2] Comparing the portfolio return with the market return
[3] Buy their long positions and sell their short positions

Although non–scientific in essence, the Dartboard competition has introduced an innovative way to test market efficiency by comparing expert predictions with random selections.

3. Delphi method for processing expert opinions

The method is developed by Project RAND in the 1960s (Jantsch 1967). It is a systematic rational method for collective expert opinion, while avoiding unwanted effects of mutual influence among the experts. The key features are that every expert opinion is anonymized by the researcher and the various expert opinions are accumulated in a certain manner so that a unified and objectified opinion of expert group is drawn. The process is known also as knowledge extraction since it makes use of the unique experience, information and insights of the expert.

Delphi method used for predictions does not require the assumption of the Bernoulli hypothesis for perseverance of the historical conditions (Bernoulli 1713). The method is used when the quantitative methods are inapplicable i.e. unstudied, highly uncertain and complex phenomena.

In its original version the method was devised for predicting the moment of occurring of a certain event. In later uses another version of the method was also developed inquiring "what would be the value of a given measurable indicator at a given future point of time".

The use of several repeating rounds of survey with feedback from the previous round achieves higher degree of consent among the experts which additionally could be measured by Kendall's coefficient of concordance. Other additions to the method may include evaluating the accuracy of each expert between rounds and weighting his/her opinion in the following round by some correction coefficient.

When applied for prediction of dynamical social phenomena such as stock markets, Delphi method could produce additional features:

- The collected data could be used for estimation of market expectations.
- If publicized, the results may vary the course of the market since the investors basically trade on their estimation of the market expectation. This would mean that simply by publicizing the results, they may become more accurate.

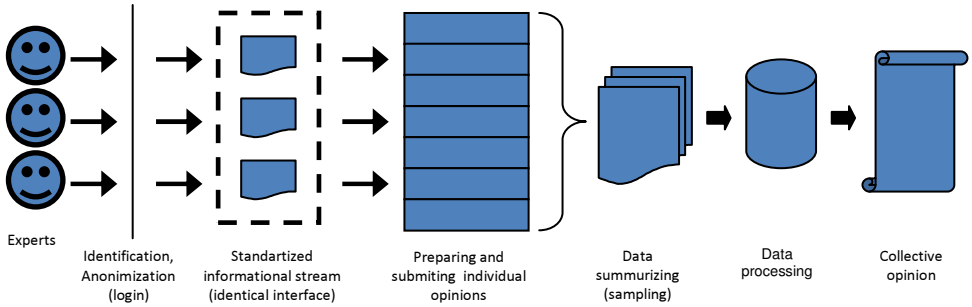
Introduction of the Internet has presented new opportunity for collecting and upgrading survey techniques (Harizanova et al. 2009). There are developments of the method in at least two directions – electronic Delphi approach and prediction markets.

The electronic Delphi approach (e-Delphi or Online Delphi) is in essence an electronic forum with non-stop, repeating anonymized survey of many experts so that there is a constant stream of new objectified opinions on the topic.

Online Delphi system introduces so many opportunities and changes to the method that it actually becomes a new methodology of system for online processing of expert opinions. Such system has a lot of advantages and solves (at least partially) most of the typical problems of Delphi approach:

- The communication with the experts is fast – sluggish communication has invariably been the biggest setback in an offline survey.
- The experts are presented with equal informational conditions when they submit their opinions.
- The results could be updated dynamically and presented instantly. An adapting mechanism for weighting the expert opinions may be incorporated dynamically.
- There is no need to follow a rhythmic schedule for the consecutive rounds of survey – every expert could log on and present opinion at own pace. The feeling of not being hustled additionally encourages the experts to participate.

Figure 1. Online system for processing expert opinions



Source: own creation

Prediction markets are a concept combining the idea of e-Delphi with stock-market-like trading of statements. Participants use virtual money to trade the statement as if it was a security on a free market. Typically there is some sort of material award to encourage participation but also to hold off any undisciplined behavior/opinions. And typically there is an end date for a given statement to be traded. The current market quote of the traded statement is the market estimation for its truthfulness. Prediction markets conceptually work not because of expert participants but because the law of large numbers. Since the traded statements are of social (non-natural) essence, the assumption is that large numbers of the subjective opinions reflect wide range of more diverse and more significant information.

4. Proposition for method of testing market efficiency by processing expert predictions

The current paper describes the principal model of a version of online system for processing expert opinions (predictions in principle). Such a research has been a long coming project of the author (Marchev Jr. 2004a, Marchev Jr. 2004b, Lomev et al. 2005) and it could only be executed in a fast-communication environment with easily accessed information streams. As an additional research objective it would be interesting to test the efficiency of an emerging market (Takala 1997) such as Bulgarian Stock Exchange. The effects from running such a competition in a small market are also of interest. The methodology of the research has several important features:

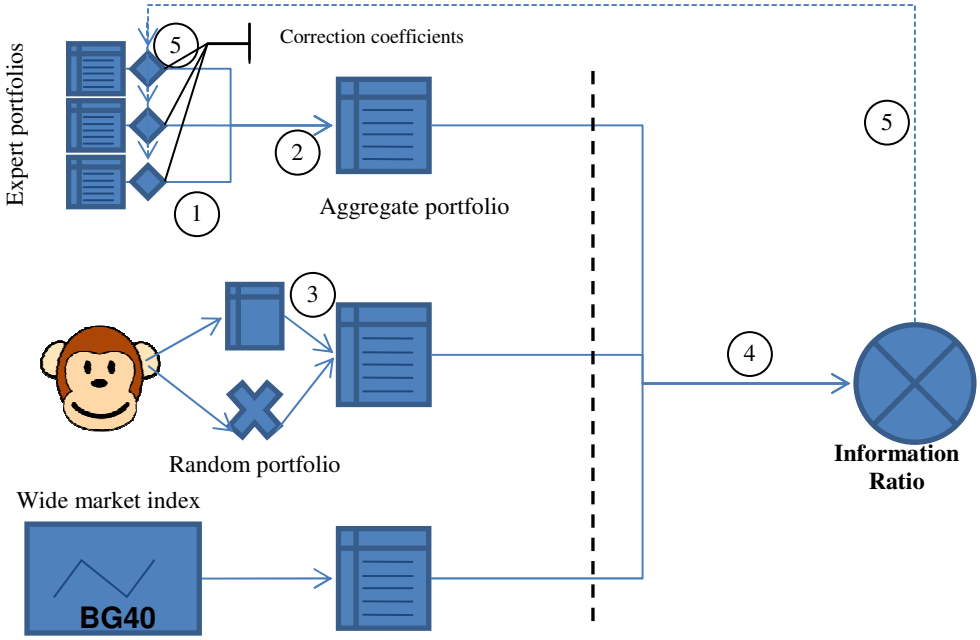
- It is conducted online (Internet), following the principles of Delphi approach such as expert anonymity, unified information stream, ability to exchange supporting arguments among experts and collecting all expert prediction portfolios in one collective portfolio.
- Every expert is put in an identical information environment, after logging in the online system. Firstly there is a standard interface page with useful information about the stock market, stocks, current news, etc. Secondly taking into account the information rich environment of the Internet and the information processing capacity of a human being, one could assume that everyone has access to incomprehensive (leaning towards infinite) volume of useful information.
- Parallel to submitting predictions, the experts may choose to leave argumentation in the electronic forum. The forum should be moderated towards anonymizing the experts' arguments. This action is reflecting one of the founding concepts of Delphi method – anonymous feedback.
- The main innovation is that the model is designed to render dynamically collected expert opinions (this has always been the challenge with applying Delphi method for forecasting financial markets) (Marchev Jr. 2004b).
- It presents the experts useful information even during the forecasting stage of the competition. When an expert submits his/her prediction portfolio, the current state of collective portfolio is available to him/her for analysis. An advantage for the expert is also that predictions could be submitted (and accessed) even when the stock market is closed for trading (much like a prediction market). These features are mainly aiming to encourage participation from the experts and present a possible solution of the problem with "Non-rhythmic" expert opinions (a typical issue with the Delphi approach).

The proposed methodology includes several steps (Fig. 2.):

1. Expert prediction submission.
2. Aggregating expert portfolios.
3. Generating random portfolio.

4. "Contest day".
5. Feedback.

Figure 2. Proposed methodology – main scheme



Source: own creation

5. Expert prediction submission

The experts submit their predictions for profitable investment portfolios for a future point of time – "contest day". There is a prediction window during which the expert predictions are collected for every "contest day". Each expert could submit "prediction portfolio" every day within the prediction window, once a day. All submitted prediction portfolios of an expert during the prediction window are kept and computed in an aggregate portfolio. Requirement for the expert to participate regularly may be imposed (e.g. at least once every 30 days).

The experts submit their "prediction portfolios", consisting of k positions each with respective weights, where k is the total number of positions traded on the market (or criteria for admission of positions may be used). For unwanted positions the weights are set to 0. Each portfolio has a nominal sum of S at submission. If the positions in a portfolio sum up to less than S, the residual is assumed a cash position

C_j. The cash position cannot be less than 0 – no borrowing allowed (1). Short positions are also possible[4], accounting the fact that the invested amount in short positions is "blocked" (counted as positive) – no margin account. The market is assumed "frictionless" e.g. no transaction costs, inflation, taxes, interest on cash positions etc. are computed.

$$P_j(t) = \sum_{i=1}^{k} w_{ij}(t) + C_j(t) = S \quad (1)$$

where :

j - serial number of expert

i - serial number of position

k - total number of possible non-cash positions

t - day of the prediction (within the prediction window)

$P_j(t)$*- value of prediction portfolio of expert j, submitted at the moment t*

$w_{ij}(t)$ *- allocated sum of position i in the portfolio of expert j, submitted at the moment t*

$C_j(t)$*- value of cash position of expert j, submitted at the moment t*

S - nominal investment amount

In its essence equation (1) is the typical way to calculate the value of a portfolio. What is important here is that the cash position C_j acts as a plug variable to sum the value of the portfolio to the investment amount *S*.

6. Expert prediction submission

This phase includes two stages – computing each expert's aggregate portfolio and collecting all aggregate portfolios in a collective portfolio. Actually the operations at this phase are done not only at the end of the contest, but also dynamically throughout the contest so the information about the current state of the collective portfolio is available. Being available means that every expert could see it since it is a valuable piece of information not only for the contest itself, but also for their real life work. So this is the motivational mechanism for participation – if an expert would like to see the current state of the collective portfolio (which reflects the expectations of all other experts), he/she must submit a prediction portfolio.

All prediction portfolios of an expert are combined in an "aggregate portfolio" by arithmetic averaging of respective positions of all prediction portfolios (2).

There is an incorporated mechanism for correcting (weighting) the expert opinions as a function of time of submitting each prediction with the newer having bigger impact on the final collective portfolio. So the aggregate positions are

[4] Short positions on Bulgarian Stock Exchange are not used in practice due to overregulation

corrected by correction factors (3). The correction factor uses exponential weighting (4).

$$\overline{w_{ij}(m)} = \frac{\sum_{1}^{m} w_{ij}(t)}{m} \tag{2}$$

where :

m - serial number of portfolio, submitted within the current prediction window

$\overline{w_{ij}(m)}$ *- value of aggregate position i in portfolio number m, submitted by expert j*

Equation (2) is a simple arithmetic average of the prediction portfolios, submitted by one expert. The index *m* here denotes the day (out of the prediction window) on which the expert submits a prediction portfolio. Note that since an expert could submit one prediction portfolio a day, the number of prediction portfolios by an expert cannot exceed the number of days in the prediction window i.e. $\{m\} \subseteq \{t\}$.

$$\overline{\overline{w_{ij}(t)}} = \frac{\sum_{t=1}^{n} \overline{w_{ij}(t)}.T(t)}{\sum_{1}^{t} T(t)} \tag{3}$$

where :

$\overline{\overline{w_{ij}(t)}}$*- corected value of aggregate position i in aggregate portfolio of expert j, at the moment t*

T(t) - correction factor for time at the moment t (see below)

It may seem that the correction for time of submission in equation (3) is done a bit too complicated and un-elegant, but it is needed to be such to have an important property – "rolling computation". It is meant that at the end of each day of the prediction window all of the submitted prediction portfolios are computed.

$$T(t) = \frac{2t}{2n + n(n-1)} \tag{4}$$

where :

t - serial number of the prediction day within the prediction window

n - total number of days in the prediction window

Equation (4) is derived from the general formula for linearly-weighted moving average. This is a necessary calculation since the predictions submitted earlier are reflecting less significant information than the predictions submitted later and are presumed less accurate.

The other stage of this phase of the contest is constructing a collective portfolio. This is done by averaging aggregate portfolios of all experts (5).

$$P_g = \sum_{i=1}^{k} w_i + C_g = S \tag{5}$$

where :

P_g *- value of collective portfolio*

w_i *- value of non-cash position i in the collective portfolio* P_g

C_g *- value of cash position in portfolio* P_g

Similarly to (1) the cash position of the collective portfolio is summing up the value of the portfolio to the investment amount *S*. The value of a non-cash position w_i in the collective portfolio (6) is an average of the current values of the corresponding positions, weighted by a correction factor $I_j(v)$. Using the equation (6) w_i could be computed after every day of the prediction window, thus making it available for the experts.

$$w_i(t) = \sum_{j=1}^{J} \frac{\overline{\overline{w_{ij}(t)}}}{j} . I_j(v) \tag{6}$$

where :

v - serial number of the current prediction window

$I_j(v)$ *- correction for expert accuracy for current prediction window v, where* $I_j(1)=1$

The correction factor $I_j(v)$ is a correction for the accuracy of every expert. Such correction is computed on the basis of the prediction accuracy of the expert from the previous prediction window. For the initial contest the value of the correction is 1 (i.e. no correction). The correction for accuracy is explained further in the paper.

7. Generating random portfolio

A "random portfolio" is selected and structured using random number generators. There is an important discussion on what sort of random generators should be used while testing market efficiency. On one hand there are the pseudo-random number generators – deterministic software producing chaotic features in sequences of numbers. On the other hand there are the real-random number generators – typically hardware device digitizing stochastic properties in sequences of numbers through some physical phenomena such as atmospheric noise or radioactive decay[5].

What is important to understand for the purpose of this paper is that generating the random portfolios is not a mere simulation but rather it is close to a game of chance. So for the current paper the most clear-cut and obvious approach is proposed. In honor of one of the most famous quote of Burton Malkiel[6] and in attempt to be as genuine as possible, real- random number generators such as darts and dice are used. Of course the author is aware that more advanced real-random number generators could be used. For a non-conclusive list of real-random number generators see (Marchev Jr. 2008).

The random portfolio is generated in two stages:

1. Random selection of fixed number of positions (at least 10) by blindfolded throwing darts at a newspaper's financial page / printed list of the positions (or other means of random selection out of a list of the positions).
2. Random definition of weight for each position using dice. Preferably regular shaped dice (platonic solids) such as icosahedron or dodecahedron but also pair of identical pentagonal trapezohedron (percentile dice) is an option since they will produce a random number from 0 to 99. Ideal would be a pair of icosahedrons with repeated sides. (0 through 9 or 10 through 90 repeated twice on a die). In the case of the latter two see (7). The aim is to incorporate more degrees of freedom, effectively meaning more „un-round" numbers. As well as introducing more instances of randomness to the portfolio[7].

[5] See for a brief introduction: http://www.random.org/randomness/

[6] See part 1 of the current paper

[7] It is easily seen that the simpler the dice the less degrees of freedom is introduced by one toss. The simplest of dice – two-sided die (i.e. a fair coin) would only produce one degree of freedom, so for a portfolio of 10 positions the weights of would be divisible to 5%.

Figure 3. Ten sided dice, pentagonal trapezohedrons

Source: http://en.wikipedia.org/wiki/File:DnD_Dice_Set.jpg

$$w_i = \frac{d_i + 1}{\sum_{1}^{i} (d_i + 1)} . S \qquad (7)$$

where :

i - serial number for postion

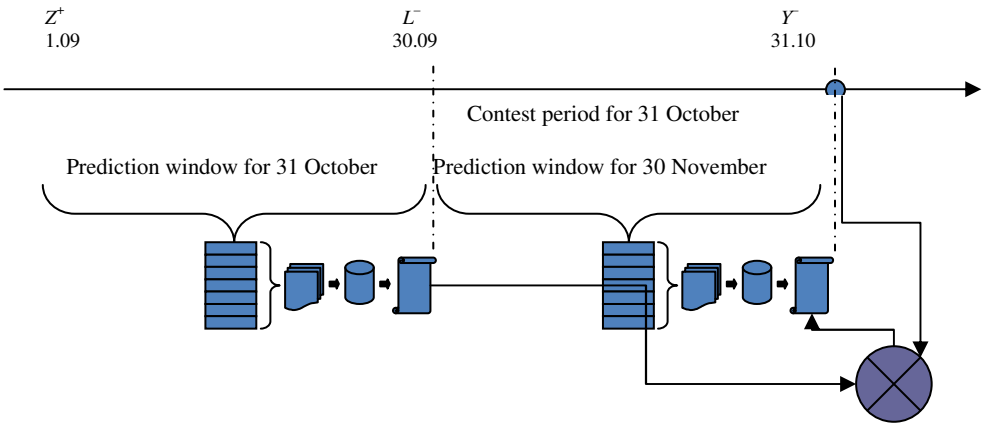
d_i *- random dwo-digit integer for position i*

w_i *- value of position i*

8. "Contest day"

The contest day (Y^-) is on the last working day of a month. The prediction window starts on the first day (Z^+) and ends on the last day (L^-) of the month prior to the month of the "contest day". The month which ends with the "contest day" is called "contest period". There is minimum one month and maximum two months of uncertainty for the experts when they prepare their predictions. All stock market values used in the research should be adjusted for splits and dividends.

Figure 4. Timeline example of the prediction window



Source: own creation

A market index for the same stock market is used as a benchmark for measuring portfolios' progress. Ideally it should be a wide market index with weighted-average of all securities traded on the market.

On "contest day" the two portfolios are competed to each other on the basis of a modified information ratio. Modified information ratio is accounting for the active return of the portfolios and variation of the historical prices which is a common measure for risk.

The portfolios are evaluated on real market quotes, considering the opening prices (L^+) on the first working day after the end of the prediction window and the closing prices (Y^-) on contest day. Rules for inputting missing values in the financial time series are necessary (8).

$$R_g = \frac{P_g(Y^-) - P_g(L^+)}{P_g(L^+)} \tag{8}$$

where :

R_g *- historic return of collective portfolio for the contest period*

Besides the value of on the contest day for the computation of equation (8) is needed the value on the opening of first working day of the contest period L^+. Note that there is a maximum of 23 working days in the contest period (with real-life stock market quotes), while there could be up to 31 days in the prediction window (on which experts could submit predictions).

$$E_g = \frac{R_g - R_M}{\delta_{R_g(t)}} \tag{9}$$

where :

E_g *- modified information ratio*

R_M *- historic return of market index for the contest period*

$R_g(t)$ *- daily dynamic return of aggregate portfolio, durring the contest period*

$\delta_{R_g(t)}$ *- tracking error of daily returns of collective portfolio, see below*

The information ratio (9) is actually a simplified version with only two values taken for calculation of the return.

$$\delta_{R_g(t)} = \sqrt{\frac{\sum_{1}^{h}(R_g(t) - R_M(t))^2}{h-1}} \tag{10}$$

where :

h - number of days in contest period

The tracking error (10) uses connotation of *t* as the serial number of a working day form the contest period.

$$R_g(t) = \frac{P_g(t)}{P_g(t-1)} - 1 \tag{11}$$

The dynamic daily return of the market index is calculated exactly the same way as for the collective portfolio, shown in equation (11).

9. "Contest day" consequences.

After the contest is over the results are not only used for analysis but also serve as correction factors for the further contests. Already explained was the correction for time of submission. The incorporated mechanism for correcting (weighting) the expert opinions as a function of accuracy was also mentioned. The principle is that

the better predictors from the previous contest have bigger impact on the final collective portfolio (12).

$$I_j(v) = (1 + E_j(v-1)).A(v-1) \tag{12}$$

where :

$E_j(v)$ *- modified information ratio of expert j for prediction window v*

$A(v)$*- correction from adaptation algorithm,* $A(1) = 1$

Again for the computation of the correction factor for accuracy the modified information ratio is used (13). This time it is calculated for every individual expert. Since the results of the previous prediction window are ready at about the end of the current prediction window, the correction for accuracy is assumed to its last computed value until the new value replaces it.

Both of the above-mentioned correction mechanisms (for time of submission and for expert accuracy) and the model as a general introduce inherent systematic errors for which an adapting mechanism is proposed (e.g. stochastic approximation) (12), see fig. 5. The value for the initial contest of the adaptation correction is 1. Important remark is that the average corrections for accuracy and adaptation on all expert aggregate portfolios should be as close to 1 as possible.

$$E_j(v) = \frac{R_j - R_M}{\delta_{R_j(t)}} \tag{13}$$

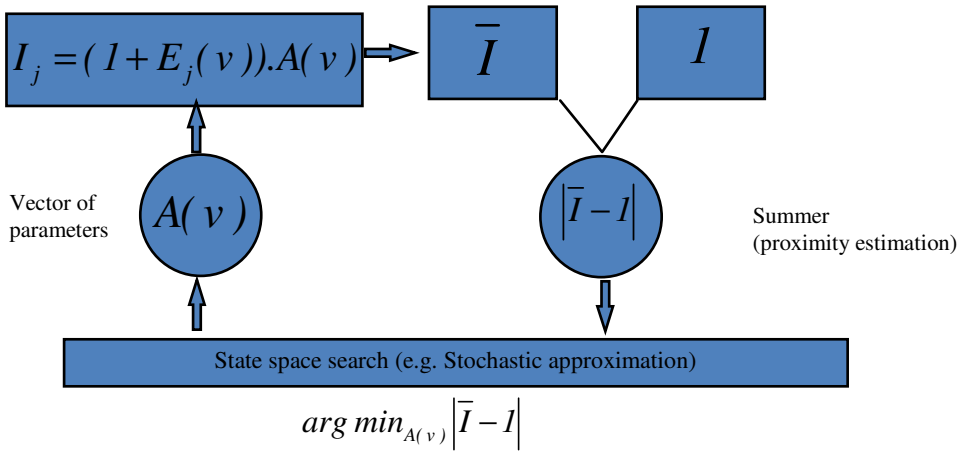
where :

R_j *- return on aggregate portfolio of expert j*

$\delta_{R_j(t)}$*- tracking error on daily return of aggregate portfolio of expert j*

All the requisite variables of equation (13) are computed exactly as for the collective portfolio from equations (8), (9), (10), (11).

Figure 5. Algorithm for computing the value of A



Source: own creation

The adapting algorithm uses the arithmetic average of the individual correction factors for accuracy $\bar{I}$ and compares it with the value of 1. Through iterative process is searched such value of *A(v)* for which the goal function $|\bar{I} - 1|$ is the closest to zero. The computed such value is included in the correction factor for each expert.

10. Empirical example

In the next few paragraphs there is an example of the implementation of the above-mentioned approach. The computations use historical data of four shares, listed on Bulgarian Stock Exchange and data of a major market index for a period of two months. Additionally the expert portfolios (opinions) are simulated using arbitrary and subjective opinion of three individuals. As much as the example is only for illustration purposes, it is not meant to be an empirical test of any sort. So the resulting numerical values of the variables are not at all important, but the whole demonstration of how the approach works. It is done mainly to satisfy a reviewer's request.

Assume there are three experts (Expert1, Expert2 and Expert3) who have submitted their prediction portfolios allocating 10000 EU to five positions (share A, share B, share C, share D and a cash position) as follows (see Table 1). In the table each row represents a prediction portfolio submitted by the corresponding expert on a given day of the forecast window. For example: on the first data row of the first section it is shown that Expert1 has submitted a prediction portfolio on the first day

of the prediction window allocating 6549 EU for share A, 432 EU for share B, 414 EU for share C and 1956 EU for share D thus leaving 649 EU (out of the 10000 EU) in cash. The random portfolio is simulated using percentile dice and computed as in equation (7)

Table 1. Simulated prediction portfolios

Expert1					
Day of the prediction window	Allocation by position				
	A	B	C	D	Cash
1	6549	432	414	1956	649
7	1328	3124	3032	780	1737
12	810	2834	541	3286	2529
15	3277	1117	1807	812	2986
19	66	1513	2152	5365	904
26	2486	1431	1280	3866	938
29	373	565	3677	3718	1667
Expert2					
Day of the prediction window	Allocation by position				
	A	B	C	D	Cash
6	2206	2558	424	3245	1567
10	2094	2898	1863	2849	296
22	2549	1641	2791	755	2264
30	1548	742	2137	1646	3928
Expert3					
Day of the prediction window	Allocation by position				
	A	B	C	D	Cash
2	249	3320	3766	1881	784
14	216	3130	1648	3405	1600
16	1210	662	4660	1845	1623
19	983	4832	1751	1561	872
27	1758	2444	3179	2520	99
Random portfolio					
Day of the prediction window	Allocation by position				
	A	B	C	D	Cash
30	1304	174	1652	3043	3826

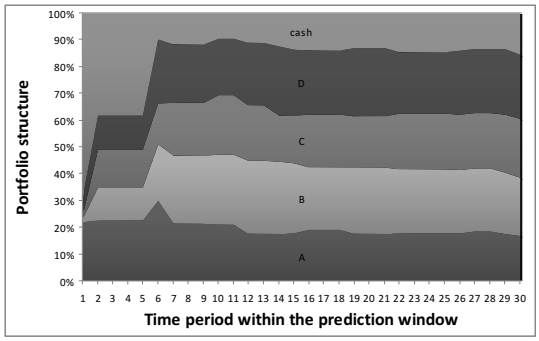
Source: own creation

Using these prediction portfolios the structure of the collective portfolio could be computed – as submitted as well corrected for day of submission (for graphical representation see fig. 6). The figure clearly shows the difference between the non-corrected and the corrected portfolio.

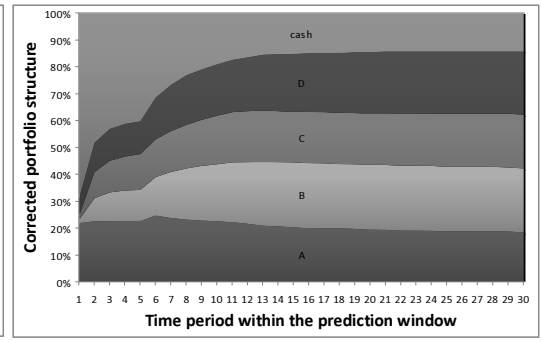
The next phase of the contest is to observe the dynamic of the portfolios (fig. 7) as well as their return on contest day (-7.2% for share A, -3.8% for share B, 4.2% for share C, 1.2% for share D and -1.2% for the market index).

Figure 6. Structure of the collective portfolio during the prediction window

a) as submitted



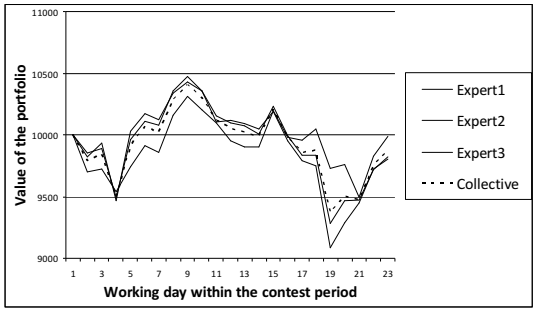
b) corrected for time of submission



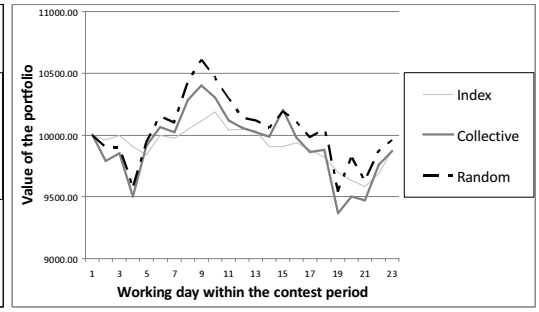
Source: own creation

Figure 7. Dynamic daily value of the portfolios during the contest period

a) expert portfolios and collective portfolio,



b) random and collective portfolios, market index



Source: own creation

The results of such simulation would be as shown in Table 2. The collective portfolio performed better than the index although still losing. The random portfolio (in this realization) was almost on par with the initial invested amount. The performance of Expert1 and Exper2 was worse than the performance of the market index. So their respective correction factors for the next prediction window would lessen the weight of their prediction in the collective portfolio. Just the opposite for Expert3 – the correction factor is above the value of 1. The whole adaptation correction factor is slightly below the value of one, correcting a systematic error of overvaluing the collective portfolio (possibly due to the decreasing values of the market index this contest period). And of course it should be reminded that this was only a simple example of the approach.

Table 2. Contest day results

	Expert1	**Expert2**	**Expert3**	**Collective**	**Random**
Value of the portfolio	9821	9841	10002	9888	10005
Value of the index	9881				
Return of the portfolio	-1.8%	-1.6%	0.0%	-1.1%	0.0%
Return of the index	-1.2%				
Information ratio	-18.9%	-14.2%	62.3%	2.9%	50.0%
Adaptive correction	0.911				
Correction for accuracy	0.739	0.782	1.479		

Source: own creation

11. Conclusions

Since it is only a proposition, the current research mostly raises questions:

1. Being able to see the predictions while predicting, do the experts make the market even more efficient?
2. Is it so on Bulgarian stock exchange, where the experts are relatively small number (and so everybody tend to believe a relatively small number of subjective opinions) and where the market is generally shallow (meaning easily manipulated) and underdeveloped (predictable)?
3. If the market follows the expert expectations, wouldn't it lead to degeneration of their predictions (due to multicolinearity)? Would the market become "sensible dependant on small changes in the initial values" (Lorentz 1993)?
4. Since the collective portfolio averages on many experts with many positions and comparing it to a random portfolio averaging even more positions, wouldn't the differences among the two portfolios and the market index be insignificant (due to regression to the mean)?

References

Bachelier, L. 1900: *Théorie de la speculation*. Gauthier-Villars.

Bernoulli, J. 1713: *Ars Conjectandi*. Nicolas Bernoulli, 1430. p.

Bernstein, P. L. 1998: *Against the gods*. John Wiley & Sons, Inc., New York et al., 199-200. p.

Dimson, E. – Mussavian, M. 1998: A brief history of market efficiency". *European Financial Management*, Vol. 4, No. 1,91-193. p.

Dorsey, W. 2003: *Behavioral trading* methods for measuring investor confidence, expectations and market trends. New York: Texere.

Fama, E. 1965: Random Walks in Stock Market Prices. *Financial Analysts Journal* 51, 1, 55-59 p.

Harizanova, M. – Mironova, N. 2009: *Communications in management*, Avangard Prima, Sofia, 20-34. p.

Jantsch, E. 1967: *Technological forecasting in perspective*. OECD, Paris.

Jasen, G. 2002: Journal's Dartboard Retires After 14 Years of Stock Picks. *Wall Street Journal*, April 2002.

Kendall, M. G. 1953: The analysis of economic time-series - part I: prices. *Journal of the Royal Statistical Society. Series A (General),* Vol. 116, No. 1, 11-34. p.

Lo, A. W. – MacKinlay, A.C. 1999: *A non-random walk down Wall Street*. Princeton University Press, Princeton, New Jersey.

Lo, A. 2004: The Adaptive Market Hypothesis; market efficiency from an evolutionary perspective.

Lomev, B. M. – Marchev, Jr., A. A. – Marchev, A. A. 2005: Research on the efficiency of some capital markets using methods of the theory of fractal time series". *Fifth Scientific conference with international attendance "Investments in the future - 2005", 20 - 22.10.2005, Varna*, Scientific and Technical Unions - Varna (only available in Bulgarian)

Lorentz, E. N. 1993: *The Essence of Chaos.* Univ. of Washington Press, Seattle.

Malkiel, B. G. 1973: *A random walk down Wall Street*. W. W. Norton & Company, New York et al., 24. p.

Marchev Jr., A. A. 2004a: Application of "Delphi" method for predicting indices based on securities' quotes, *Burgas Free University, Second Scientific conference "Modern management practices II", Sveti Kirik i Julita, 22.02. - 25.02.2004 collected papers*, Publisher "Briag" – Burgas.

Marchev Jr., A. A. 2004b: On-line system for expert predictions of financial markets. *Scientific conference "Modern management practices", 08.10. - 10.10.2004, Bankia*.

Marchev Jr., A. A. 2008: What to do when pseudo- is not good enough (*poster*), European mathematical society summer school "Risk theory and related topics". European mathematical society, Institute of mathematics of the Polish academy of science, 29.09. - 08.11.2008, Bedlewo, Poland.

Rasp, J. – Hamrick, J. 2003: The Experts and the Darts: Using the Wall Street Journal "Dartboard" Data in Statistics Classes.

Takala, L. 1997: Stock exchange development and the use of information: A step towards market efficiency? The Czech Republic, *Emerging capital markets in a transitional environment*. AnCo Commercial and advertisement publishing house, Sofia.

Rating model opportunities for emerging markets

Alexander Karminsky *

Ratings are in high demand in market-driven economies. The Basel II Accords have sparked increased interest in the development of approaches based on internal ratings systems. They have initiated the elaboration of models for remote ratings forecast based on external ones as part of Risk Management Systems and Early Warning Systems.

This article evaluates peculiarities of current ratings systems and specific issues of development of econometrical rating models (order probit) for emerging market companies. Financial indicators, market-value appraisals and macroeconomic indicators of different countries are used as explanatory variables. Standard & Poor's and Moody's ratings are considered as modeling ratings. The sample was based on data from an information agency as well as the rating agencies, and related to companies from almost 40 countries.

Keywords: financial risk, corporate ratings, econometric model

1. Introduction

Ratings are in high demand in market-driven economies. As a business, the rating process has a moral component, since the agency does not bear legal responsibility for its conclusions. Its reputational capital serves as a regulatory element however (Partnoy 2002). In addition to independent appraisals of investment risk in the form of the rating agency's opinion, ratings also function as licensing of a sort.

The Basel II Accords (Basel 2004) have sparked increased interest in ratings and their models. The development of approaches based on internal ratings systems holds a practical interest, especially for developing markets. The topic has received increased attention in connection with the global recession that began in 2007.

In this work, an analysis is made of the possibilities for rating modeling as applied to industrial companies and banks of developing countries (according JP Morgan classification). Emphasis is placed on the elaboration of econometric models. As explanatory variables, financial indicators, which characterize the activities of a company, are incorporated. Market indicators are used to reflect the

* Dr. Alexander Karminsky, Professor, State University – Higher School of Economics, Faculty of Economics (Moscow)

Acknowledgement: I would like to thank Anatoly Peresetsky for his helpful comments.

dynamics of stock quotations, and macroeconomic variables and dummies of industrial and country affiliations are also used.

Ratings of Moody's Investors Service and Standard & Poor's agencies are considered as modeling ratings, which also allows an evaluation of the specific approaches of each of those agencies to be made. Samples were made up of data from those agencies and the Bloomberg information agency.

Analysis of the predictive power of the derived econometric models allows an appraisal of their remote use in risk management systems to be made. Particular attention is given to variables in ratings of companies in accordance with their affiliations with developing countries or with particular industries.

It is demonstrated that the ratings companies and banks from developing countries with the same financial indicators have significantly lower ratings than those in developed countries ("advanced economy" by new IMF understanding). It is also shown that industry affiliation influences ratings.

The work consists of six sections. The second section addresses the particularities of ratings as a measurement of risk in Russia and the countries of Central and Eastern Europe. Section 3 includes an overview of different approaches to rating modeling. In Section 4, there is an examination of the types of models used and the formation and statistical characteristics of the samples.

Models of corporate ratings and bank ratings relating to developing markets and a comparison of the ratings of Moody's and S&P are systematized in Sections 5. Conclusions are presented in the final section.

2. Development of ratings services in developing countries

As an example of the ratings system using, we may consider the ratings in Russia where they are rather particular. We may observe several waves of interest in these possibilities.

At the initial stage of the establishment of market relations, ratings in Russia were predominantly used for banks. The entry of the international rating agencies and the ratings they made in Russia (beginning in 1996) had little impact before the 1998 financial crisis and immediately after it.

The opportunities for foreign borrowings, including by industrial companies and corporations, beginning in 2003, gave impetus to their development. The number of ratable entities in Russia has more than tripled since then, reaching more than 300 at the beginning of 2009. (About half of them are banks and more than a third are companies.) The process was encouraged when Russia received investment-level ratings in 2005-2006.

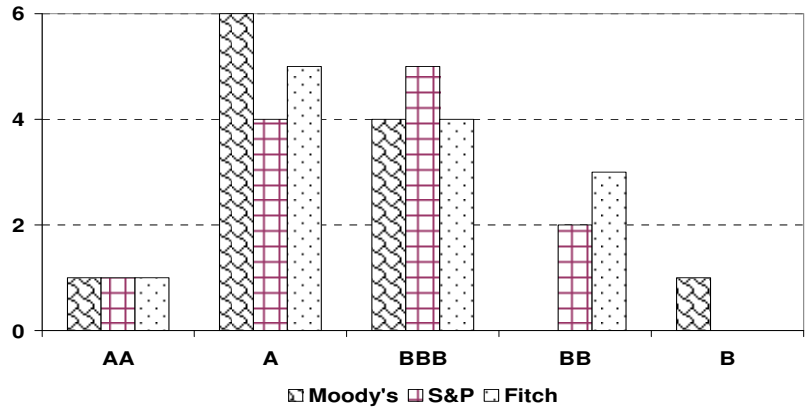
The recession of 2008-2009 has had an effect on the rating process. In particular, a number of ratings were withdrawn. Russia's sovereign ratings were lowered by Standard & Poor's and Fitch Ratings by one grade, although ratings

remained on the investment level at BBB. The insignificant lowering of the sovereign ratings did not dampen interest in them from economically active entities, as happened in 1998.

A large portion of the bank ratings was assigned by Moody's Investors Service (hereinafter, "Moody's"), the Standard & Poor's agency (hereinafter "S&P") leads in ratings of industrial companies and their financial instruments. The distribution of international agencies' corporate ratings by grades shows that the level of ratings of Russian companies is comparatively low. Less than 20 companies have investment-level ratings. The average level of ratings is between BB- and BB for all three agencies, while the average level for S&P is almost BB- and for Moody's, it is Ba2, which is equivalent to BB. For the Fitch Ratings agency, the average level is between these grades.

The ratings of banks and industrial companies in Central and Eastern Europe (CEE) have much in common as they concern developing countries. The level of development of these countries lags considerably behind that of the original members of the European Community. However, these countries were oriented toward membership in the EU, and many of them have become members of EU.
Data analysis shows (Fig.1) that the sovereign ratings of the countries of the CEE are in the lower part of the investment range (A and lower), with the exception of the higher ratings of Slovakia, the Czech Republic (A+) and especially Slovenia (AA). Serbia, Macedonia and Bosnia and Herzegovina have noninvestment ratings.

Figure 1. Distribution of CEE countries ratings

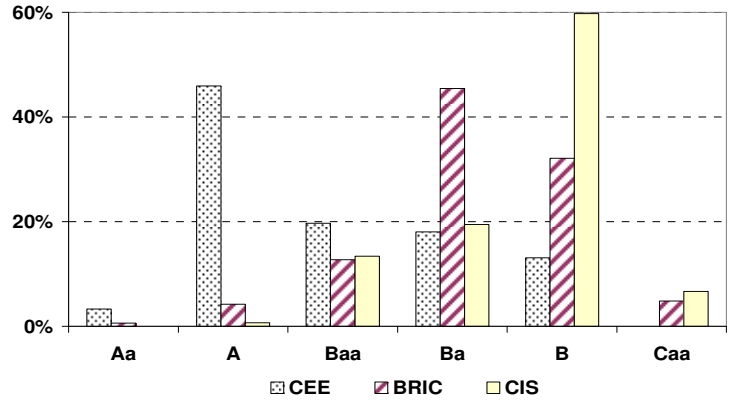


Source: own creation

At the same time, companies and banks mainly have ratings on the speculative level, and the number of rated entities is not high (Fig. 2). This is largely explained by European Union support for these countries, although that support has been limited since the recession began. On the other hand, there are few large companies in these

countries and the existing ones happen to be subsidiaries of transnational companies. This may potentially lower the level of the ratings, since their market capitalization (or volume of assets) is one of the main explanatory variables for them.
Thus, in spite of higher sovereign ratings, company ratings in these countries are on the level of Russian companies, being predominantly in the upper range of uninvestment ratings.

Figure 2. Distribution of banks developing countries ratings



Source: own creation

In spite of the comparative growth in the number of ratings, there are still clearly too few of them in developing countries. In addition, rating methods are largely inexplicit, and expertise plays a significant role in them which hinders the use of ratings for risk evaluation and decision making even on the state level. This is the reason for interest in the creation of internal ratings and model ratings that can be used as preliminary evaluations in making management decisions.

3. Overview of the literature and particulars of the methodology

Changes in ratings play an important role in transactions with interest-rate risks. In spite of a certain decrease in the normative significance of ratings, their presence and popularity have grown since the mid-1970s (Partnoy 2002). That is due largely to the regulatory significance of ratings, in addition to their market significance (Karminsky-Peresetsky 2009).

Although the initial application of ratings was to debt obligations, corporate ratings are now steadily gaining significance (Altman and Suggitt, 2000) for the organization of syndicated credits, the rating of corporate bonds and other purposes (Altman-Saunders 1998; Servigny-Renault 2004; Partnoy 2002).

The recession has exacerbated the problem of ratings. A number of researchers have shown (Altman-Rijken 2004; Pederzoli-Torricelli 2005; Curry et al. 2008) that ratings are pro-cyclical. There is a certain lag between the recording of financial results and the changing of a rating. While this creates stability in the rating process and averts reactions to comparatively insignificant events, it does not always provide for a timely reaction by ratings agencies to significant events. The widely-discussed collapse of several major companies is an example of this (Servigny-Renault 2004).

At the same time as (Amato-Furfine 2004) demonstrated based on the activities of U.S. firms and the data of the S&P rating agency between 1980 and 2000, credit ratings rise less during times of recession. However, ratings do not display excessive sensitivity to business cycles. Without discounting this factor, especially for new financial instruments (at present predominantly structured transactions), it should be noted that degradation can be connected with the dynamics of the market as a whole. With regards to banks, the absence of such degradation during the transition to ordinal scales was shown in (Karminsky-Peresetsky 2007).

The reason for the relative volatility of ratings is the specifics of the assignment of country ratings, especially for developing countries, considering the specifics of their reexamination and the modification of the rating agencies' methodology (Moody's 2007). In (Kaminsky-Schmucler 2002; Reinhart 2002), it is shown that there are three possible channels of instability that arise from changes to a country rating during a recession:

- directly through the value of debt obligations and stock on the market,
- through contagion and generated global unstability,
- due to markets in countries with lower ratings because of their greater liability to fluctuations.

In (Reinhart 2002), it is also shown that changes in sovereign ratings influence the spread and income of bonds, complicate access to resources of developing markets, hastens the transition from currency crisis to banking crisis and can act on recession. Some of the specifics of the rating process during crisis, including its effects on developing countries, are also examined in (Joo and Pruitt, 2006). Significant attention has been given to an analysis of indicators of financial and banking crises, especially with regards to developing countries (Kaminsky-Schmucler 2002; Rojas-Suarez 2002).

Evaluating sufficiency of capital as a measurement of risk based on internal ratings, as is foreseen by the IRB approach put forward as part of Basel II (Basel 2004), can use probability of default as a model as well as ratings. It can also use an evaluation of transition matrices and the mechanism of Markov chains (Frydman-Schuermann 2008) or econometric models, including scoring (Altman-Saunders 1998; Altman 2005; Feng et al. 2008).

A number of articles have been devoted to the elaboration of internal ratings and early warning systems. An overview of methodological specifics of elaborating models is made in (Altman-Saunders 1998; Karminsky et al. 2005).

In (Carey-Hrycay 2001), specifics are examined of the joint use of several methods for the evaluation of the probability of default on debt instruments according to an internal rating scale. Mapping to a standardized scale and scoring models are used. The presence of a data series of long duration is critical. A number of the specifics of the elaboration of internal ratings systems are also examined in (Jacobson et al. 2006; Servigny-Renault 2004; Hanson-Schuermann 2006). In the last of these articles, the confidence interval technique is used to refine rating gradation.

Selection of the explanatory variables is methodologically important for the elaboration of corporate ratings models. The indicators that are employed by the rating agencies can be also used (Moody's 2009; S&P 2008) by other researchers (Rojas-Suarez 2002; Servigny-Renault 2004; Guttler-Wahrenburg 2007; Curry et al. 2008). Typical indicators are the size of the company, its profitability, stability, liquidity and structure of the business, as it is expressed through companies' balance-sheet figures. In recent years, the use of such factors as state support for companies, and support from the parent company or group of companies, has become more prominent (Moody's, 2007; S&P, 2009).

The use of macroeconomic indicators has also become more typical recently (Carling et al. 2007; Curry et al. 2008; Peresetsky-Karminsky 2008). Among the most common indicators are inflation index, real GDP growth, industrial production growth and, for export-oriented countries, oil prices and changes in the cross-rate of currencies. Separate mention should be made of market indicators (Curry et al. 2008), which is especially important for publicly held companies (market value of companies, volatility of stock prices, systemic parameters, etc.).

It should also be noted that alternate indicators can also be used for developing countries (Altman 2005; Rojas-Suarez 2002; Karminsky et al. 2005) that are characteristic of developing markets and predominantly uninvestment ratings. These include value of resources, percent margin, pace of asset growth and growth of interbank debt, including on an international level.

Variation over time, both of the dimensions of the risk and the approaches of the rating agencies points to the use of a time factor in models including that of panel data. Some of the specifics of these approaches are found in (Elton et al. 2004; Frydman-Schuermann 2008).

The particular significance of industry affiliation and possible differences among ratings of companies of varied profiles and regions can be noted for corporate ratings (Niemann et al. 2008). This is connected with the specifics of business in various segments of production activity. Industry-specific models and the use of fictitious variables depending on the industry and the location of companies are possible.

A number of articles have noted differences in the ratings of various agencies (Packer 2002; Bae-Klein 1997; Kish et al. 1999). Corresponding factors of national and international agencies were analyzed in these works. In practically all the research, the two main rating agencies, Moody's and S&P, were considered.

The global financial recession exposed a number of problems of the ratings business and the entire financial management system. The financial system grew markedly in the first decade of the 2nd millenium (IMF 2009). New financial instruments were created, designed for higher profit without adequate quality of risk management. However, the inability of both regulatory organs and rating agencies to evaluate the global outlook and the threat from the asset price bubble.

Faith in light regulation based on the discipline of the financial market and hope for the successful distribution of risk through financial innovation does not preclude their concentration. An IMF analysis shows problems on three levels:

- Financial regulatory and monitoring organs proved to be incapable of exposing the higher concentration of risk brought on by the rapid growth in financial innovation.
- No account was taken of growing macroeconomic imbalances that contributed to the growth of systemic risk in the financial system and real estate market, as well as in the shadow financial system.
- International financial organizations and the monitoring and control system that was in place were unable to cooperate reliably on the international level to identify vulnerable areas in transnational relations.

Heightened possibilities of infection during the liquidity deficit are noted in (Karas et al. 2008) in regard to the Russian banking stystem and developing markets. It is shown there that regulating the liquidity of individual banks is not sufficient to avoid a systemic crisis. Rather, resource management by the lender of the last resort is necessary to restore the coordination of the interbank market.

The crisis emphasized the need for clearer signals in economic policy and the development of international cooperation on a number of economic and financial questions, including ratings. Among the steps suggested was taking leadership in responsive measures to systemic global risk. Establishing an early warning system should be included as one of these preventive steps.

Development and regulation of early warning systems require improved independent evaluation. Rating agencies should also be employed for this purpose. Policy has to be coordinated in various areas. Some of the elements included in these should be: supervision of the rating agencies, bookkeeping practice and auditing. Those initiatives should be coordinated both within a country and on the international level (IMF 2009).

Among the problems that arise in connection with the financial crisis, rating agencies' lag in the methodology of assigning ratings can be pointed out, along with calculation of the systemic risks of the global financial system and a lag in the

evaluation of complex financial instruments. A more active role of agencies in developing methodology, including areas emphasized by the Basel committee, can be noted (Basel 2009).

4. Data and models

4.1. Models and rating scales

Multiple choice probit models are used for follow-up studies. Earlier, they had been used for the elaboration of bank rating models. Further, three numerical scales are used. They correspond to the classes of ratings and gradations of ratings and a mixed scale that is tied to the limitations on the volume of the sample. The mapping of these scales to the numerical scales has 8, 18 and 12 levels, respectively. A higher rating corresponds to a lower number.

4.2. The Sample: financial and market variables, macrovariables and ratings

To elaborate ratings models for industrial enterprises, a sample was made on the principle of affiliation with companies in a number of industries (oil and gas, metals, retail trade, energy, telecommunications and heavy industry) that are potential competitors of Russian production companies of the same profile.

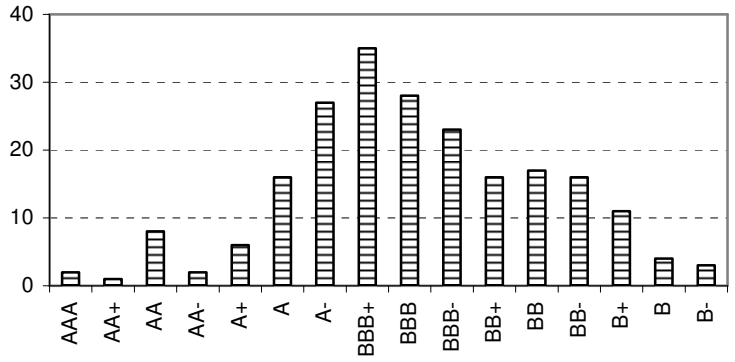
Conditions for selection in the sample were: the presence of an S&P rating at the time the sample was made, affiliation with a selected industry, accessibility of financial data and market indicators, and tradability as indicated by a liquid market for the company's stock. In the sample, there were 215 companies from 39 countries with S&P ratings as of spring, 2008.

The sovereign and corporate credit ratings of companies were taken from the websites of the S&P and Moody's agencies as they were presented as of February, 2009. Financial and market indicators were taken from the Bloomberg information system. A minimum of one financial indicator is assumed in each of the following groups: market valuation, size, profitability, market risks, balance-sheet and cash flow.

Distribution of the companies represented in the sample by rating gradation as of spring, 2008 is presented in Fig. 3. The small number of companies with ratings in categories AAA, AA and B justify the use of the mixed scale with differentiation of gradation in classes' A to B.

More than half the companies were represented by five countries: U.S. (74), Russia (31), Canada (15), Great Britain (13) and Japan (10). Division of the countries into developed or developing economies was done according to the methodology of the International Monetary Fund (IMF). The sample contains a greater number of companies from developed countries (152) than from developing ones (63).

Figure 3. Distribution of companies in the sample by gradation of rating



Source: own creation

The time lag between financial indicators and ratings was determined to be 1.5 years (financial indicators were taken from autumn, 2007 and ratings from the beginning of 2009). This agrees with the conclusions reached in (Karminsky-Peresetsky 2007) for bank ratings, and with a number of other articles, such as (Altman-Rijken 2004). Financial and market data for only one accounting period can be considered a flaw in the sample.

The sample of banks included (Peresetsky-Karminsky 2008) about 1000 observations in 2002-2005 of 380 banks in 42 developing and developed countries. The distribution of the sample by gradation and region, and the statistical distinctions of the sample may be found in the work indicated. The sample included 37 banks from CEE.

4.3. Explanatory variables and descriptive statistics

The lists of the main financial and market indicators used for the elaboration of the rating models, as well as a description of them and their expected influence on the ratings of industrial companies and banks are partly presented in Tables 1. Supplemental indicators have been used in a number of models as described in the text. According to the Bloomberg, market risks correspond to the Beta coefficient and Volatility.

Country distinctions will be calculated with macrovariables: annual rate of inflation, real GDP growth, CPI (Corruption Perception Index), and sovereign rating, which can be seen as a proxy for the institutional environment in which a company works. The first two indicators were taken from World Bank 2007 data, the corruption index comes from Transparency International, and sovereign ratings were taken from rating agencies data. The expected influence of inflation and the

level of corruption are negative. The remaining indicators are positive. Higher CPI corresponds to lower corruption. This indicator with other macrovariables may be used as explonatory variables instead of sovereign rating.
A number of dummies will also be used in the models. The relationship of a company to a country with a developed economy (1- developed, 0- developing) and Russia were introduced for the analysis of the influence of affiliation to the groups on the level of the rating. Companies' liability to risk in dependence of their affiliation with various industries will be traced through the introduction of dummies affiliated with the following sectors: telecommunication, oil and gas, metal and mining, consumer, utilities, and manufacturing and chemicals.
The variables used in the models will be commented on in the analysis of the resulting tables.

5. Econometric rating models

5.1. Models of corporate ratings

The models examined in this work depend exclusively on open information. Accordingly, we will examine the possibilities derived from the use of indicators based on company financial accounts prepared to international standards and supplemental possibilities provided by macroeconomic variables and market elements.

Among the questions that we face in elaborating rating models, we emphasized the following:

- Do the ratings of enterprises depend on their affiliation with a group of countries (developing countries, Russia)?
- Do ratings depend on affiliation with an industry?
- Is it possible to incorporate a high enough level of information in sovereign ratings using macrovariables?

In the elaboration of the base rating model for the Standard & Poor's agency, indicators were chosen from each group of financial indicators. The indicator of size of the company was included in all models. Below (in logarithmic scale), capitalization of the company is taken as the same. As criteria for comparison at the first stage, statistical characteristics of the quality of the models (Pseudo-R^2, t-statistics) were used, to which predictive characteristics were added at the next stage.

Coefficient signs match prior expectations. As one of the areas for the improvement of the quality of the model, transition to a scale of gradations or a mixed scale may be considered. This will ensure the models are more accurate. Relevant models are presented in Table 1.

Table 1. Models of ratings on scales of gradations and a mixed scale

	S&P		S&P, market		Moody's
Scale	**Grades**	**Mixed**	**Grades**	**Mixed**	**Mixed**
Volatility of value			0.022***	0.068***	
			(0.0060)	(0.011)	
Share value/			-0.015**	-0.26***	
Cash flow			(0.0075)	(0.0078)	
Capitalization	-0.517***	-0.509***	-0.528***	-0.588***	-0.502***
(logarithm)	(0.151)	(0.153)	(0.154)	(0.158)	(0.158)
EBITDA/Interest	-0.0062*	-0.0062*	-0.0089***	-0.0086***	-0.017**
expenses	(0.0034)	(0.0035)	(0.0033)	(0.0033)	(0.0070)
Return on assets	-0.035***	-0.033**	-0.042***	-0.041***	-0.032**
	(0.014)	(0.014)	(0.015)	(0.015)	(0.014)
Long-term debt/	-0.012*	-0.012**			0.0095*
Capital	(0.0045)	(0.0047)			(0.0049)
Inflation level	0.379***	0.391***	0.443***	0.561***	0.345***
	(0.063)	(0.065)	(0.070)	(0.077)	(0.069)
GDP growth	-0.186***	-0.184***	-0.185***	-0.252***	-0.96
	(0.060)	(0.060)	(0.185)	(0.053)	(0.076)
Metal and mining			-0.456*	-0.835***	
			(0.258)	(0.270)	
Oil and gas	-0.619***	-0.625***	-0.866***	-0.954***	-0.413*
	(0.197)	(0.198)	(0.212)	(0.215)	(0.228)
Utilities	-1.217***	-1.223***	-1.127***	-0.973***	-1.403***
	(0.0224)	(0.0225)	(0.234)	(0.238)	(0.243)
Developed countries	-0.636**	-0.611**			0.086
	(0.308)	(0.310)			(0.355)
Pseudo-R^2	0.159	0.169	0.166	0.219	0.148
Accuracy of forecast	34	31	35	39	33
Error up to 1 grades	52	57	51	50	57
Error up to 2 grades	14	12	13	10	9

*, **, *** signify 10%-, 5%- and 1%-level of significance, respectively.
Source: own creation

For models using market indicators, as before, volatility of a company's share value exerted a negative influence. The influence of the ratio of share value to cash flow is positive. Among the balance indicators, size of the company (capitalization) retains its positive influence, as do probability of loan repayment with earnings and return on assets. The influence of the long-term debt as ratio to capital is also positive, but the sign in the model was due to the rather high negative correlation of this variable with the three previous ones.

The inclusion of macroeconomic indicators and consideration of the factors of industry and country affiliation raise the quality of the base model. The influence of macroeconomic factors on the rating is expected: negative for inflation and positive for the GDP growth indicator, which determines the level of stability of the external business environment. Developing countries dummy is among the explanatory variables and have positive influence.

Affiliation with developed countries in our research was not obviously a positive factor, which, of course, is connected with the correlation present between this element and the macroenvironment indicator. Russian companies are not significantly distinguishable from companies in other developing countries.
Among industries, the positive influence on the rating of affiliation with the oil and gas or utilities industries can be noted. Moreover, an analysis of average deviations shows that that tendency was stable.

The use of a stock market indicator (the S&P market model) is raising of the quality of the models for publicly-traded companies, that is, those with market quotations. In particular, there were in our distribution indicators of value volatility, level of systemic risk, ratio of share value to cash flow and a number of others, as well as an indicator of market discipline in the country where the company is located.

Systemic risk was insignificant in practically all the models examined. Volatility of value negatively influenced the level of the rating. Growth of share value in relation to cash flow has a positive influence on the rating. The positive influence of capitalization, the ratio of gross earnings to interest expenses and return on assets is preserved, as is the influence of macroeconomic indicators.

An analysis of the predictive power of the models was conducted by making a comparison of the true ratings of enterprises with their model values. The accuracy of the forecasts was approximately equal for the models examined. Accuracy of forecasts with an error of 1 gradation was about 90% and the accuracy of forecast was 39% and 33%, respectively. The accuracy of forecast with an error of no more than 2 gradations was higher than 99%, which is better than for models in the scale of classes.

5.2. Rating models for banks

Modeling ratings of financial stability for banks (bank financial strength rating, BFSR) depends on the same scheme. Selection of explanatory variables is made by both statistical and predictive criteria. The results of the appraisal of various models for choosing a bank profitability indicator are presented in (Peresetsky-Karminsky 2008). Unlike models for rating bank deposits, BFSR is not grouped in classes, but presented in grades.

A logarithm of assets is used as a variable to characterize the size of the bank. The debt-to-equety ratio characterizes the adequacy of the capital. The share of

overdue credits was chosen as a proxy for quality of assets. That variable has practically no correlation with other indicators except the provision for impairment of loans.

The following conclusions can be drawn from an analysis of the signs of the coefficients:

- Banks located in developing countries have a lower BFSR than banks in developed countries, and Russian banks are undervalued even in relation to banks in developing countries.
- The volume of assets influences the BFSR positively.
- Parameters reflecting efficiency (ratio of personnel expenses to operating income), the quality of assets (ratio of overdue debt to all debt) and ratio of client funds to own capital are significant; growth of personnel expenses and the portion of bad debts or financial leveraging have a negative influence on BFSR.
- The ratio of interest expenses to interest income and return on assets yielding interest income are indicators of profitability; net interest margin has the expected signs that reflect the fact that growth of profitability indicators encourages stability.

It should especially be noted that the coefficient before the value of interest liabilities is positive, which means that banks that pay more to obtain funds have a lower BFSR indicator. This conclusion coincides with a previous analysis (Karminsky-Peresetsky 2007). It is the value of resources that to a large extent determines the level of stability and efficiency of banks' activities.

The forecast of ratings based on the financial results of Russian banks in 2005, which were not used for the elaboration of the model, gave an exact forecast on the level of 56% and over 90% with 1 grade error scope.

The significance of the developing markets dummy should also be brought to attention, as well as differences in evaluations of Russian banks in comparison with banks of developing markets. This indicates the potential for Russian banks' ratings growth because of both the rise of the level of evaluation of business in Russia as a whole and the rise of the quality of the operating and regulatory environments, which have been a substantial hindrance to the rice of the evaluation of Russian banks and industrial companies. Such improvements were partially implemented in 2005-2008.

5.3. Comparative analysis of the corporate ratings of the two agencies

We also made a statistical comparison of the ratings of the S&P and Moody's agencies. We used a subsample containing observations of companies that have ratings from both agencies simultaneously. That sample amounted to 178 companies.

Three measures of difference were used for the comparison:

Δ - the difference between S&P and Moody's ratings,

FDS = $|\Delta|$ - the difference module,

SPLIT - a binary function that takes the value of 0 when the ratings coincide and 1 otherwise.

For each measure, we elaborated econometric models to determine the factors that significantly influence agencies' opinions. The results of the comparison are presented in Table 2.

Table 2. Comparison of the ratings of the S&P and Moody's agencies

	Model number				
	Difference Δ		**Difference module**		**SPLIT**
	1d	2d	1a	2a	1s
Return on assets	0.028**	0.022*			
	(0.011)	(0.011)			
Instant liquidity	-0.462***	-0.507***			
	(0.140)	(0.150)			
Fixed assets/Assets		-0.924**			
		(0.383)			
Share value/		-0.0098			
Cash flow		(0.0073)			
Value volatility			-0.007*	-0.006	-0.0038*
			(0.0036)	(0.0036)	(0.0023)
Inflation level	-0.221***	-0.156***			
	(0.056)	(0.060)			
Corruption index	-0.303***	-0.305***	-0.81		
	(0.060)	(0.058)	(0.053)		
Consumer sector	-0.857***	-1.084***	-0.169		
	(0.284)	(0.288)	(0.216)		
Developed countries			-0.572**	-0.838***	-0.309***
			(0.237)	(0.163)	(0.086)
Russia			-0.649***	-0.098	
			(0.243)	(0.341)	
Other insignificant				+	
Pseudo-R^2	0.182	0.217	0.127	0.140	0.061

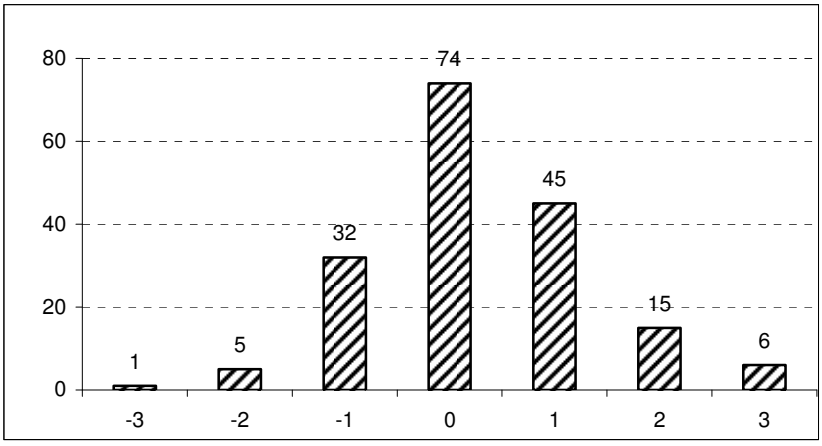
*, **, *** signify 10%-, 5%- and 1%-level of significance, respectively.

Source: own creation

The following conclusions can be drawn from this analysis of the table:

1. The most substantial differences are in the rating of companies from developing countries, which is expressed either directly through the dummy or indirectly as the influence of corruption.
2. Among the most significant and positive factors influencing the ratings of the Moody's agency, return on assets can be pointed out. For the S&P, factors such as instant liquidity, share of fixed assets in assets, level of inflation and corruption were more significant.
3. No substantial difference in the ratings of Russian companies was uncovered except indication of the large differences in the ratings on both the positive and negative sides (model 1a).
4. Growth of the volatility of companies' share value creates multidirectional differences, although not at a very high level of significance – 10% (models 1a and 1s). That indirectly confirms the previous conclusion.
5. The S&P agency takes a more critical stance toward companies from the consumer sector (models 1d and 2d). On average, the divergence between the agencies' ratings, expressed as their difference of the ratings, was 0.26 grades for our sample and is characterized by the bar chart in Fig. 4.

Fig. 4. Sample distribution of the difference between the ratings of the S&P and Moody's agencies



Source: own creation

6. Conclusion

Probability evaluations based on econometric models should be an integral part of internal rating systems, which determines the potential practical significance of such models, especially in developing countries. In this article, rating models of corporate ratings were elaborated based on multiple choice models. Financial indicators of corporations, dummies of regional and industry affiliation, market and macroeconomic indicators were used as explanatory variables.
It was shown that:

- A set of explanatory financial indicators is sufficient and are interpreted well to rating models.
- When other conditions are equal, industrial companies in developing countries receive lower ratings in comparison with companies in developed countries.
- The predictive power of corporate ratings models is somewhat better than bank deposit rating models.
- The degree of influence of country affiliation, return on assets, instant liquidity and inflation level are prominent factors that differentiate the approaches of the two agencies.

References

Altman E., 2005: An emerging market credit scoring system for corporate bonds. *Emerging Markets Review*, 6, 311-323. p.

Altman, E. - Rijken, H. 2004: How rating agencies achieve rating stability. *Journal of Banking & Finance*, 28, 2679-2714.p.

Altman, E. - Saunders, A. 1998: Credit risk measurement: Developments over the last 20 years. *Journal of Banking & Finance*, 21, 1721-1742. p.

Altman, E. - Suggitt, H. 2000: Default rates in the syndicated bank loan market: A mortality analysis. *Journal of Banking & Finance*, 24, 229-253. p.

Amato, J. - Furfine, C. 2004: Are credit ratings procyclical? *Journal of Banking & Finance*, 28, 2641-2677. p.

Bae, S.C. - Klein, D. P. 1997: Further evidence on corporate bonds with event-risk covenants: Inferences from Standard and Poor's and Moody's bond ratings. *The Quarterly Review of Economics and Finance*, 37, 709-724. p.

Basel, 2004: *International convergence of capital measurement and capital standards*. A revised framework. Basel, Bank for International Settlements, Basel Committee on Banking Supervision.

Basel, 2009: *Revisions to the Basel II market risk framework* - final version. Basel, Bank for International Settlements, Basel Committee on Banking Supervision.

Carey, M. - Hrycay, M. 2001: Parameterizing credit risk models with rating Data. *Journal of Banking & Finance*, 25, 197-270. p.

Carling, K. - Jacobson, T. - Linde, J. - Roszbach, K. 2007: Corporate credit risk modeling and the macroeconomy. *Journal of Banking and Finance*, 31, 845-868. p.

Curry, T. - Fissel, G. - Hanweck, G. 2008: Is there cyclical bias in bank holding company risk ratings? *Journal of Banking & Finance*, 32, 1297-1309. p.

Elton, E. - Gruber, M. - Agrawal, D. - Mann, C. 2004: Factors affecting the valuation of corporate bonds. *Journal of Banking & Finance*, 28, 2747-2767. p.

Feng, D. - Gourieroux, C. - Jasiak, J. 2008: The ordered qualitative model for credit rating transitions. *Journal of Empirical Finance*, 15, 111-130. p.

Frydman, H. - Schuermann, T. 2008: Credit rating dynamics and Markov mixture models. *Journal of Banking & Finance*, 32, 1062-1075. p.

Guttler, A. - Wahrenburg, M. 2007: The adjustment of credit ratings in advance of defaults. *Journal of Banking & Finance*, 31, 751-767. p.

Hanson, S. - Schuermann, T. 2006: Confidence intervals for probabilities of default. *Journal of Banking & Finance*, 30, 2281-2301. p.

Jacobson, T. - Linder, J. - Roszbach, K. 2006: Internal ratings systems, implied credit risk and the consistency of banks' risk classification policies. *Journal of Banking & Finance*, 30, 1899-1926. p.

Joo, S. - Pruitt, S. 2006: Corporate bond ratings changes and economic instability: Evidence from the Korean financial crisis. *Economics Letters*, 90, 12-20. p.

IMF 2009. *Global Financial Stability Report. Navigating the Financial Challenges Ahead* / October 2009, International Monetary Fund.

Kaminsky, C. - Schmucler, S. 2002: Rating agencies and financial markets. In *Ratings, rating agencies and the global financial system*. Editors R. Levich, G. Majononi and C. Reinhart. Boston, Kluwer Academic Publishers. 227-250. p.

Karas, A. - Schoors, K. - Lanine, G. 2008: Liquidity matters: Evidence from the Russian interbank market // *BOFIT Discussion papers*, No. 2008/19.

Karminsky, A. - Perestsky, A. 2007: International agencies rating models. *Prikladnaya ekonometrica* (Applied econometrics), 1, 5, 1-17. p.

Karminsky, A. - Perestsky, A. 2009: Ratings as financial risks measure. Evolution, purposes, applications. *Journal of New economic association*, No. 1-2.

Karminsky, A. - Perestsky, A. - Petrov, A. 2005*: Ratings in economy: methodology and practice*. Financi i Statistika (Finance and Statistics Publishing House), 240. p.

Kish, R. - Hogan, K. - Olson, G. 1999: Does market perceive a difference in rating agencies? *The Quarterly Review of Economics and Finance*, 39, 363-377. p.

Moody's 2007: *Incorporation of joint-default analysis approach into Moody's bank ratings: a refined methodology*. Moody's Investors Service, www.moodys.com.

Moody's 2009: *Moody's Financial Metrics. Key Ratios by Rating and Industry for Global Non-Financial Corporations*. Moody's Investors Service, www.moodys.com.

Niemann, M. - Schmidt, J. - Neukirchen, M. 2008: Improving performance of corporate rating prediction models by reducing financial ratio heterogeneity. *Journal of Banking & Finance*, 32, 434-446. p.

Packer, F. 2002: Credit ratings and the Japanese corporate bond market. In Levich, R. – Majononi, G. – Reinhart, C. (eds): *Ratings, rating agencies and the global financial system*. Kluwer Academic Publishers, Boston, 177-201. p.

Partnoy, F. 2002: The Paradox of Credit Ratings. In Levich, R. – Majononi, G. – Reinhart, C. (eds): *Ratings, rating agencies and the global financial system*. Kluwer Academic Publishers, Boston, 65-84. p.

Pederzoli, Ch. - Torricelli, C. 2005: Capital requirements and business cycle regimes: Forward-looking modeling of default probabilities. *Journal of Banking & Finance*, 29, 3121-3140. p.

Peresetsky, A. - Karminsky, A. 2008: Models for Moody's bank ratings. *BOFIT Discussion Papers,* 2008/17.

Reinhart, C., 2002: Sovereign ratings before and after financial crisis. In Levich, R. – Majononi, G. – Reinhart, C. (eds): *Ratings, rating agencies and the global financial system*. Kluwer Academic Publishers, Boston, 251-268. p.

Rojas-Suarez, L., 2002: Rating banks in emerging markets: What credit rating agencies should learn from financial indicators? In Levich, R. – Majononi, G. – Reinhart, C. (eds): *Ratings, rating agencies and the global financial system*. Kluwer Academic Publishers, Boston, 177-201. p.

Servigny, A. – Renault, O. 2004: *Measuring and managing credit risk*. McGraw-Hill, N.Y. etc., 467 p.

S&P 2008. *Global economic decline increases Russia's country risk*. Standard & Poor's. www.standardandpoors.ru.

S&P 2009. *Improved methodology for assigning ratings to organizations linked to the state.* Standard & Poor's. www.standardandpoors.ru.

The use of infection models in accounting and crediting

András Csernenszky[1] *–Gyula Kovács*[2] *- Miklós Krész*[3] *– András Pluhár*[4] *– Tamás Tóth*[5]

Recently one of the main directions of data mining is the study and use of network data. Our research is concentrated on the network data about the links of big and medium size of companies that can be deduced from the bank transactions. The main goals are to develop models to predict churn and bankruptcy. We implement the Domingos-Richardson cascade model, and for the parametrization and evaluation we use the database of the OTP Bank. The results suggest that the developed system is capable of supporting a wide range of applications of network problems such as churn, bankruptcy, campaign management, information diffusion etc.

Keywords: Graph mining, Influence models, Consumer value

1. Introduction

Earlier results on network processes suggest that it is worthwhile to examine the transactions between corporate clients. We believed that if something happens with a company's supplier or purchaser, it obviously has some effect on its business partners. By using a transaction data between corporate clients, we modeled the spreading of Basel II default events on a bank's corporate portfolio. To simulate this process, we used the DR cascade model with appropriate parameters. Our results show that the bankruptcy forecasting can be greatly improved by this method, provided with a careful parameterization. The research is based on the OTP's corporate transaction database and the actual computations were done by the commercial graph-mining software of Sixtep Ltd.

[1] András Csernenszky, OTP Bank Hungary
[2] Gyula Kovács, Sixtep Ltd
[3] Miklós Krész, University of Szeged
[4] András Pluhár, University of Szeged
[5] Tamás Tóth, OTP Bank Hungary

2. Network research in the Corporate Banking

There is a lot of area for network modeling in the operations of a Corporate Bank. For the market development these are the acquisition, estimating customer value, forecasting attrition (performing at the OTP), product development and sales support and campaign optimization. There are also some applications for the credit risk department such as forecasting arrears, using networks for segmenting risk-groups, money laundering investigations, mapping the client's total customer-supplier relationship for business purposes. Here we concentrate on risk decisions especially the bankruptcy forecasting (till 2009 September performing at the OTP) – this is the main topic of this paper.

3. The Independent Cascade model

First we have to understand the role of modern network theory in Economics. The theory of graph is a well developed subject with plenty of beautiful theoretical results, applications and algorithms (Bollobás 1998). It was noted only recently, that some very large, but important graphs, the so-called *Small World graphs*, have characteristics that have not been explored in the classical studies, see Albert - Barabási (2002), or Newman (2003) for comprehensive introduction.

These graphs may arise by mapping the links among people or companies that already indicates their significance in the investigation of epidemics, spread of information or economical troubles and so on (Boguna et al. 2003), (Diekmann - Heesterbeek 2000). However, these works are based on SIS or SIR models; in which repeated infections and recoveries are both possible. In our case, predicting default, no recovery is possible, and a node might get infected *without* outer influence with an apriori probability, depending on the properties of the node itself.

So let us introduce our main tool, a model that tries capturing the process when such an effect propagates on a network. It was invented by Domingos – Richardson (2001). Originally it was proposed to support marketing decisions and determining client values. Nevertheless, Kempe et al. (2003, 2005) showed it is equivalent to a model given by Granovetter (1978), which shows these models can grasp a great variety of phenomena.

However, it is not quite obvious that it might be readily used in solving finance decisions. First of all, one has to define and built up a network (graph) with weighted edges that estimate the probability of one node infecting another. We shall give some of the details of that work later. The other problem is the arising computational issues. In order to get the approximated default probabilities, and consequently the expected value of default, one has to run a large number of Monte Carlo simulations on enormous size of network. As there had not been such applications available on the market, we decided to develop such code.

We give a sketch of the Independent Cascade model (Domingos-Richardson 2001). For a network *G* and probabilities assigned to the edges of *G*, and a set of *active (infected)* vertices *A*, the initial infection will affect the other vertices in the future. The time is modeled by discrete steps. In each steps the vertices that got the infection in the previous step might infect healthy vertices connected to them with probabilities assigned to the connecting edges. (The infection by different vertices is considered to be independent.) The process goes until there are newly infected vertices, otherwise halts. In a pseudo code:

1. Infected dataset = Active dataset A
2. The new infected vertices are these ones which are infected by the edges where one of the vertices is "active." The probability that a healthy vertex v stays healthy at the given step is the product $\prod_{v \in A} q_{uv}$, where u is a neighbor of v infected in the previous step. Here $q_{uv} = (1 - p_{uv})$ and p_{uv} is the infection parameter of the edge uv.
3. Vertices infected in the previous period = Active dataset;
4. If there is no new infection then STOP, else back to 2.

Note, that we have generalized the original Independent Cascade model such that not only an initial infected set can be given, but a probability distribution describing the a priori infection probability of the vertices.

4. Parameter estimation

Dataset:

There are several possibilities to compute the parameters within the model in order to get the best estimation for the bankruptcy probability of a corporate client. Our first task is to define the transaction dataset, in which lots of hard question have to be answered. Which is better to analyze all corporate clients with all of its transactions or just the debtors? Can we build up a meaningful weighted network drawing information out of our data warehouse? Certainly the vertices should be the clients, edges (some of the) transactions and the weights have to correspond to the transfers on those edges. Which length of time period should we monitor looking for transaction, and upon which assumptions should we declare there is an edge between the vertices representing two clients? How old basel2 default events have effects on a company's business partners, how long does it hold, and how big is it?

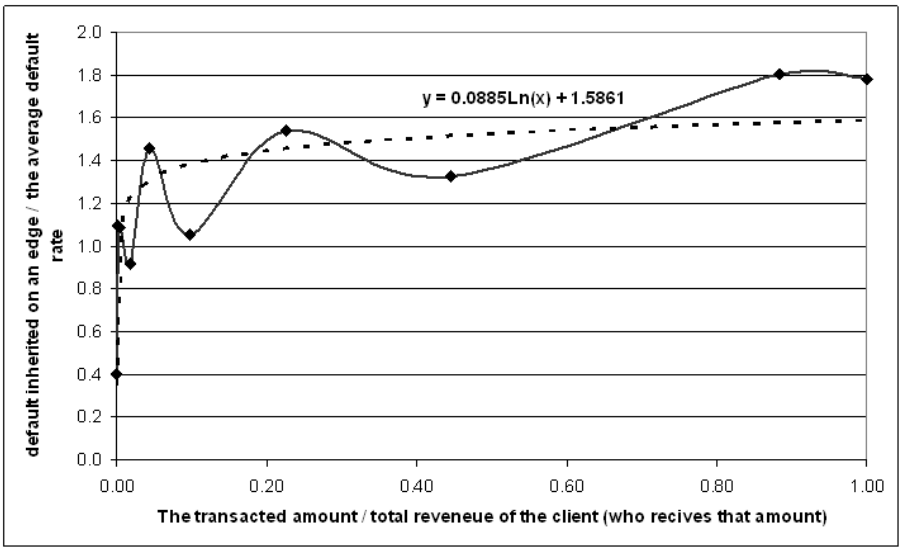
Our research clearly showed the followings. It was better for us to use just the debtors network because its homogeneity (it both means that the credit portfolio is divers from the account holders portfolio, and only just here has the default definition any sense). The other parameters were more flexible, but after some trials we

have reached the consensus that the business relation-graph should be based on at least 6 month, but not more than one year transactions data. Furthermore at most 3 months basel2 default event observation needed for the optimal result. In practice it means that to get a fresh report for November 2009, then the edges of the transaction graph should be based on the data from February 2009 to July 2009, and the default events are to be observed from August 2009 to October 2009. For the parameter estimations we used data from the previous year to re-measure its effect on a one year time-interval.[6] In the finally examined database there were 21 696 corporate client with 34 388 connections.

Vertices:
There are at least three approaches about what should be the initial influence distribution of the vertices. The first is to simply write 1 if the company is bankrupted in a given time period and 0 otherwise. (This corresponds to the original Independent Cascade model, where a vertex is either healthy or infected.) The other is to write the apriori default probabilities, the score values, to the vertices. For that case we generalized the model allowing fractional infection in the input. Our research confirmed that the second one yields better solution, increasing the efficiency on the reported target segment by 25 percent.
Another possible solution is to use the first method (0 and 1 to the vertices) and after multiplying the influences by the apriori probabilities. It has somehow different meaning; however it still increases the efficiency with 20 percent. In spite of its better classification we use for reporting the second method, because here the companies with higher influence rates have more (direct) bankrupted business partner and hence the interpretation and the acceptance of the results are easier.

[6] Our transaction graph comes from the data between January 2008 and Jun 2008, the default events from July 2008 to September 2008, while the measuring period is from October 2008 to September 2009.

Figure 1. Curve-fitting by a one variable function



Edges:

To get a useful model we also have to deal with the edge parameters. Here the main question is: 'How to model the probability of influence on the edges?' An obvious answer is to assign the average probability of influence on the edges. It means that for the bankrupted company we considered those companies that were directly connected to it and also bankrupted within the following twelve months, estimated the influence probability with the ratio of the bankrupted and all companies.

However, for the estimation of the infection probability there are better functions than a linear function of that ratio. Instead of the ratio, we took the ratio of the transacted amount among the clients per the total revenue of the client (the receiver) on that period. By scaling the function on the axis and the re-measured effect is on the other axis, we found that a logarithmic curve fits best to it;[7] see it on Figure 1.

To sum up the results, we can see the effects of different treats in Figure 2. For example if we write 0-1 to the vertices, but we use a one variable sigmoid function to estimate the influence of the edges we get 2.72 times more defaulted company at the top 10% percent of the portfolio than the average default rate.

[7] We tried other curve fitting methods like linear, polynomial, exponential, power fitting, etc…, but the R^2 here was the largest with 83%.

Figure 2. The classification power of various treatments

The top 10%'s default-rate per the average default-rate	Constant (as influence value) on the edges	One variable sigmoid function (as influence value) on the edges
Default/ non default flag (0,1) on the vertices	2,18	2,72
Writing score values on the vertices, and 1 if it is in default	3,82	4,25

Source: own creation

5. Results:

Using the parameterized Independent Cascade model, we found segments where the expected bankruptcy is 3-4 and even 10 to 12 times of the average. Of course there are some other standard variables that are used to assess the risk; such as the size of a given company, the importance its sector in the Hungarian economy (and even weather it is a municipality or not). Incorporating all these predictions, a monthly report is being built, which is installed to the OTP monitoring system that is used by the credit monitoring and the credit controllers department. According to the preliminary results, by the improved monitoring system the bank can significantly reduce the loss on bankruptcies.

References

Albert, R. – Barabási, A. L. 2002: Statistical mechanics of complex networks. *Reviews of Modern Physics*, **74**.

Boguna, M. - Pastor-Satorras, R. – Vespignani, A. 2003: Absence of epidemic threshold in scale-free networks with degree correlations. *Phys. Rev. Lett*, Vol. **90** Issue **2.** pp. 028701-1—4.

Bollobás, B. 1998: Modern Graph Theory. *Springer*, New York.

Diekmann, O. - Heesterbeek, J. A. P. 2000: Mathematical epidemiology of infectious diseases: Model Building, Analysis and Interpretation. *John Wiley & Sons*, New York.

Domingos, P. - Richardson, M. 2001: Mining the Network Value of Costumers.*7th Intl. Conf. on Knowledge Discovery and Data Mining*.

Granovetter, M. 1978: Threshold models of collective behavior. *American Journal of Sociology* 83/6, 1420-1443. p.

Kempe, D. - Kleinberg, J. – Tardos, E. 2003: *Maximizing the Spread of Influence though a Social Network*. Proc. 9th ACM SIGKDD Intl, Conf. on Knowledge Discovery and Data Mining.

Kempe, D. - Kleinberg, J. – Tardos, E. 2005: Influential Nodes in a Diffusion Model for Social Networks. Proc. 32nd International Colloquium on Automata, *Languages and Programming* (ICALP).

Newman, M.E.J. 2003: *The structure and function of complex networks*. Preprint cond-mat/0303516.

Dependence structure analysis using Copula-GARCH model: An application to UK and US stock markets

Ling Xiao[1] *- Gurjeet Dhesi*[2]

The relationship between different international stock markets is of importance for both financial practitioners and academicians in order to manage risks. Especially after the financial crisis, the pronounced financial contagion draws the public attention to look into such associations. However, measuring and modelling dependence structure becomes complicated when asset returns present nonlinear, nongaussian and dynamic features. In this paper, we firstly investigate volatility spillover effect between FTSE100 and S&P500 stock indices. Strong lagged volatility of stock market itself and asymmetric spillover effect between UK and US stock markets are found out based on the multivariate GARCH-BEKK model. We also take a pilot study based on two step Copula-GARCH model to examine the correlation and tail dependence of returns. Some interesting results of co-movement between UK and US stock markets are discussed.

Keywords: Copula-GARCH, dependence structure, time-varying, volatility spillover effect

1. Introduction

Since the worldwide 2007- 2009 financial crisis or maybe even earlier, academicians and financial practitioners could not stop doubting the effectiveness of risk measurement and management of the financial market(s). It has been acknowledged that associations (co-movement) cross international stock market returns are of importance to measure and manage risks (Longin and Solnik, 1995). Dependence structure describes the relationship between risks and provides an estimation of risks. A deep understanding of dependence structure would help financial practitioners to

[1] Ling Xiao, Business Studies Department, London South Bank University (London)
[2] Dr Gurjeet Dhesi, PhD student, Senior Lecturer, Business Studies Department, London South Bank University (London)
Acknowledgement: The authors acknowledge Dr Lee Rose and Professor Jon Warwick for their valuable comments and informative feedback. We also thank the Business Studies Department, Faculty of Business at London South Bank University for financial support. .

make their financial decisions and financial regulators to control the financial contagion.

Traditional approaches measure dependence by linear correlation (namely Pearson's correlation coefficient) because linear correlation measure is straightforward to calculate. However, the use of linear correlation is only a symmetric, linear dependence metric (Embrechts *et al*. 2001). All the classical theories of risk measurement and management have been built on the primary assumption of multivariate normal independent identically distributed (i.i.d) return distributions. On contrary to the classical portfolio theory, modern asset returns present nonlinear, nongaussian and dynamic features. In addition, the ongoing unpredictable and changeable financial market scenario suggests that this kind of phenomena is likely to happen much more frequently. Accordingly, linear correlation measurement may induce misinterpretation when applied to nonlinear associations and tail dependency systems.

A powerful tool borrowed from mathematics namely copula functions (hereafter, referred as copulas) are capable to summarize the dependence structure between risks without the standard linear dependence and multivariate i.i.d return distribution restrictions. Copulas are defined as: functions that join one-dimensional distribution functions together to form multivariate distribution functions (Sklar 1959). Copulas not only solve the non-linear, non-elliptical problems, but also extend the bivariate world to multivariate dimension which gives us the ability to take all the risks into account as a whole unit. There were very few practical applications of copulas before the 1990s. As a consequence, the investigation of copulas has been restricted within the field of mathematics until the late 1990s when a number of statisticians showed an interest in implementation. Nelson (1999), who gave a thorough description of copulas from a mathematic perspective, is one notable example. Before too long, some academicians detected the potential of applying copulas in the field of finance. As soon as copulas and their applications were introduced; they became a favourite research trend of dependence in finance. Embrechts *et al* (1999, 2001) are among pioneers who modeled dependence with copulas and employed them to risk management. In terms of asset pricing with copulas, Cherubini *et al.* (2004) made a seminal contribution to the advent of pricing multivariate option by using copulas.

Jondeau, Rockinger (2006) advocated Copula-GARCH models which brought the advent of dynamic copula period. Copula-GARCH models are the class of models where some of copula parameters are potentially time-varying, in an autoregressive manner, conditional on the set of past information. Unlike the common normal and Archimedean copulas which have the constant parameters,

Copula-GARCH models have time-varying dynamic parameters which would be naturally more suitable for financial time series data. Based on Copula-GARCH models, time varying conditional correlations will be provided rather than a constant correlation coefficiency over time period. In this case investors and policymakers are able to obtain the boundary of market fluctuation which would be helpful for them to analyse and diversify risk. Patton (2006a) proposed a practical two-stage maximum likelihood method to estimate copulas. Patton (2006b) introduced static and time-varying symmetric Joe-Clayton (SJC) copulas which allow different tail dependence. Aas *et al* (2009) introduce copula vines: a decomposition of a multivariate copula to a product of bivariate copulas.

In order to understand the linkages across the international stock markets, in this paper we tentatively investigate volatility spillover effect between FTSE100 and S&P500 stock indices. Strong lagged volatility of each stock market itself and asymmetric spillover effect cross different stock markets are found out based on the multivariate GARCH-BEKK model (Engle-Kroner, 1995). We employ a two step Copula-GARCH model to examine the dependence structure of daily stock markets returns. Firstly, we filter log-return daily data using univariate GARCH model to obtain standard residuals and construct the marginal distributions. Secondly, a couple of static and time-varying copulas are selected to join the estimated marginal distributions. The Akaike information criteria (AIC) and Bayesian information criteria (BIC) method are then adopted to determine which copula provides best fitness to the market data. Finally, some interesting results of co-movement between different stock markets are discussed.

2. Methodologies

2.1 Multivariate GARCH model (MV- GARCH)

It is a common belief that financial volatilities move together across markets over times. The GARCH model is regarded as an important implement to describe volatilities clustering of financial returns. MV-GARCH models extend the univariate GARCH model to multivariate dimensions which enable us to investigate spillover effects between different markets efficiently. Therefore we employ MV- GARCH model to analyse volatility movements between the UK and US stock markets. Engle and Kroner (1995) made improvement on the work of Baba, Engle, Kraft and Kroner (1988) and created a new simplified multivariate BEKK model which has been widely adopted quickly. The BEKK model can be represented as follows:

Mean equation

$R_t=C+\Theta R'_{t-1}+e_t \quad e_t \sim N(0, H_t)$ ……..(1)

R_t is a T*1 vector, R_t is the daily logarithm return of each stock index. H_t is the variance-covariance matrix and it is constrained to be positive because the BEKK model uses a quadratic form for the parameter matrices to ensure a positive definite variance / covariance matrix H_t which is expressed as:

Variance-covariance

$$Ht=\Omega\Omega'+A\,(e_{t-1}e'_{t-1})A'+BH_{t-1}B' \qquad (2)$$

The error term e_t is defined as $e_t \sim N(0, H_t)$. It is suggested that student t distribution maybe the preferable one when we are dealing with the financial time series returns. Accordingly we redefine $e_t \sim ID(0, H_t, v)$ where *v* is degree of freedom and thus T-BEKK model is employed in this empirical study.

2.2 Copula-GARCH

2.2.1 Copulas

Copulas are defined as: functions that join one-dimensional distribution functions together to form multivariate distribution functions (Sklar 1959). Copulas not only solve the non-linear, non-elliptical problems, but also extend the bivariate world to multivariate dimension which gives us the ability to contain all the marginal distribution as a whole dependence structure. Generally speaking, there are two main types of copulas namely: implicit (elliptical) copula and Archimedean copulas. Each of them have their own family members and distinguishing properties. The elliptical copulas are known to perform better on systematic dependence structure problems. They constitute of Gaussian copula and student t copula. Archimedean copulas namely: Clayton copula and Gumbel copula captures lower tail dependence and upper tail dependence respectively (Alexander 2008). Empirical studies show that Archimedean copulas fit market data better than the elliptical copulas (Melchiori *et al*. 2003; Chen *et al. 2007,* Koziol-Kunisch 2005). We select Gaussian, Student t, Clayton and SJC copulas as candidates in the empirical study. The reasons are as following: firstly Gaussian copula is the standard copula tool although it does not consider the possibility of tail dependence which measures the joint probability of extreme events. Secondly, since the financial series data are known to have fat tails, Student t, Clayton and SJC copulas which all consider tail dependence are selected.

2.2.2 Estimation of Copula-GARCH

The method that we employ utilizes two stages maximum likelihood method to estimate Copula-GARCH. We start by constructing marginal distribution by fitting standard residuals data with unviariate GARCH model. Since financial time series return are known to be heteroscedastic and often autocorrelated, proper filtration is required. There are many types of conditional volatility GARCH models; in general they have been categorized as symmetric GARCH or asymmetric GARCH models. The latter satisfies daily data for equities(indices) and commodities at daily frequency because equity market volatility increases are more significant following a large negative return than when they are following a positive return of the same size which is so-called leverage effect(Alexander 2008). The GJR (Glosten, Jaganathan and Runkle, 1993) model modifies the conditional variance equation by introducing an extra factor—'leverage' parameter. Thus we use the GJR model which could reflect the asymmetric effect of how negative shock has greater impact on volatility than positive shock. The mathematical equations of the GJR are:

$$\varepsilon_t = h_t \bullet z_t, z_t \sim \mathrm{N}(0,1) \tag{3}$$

$$h_t = \varpi + \alpha \bullet \varepsilon_{t-1}^2 + \beta \bullet h_{t-1} + \gamma \bullet \varepsilon_{t-1}^2 \bullet \delta_{t-1} \tag{4}$$

where

δ_t is equal to 1 when $\varepsilon_t < 0$

δ_t is equal to 0 when $\varepsilon_t > 0$

The mean equation for AR (1)-GJR (1) is given by

$$r_t = c + s \bullet r_{t-1} + \varepsilon_t \tag{5}$$

Subsequently copulas are now used to construct the joint distributions. We follow Chen *et al* (2007) to select Kendall's *tau* and Spearman's *rho* as candidates of rank correlation statistics in our study. Then the correlation parameters corresponding to each copula are calculated based on the estimated Kendall's *tau*. Copulas estimation asks for uniform distribution data, the sample data should be transformed firstly. With the uniform distributed data, Maximum Likelihood Estimation (MLE) is then applied to estimate copula parameters:

$$L(\theta_c) = ArgMax_{\theta_c} \sum_{t=1}^{T} \ln C_t(F_1(x_{1t)}), F_2(x_{2t}),...., F_n(x_{nt}); \theta_c) \quad (6)$$

where $F_n(x_{nt})$ are marginal distributions.

3. Data description

We choose the daily closed price from 5th Jan, 2004—31st, Sep, 2009 of FTSE100, and S&P500[3] stock indices as our observations. Logarithmic daily return of stock market (as a percentage) is defined as:

$$\gamma = 100 \ln(\frac{P_t}{P_{t-1}}) \quad (7)$$

P_t is the daily closed price at time period t, basic statistics descriptive is shown in Table 1. According to the Jarque-Bera statistics, it is obvious that both of these two stock indices are non normal distributed. In particular, the FTSE100 is mildly right skewed while the S&P500 is negatively skewed. In addition, as evidenced by the large kurtosis value, the logarithmic daily return series show a strong leptokurtic feature. Furthermore, the results of Ljung-Box test reveal that returns series are also serially autocorrelated. The relationship between the mean and the median of FTSE100 provides possible evidence against the positive skewness. According to many standard text books the mean is usually located to the right of the median in the presence of positive skewness. However, in practise this rule often fails to apply. In particular for this FTSE100 return series distribution, areas to left and right of mean are not equal, and where one tail is long but when the other is heavy.

[3] Financial times stock index (FTSE100), Standard and Poor500 (SP500).

Table 1. Summary of Descriptive Statistics

	FTSE100	**S&P500**
Mean	0.008353	-0.006410
Median	0.054070	0.078815
Std. Dev.	1.395950	1.476607
Skewness	0.122954	-0.291525
Kurtosis	14.42809	13.23028
Jarque-Bera	7295.278	5862.420
Probability	0.000000	0.000000
Q(4)	35.37[0.39e-6]	34.08[0.72e-6]
Q(12)	89.56[0.00e-6]	54.34[0.24e-6]
Q^2 (4)	547 [0.00]	547.6 [0.00]
Q^2(12)	1176.3 [0.00]	1920.7 [0.00]

Source: own creation

4. Empirical results and analysis

4.1 The T-BEKK model

In the BEKK model, the conditional variance is not only a function of all lagged conditional variances and squared returns, but also a function of conditional covariance and cross-product returns. The diagonal elements in the parameter matrix B measures the effect of lagged volatility; the off-diagonal elements capture the cross market effects (Zhang 2009). We firstly carry out the bivariate T-BEKK estimation to examine the interactions between the UK and US stock markets.

Table 2. BEKK Estimation

	FTSE100			S&P500	
	Estimates	**Standard Error**		**Estimates**	**Standard Error**
$\mathbf{C_{11}}$	0.1929	0.21	$\mathbf{C_{21}}$	-0.0576	0.026
$\mathbf{C_{12}}$	0	0	$\mathbf{C_{22}}$	0.0002	0.008
$\mathbf{A_{11}}$	0.3192	0.0046	$\mathbf{A_{21}}$	-0.4325	0.005
$\mathbf{A_{12}}$	0.2281	0.0027	$\mathbf{A_{22}}$	-0.0198	0.0031
$\mathbf{B_{11}}$	0.841	0.0033	$\mathbf{B_{21}}$	0.1119	0.002
$\mathbf{B_{12}}$	-.0.0383	0.0041	$\mathbf{B_{12}}$	0.9953	0.001
Degree of Freedom		10.589	**Log-likelihood**		-3510

Source: own creation

First of all, the diagonal elements (B_{11}, B_{22}) all pass the 5% significance test which means strong lagged volatility exists in each stock market. Secondly, the volatility spillover effect is not symmetric. B_{12} (0.0383) and B21 (0.1119) suggest S&P500 transmit nearly 11.2% volatility shock to FTSE100 while only one third the other way around. It implies that S&P500 is dominating the volatility transmission effect between FTSE100 and S&P500.

4.2 Copula-GARCH

4.2.1 AR (1)-GJR model

The results of two marginal distributions are presented in Table 3. The value of the parameters for the variance in Eq. (4) and the return Eq. (5) display similar patterns. The leverage effect parameter γ are significant with estimates (0.16, 0.12) for FTSE100 and &P500 respectively. All the values are significantly positive, which indicates the existence of these effects. However, the leverage effects presenting in the U.S stock market seems to be weaker. It is worth to point out that persistence parameter β is significant for both indices which suggests that variances H_t rely on previous time period variance H_{t-1}. Furthermore, the null value of ARCH (1) estimates can be explained by the fact that in the GJR model, the reaction effects

have been taken into the leverage side.

Table3. Conditional marginal models estimation

	FTSE100		**S&P500**	
	Estimates	**Standard Error**	**Estimates**	**Standard Error**
C	0.021223	0.023123	0.0049888	0.026
AR(1)	-0.04504	0.0314	-0.07327	0.008
K	0.013301	0.0043661	0.012714	0.005
GARCH(1)	0.91904	0.015548	0.92962	0.0031
ARCH(1)	0	0.020629	0	0.0041
Leverage(1)	0.16191	0.032267	0.11397	0.001
Degree of Freedom	7.9116	2.0622	8.2209	1.4257

Notes: C is parameter c, AR (1) is the s, k is *w*, ARCH (1) is α, GARCH (1) is βand *r* is leverage effect respectively in Eq.(4) and Eq.(5) .

Source: own creation

4.2.2 Copula-GARCH estimation

With the conditional marginal model we are now in a position to estimate copulas. In this section we present the results of the six copulas estimation in Table 4. Compared with the copulas with constant parameters, the time-varying copulas have the lower negative log-likelihood which demonstrates that the time-varying copulas performance better. In other words, it tells us that the dependence structure is changing over the time since time-varying copulas improve the performance of the static copulas. Secondly, relied on the AIC, BIC value we conclude that the SJC copula is the most appropriate one.

Table 4. Log-likelihood Copula Estimation

FTSE100/SP500	Copula Type	Log-Likelihood	AIC	BIC
Constant	**Gaussian**	-216.346	-432.691	-432.687
	Student t	-218.65	-437.297	-437.289
	Clayton	-165.937	-331.873	-331.869
	SJC	-218.493	-436.981	-436.97
Time-varying	**Gaussian**	-218.443	-436.881	-436.87
	SJC	-221.998	-443.987	-443.964

Source: own creation

5. Conclusion and Future works

We examine the dependence structure between the UK and US stock market from two aspects: T-BEKK has been employed to investigate the volatility spillover effects between different markets. We find evidence that a volatility spillover effect significantly exists between the UK and US stock markets. Moreover the spillover effect is asymmetric; the S&P500 dominates the volatility effect. Furthermore Copula-GARCH methodology has been used to describe the dependence structure between different stock market. We conclude that the time-varying copulas perform better amongst all the six copula candidates. Especially the SJC time-varying copula accommodates difference in upper and lower tail dependences and so improves the log-likelihood of estimations. Future work will be done in two strands. Firstly, volatility spillover effects and dynamic conditional correlations among international stock markets are being investigated by a variety of multivariate GARCH models. Secondly, based on this pilot study work on GARCH-Copula applications to dynamic linkages between international financial markets is to be further developed. Moreover in light of the long, heavy tail feature of the return series tail dependences presented among different markets will be examined by SJC copula. Details and results would be seen in our forthcoming publications.

References

Alexander, C. 2008: "MarketRisk Analysis" Volume II. *Practical financial econometrics* 2008 redition, Jon Wiley and Sons Ltd.

Aas, K. - Czado, C. - Frigessi, A. - Bakken, H. 2009: Pair-copula constructions of multiple dependence. *Insurance, mathematics and economics,* 44, 182-198. p.

Bollerslev, T. - Engle, R.F. – Wooldridge, J.M. 1988: A capital asset pricing model with time varyng covariance. *Journal of Political Economy*, 96, 116-131. p.

Chen, Y-H. – Tu, A.H. - Wang .K. 2007: Dependence structure between the credit default swap return and the kurtosis of the equity return distribution: Evidence from Japan. *Journal of International Financial Markets, Instituions&Money, 18., 259-271 pp.*

Cherubini, U. – Luciano, E. – Vecchiato, W. 2004: *Copula Methods in Finance* Chichester: John Wiley and Sons.

Clayton, D. G. 1978: A Model for Association in Bivariate Life Tables and its application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence. Biometrika N. 65, 141-151. p.

Embrechts, P. – McNeil, A. J. – Straumann, D. 1999: Correlation and Dependence in Risk Management: Properties and Pitfalls Preprint ETH Zurich, available from http:// www.math.ethz.ch/embrechts

Embrechts.P. Lindskog.F. and McNeil, A. J. 2001: Modelling Dependence with Copulas and Application to Risk Management Preprint ETH Zurich, available from http:// www.math.ethz.ch/finance

Engle, R. F. – Kroner, K. F. 1995: Multivariate simultaneous generalized ARCH. *Econometric Theroy*, 11, 122-150. p.

Gumbel, E. J. 1960: Bivariate Exponential Distributions. *Journal of the American Statistical Association,* 55, 698-707. p.

Jondeau, E. – Rockinger, M. 2006: The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance,* 25, 827-853. p.

Glosten,L.R., R. On the relationship between the expected value and the volatility of the nominal express returns on stocks. *Journal of Finance,* 48, 1779-1801. p.

Koziol, P. – Kunisch, M. 2005: Analysis of copula functions and applications to credit risk management. *Financial Modelling Workshop,* Ulm, 2005

Longin, F. - Solnik, B. 1995: Is the correlation in international equity returns constant: 1960-1990? *Journal of International Money and Finance*, 14, 1, 3-26. p.

Melchiori, M. R. 2003: Which Archimedean copula is the right one? Yield Curve,

2003.

Nelson, R. B. 1999: *An introduction to Copulas*. Springer, New York.

Patton, A. J. 2006a: Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, 21, 2, 147-173. p.

Patton, A. J. 2006b: Modelling asymmetric exchange rate dependence. *International Economic Review*, 47, 2, 527-556. p.

Sklar, A. 1959 : Function de repartition a n dimensions et leurs marges. *Publications de I' Institut deStatistique de l'University de Paris* 8, 229-31. p.

Zhang, S. Q. – Paya, I. – Peel, D. 2009: Linkages between Shanghai and HongKong stock indices. *Lancaster University Management School Working Paper.* 2009.

The impact of the US Stock Market on the Romanian Stock Market in the context of the Financial Crisis

Costel Nistor[1] *– Razvan Stefanescu*[2] *- Ramona Dumitriu*[3]

This paper explores the stock market interlinkages between the United States and Romania during the actual financial crisis. For this purpose we analyze, in a Vector Autoregressive framework, daily values of Dow Jones and BET, being two reference indexes for the US and the Romanian Stock markets. By comparing with the results for a more tranquil period of time, we conclude that in the context of the financial crisis the Romanian stock market became more sensitive to the US stock market evolution.

Keywords: Financial Linkages, Romanian Stock Market, Crisis, Vector Autoregressive Model

1. Introduction

It is widely admitted that in the recent years stock markets from around the world became more integrated. Several circumstances led to this evolution: the practices of international portfolio diversification, the international financial markets deregulation, the abandon of Bretton Woods Monetary System, the financial innovations proliferation, the new technologies in communication and information, the European integration a.s.o. (Sharpe 1964, Lintner 1965, King et al. 1994, Kasa 1992, Roca 2000, Kaminsky et al. 2001, Forbes and Chinn 2004). The financial linkages between the stock markets are materialized not only in the changes of returns, but also in the transfer of volatilities (Kyle 1985).

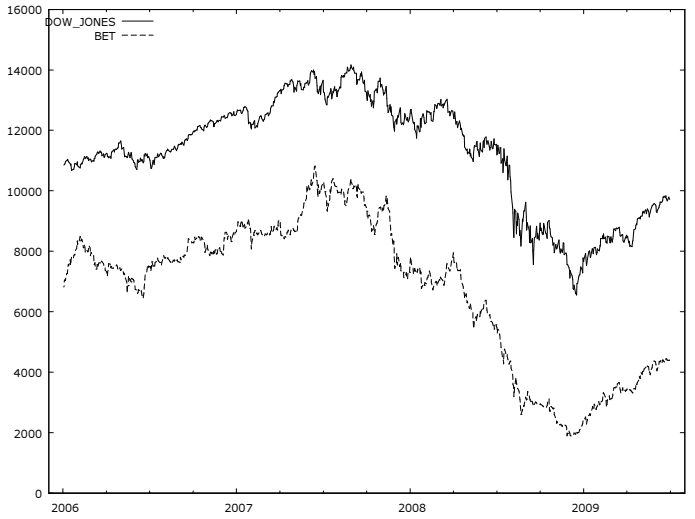
Some studies approached the particularities of the financial linkages between the stock markets due to specific circumstances. There were revealed the considerable influences of the US stock market on the financial markets from other countries

[1] Costel Nistor, PhD, assoc. prof, University "Dunarea de Jos" Galati, Faculty of Economics, Romania
[2] Razvan Stefanescu, PhD, lecturer, University "Dunarea de Jos" Galati, Faculty of Economics, Romania
[3] Ramona Dumitriu, PhD student, lecturer, University "Dunarea de Jos" Galati, Faculty of Economics, Romania

(Janakiramanan and Lamba 1998, Hsiao et al. 2003, Gilmore & McManus 2002). The linkages between the emerging markets and the financial markets from the developed countries depend on the role of the foreign investors (Enn and Shim 1989, Ferson and Harvey 1995, Masih and Masih 1997, Longin and Solnik 2001). In the Eastern European emerging markets case the perspective to become members of the European Union raised the international investors' interest. In this context they became more sensitive to the foreign stock markets evolutions (Rockinger and Urga 2000). Some researches approached the impact of the crises on the linkages between the international stock markets. Lin et al. (1994) found that during the periods of high volatility the international stock markets were more cointegrated than in the tranquil periods. Yang et al (2005) found that dynamic linkages between US, Germany and four East European (Russia, Poland, Hungary and Czech) stock markets were strengthen after 1998 Russian financial crisis. Zhang (2009) found that effects of the US stock market on the major Asian stock markets were greater after the Asian financial crisis.

In the last years the stock market from Romania experienced significant changes. The perspective of adhesion to the European Union attracted foreign investors and the Bucharest Stock Exchange (BSE) became more integrated with the international financial markets. Between 2006 and 2008 the Romanian stock market experienced an ascendant trend. However, since 2008, in the context of the global crisis, the stocks prices have fallen.

Figure 1. Evolution of DOW JONES and BET from January 2006 to September 2009



Source: own creation

In this paper we study the linkages between the Romanian and the US stock markets in the financial crisis context. We use two main indexes of these stock markets: the well known DOW JONES index from the New York Sock Exchange (NYSE) and BET, a reference index from BSE. The quite similar trends followed by the two indexes in the recent years suggest significant between them (Figure 1). We investigate these linkages in the period of crisis and in a more tranquil period using a VAR framework.

The rest of this paper is organized as follows. In the second part we describe the data and the methodology used in our analysis. In the third part we present the empirical results and in the fourth part we conclude.

2. Data and Methodology

In our analysis we employ daily close values of DOW JONES, provided by Yahoo Finance, and BET, provided by BSE. Because of the time differences between the two countries we study the relationship between the two indexes in two forms: in the first we take the values from the same day for the two variables while in the second we use one day lagged values for DOW JONES. The time period of the data is from the 3rd January 2006 to the 30th of September 2009. For both indexes we compute returns as:

$$R_t = (\ln P_t - \ln P_{t-1}) * 100 \quad (1)$$

where:

- R_t is the return of an index in the day t;
- P_t, P_{t-1} are the values of an index in the day t, respectively t-1.

We use the following variables:

- RBET – the return of BET;
- RDOWJ – the return of DOW JONES;
- LRDOWJ – the one day lagged return of DOW JONES;
- VRBET – the conditional variance of RBET;
- VRDOWJ – the conditional variance of RDOWJ;
- LVRDOWJ – the one day lagged conditional variance of RDOWJ;
- d_VRBET – the first differences of VRBET;
- d_LVRBET – the first differences of LVRBET.

We separate our sample of data in two sub-samples:

- a sub-sample from the 3rd January 2006 to the 11th of April 2008, corresponding to a tranquil period of time;
- a sub-sample from the 12th of April 2008 to the 30th of September 2009, corresponding to a period of time when the stock markets were affected by the financial crisis.

Table 1. Descriptive statistics of RBET and RDOWJ for the two sub-samples

Indicator	RBET		RDOWJ	
	Sub-sample 1	**Sub-sample 2**	**Sub-sample 1**	**Sub-sample 2**
Mean	0.0262048	-0.181497	0.0242919	-0.0774979
Median	0.00744916	-0.116589	0.0651782	-0.0824473
Minimum	-9.57338	-13.5461	-3.34876	-8.20051
Maximum	4.83962	10.0907	3.48749	10.5083
Std. Dev.	1.58528	3.05046	0.896141	2.32671
C.V.	60.4958	16.8072	36.8904	30.0228
Skewness	-0.640705	-0.529297	-0.298019	0.302038
Ex. kurtosis	3.45954	2.50959	2.01457	3.08540
Jarque - Bera test for normality	331.754	100.77	107.585	134.266
p-value for Jarque - Bera test	0.00001	0.00001	0.00001	0.00001

Source: own creation

In the Table 1 there are presented the descriptive statistics of the two indexes returns for both sub-samples. There are significant differences between the means and the standard deviations for the two sub-samples.

We investigate the stationarity of the variables using two tests: the classical Augmented Dickey – Fuller Test and a test proposed by Saikkonen and Lütkepohl (2002) and Lanne et al. (2001) which allow us to take into account the eventual structural breaks. We employ a VAR model to analyze the transmission of the shocks from NYSE to BSE. This model allows the test of the Granger causality between DOW JONES and BET.

In order to analyze the linkages between the volatilities of RBET and RDOWJ we compute, using ARCH - GARCH models, the conditional variances of these variables. Then we study the interactions between them by a VAR model.

3. Empirical Results

3.1. Results for the first sub-sample

Based on the graphical representation we used in the analysis of stationarity for both variables only intercept as deterministic term. In the Table 2 there are presented the results of the Augmented Dickey – Fuller Tests which indicate that both variables are stationary.

Table 2. Augmented Dickey-Fuller Test for the observations from the first sub-sample

Variable	**Lagged differences**	**Test statistics**	**Asymptotic p-value**
RBET	4	-11.0712	0.00001***
RDOWJ	1	-17.9529	0.00001***

Note: The number of the lagged differences was chosen based on Akaike Information Criteria.
Source: own creation

The results of unit root tests with structural breaks are presented in the Table 3. They also indicate the stationarity of RBET and RDOWJ.

Table 3. Unit root tests with structural breaks for the observations from the first sub-sample

Variable	**Shift Function**	**Break Date**	**Lagged differences**	**Test statistics**
RBET	Impulse dummy	498	4	-11.5958***
	Shift dummy	97	4	-4.8832***
RDOWJ	Impulse dummy	532	2	-13.6438***
	Shift dummy	486	1	-6.0759***

Note: The number of the lagged differences was chosen based on Akaike Information Criteria.
Source: own creation

The two equations of a VAR model with RBET and RDOWJ as dependent variables are presented in the Table 4. It shows a low interaction between the variables and an insignificant influence of RBET to RDOWJ.

Table 4. VAR system for the first sub-sample
Equation 1: RBET

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	0.00535269	0.0624756	0.0857	0.93175
RBET_1	0.0282291	0.0453859	0.6220	0.53420
RDOWJ_1	0.557409	0.0769292	7.2457	<0.00001***

Mean dependent var	0.020592	S.D. dependent var	1.580809
Sum squared resid	1309.626	S.E. of regression	1.501362
R-squared	0.101083	Adjusted R-squared	0.097988
F(2, 581)	26.27619	P-value(F)	1.19e-11
rho	-0.005273	Durbin-Watson	2.008053

Equation 2: RDOWJ

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	0.0262874	0.0372152	0.7064	0.48025
RBET_1	-0.0104802	0.0312211	-0.3357	0.73724
RDOWJ_1	-0.0851081	0.0470438	-1.8091	0.07095*

Mean dependent var	0.023817	S.D. dependent var	0.896835
Sum squared resid	465.3228	S.E. of regression	0.894930
R-squared	0.007660	Adjusted R-squared	0.004244
F(2, 581)	1.686554	P-value(F)	0.186062
rho	-0.000112	Durbin-Watson	1.997710

Tests of the residual values

Type of Test	**Test Statistic**	**P-value**
Test for multivariate normality of residuals Doornik-Hansen Chi-square(4)	155.563	0.0001
ARCH-LM Test for residual values of first equation	76.5021	0.0001
ARCH-LM Test for residual values of second equation	40.3567	0.0007

Decomposition of variance for RBET

Period	Std. error	RBET	RDOWJ
1	1.4975	100.0000	0.0000
2	1.57931	90.1059	9.8941
3	1.57959	90.0774	9.9226
4	1.57959	90.0774	9.9226
5	1.57959	90.0774	9.9226
6	1.57959	90.0774	9.9226
7	1.57959	90.0774	9.9226
8	1.57959	90.0774	9.9226
9	1.57959	90.0774	9.9226
10	1.57959	90.0774	9.9226
11	1.57959	90.0774	9.9226
12	1.57959	90.0774	9.9226
13	1.57959	90.0774	9.9226
14	1.57959	90.0774	9.9226
15	1.57959	90.0774	9.9226

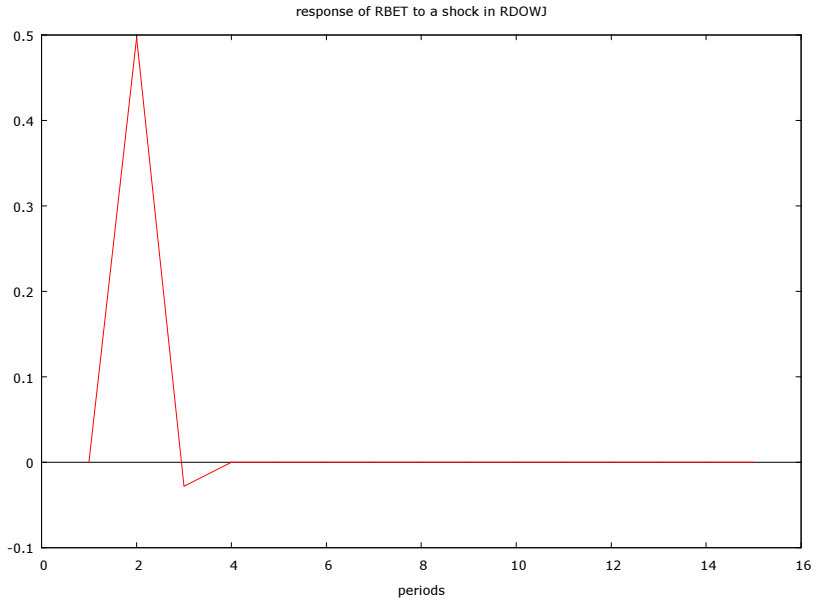
Decomposition of variance for RDOWJ

Period	Std. error	RBET	RDOWJ
1	0.892628	0.3169	99.6831
2	0.896068	0.3642	99.6358
3	0.896069	0.3643	99.6357
4	0.896069	0.3643	99.6357
5	0.896069	0.3643	99.6357
6	0.896069	0.3643	99.6357
7	0.896069	0.3643	99.6357
8	0.896069	0.3643	99.6357
9	0.896069	0.3643	99.6357
10	0.896069	0.3643	99.6357
11	0.896069	0.3643	99.6357
12	0.896069	0.3643	99.6357
13	0.896069	0.3643	99.6357
14	0.896069	0.3643	99.6357
15	0.896069	0.3643	99.6357

Source: own creation

The impulse-response analyses indicate that a shock in RDOWJ leads to a raise of BET and the effects are persistent for some days (Figure 2).

Figure 2. Impact of a shock in RDOWJ on RBET for the first sample



Source: own creation

The Granger causality test indicates a unidirectional causality from RDOWJ to RBET (Table 5).

Table 5. Tests of Granger causality between the variables for the first sub-sample

Null hypothesis	**F-statistic**	**P-value**	**Causal inference**
H0: "RBET" do not Granger-cause "RDOWJ"	0.2390	0.6251	"RBET" do not Granger-cause "RDOWJ"
H0: "RDOWJ" do not Granger-cause "RBET"	63.9889	0.00001***	"RDOWJ" Granger-cause "RBET"

Source: own creation

We analyzed, in the same VAR framework, the relation between RBET and the lagged values of RDOWJ. The results indicated insignificant interactions between RBET and LRDOWJ.

We compute the conditional variances of the two variables based on the GARCH models (Table 6 and Table 7).

Table 6. GARCH model with RBET as dependent variable for the first sub-sample

Variable	**Coefficient**	**Std. Error**	**z-stat**	**p-value**
const	0.0474254	0.0522473	0.9077	0.36403
alpha(0)	0.285152	0.113702	2.5079	0.01215**
alpha(1)	0.277888	0.0764693	3.6340	0.00028***
beta(1)	0.641435	0.0687922	9.3242	0.00001***

Mean dependent var	0.026205	S.D. dependent var	1.585278
Log-likelihood	-1064.137	Akaike criterion	2138.274
Schwarz criterion	2160.132	Hannan-Quinn	2146.792

Source: own creation

Table 7. GARCH model with RDOWJ as dependent variable for the first sub-sample

Variable	**Coefficient**	**Std. Error**	**z-stat**	**p-value**
const	0.0592195	0.0304368	1.9457	0.05170*
alpha(0)	0.0120848	0.00914273	1.3218	0.18624
alpha(1)	0.0596406	0.0147079	4.0550	0.00005***
beta(1)	0.92541	0.0187985	49.2280	<0.00001***

Mean dependent var	0.024292	S.D. dependent var	0.896141
Log-likelihood	-710.5048	Akaike criterion	1431.010
Schwarz criterion	1452.868	Hannan-Quinn	1439.528

Source: own creation

We studied, in a VAR framework, the interactions between the conditional variances of RBET and RDOWJ. Because VRDOWJ proved to be not stationary we use the first differences of the two variables. However, we found no significant relation between d_VRBET and d_VRDOWJ.

3.2. Results for the second sub-sample

The graphical representation suggests, for both variables, the use of one intercept as deterministic term in the analysis of stationarity. In the Table 8 there are presented the results of the Augmented Dickey – Fuller Tests which indicate that both variables are stationary.

Table 8. Augmented Dickey-Fuller Test for the observations from the second sub-sample

Variable	**Lagged differences**	**Test statistics**	**Asymptotic p-value**
RBET	1	-12.5798	0.00001***
RDOWJ	1	-15.9689	0.00001***

Note: The number of the lagged differences was chosen based on the Akaike Information Criteria.
Source: own creation

The results of the unit root tests with structural breaks are presented in the Table 9. They also indicate the stationarity of RBET and RDOWJ.

Table 9. Unit root tests with structural breaks for the observations from the second sub-sample

Variable	**Shift Function**	**Break Date**	**Lagged differences**	**Test statistics**
RBET	Impulse dummy	145	1	-12.2033***
	Shift dummy	93	1	-3.3664**
RDOWJ	Impulse dummy	93	1	-10.1078***
	Shift dummy	122	1	-3.4619**

Note: The number of the lagged differences was chosen based on the Akaike Information Criteria.
Source: own creation

In the Table 10 there is presented a VAR model with RBET and RDOWJ as dependent variables. It results again a low interaction between the two variables and an insignificant influence of RBET on RDOWJ.

Table 10. VAR system for the second sub-sample

Equation 1: RBET

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	-0.150854	0.160168	-0.9418	0.34698
RBET_1	-0.0267816	0.0767321	-0.3490	0.72730
RDOWJ_1	0.416605	0.0853175	4.8830	0.00001***

Mean dependent var.	-0.177951	S.D. dependent var	3.054489
Sum squared resid.	2734.002	S.E. of regression	2.913879
R-squared	0.095566	Adjusted R-squared	0.089949
F(2, 322)	12.49621	P-value(F)	5.93e-06
rho	0.016311	Durbin-Watson	1.965388

Equation 2: RDOWJ

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	-0.0883392	0.12495	-0.7070	0.48008
RBET_1	-0.0151269	0.073738	-0.2051	0.83759
RDOWJ_1	-0.134569	0.0694478	-1.9377	0.05353*

Mean dependent var.	-0.075242	S.D. dependent var.	2.329937
Sum squared resid.	1722.917	S.E. of regression	2.313152
R-squared	0.020440	Adjusted R-squared	0.014356
F(2, 322)	2.391524	P-value(F)	0.093114
rho	-0.022570	Durbin-Watson	2.045078

Tests of the residual values

Type of Test	**Test Statistic**	**P-value**
Test for multivariate normality of residuals Doornik-Hansen Chi-square(4)	96.6272	0.00001
ARCH-LM Test for residual values of first equation	67.801	0.00001
ARCH-LM Test for residual values of second equation	13.5091	0.0190478

Decomposition of variance for RBET

Period	Std. error	RBET	RDOWJ
1	2.9004	100.0000	0.0000
2	3.04538	91.9159	8.0841
3	3.0497	91.7290	8.2710
4	3.04975	91.7272	8.2728
5	3.04975	91.7271	8.2729
6	3.04975	91.7271	8.2729
7	3.04975	91.7271	8.2729
8	3.04975	91.7271	8.2729
9	3.04975	91.7271	8.2729
10	3.04975	91.7271	8.2729
11	3.04975	91.7271	8.2729
12	3.04975	91.7271	8.2729
13	3.04975	91.7271	8.2729
14	3.04975	91.7271	8.2729
15	3.04975	91.7271	8.2729

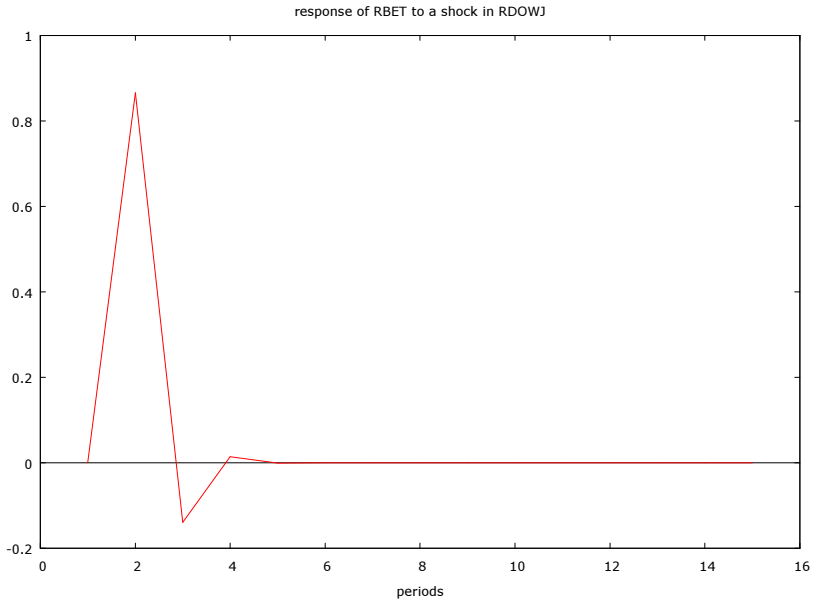
Decomposition of variance for RDOWJ

Period	Std. error	RBET	RDOWJ
1	2.30245	18.5140	81.4860
2	2.32614	18.7192	81.2808
3	2.32634	18.7224	81.2776
4	2.32634	18.7224	81.2776
5	2.32634	18.7224	81.2776
6	2.32634	18.7224	81.2776
7	2.32634	18.7224	81.2776
8	2.32634	18.7224	81.2776
9	2.32634	18.7224	81.2776
10	2.32634	18.7224	81.2776
11	2.32634	18.7224	81.2776
12	2.32634	18.7224	81.2776
13	2.32634	18.7224	81.2776
14	2.32634	18.7224	81.2776
15	2.32634	18.7224	81.2776

Source: own creation

The impulse-response analyses indicate again that a shock in RDOWJ leads to a raise of BET and the effects are persistent for some days (Figure 3).

Figure 3. Impact of a shock in RDOWJ on RBET for the second sample



Source: own creation

The Granger causality tests indicate again a unidirectional causality from RDOWJ to RBET (Table 11).

Table 11. Tests of Granger causality between the variables for the second sub-sample

Null hypothesis	**F-statistic**	**P-value**	**Causal inference**
H_0: "RBET" do not Granger-cause "RDOWJ"	0.3563	0.5508	"RBET" do not Granger-cause "RDOWJ"
H_0: "RDOWJ" do not Granger-cause "RBET"	31.1231	0.00001***	"RDOWJ" Granger-cause "RBET"

Source: own creation

The VAR analysis indicates an insignificant interaction between RBET and LRDOWJ. We compute the conditional variances of the two variables using the GARCH models (Table 12 and Table 13).

Table 12. GARCH model with RBET as dependent variable for the second sub-sample

Variable	**Coefficient**	**Std. Error**	**z-stat**	**p-value**
const	-0.0851979	0.155268	-0.5487	0.58320
alpha(0)	5.44242	0.895769	6.0757	0.00001***
alpha(1)	0.489471	0.1764	2.7748	0.00552***

Mean dependent var.	-0.181497	S.D. dependent var	3.050458
Log-likelihood	-807.7566	Akaike criterion	1623.513
Schwarz criterion	1638.661	Hannan-Quinn	1629.558

Source: own creation

Table 13. GARCH model with RDOWJ as dependent variable for the second sub-sample

Variable	**Coefficient**	**Std. Error**	**z-stat**	**p-value**
const	0.0566057	0.0822904	0.6879	0.49153
alpha(0)	0.00390167	0.0219536	0.1777	0.85894
alpha(1)	0.0956819	0.0200921	4.7622	0.00001***
beta(1)	0.904318	0.0180734	50.0359	0.00001***

Mean dependent var.	-0.077498	S.D. dependent var	2.326707
Log-likelihood	-670.2819	Akaike criterion	1350.564
Schwarz criterion	1369.498	Hannan-Quinn	1358.120

Source: own creation

The interactions between the conditional variances were studied in a VAR framework. The two equations presented in the Table 14 indicate a unidirectional influence from d_VRDOWJ to d_VRBET.

Table 14. VAR system on conditional variances for the second sub-sample

Equation 1: d_VRBET

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	-0.0217447	0.585055	-0.0372	0.97038
d_VRBET_1	-0.615212	0.123923	-4.9644	0.00001***
d_VRBET_2	-0.262329	0.106925	-2.4534	0.01469**
d_VRDOWJ_1	0.168472	0.956014	0.1762	0.86023
d_VRDOWJ_2	-0.567585	1.00297	-0.5659	0.57186

Mean dependent var	-0.0105011	S.D. dependent var	12.275
Sum squared resid	34301.3	S.E. of regression	10.3858
R-squared	0.293014	Adjusted R-squared	0.25114
F(4, 318)	6.17334	P-value(F)	0.00001***
rho	-0.0482122	Durbin-Watson	2.09618

Equation 2: d_VRDOWJ

Variable	**Coefficient**	**Std. Error**	**t-ratio**	**p-value**
const	-0.0113635	0.0577988	-0.1966	0.84426
d_VRBET_1	0.0179907	0.0103024	1.7463	0.08173*
d_VRBET_2	0.00947066	0.00829925	1.1411	0.25467
d_VRDOWJ_1	-0.13762	0.0669609	-2.0552	0.04067**
d_VRDOWJ_2	0.164727	0.144604	1.1392	0.25550

Mean dependent var.	-0.012173	S.D. dependent var.	1.06932
Sum squared resid.	335.48	S.E. of regression	1.02712
R-squared	0.088839	Adjusted R-squared	0.068241
F(6, 315)	2.11513	P-value(F)	0.0787*
rho	0.0011	Durbin-Watson	1.99764

Tests of the residual values

Type of Test	**Test Statistic**	**P-value**
Test for multivariate normality of residuals Doornik-Hansen Chi-square(4)	507.446	0.00001
ARCH-LM Test for residual values of first equation	5.02361	0.00002
ARCH-LM Test for residual values of second equation	2.10824	0.087371

Decomposition of variance for d_VRBET

Period	**Std. error**	**d_VRBET**	**d_VRDOWJ**
1	10.3051	100.0000	0.0000
2	12.1046	99.9799	0.0201
3	12.1906	99.6439	0.3561
4	12.2253	99.4777	0.5223
5	12.2523	99.4413	0.5587
6	12.2556	99.4393	0.5607
7	12.2556	99.4393	0.5607
8	12.2559	99.4393	0.5607
9	12.256	99.4393	0.5607
10	12.256	99.4393	0.5607
11	12.256	99.4393	0.5607
12	12.256	99.4393	0.5607
13	12.256	99.4393	0.5607
14	12.256	99.4393	0.5607
15	12.256	99.4393	0.5607

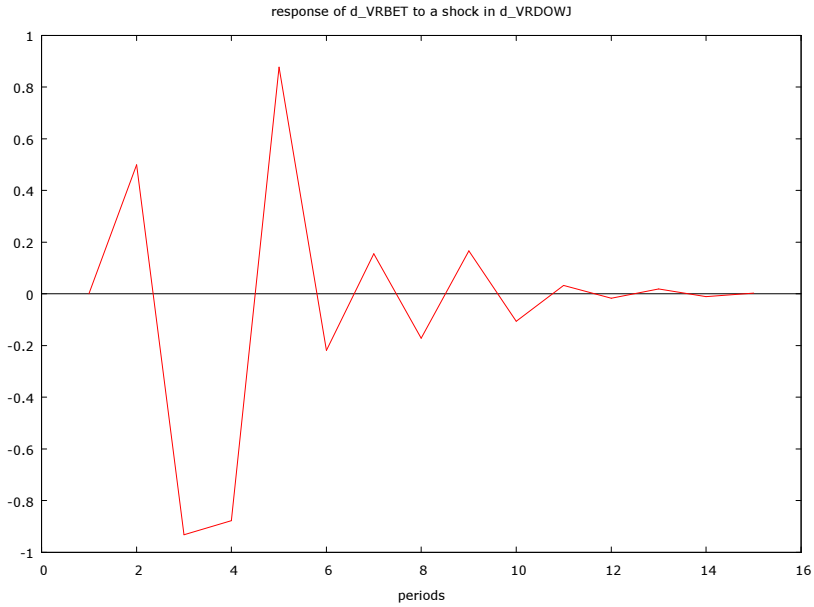
Decomposition of variance for d_VRDOWJ

Period	Std. error	d_VRBET	d_VRDOWJ
1	1.01914	0.2161	99.7839
2	1.04647	3.5683	96.4317
3	1.0648	3.6743	96.3257
4	1.06651	3.6626	96.3374
5	1.06748	3.6823	96.3177
6	1.06762	3.6881	96.3119
7	1.06765	3.6882	96.3118
8	1.06766	3.6883	96.3117
9	1.06766	3.6884	96.3116
10	1.06766	3.6884	96.3116
11	1.06766	3.6884	96.3116
12	1.06766	3.6884	96.3116
13	1.06766	3.6884	96.3116
14	1.06766	3.6884	96.3116
15	1.06766	3.6884	96.3116

Source: own creation

The impulse response analysis indicates that a shock in d_VRDOWJ has a persistent impact on d_VRBET (Figure 4).

Figure 4. Impact of a shock in d_VRDOWJ on d_VRBET for the second sub-sample



Source: own creation

4. Conclusions

In this paper we studied the impact of the US stock market on the Romanian stock market in the financial crisis context. Our analysis covered a period of time from the 12th of April 2008 to the 30th of September 2009, when the stock markets were affected by the global crisis. For comparison we perform a similar analysis in a more tranquil period of time, from the 3rd January 2006 to the 11th of April 2008.

We approached the NYSE evolution impact on the stock prices and on the volatility from BSE. Regarding the stock prices we found for both periods of time a unidirectional causality from the US stock market to the Romanian stock market. However, the VAR impulse – response analysis suggests the impact of NYSE evolution on BSE was more consistent during the crisis period than during the tranquil period. We also found the volatility of the US stock market had a significant influence on the Romanian stock market only in the financial crisis context.

The results indicate that Romanian stock market became quite integrated with the US stock market. They also confirm the theory that during the financial crisis the

emerging markets are more sensitive to the financial markets evolution from the developed countries.

References

Eun, C. – Shim, S. 1989: International transmission of stock market movements. *Journal Financial and Quantitative Analysis*, 24, 2, 241-256. p.

Ferson, W. E. - Harvey, C. R. 1995: Predictability and Time-varying Risk in World Equity Markets, *Research in Finance*, 13, JAI Press, 25-85 p.

Forbes, K. J. – Chinn, M. D. 2004: A Decomposition of Global Linkages in Financial Markets over Time. *The Review of Economics and Statistics*, 86, 3, 705–722. p.

Gilmore, C. G. – McManus, G. M. 2002: International Portfolio Diversification: US and Central European Equity Markets. *Emerging Markets Review*, Vol. 3, 1, 69-83. p.

Hsiao, F. S.T. – Hsiao, M. W. – Yamashita, A. 2003: The Impact of The US Economy on The Asia - Pacific Region: Does It Matter? *Journal of Asian Economics*, Vol. 14, 2, 219–241. p.

Janakiramanan, S. - Lamba, A. S. 1998: An Empirical Examination of Linkages Between Pacific - Basin Stock Markets. *Journal of International Financial Markets*, Institutions and Money, Vol. 8, 2, 155-173. p.

Kaminsky, G. – Lyons, R. – Schmukler, S. 2001: *Mutual Fund Investment in Emerging Markets: An Overview*. In S. Claessens and K. Forbes (Eds.): International Financial Contagion, Kluwer Academic Publishers, Boston, 158-185. p.

Kasa, K. 1992: Common Stochastic Trends in International Stock Markets. *Journal of Monetary Economics*, 29, 1, 95-124. p.

King, M. A. - Sentana, E. - Wadhwani, S. B. 1994: Volatility and Links between National Stock Markets. *Econometrica*, 62, 4, 901-33. p.

Kyle, A. S. 1985: Continuous Auction and Insider Trading. *Econometrica*, Vol. 53, 6, 1315-1335. p.

Lanne, M. - Lütkepohl, H. - Saikkonen, P. 2001: *Test procedures for unit roots in time series with level shifts at unknown time*, Discussion paper, Humboldt-Universität, Berlin.

Lin, W. - Engle, R. F. – Ito, T. 1994: Do bulls and bears move across boarders? International transmission of stock returns and volatility. *Review of Financial Studies* **7**, 3, 507-538. p.

Lintner, J. 1965: The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, 47, 1, 13-37. p.

Longin, F. - Solnik, B. 2001: Extreme Correlation of International Equity Markets, *The Journal of Finance,* 56, 2, 649–676. p.

Masih, A.M.M. – Masih, R. 1997: Dynamic linkages and the propagation mechanism driving major international stock markets. *Quarterly Review of Economics and Finance*, 37, 4, 859-885. p.

Roca, E. D. 2000: *Price Interdependence Among Equity Markets in the Asia-Pacific Region: Focus on Australia and ASEAN*, Ashgate, England.

Rockinger, M. – Urga, G. 2000: The evolution of stock markets in transition economies, *Journal of Comparative Economics*, 28, 3, 456-472. p.

Saikkonen, P. - Lütkepohl, H. 2002: Testing for a unit root in a time series with a level shift at unknown time. *Econometric Theory*, 18, 313-348. p.

Sharpe, W. 1964: Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19, 3, 425-442. p.

Yang, J. - Hsiao, G. – Wang, Z. 2005: The Emerging Market Crisis and Stock Market Linkages:

Further Evidences. *IEPR Working Papers 05.27*, University of Southeastern California, California.

Zhang, Y. 2009: Linkages of Stock Prices in Major Asian Markets and the United States**,** International Conference on Econometrics and the World Economy. *The Center for Advanced Economic Study (CAES) Fukuoka University*, Fukuoka, Japan.

Originate-to-Distribute Model and UK Financial Institutions

Po Yu Lin[1] - Lee Rose [2]– Gurjeet Dhesi[3]

The subprime mortgage crisis has brought attention to the business model, (namely the originate to distribute (OTD) model), that a vast majority of banks have adopted. In the OTD model the originator of a loan sells it to third parties through a securitisation process. The OTD model can be an efficient risk sharing tool for financial institutions to diversify their portfolio. However, the conflict of interest between different parties and potential incentive problem has eroded the original intention of the model. Furthermore, the financial contagion effect evolving from the OTD model of lending, contributes towards a significant proportion of the credit crisis in 2008. Because of data limitation we choose to examine the OTD business model of Lloyds Banks and Northern Rock. We adopt Purnanandam's (2008) approach using the difference-to-difference method to analyse the participation of UK OTD mortgage market pre-subprime crisis and post-subprime crisis periods. We show the difference of two UK commercial banks' participation in mortgage market prior to the credit crisis and post-disruption period. We find that the ability of the transfer of credit risk through the OTD model encouraged the origination of inferior quality loans by the banks. We also find that the OTD model affects banks' attitude towards risk from risk aversion to risk taking investment behaviour. We conclude that the OTD model is a positive financial innovation. However the screening incentive needs to be revised and monitored.

Keywords: Originate- to-Distribute, subprime crisis, incentive problem

1. Introduction

The resilience of UK financial system has been severely challenged in 2008. The collapse of US subprime mortgage market has sent a shock wave into the global financial system. The severity of this financial crisis is substantial. Reputable financial institutions have been bankrupted (Lehman Brother), nationalised

[1] Po Yu Lin, London South Bank University
[2] Lee Rose, London South Bank University
[3] Gurjeet Dhesi, London South Bank University

(Northern Rock) or acquired (Merrill Lynch) by others. Fund managers were forced to liquidate their position to meet the clients' demand as investors tried to rush out of the capital market. Public criticisms and academic debates are focused on the greediness of banks. However, the question reminds unanswered as how can a collapse of single domestic market leads to a global financial chaos. Ashcraft - Schuermann (2008) provide an overview of subprime mortgage securitisation process and several key informational frictions that arise.

Financial innovation enables financial institutions to shift its investment behaviour from originate-to-hold to originate-to-distribute (OTD), where the originator of a loan sells it to various third parties. In the past, banks receive funds from depositors and make loans to borrowers. The function of intermediary based on different liquidity preferences allows the flow of funds smoothly within the banking system. However, in recent years, because of the ability of transferring credit risk through the originate-to-distribute business model, banks are encouraged to originate excessive loans without appropriate monitoring and screening process. Purnanandam (2008) found that lack of screening incentive created by the separation of origination from the ultimate bearer of the default risk has been one of the major contributing factors to the US subprime mortgage crisis. However, the originate-to-distribute model of lending has its own strength. Diversification of banks' portfolio through originate-to-distribute model allows banks to reach optimal risk-sharing and therefore enhance the resilience to possible financial shock and reduce regulatory capital requirement by moving its positions to off-balance sheet. It is however, the incentive structure of the originate-to-distribute model causes irresponsible investment behaviour.

Banks possess an unique function in the economy, they acts as an intermediary upon various liquidity preferences. Banks are able to assess borrowers' credibility shall they require a loan. It is banks' responsibility to conduct appropriate due diligence before approval any loans. It is also a standard procedure in terms of risk assessment for banks to perform screening and monitoring functions. However, sometimes they might not fully use these functions. With the application of Originate-to-Distribute business model, financial institutions are able to transfer credit risk to third parties while maintaining their profits. The incentive of performing due diligence and other risk assessments has been reduced significantly. The quality of mortgage originated by the financial institutions who widely adopt OTD model deteriorated as screening function is poorly performed.

This business model works well from 2001 to 2006 as the market was enjoying an explosive growth. Assets value rose significantly, in particular the property price has gone up dramatically since then. Mortgagors were able to roll over the loan based on the appreciation of their property value. However, the sign of market reversal first appears in the end of 2005 as loan quality had been worsening for five years in a row at that point. (Demyanyk - Van Hemert 2008) the cost of

excessive lending is severe as the world economy is suffering the worst economy downturn since the Great Depression. It is clear that banks were unable to understand the true risks of these loans that they originated. The flowchart provided below is a simplified OTD model and there are a number of conflicts involved in different stage. Subprime lenders are willing to provide mortgage to subprime borrowers as they can generate fee without bearing the risk of default through the OTD business model. Investment bank securities subprime mortgage and sell it to pension fund and asset management. Arguably, if the credit rating agencies performed its function correctly, the severity of this financial crisis might not be as damaging as it is now. It is difficult to assess the risk of a structured product due to its complexity and lack of information, that is why investors and fund managers are heavily reply on credit rating agencies, unfortunately, because of the nature of its fee structure, the credit rating agencies were unable to provide independent credit assessment as they are being paid by investment banks who securities subprime mortgages.

While the OTD model being on the centre of debate, this financial innovation delivered numerous benefits to the financial system. This paper will focus on its incentive structure and the participation of OTD market within UK financial system. The paper is structured as follows: chapter 2 presents an overview of OTD model, in particular, the conflict of interest between different parties. It must be noted that the availability of data are lower in UK compare to US. Chapter 3 describes the data and summary statistics. In chapter 4 we present the empirical evidence for the participation of OTD market and chapter 5 conclude the paper.

2. Overview of the OTD model

Financial innovation evolves in response to capital constraint. Over the last decade, bank credit has shifted from traditional originate to hold to originate to distribute business model, where banks originate loans and sell it to third parties through securitisation process. There are six major players in the model and each of them possesses different key role. They are borrowers, originators, arrangers, distributors, investors and credit rating agencies. The relationship between each party is illustrated in figure 1 below. First thing to note from the chart was that the model is a bottom-up approach, start from borrowers and move up to investors. Each party is only interest in its next party. This next nearest phenomenon creates potential incentive problem which prohibit each party from performing its origin screening and due diligence functions.

2.1. Conflict of interest between Borrowers & Originators

Banks are able to act as financial intermediary base on individual's different liquidity preference. Borrowers are people who needs mortgage to finance their home purchase or refinance their existing mortgage. Borrowers might not be financially sophisticated. They might not be able to act for their best interest because of lack of financial knowledge. If mortgage originators do not provide appropriate financial advisory, borrowers might have a financial product which is not suitable for their financial situation.

Borrowers' credit history will normally be assessed by mortgage originators to determine whether the mortgage should be approved. However, In the case of the subprime mortgage crisis, borrowers' credit history has not been fully assessed because mortgage originators were able to sell their pool of mortgage funds to third party. The incentive for screening procedure on mortgages quality has been significantly reduced because of this credit risk transition.

2.2. Conflict of interest between Originators & Arrangers

Arrangers are usually investment banks or large commercial banks with investment or asset management division. Arrangers are in the centre of the securitisation process. When arrangers buy pool of mortgage loans from originators, the first responsibility is to perform due diligence on originators. This includes review on financial statement of originators and underwriting guidelines. However, arrangers might not fully perform its screening function as they are able to pass it to its next party. Moreover, as originators have superior information on the quality of the mortgages over arrangers, it is difficult for arrangers to fully assess and understand the true risks behind the pool of mortgage loans.

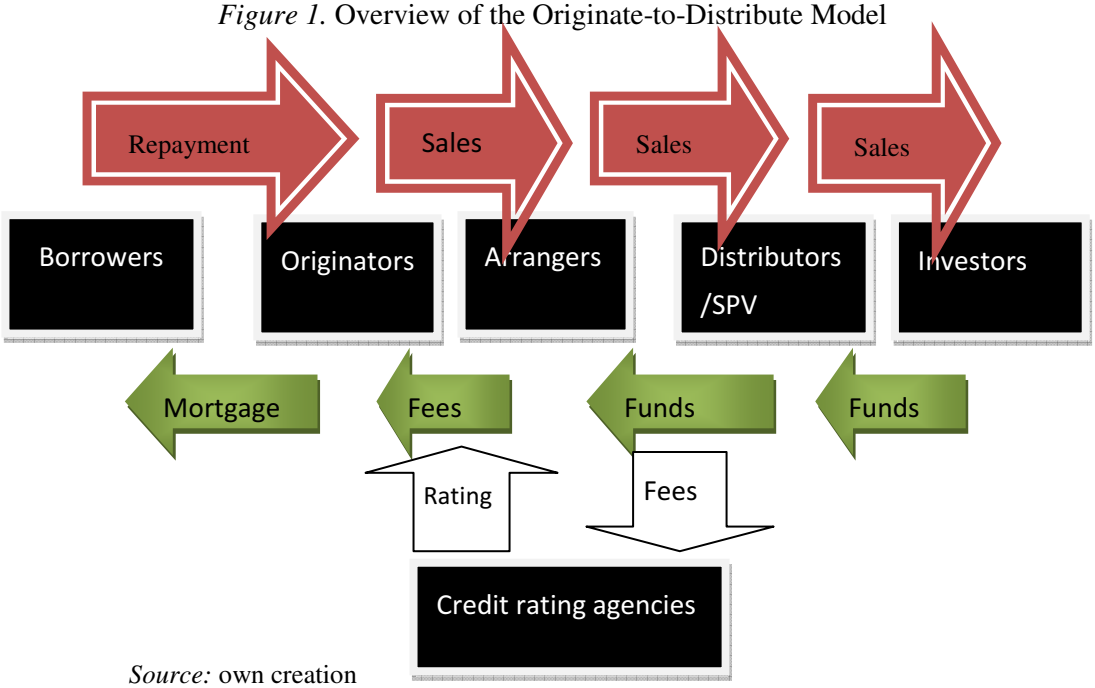
Arrangers then repackage pool of mortgage loans into different debt securities such as collateral debt obligation (CDO). Arrangers assign risk and return accordingly to each CDO and obtain credit rating from the rating agencies for a fee. The nature of this fee structure between the rating agencies and arrangers creates potential conflict of interest. The credit rating agencies were unable to conduct independent credit assessment because they are being paid by arrangers. It is clear that if the credit rating agencies issue unfavourable credit report to the arrangers, the credit rating agencies might risking lose business. The arranger creates a bankruptcy-remote entity to purchase these debt securities in order to protect it from default risk. This special purpose vehicle (SPV) is designed to finance the purchase of pool of mortgage loans without putting the entire firm at risk.

Investors here refer to institutional investors such as mutual funds, hedge funds, pension funds and insurance companies. Investors are heavily relying on

credit rating agencies' opinions to make investment decision. This is because it is difficult for investors to estimate the underlying credit risk of a structured product due to the nature of its complexity. However, as mention above, credit rating agencies were unable to issue objective opinion due to the conflict of interest created by issuer-paid fees model. The structured finance deal contributed a significant proportion of rating agencies' revenue. The information asymmetry between investors and the credit rating agencies has exposed investors to the risk of dishonest of the rating agencies.

There are different degrees of conflict of interest in different stage of the OTD model. The incentive of screening on loan quality and perform due diligence on originators are crucial. Moreover, disclosure of rating and downgrade criteria can be helpful for public to observe.

Figure 1. Overview of the Originate-to-Distribute Model



Source: own creation

3. Methodology and Empirical result

UK banking system is constituted by a small number of large financial conglomerates. The big four banks are accounted for 80% of the market share. The remaining 20% of market shares are accomplished by local banks and building societies. We use Lloyds TSB and Northern Rock for this study. We obtain data from the bank annual report between 2005 to 2008, this includes bank's balance

sheet, income statement and off-balance sheet activities. We define 2005 to 2006 as pre-crisis period and 2007-2008 as post-crisis, use Purnanandam (2008)'s difference-in-difference methodology.

We extract several key information from balance sheet, income statement and off-balance sheet activities. We obtain information on bank's total assets, mortgage loans, total deposit, demand deposit and other variable from annual report. We use the loan subject to securities as the measure of a bank's involvement in the OTD market and loan lost provision as the measure of mortgage lost. We measure bank's liquidity as the ratio of demand deposit over total deposit. This is to analyses bank's attitude towards the OTD market under capital constraint. We use available for sale debt securities on balance sheet to measure the result of involving in the OTD market. Our interest is in creating a meaning proxy measurement of the credit risk transfer that a bank engages in. Our hypothesis is that banks have the incentive to issue inferior mortgage loans because they are able to generate profit without bearing the credit risk of borrowers through the OTD business model. This is preliminary study as UK data is difficult to obtain compare to the US.

Our sample consists of Lloydss TSB bank and Northern Rock with available data in mortgage issuance activity from 2005 to 2008. We define 2005 and 2006 and pre crisis period and 2007 and 2008 as post crisis period. We start our investigation with loan lost on banks' balance sheet. We are interested in relating the loan lost to a bank's involvement in the OTD market before the subprime mortgage crisis. We initial our test by fit the following regression model with four years data from 2005 to 2008.

$$\boldsymbol{loanlost_{it}= \beta_0+\beta_1 preotd_i+\beta_2 dummyaftert+\beta_3 available+\varepsilon_{it}}$$

The dependant variable is loan lost on banks' balance sheet from year of 2005. $preotd_i$ is banks' loan subject to securitises, which measures the extend of bank i 's participation in the OTD market prior to the subprime mortgage crisis. The coefficient of this variable measures the average quality of loans issued by such banks. *Dummyafter* is a dummy variable that equals one for post crisis period and zero for pre-crisis period. This allows us to capture the severity of the subprime mortgage crisis and banks' damage according to different level of the OTD market involvement. *Available* is the available for sale debt securities. This variable measures banks' involvement in the OTD market.

The results of the OLS are provided in appendix. We regress loan lost with *preotd, dummyaftr and available*. We find *preotd* is negative and insignificant. This suggests that there is no evidence that the sample banks are heavily participate in the OTD market. This result matches the conclusion of European Central Bank's working paper 2008 December volume which was that European banks are less involved in the OTD market, however, they have been seriously damaged by the

subprime mortgage crisis because they bought large amount of financial derivative originated from the US market. We find positive and significant coefficient on the *dummyafter* variable. This indicates that on average, the loan lost increased significantly after 2006 compare to before. This result confirms the large write-offs in late 2007 because of the subprime mortgage crisis. The coefficient on *available* variable is positive and significant. This suggests that banks with large available for sale debt securities have suffered from larger loan lost.

4. Conclusion

While the public and media blamed the OTD model as the source of the subprime mortgage crisis, this business model helps financial institutions to achieve better risk sharing and managing banks' portfolio. It is more important to introduce appropriate screening system within the model rather than completely give up this financial innovation. In this paper we focus on UK banks' participation in the OTD market. We show that UK banks were not heavily involved in the OTD market. However, the loan lost on banks' balance sheet was still severe. This is because the majority of UK banks are in the last of the OTD model chain as investors. They invest significant in financial products that were originated from UK mortgage market. This investment behaviour and attitude towards risk might be driven by low interest rate from 2001 to 2004, as well as the booming of the property market.

Our evidence shows that there are different degrees of conflict of interest within the OTD model, which confirms the public criticism that information asymmetry and lack of screening incentive have been a significant contribution to the subprime mortgage crisis. Each party in the model does not have the incentive to perform its original screening and auditing function. This is one of the main direct causes of the subprime mortgage crisis.

Our findings have important implication for UK financial market and regulators. We show that as the majority of UK financial institutions are the investor in the OTD business model, their investment activity is crucial. The risk exposure to cross boarder countries, in particular, the US market is very much concentrated. These finding can help financial institutions to review the diversification of their portfolio. From the regulator's perspective, enhancing transparency in trading activity and certain degree of public disclosure could improve the OTD mode.

References

Ashcraft - Schuermann 2008: *Understanding the Securitization of Subprime Mortgage Credit*. Wharton Financial Institutions Center, Working Paper No. 07-43.

Croupy - Jarrow - Turnbull 2008: *The Subprime Credit Crisis of 07*. Enterprise Risk Management, Working Paper Series.

Demyanyk - Van Hemert 2008: *Understanding the Subprime Mortgage Crisis*. Working Paper Series, New York University, New York.

European Central Bank 2008: *The Incentive Structure of the Originate to Distribute Model,* European Central Bank..

Purnanandam, A. 2008: *Originate-to-Distribute Model and The Sub prime Mortgage Crisis.* Ross School of Business, University of Michigan, Ann Arbor.

Appendix

Table 1.

Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
year	8	2006.5	1.195	2005	2008
totalassets	8	2305501	144668	82709	436191
residentmortgage	8	87392	17798	62290	114643
Mortgage/ta	8	.514	.255	.26	.83
customerdeposit	8	50952	33543	11563	96220
totaldeposit	8	85502	70403	11563	172364
Cd/td	8	.77	.246	.53	1
loanlost	8	1091	978	57	2876
liquiditasset	8	52991	50215	1483	113405
liquidity	8	.173	.123	.01	.33
availableforsale	8	12072	8530	1483	26457
preotd	8	34078	14715	10048	49326
dummyafter	8	.5	.5345	0	1

LONDON SOUTH BANK UNIVERSITY

Table 2.

$$loanlost_{it} = \beta_0 + \beta_1 preotd_i + \beta_2 dummyafter_{t+} \beta_3 available + \varepsilon_{it}$$

Source	SS	df MS	Number of obs	= 8
		F(3, 4)	= 27.97	
Model	6395793	3 2131931	Prob > F	= 0.004
Residual	304855	4 76214	R-squared	= 0.955
		Adj R-squared	= 0.920	
Total	6700647	7 957236	Root MSE	= 276.07

loanlost	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
preotd	-.023	.009	-2.61	0.060	-.0474	.001
dummyafter	802	234	3.43	0.026	153	1451
availableforsale	.085	.013	6.28	0.003	.047	.123
_cons	447	361	1.24	0.283	-554	1449

LONDON SOUTH BANK UNIVERSITY

Public projects and their major sources of financing in Poland after the accession into the European Union: The case of the Silesian voivodship

Jan Czempas [1]- Piotr Tworek[2]

In the period of time following the EU's accession of Poland (1 May, 2004), investment processes in individual voivodships have revived considerably. Infrastructure has become the main direction for investments. Local government units have a variety of options available in order to acquire funds for financing of public projects. The key aim of the paper is to discuss such issues, i.e. the authors present the scale of investments in municipalities, communes, poviats and self-governing voivodships as well as the extent to which EU funds are engaged in financing of investment projects carried out by these entities, based on the example of the Silesian Voivodship. The paper is an attempt at evaluating the dynamics of changes which occurred in the years of 2004-2008, within the source of financing for investments which plays an increasingly important part in budgets of local government units i.e. EU funds. The article presents a theoretical and empirical approach to the issues.

Keywords: Public projects, Infrastructure, Local governments' projects, Financing of investments, European Union's funds, Silesian Voivodship, Poland

1. Introduction

The Poland's accession to the European Union opened up new opportunities in terms of financing of projects in the public sector. In Poland public and private investors alike can use numerous ways of acquiring EU funds. (Currently, apart

[1] *Jan CzempaS, The Karol Adamiecki University of Economics, Department of Investments and Real-Estate, Poland*
[2] *Piotr Tworek, The Karol Adamiecki University of Economics, Department of Investments and Real-Estate, Poland*

from traditional forms of financing for public investments, investors may also take advantage of a number of EU programmes, which are coordinated in Poland by the Ministry of Regional Development[3]). Also, at the level of voivodships in Poland, which are 17, EU funds are distributed among the entities involved. (Taking into account the size of the Silesian Voivodship, i.e. the second largest voivodship in terms of population, the EU funds are given to the local government authorities). It should be noted that in Poland, there's a formal administrative division (within voivodships) into towns with the rights of a poviat, municipal communes, rural communes, and municipal rural communes. These are so-called local government units, that's the basic units functioning within the administrative division of the country, which are headed by representatives of local government authorities (town presidents, mayors, commune leaders). Therefore, this paper aims at discussing the selected issues related to financing of public sector projects in Poland, based on the example of the Silesian voivodship, i.e. as seen from the perspective of the 19 biggest towns of the Silesian Voivodship. In particular, it discusses the use of EU funds for investments carried out in specific towns of the Silesian Voivodship. The deliberations in this article are both theoretical and empirical[4]. (This publication is the outcome of the authors' studies and research, conducted in two project teams doing statutory research[5] in the Department of Investments and Real-Estate, The Karol Adamiecki University of Economics in Katowice, Poland).

[3] See more: Ministry of Regional Development. Website: http://www.mrr.gov.pl/

[4] In particular, point 2 of the paper is of the empirical character, i.e. its author is J. Czempas. Point 1, in turn, is of the theoretical character and it aims at e.g. definition of the term of "public project", including the discussion of the term of „infrastructure" and the importance of public investment in Poland. This point of the paper is written by P. Tworek. (In particular, this fragment of the article provides the background for the solution of the key issue, which is the financial support for public sector investments in the Silesian Voivodship in the years of 2004-2008, i.e. with the focus on utilisation of the EU funds).

[5] See: Financing of investment projects in the public sector – availability of sources and their effectiveness. Stage I. Project leader: Prof. K. Marcinek. The K. Adamiecki University of Economics in Katowice. Statutory research, Katowice 2006; Financing of investment projects in the public sector – availability of sources and their effectiveness. Stage II. Project leader: Prof. K. Marcinek. The K. Adamiecki University of Economics in Katowice. Statutory research, Katowice 2007.

2. Public projects and their major sources of financing in Poland (in theory and practice)

In the scientific literature and in practice, public investments tend to be examined from the angle of infrastructural projects[6]. As emphasized by B. Winiarski (2000) "[...] a traditional area for public investment is economic, technical and social infrastructure". Hence the assumption that an investment policy of the state and the local government is carried out in Poland, in most cases, through public projects, i.e. infrastructural undertakings[7]. As defined by The New General Encyclopaedia of PWN "[...] infrastructure is the essential equipment and service-providing institutions necessary for an economy and a society to operate" (Nowa Encyklopedia Powszechna 1995) . ("The economic infrastructure covers services related to transport, communications, power industry, irrigation etc, e.g. harbours, railway tracks, power plants, dams)"[8]; "the social infrastructure covers services related to the law, security, education and schooling, culture, welfare and health care, residential construction industry etc., e.g. schools, hospitals, courts of law, prisons, public administration institutions"[9]). The classic division of infrastructure into the technical infrastructure and the social infrastructure, with the specific examples, are given in table 1.

[6] Both in theory and in practice the division of infrastructure into the technical infrastructure and the social infrastructure is becoming increasingly common. There's also the business infrastructure, which is synonymous to the economic, technical and production infrastructure.

[7] The article assumes, as its starting point for any further deliberations, that the most representative example of public investments are infrastructural projects, therefore most of the issues were examined from this perspective. Consequently, such notions as a public project, an infrastructural investment (infrastructure), to be understood as a public investment, are frequently used in the text.

[8] Ibidem, p. 55.

[9] Ibidem, p. 55.

Table 1. Selected technical and social infrastructure by territory and competence

Areas of infrastructure	State authorities	Territorial scope	Local scope
Transport and communications	Motorways, railway tracks and stations, airports, sea harbours	Voivodship roads, poviat roads	Communal roads
Power industry	Power plants, transmission grid, power distribution network	Gas industry – transmission pipelines	Power supply to subtransmission grids, gas industry, street lighting, heat generation industry
Water management and environmental protection	Reservoirs and dams	Water supply and sewage systems	Water supply and sewage systems, waste disposal, town cleaning services
Schooling and education	Institutes of tertiary education, scientific units	Secondary and post-secondary schools	Kindergartens, primary schools, pre-secondary schools
Health care	University clinics, medical centres	Voivodship hospitals, specialist outpatient units	General outpatient units, welfare centres
Culture	Specialist and scientific libraries, museums, national galleries	Specialist libraries, galleries, voivodship concert halls, theatres	Public libraries, culture centres
Sport and leisure	Big stadiums and sports halls, monument parks, zoological gardens		Greenery, sports fields, gymnasiums

Source: Janowska (2002)

Essentially, all the examples of infrastructure listed in table 1 are specific public projects with a regional, local or even supranational scope. It may be assumed, therefore, that infrastructural projects create public goods. (A distinctive feature of public goods is the fact that it isn't possible to make the access to public goods conditional on payment of a price for their use (Denek et al. 2005) Consequently, it may be assumed that public investments are mostly real investments. The importance of public investments for the economy is enormous. It results, first of all, from the function which public projects perform in a state's economic system. The state or local government units take on investors'

responsibilities, making expenditures in the public sector and increasing the fixed resources located there (Winiarski). Based on that, it's easy to define "a public investment", i.e. "a public project", which should be understood as deliberate measures taken in order to bring about a specific effect in form of tangible goods (i.e. an infrastructural object), which may be obtained due to engagement of specific forces and means (human, tangible and financial resources), and the results of these measures demonstrate, first of all, the features of public goods[10]. Public investment, therefore, is simply a specific form of capital management, i.e. basically, a form of public funds management. The best example of such investments are communal investments carried out by communes in Poland. As emphasized by Bończyk-Kucharczyk et al. 1998 "[...] the importance of the investments carried out by communes for other investments, for the local development of businesses and for the economic development as such depends on the following factors:

- well-planned infrastructural projects bring about subsequent investments, e.g. territorial development for the construction industry may attract building investors,
- projects financed by a commune generate jobs in local businesses and, consequently, improve the economic outlook on the construction services market,
- public projects – by creating a local labour market – are the most effective method for prevention of unemployment,
- infrastructural projects, such as communal investments, raise the living standards of the commune's population,
- without certain investments, development of some business areas is impossible to achieve,
- lack of projects, particularly into the business infrastructure and the social infrastructure, imposes a dangerous barrier to the economic development,
- lack of necessary investments into environmental protection and other activities are a barrier to sustainable development, may lead to permanent deterioration of the inhabitants' quality of life and the failure to meet the basic needs, and result in social discontent"[11].

[10] In the literature of the subject, there are a number of attempts at the definition of „a public project". A very common definition, for example, is the definition of a public investment as a gross investment into the tangible capital of a public sector.

[11] E. Bończyk-Kucharczyk, K. Herbst, K. Chmura: Jak władze lokalne mogą wspierać przedsiębiorczość. Fundacja Inicjatyw Społeczno-Ekonomicznych. Polska Fundacja Promocji Małych i Średnich Przedsiębiorstw, Warszawa 1998, p. 34. After: S. Słupik: Inwestycje infrastrukturalne jako niezbędny warunek rozwoju gospodarczego gminy. In:

In general, management of public investments follows slightly different rules than management of commercial projects, which results e.g. from the need to have them administered by the state (local government) authorities, and the main difference lies in the way the economic effectiveness of public projects is measured. (Completely different evaluation criteria[12], as well as non-commercial goals, make public projects specific). What's important - their implementation provides conditions, first of all, for undertaking further commercial investments, that's why they create circumstances for development of small and medium-sized entrepreneurship, among the others. Therefore, there may be two types of functions to be performed by public investments, namely: the microeconomic function and the macroeconomic function. Macroeconomic functions are particularly important here, as they carry benefits for a big group of entities (including businesses), and the most vital ones are (first of all) creating conditions for a growth in national economic output, generating new jobs (elimination of unemployment), creating conditions for changes in the country's economic structure. It should also be noted that in the theory of economics special importance is attached to the allocating function. (Allocation means distribution of available factors of production throughout various types of activities (Szczodrowski 2003). There's a very clear correlation between public investments and private investments, as the allocation function of public investments may, in a way, be connected with the launch of commercial investments in the private sector, which belongs to the essential effects of public investment. Here, we can also talk about the stabilising function, which should be associated with the impact public investments have on economic processes, in macroeconomic terms, in order to ensure that specific goals will be reached, and these goals may include e.g. stabilisation. At the same time, public investments stimulate development, which means that such undertakings induce development on a macro, mezzo and micro-scale. This's particularly strongly emphasised by B. Winiarski, according to whom "[...] public investments are, first of all, a stimulant of investment processes in an economy" (Winiarski 2000). (As pointed out by M. Ratajczak, in turn, "[...] infrastructure may perform the role of allocation, location and spatial integration factors and may serve as a tool for stimulating social and economic development" (Ratajczak 1990).

Public project management also involves raising funds for implementation of public projects. In Poland infrastructure may be funded using a number of various sources. The best-known division in the scientific literature is the division into one's own sources and outside sources as well as into internal and external sources. (For instance, communes in Poland may finance the projects with the funds obtained

Rozwój oraz polityka regionalna i lokalna w Polsce. Edit. by J. Kaja, K. Piech. Szkoła Główna Handlowa w Warszawie, Warszawa 2005, p. 259.

[12] In Poland the Cost-Benefits Analysis – CBA method is recommended.

through the issue of municipal bonds, i.e. the town of Poznań, i.e. the Wielkopolska Voivodship, issued such bonds for approx PLN 500 million). In Poland financing of investment activities of local government units is governed by the provisions of The Act on Income of Local Self-Government Units *(Ustawa o dochodach jednostek samorządu terytorialnego)* (The Act on Income 2003). Furthermore, the entire array of other regulations apply to these issues, in particular the provisions of The Act on Public Finances *(Ustawa o finansach publicznych)* (The Act on Public Finances 2005), which is directly related to various financing options. In broad terms, the most essential sources of financing for communes' infrastructural projects in Poland are their own budgetary funds (of local government units), subsidies and grants from the state treasury, bank credits and soft loans, leasing, issue of municipal bonds. Moreover, here we should also mention private funds employed in projects carried out within Public-Private Partnership (PPP)[13], or EU funds. Under the Polish law, the income of local government units comprises: "1) own income, 2) general subsidies, 3) designated grants from the state budget" (The Act on Income of Local Self-Government Units 2003).

In particular "the income of local government units may come from:

1. non-repayable funds from foreign sources,
2. funds from the EU budget,
3. other funds specified under separate regulations" (The Act on Income of Local Self-Government Units 2003).

In turn, "a general subsidy is made up of the following parts:

1. for communes: compensatory, equalising,
2. for poviats: compensatory, equalising,
3. for voivodships: compensatory, regional,
4. educational one - for communes, poviats and voivodships" (The Act on Income of Local Self-Government Units 2003).

(In compliance with the legal regulations effective in Poland, "an entity which is a local government unit makes a decision of what the funds from a general subsidy should be spent on" (The Act on Income of Local Self-Government Units 2003). However, "designated grants from the state treasury may constitute the income of the local government units, to be spent on:

1. government administration tasks and other tasks ordered by the legal acts,
2. tasks performed by local government units under the agreements concluded with the government administration authorities,
3. removal of direct threats to public security and order, consequences of floods, landslides and other natural disasters,

[13] In Poland these issues are governed by the provisions of the Act on Public-Private Partnership (*Ustawa z dnia 19 grudnia 2008 r. o partnerstwie-publiczno-prywatnym*). See: Act on Public-Private Partnership of 19 December, 2008 (Journal of Laws of 5 February, 2009).

4. financing or supplementary financing of own tasks,
5. performance of tasks resulting from international agreements" (The Act on Income of Local Self-Government Units 2003).

Currently, in Poland the EU funds make a significant source of financing for investments in the public sector. Before the Poland's accession to the European Union, such funds as PHARE, SAPARD and ISPA used to play a special role. After the accession, however, a decisive role started to be played by the Structural Funds and the Cohesion Fund. In the period of 2007-2013, within the framework of the so-called New Financial Perspective, the key sources of financing are the Cohesion Fund, the European Social Fund and the European Regional Development Fund.

3. The importance of European Union's funds in financing of local governments in the Silesian voivodship in Poland in the period of 2004-2008

The principal aim of this part of the article is the analysis and evaluation of the importance of EU funds in financing of investments in 19 big towns of the Silesian Voivodship in the years of 2004-2008. Using the relative increments, the dynamics of total EU funds and the change in importance of these amounts in the local government units' budgets for performance of annual investment tasks were compared. The data illustrating these issues are presented in tables 2 and 3.

Based on the figures given in tables 2 and 3, the following conclusions may be drawn:

- in consecutive years, investment expenditure incurred in the towns researched has increased regularly – from PLN 834 million in 2004 to almost PLN 2 billion in 2008 (PLN 1.917 billion – an increase by almost 130 percent);
- in the same period of time, the EU funds engaged in financing of the investments in these towns have gone up by over 200% – from PLN 292 million in 2004 to 882 million five years later;
- the importance of the EU funds has grown significantly every year – initially a third (35 percent), and then even more than 50 percent (in 2006) of total funds planned for financing of investments came from this source. In the last year of the period analyzed, this figure rose to 46.1 percent;
- EU funds have become an important external and non-repayable source of financing for investments - the annual growth dynamics in the first three-year period was definitely higher than in the following two years – it was 48.16 percent and 79.30 percent respectively, and then it was considerably reduced in the two recent years to 4.34 percent in 2007, compared to 2006, and almost 9 percent, when compared to the next two years;
- the towns which allocated the most funds for investments were: Katowice, Rybnik, Gliwice, Ruda Śląska, Dąbrowa Górnicza, Zabrze. Depending on the type

of the tasks performed, their sources of financing varied. In 2004 in such towns as Zabrze and Dąbrowa Górnicza, the participation of their own funds or grants in financing of investments went up to even a hundred percent;

- the importance of EU funds for investments has always differed significantly – in some towns as much as 80 percent of the funds came from this source (mainly Rybnik);
- the EU funds appeared in the budgets quite irregularly and, as may be suspected, accidentally – which may be confirmed by the amounts received by the local government units in the subsequent years: negligible amounts in the first year, followed by an increase of several thousand percent, to get to the level close to zero in the third year of the analysed period. The reasons for this should be looked for in a long-lasting process of preparation of applications for financing, in particular when investment plans were be quite capital-intensive;

Table 2. Investment expenditure of poviat towns in Silesian voivodship and EU funds (in PLN million) and their percentages in the year of 2004-2008

Tab. 1 Investment expenditure of poviat towns in Silesian voivodship and EU funds (in PLN million) and their percentages in the years of 2004-2008

No.	Town	Population in 2008	Investment expenditure					EU funds					Participation of EU funds in investment expenditure				
			2 004	2 005	2 006	2 007	2 008	2 004	2 005	2 006	2 007	2 008	2 004	2 005	2 006	2 007	2 008
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*
1	BIELSKO-BIAŁA	175 677	77,55	96,48	86,50	104,96	164,04	57,94	50,56	14,04	16,51	23,63	74,71	52,41	16,24	15,73	14,41
2	BYTOM	183 829	10,52	12,47	19,39	59,00	106,32	4,43	3,01	13,12	39,03	80,57	42,07	24,14	67,68	66,15	75,78
3	CHORZÓW	113 314	13,67	17,10	35,03	36,68	40,16	0,01	2,88	7,41	17,64	11,89	0,07	16,82	21,16	48,09	29,60
4	CZĘSTOCHOWA	240 612	49,78	120,63	166,32	153,69	136,48	3,87	43,08	116,79	110,30	78,51	7,78	35,71	70,22	71,77	57,52
5	DĄBROWA GÓRNICZA	128 315	45,48	54,27	45,24	41,34	51,05			2,55	6,73	6,89			5,63	16,28	13,49
6	GLIWICE	196 669	39,84	86,61	107,94	182,89	155,20	0,82	9,51	31,33	91,24	4,49	2,05	10,98	29,03	49,89	2,89
7	JASTRZĘBIE-ZDRÓJ	93 554	27,11	44,07	31,25	31,71	19,59	0,25	9,32	4,15	11,53	2,31	0,91	21,14	13,27	36,36	11,81
8	JAWORZNO	95 228	13,12	15,69	46,90	60,40	94,50	2,39	4,41	16,92	5,88	1,10	18,21	28,12	36,08	9,73	1,16
9	KATOWICE	309 621	227,72	223,89	303,81	208,58	264,81	146,15	181,06	244,26	15,66	181,47	64,18	80,87	80,40	7,51	68,53
10	MYSŁOWICE	74 998	12,70	26,48	31,16	21,68	27,24	2,00	4,00			3,56	15,74	15,12			13,06
11	PIEKARY ŚLĄSKIE	58 832	6,28	12,26	20,41	2,22	9,66	1,88	6,61	7,87	0,15	2,26	29,94	53,91	38,57	6,74	23,42
12	RUDA ŚLĄSKA	143 930	53,19	76,74	156,72	169,08	79,59	13,24	17,61	93,46	120,94	43,85	24,89	22,94	59,63	71,53	55,09
13	RYBNIK	141 177	78,36	119,27	160,30	232,66	177,54	40,79	69,85	112,83	188,43	131,18	52,05	58,56	70,38	80,99	73,89
14	SIEMIANOWICE ŚLĄSKIE	71 118	22,53	17,48	12,53	13,11	18,26	0,01	0,12	2,27	0,21	1,55	0,02	0,69	18,13	1,58	8,48
15	SOSNOWIEC	221 259	32,85	39,96	104,54	137,72	71,93	7,05	8,54	75,63	89,14	20,76	21,47	21,38	72,34	64,72	28,87
16	ŚWIĘTOCHŁOWICE	54 360	6,18	4,15	7,28	19,05	15,66				1,32	4,91				6,95	31,38
17	TYCHY	129 475	24,02	33,66	47,76	81,51	216,71	9,18	15,28	7,63	21,80	135,31	38,20	45,40	15,97	26,74	62,44
18	ZABRZE	188 401	81,16	81,45	70,10	89,93	160,67	0,00	3,27	8,69	15,38	80,07	0,00	4,02	12,40	17,10	49,84
19	ŻORY	62 044	11,94	13,97	42,60	92,95	107,59	1,99	3,51	16,72	57,46	67,78	16,66	25,13	39,25	61,82	63,00
	Total towns	2 682 413	*834,01*	*1 096,63*	*1 495,78*	*1 739,15*	1 917,00	*291,99*	*432,62*	*775,67*	*809,32*	*882,09*	35,01	39,45	51,86	46,54	46,01
	minimum	*54 360*	*6,18*	*4,15*	*7,28*	*2,22*	9,66	*0,00*	*0,12*	*2,27*	*0,15*	*1,10*	*0,00*	*0,69*	*5,63*	*1,58*	*1,16*
	maximum	*309 621*	*227,72*	*223,89*	*303,81*	*232,66*	264,81	*146,15*	*181,06*	*244,26*	*188,43*	*181,47*	*74,71*	*80,87*	*80,40*	*80,99*	*75,78*
	max - min	*255 261*	*221,54*	*219,74*	*296,53*	*230,44*	255,14	*146,15*	*180,94*	*241,98*	*188,28*	*180,37*	*74,71*	*80,18*	*74,77*	*79,41*	*74,61*

Source: authors' own calculations based on annual statements of the Regional Accounting Chamber in Katowice

Source: own calculations based on annual statements of the Regional Accounting Chamber in Katowice

Table 3. Dynamics of investment expenditure in big towns of Silesian Voivodship and EU funds in 2004-2008 (in %)

Tab. 2 Dynamics of investment expenditure in big towns of Silesian Voivodship and EU funds in 2004-2008 (in%)

No.	Towns	Dynamics									
		Investment expenditure					EU funds				
		2008/2004	2005/2004	2006/2005	2007/2006	2008/2007	2008/2004	2005/2004	2006/2005	2007/2006	2008/2007
1	2	3	4	5	6	7	8	9	10	11	12
1	BIELSKO-BIAŁA	111,53	24,42	-10,34	21,34	56,28	-59,21	-12,73	-72,22	17,55	43,16
2	BYTOM	910,59	18,57	55,45	204,26	80,21	1 720,24	-31,96	335,80	197,36	106,43
3	CHORZÓW	193,77	25,12	104,80	4,71	9,49	121 376,33	29 301,23	157,70	137,93	-32,62
4	CZĘSTOCHOWA	174,19	142,35	37,87	-7,59	-11,20	1 926,27	1 011,82	171,12	-5,56	-28,82
5	DĄBROWA GÓRNICZA	12,23	19,33	-16,65	-8,63	23,49				164,20	2,33
6	GLIWICE	289,53	117,38	24,62	69,44	-15,14	450,09	1 066,06	229,41	191,20	-95,08
7	JASTRZĘBIE-ZDRÓJ	-27,75	62,52	-29,09	1,47	-38,21	836,16	3 670,10	-55,49	178,00	-79,93
8	JAWORZNO	620,19	19,55	198,95	28,79	56,45	-53,98	84,57	283,67	-65,26	-81,30
9	KATOWICE	16,29	-1,68	35,69	-31,35	26,96	24,16	23,88	34,90	-93,59	1 058,93
10	MYSŁOWICE	114,41	108,41	17,68	-30,41	25,62	77,84	100,10	-100,00		
11	PIEKARY ŚLĄSKIE	53,82	95,08	66,55	-89,10	334,33	20,30	251,20	19,16	-98,09	1 408,66
12	RUDA ŚLĄSKA	49,62	44,27	104,21	7,89	-52,93	231,13	32,96	430,82	29,40	-63,74
13	RYBNIK	126,55	52,20	34,40	45,14	-23,69	221,62	71,25	61,53	67,01	-30,38
14	SIEMIANOWICE ŚLĄSKIE	-18,93	-22,44	-28,30	4,59	39,37	29 446,63	2 182,48	1 797,25	-90,87	647,56
15	SOSNOWIEC	118,99	21,65	161,62	31,74	-47,77	194,45	21,14	785,26	17,87	-76,71
16	ŚWIĘTOCHŁOWICE	153,49	-32,80	75,43	161,60	-17,81					271,24
17	TYCHY	802,13	40,11	41,91	70,67	165,85	1 374,47	66,52	-50,09	185,77	520,76
18	ZABRZE	97,98	0,36	-13,94	28,28	78,67			165,79	76,86	420,76
19	ŻORY	801,11	16,98	205,00	118,19	15,75	3 307,97	76,48	376,31	243,67	17,97
	Total towns	129,85	31,49	36,40	16,27	10,23	202,10	48,16	79,30	4,34	8,99

Source: authors' own calculations based on annual statements of Regional Accounting Chamber in Katowice

Source: own calculations based on annual statements of Regional Accounting Chamber in Katowice

- it's surprising that in 2004 and 2005 there were such big towns in the Silesian Voivodship which did not use the EU funds at all. It may be suspected that they had not had their projects prepared in advance, in order to finance them, completely or partially, from this important source, which was becoming increasingly relevant. Another reason for this may be initial poor preparation of the staff working on applications;
- a surprising thing are significant differences in the importance of EU funds in investment budgets of such communes as the regional capital town of Katowice, Bielsko-Biała, Żory, Zabrze, Gliwice, Chorzów. This may be due to implementation of investments with various levels of necessary engagement of such non-repayable outside funds;
- attention should also be paid to the towns in which the engagement of EU funds in the entire analysed period was very high: Rybnik, Tychy, Bytom. It may be supposed to be connected with appropriate preparation of applications already in the

first years of Poland's presence in the EU. (The ability acquired by these towns already at the stage of application for pre-accession funds).

Summing up, there are grounds to assume that most of the towns in the Silesian Voivodship have made an appropriate use of the new opportunities resulting from the Poland's accession to the EU.

4. Conclusion

The recovery of investment processes in Poland (in the Silesian Voivodship) was mostly due to the EU accession. Owing to the accessibility of EU funds, numerous public sector projects could be carried out. Moreover, the observations made so far indicate that the role of EU funds in financing of investments in communes increases regularly. However, it should be kept in mind that UE funds are just one out of many sources of financing for undertakings conducted by local authorities – the sources which are more and more frequently combined with one another. The limitation of budgetary funds and growing competition when applying for non-budgetary funds make communes face the need to use increasingly complex financing instruments which, on one hand, help them to increase the probability that an investment is going to be implemented and, on the other hand, allow them to adequately spread the risk related to individual sources of financing. It should also be noted that EU funds can only be used by communes as supplementary financing for the projects performed, and the remaining funds required for financing of investments had to be raised by communes from the state budget, generated using their own resources or a credit. Only rich communes could afford to gradually repay the credit or use free resources from their own income.

It should also be mentioned that currently in Poland a number of public projects related to the organization of EURO 2012 football championship with Ukraine are being implemented. For instance, the value of the "Construction of the National Stadium in Warsaw" project itself will amount to approx. PLN 1.2 billion.

References

Bończyk-Kucharczyk, E. – Herbst, K. – Chmura, K. 1998: Jak władze lokalne mogą wspierać przedsiębiorczość. Fundacja Inicjatyw Społeczno-Ekonomicznych. Polska Fundacja Promocji Małych i Średnich Przedsiębiorstw, Warszawa.

Denek, E. – Sobiech, J. – Wolniak, J. 2005: Finanse publiczne. Wydawnictwo Naukowe PWN, Warszawa.

Janowska, H. 2002: Strategie finansowania gminnych inwestycji infrastrukturalnych w Polsce. Uniwersytet Szczeciński, Szczecin.

Financing of investment projects in the public sector – availability of sources and their effectiveness. Stage I. Project leader: Prof. K. Marcinek. The K. Adamiecki University of Economics in Katowice. Statutory research, Katowice 2006.

Financing of investment projects in the public sector – availability of sources and their effectiveness. Stage II. Project leader: Prof. K. Marcinek. The K. Adamiecki University of Economics in Katowice. Statutory research, Katowice 2007.

Ministry of Regional Development. Website: http://www.mrr.gov.pl/

Nowa Encyklopedia Powszechna PWN. Wydawnictwo Naukowe PWN, Warszawa 1995.

Ratajczak, M. 1990: Infrastruktura a międzynarodowa współpraca gospodarcza w Europie. Wydawnictwo Książka i Wiedza, Warszawa.

Słupik, S. 2005: Inwestycje infrastrukturalne jako niezbędny warunek rozwoju gospodarczego gminy. W: Rozwój oraz polityka regionalna i lokalna w Polsce. Edit. J. Kaja, K. Piech. Szkoła Główna Handlowa w Warszawie, Warszawa.

Szczodrowski, G. 2003: Finanse publiczne. Wyższa Szkoła Zarządzania i Przedsiębiorczości im. B. Jańskiego w Łomży, Łomża-Elbląg.

The Act on Income of Local Self-Government Units of 13 November, 2003 (2003 Journal of Laws No. 203, item 1966).

The Act on Public Finances of 26 November 1998 (Journal of Laws No. 155, item 1014).

The Act on Public-Private Partnership of 19 December, 2008 (Journal of Laws of 5 February, 2009).

Winiarski, B. 2000: Polityka gospodarcza. Wydawnictwo Naukowe PWN, Warszawa.

Segmentation of teenagers in terms of their buying behavior and their attitudes towards commercials

Szabolcs Prónay[1] *- Krisztina Karsai*[2]

In my research, I intend to explore the consumption of teenagers and their attitude towards commercials and ads. My aim is to establish segments among teenagers according to their relation to commercials, their buying behaviour and their financial state. The segmentation of teenagers is peculiar and sometimes problematic, because the common segmenting methods – based on status and values – usually does not result in stable and understandable segments. The parents and the friends are the common opinion leaders of teenagers, but we should not forget the infulence of the media and the commercials within it. I conducted my research using a sample of 911 teenagers between the age of 14 and 19 from Csongrád county in Southern Hungary. In the first step I examined the teenagers relation to commercials and ads. Factor analysis were used for revealing the mean components of the commercial valuables. These mean components were: Commercial loving; Commercial reality; Commercial manipulation; Commercial status. Using these factors, five segments were made up according to the teenagers relation to commercials. In general they think commercials are manupulative and sometimes even stupid. However we can find some commercial loving segments as well. In the second step, I set up a goal of establishing complex segments based on financial, commercial and buying variables. I used factor analysis for the teenagers buying behaviour, which yielded seven factors. Applying this seven buying factors, the commercial factors and the financial variables, six segments were identified. These segments differ in location, in type of school, in financial state, in buying behavior and also in their relation to commercials. According to the research – although traditional segmentation often yield confusing results between teenagers –an acceptable way of grouping the youngsters could be found.

*Keywords: segmentation, teenagers, commercials, buying behaviou*r

[1] Szabolcs Prónay, University of Szeged, Faculty of Economics and Business Administration (Szeged)
[2] Krisztina Karcsai, University of Szeged, Faculty of Law (Szeged)

1. Introduction

The young. Interest has always been shown towards this particular age group. They are not adults yet, but not any longer children. They are in the stage of their life which most of them get over, live through struggling and later recall it in a romantic way. This is the time when our personality develops and we find our place in the world. Nowadays this process is more complex than ever. This world is not constant, not stable; the several hundred year old guidelines are less and less visible, which determined how to become adult. The environment is changing rapidly, ideals rise and fall more quickly than ever, and in this world the adults, who could set an example to the young, perhaps feel more lost than these young people. Today it is not any longer the father who shows his son how a machine works; it is not the mother who teaches her daughter about dressing. The roles seem to be interchanged. The young people teach the adults how to use technical appliances, they know what is fashionable and they try to keep up with the constant change. Now the young people are the ideals. The desire for eternal youth is more palpable than ever. But are they capable of this? I wonder whether it is a too large task for them to determine the trends and develop their personality at the same time in this world. In addition, do it in a way that they keep searching in the digital universe left alone without a stable, well-defined model and limits.

In our study we examine the consumer behaviour of this particular age group. We would like to know how they behave as consumers. What is their attitude towards shopping, brands and commercials? To our point of view, to understand the consumption of the young first we have to understand that although they are close to each other in age, they cannot be treated as a homogeneous group. Their situation and lifestyle cab be quite diverse, thus it is practical to examine the broken down in segments to understand their consumption..

2. The young

If we talk about generations in terms of marketing, one of the central elements of generation studies turns up most of the time, the Yankelovich report appearing from time to time. One of these central ideas is that the new generations typically brush aside, or at least re-interpret the values owned by previous generations. That is why it is important to understand the particular values and lifestyle of each generation, because they essentially determine their consumer behaviour (Smith–Clurman 2003). Consequently, the examination of today's young consumers, the examination of their values and lifestyle, and above all the examination of their consumer behaviour is a particular area, which has a serious academic and practical relevance.

We know about the young that they are in a quite particular situation. This is a rather complex and difficult period of life, when the individual's own personality is developed (Törőcsik 2006), during which the parent-child relationship weakens, while the contemporary relationship become stronger (Benedek 2001). This period is often called "second socialisation" (Tárkányi-Józsa 2006). However, in terms of marketing, we could call this period "first (or primary) socialisation" as well, since the individual makes his/her first independent consumer decisions at this time. That is why this period has a special significance for marketing experts.

2.1. Defining the young

In our study we deal with the consumption of the young. However, first we have to determine who we mean by "the young". This is not a simple task at all, because the professional literature is not uniform in this respect. Besides the fact that the concept of "the young" is not clear, it makes the situation more complicated that it is rather difficult to separate the "Y generation" from the also little defined concepts of "teenagers" and "young adults".

According to Törőcsik (2006), we can define the young according to age, furthermore, in terms of biology, psychology, sociology and law, but he also points out the difficulties of separating: "Though we mean a period of development and definiteness of age by that, defining the group is not a simple task, we have to draw the lines down to the children and up to the adults, which both hides several problems" (Törőcsik 2006, pp.193).

The young are often mentioned as Y generation. The Y-Gen (or Gen-Y) is a well-known sociological and marketing concept. They are determined primarily according to age; those who were born roughly between the late '70s and the turn of the millennium are meant by that. The concrete definitions are different by author: 1980-2000 (Cheung et al. 2008); 1980-1997 (Johnson 2006); 1977-1994 (Evans 2006; Noble 2009). Besides the date of birth, this generation is defined according to their common experiences and values. It is them who consider the remote control, the microwave oven and the pushbutton phone original equipment, and regard the music of the '80s old-fashioned, who can remember 11 September sharply, but hardly remember the collapse of the Berlin Wall (Cheung et al. 2008).
Noble et al. (2009) call the attention that it can lead to distortion if we treat such a wide age group collectively. They focus their studies on secondary school students (college-aged market).

Taken the caution mentioned above, in addition, to avoid literary misunderstandings, we define the young as a subset of the Y generation in terms of age, in a synonymous way with teenagers. Thus we call the people who were born between 1995 and 1989 (that is between the age of 14 and 20) young.

2.2. The values and lifestyle of the young

The young, as the other age groups, are characterised by particular topics, values and lifestyle, the bases of which are those common experiences, events, social and personal situations that they were, in one way or other, involved in collectively. Igaz Ágnes's idea exemplifies it well: "They do not call the Russian "the red" any more, for them Michael Jackson always was a white-skinned pop star, and the "Ctrl + Alt + Del" formula belongs to the fundamental operations."[3]

The values typical of the young and their particular interpretation based on Kazár (2009, pp. 8-10) and Noble et al. (2009, pp. 626) can be summarised briefly according to the following:

- *Rapidity*: They live in the digital world, they manage everything with pushing a button, and they measure the time not in hour but in second and the speed not in kilometre per hour but in megabyte per second. They are impatient, they do their activities parallel (they are chatting during downloading, while they are eating fast food).
- *Freedom*: They surf on the Internet without any limits, select from the vast number of TV channels, download any kind of music or film and try to become independent avoiding parental control more and more.
- *Self-acknowledgement and self-definition:* They would like to find and express their own personality. They create avatars, write blogs, keep twitter and have an iwiw site. With the help of these they try to show the outside world who they (at least virtual) really are.
- *Individualism*: They try to give themselves an individual look; they wish to stand out from the crowd. They form both their virtual and real living space to their own image. The personalised services are obvious for them.
- *Sense of belonging*: Besides individualism, or in contrast with it, they desire for the appreciation their peers. They are enthusiastic fans, keystones of on-line communities, who adjust to their reference group in their consumption as well.
- *Fashion*: Every form of self-representation is extremely significant for them. They are innovative and open-minded to new fashion trends; they are able to throw away older things without consideration. There is a great proportion of fashion innovators within the group, who "consider themselves more exciting, more enjoyable, more modern, more formal, more colourful and vainer individuals than the fashion followers." (Kovács K. 2009 pp 63).
- *Security*: Security is not the opposite of wanderlust for them, but it expresses that they need those reliable brands and stable groups who they can rely on in the course of developing their personality.

[3] http://www.kreativ.hu/cikk.php?id=13019 downloaded: 17. 12. 2009

The above mentioned values also refer to their lifestyle, which goes on in a fast and dynamic rhythm, and to which infocommunicational devices are connected in a greater share compared to the other age groups (Kollár 2007). In their free time they mostly listen to music, in addition they spend a lot of time watching TV or at home with their family. In their life, the permanent relationship with a partner and the circle of friends have a significant role. They live their life in a more open-minded and more adventurous way than the other age groups; however, it may be surprising that the anxiety about violence and crime has a surprisingly great role in their life (Sulyok 2003). This implies that the life of today's young people covers certain doubts and problems. Tari (2008) draws the attention to that this is the generation that is left alone, "they have been wandering the Internet for 10 years, the adults do not pay attention to them, they can do what they want." (Tari, 2008, pp.56). This state of being left alone is especially disconcerting in such a critical stage of personality development. The on-line freedom also means that avoiding conflicts and accepting one's own personality can be avoided by one click and closing the window. Consequently this age group has been given a great deal of threat besides a number of opportunities in this sense.[4]

3. Consumption of the young

The exact proportion of the young in the population can be estimated by the 2001 census. At this time the proportion of the people born between 1982 and 1996 were 18,4% of the population. This data itself suggests that it is worth dealing with a group sized like this. Although the rate of people having jobs and permanent income is far lower compared to the older generations, apart from this they do not live only on their pocket money. Besides their monthly pocket money and earnings, the "income" coming in as occasional gift, casual work, savings, grant, student loan and social support can constitute their income, but the circle of possibilities can be extended by the supporting parental background (Törőcsik 2006; Tárkányi-Józsa 2006). According to the surveys conducted in the United States the annual spending power of the young can be put at 600 billion dollars (Noble et al 2009). However, an even larger shopping potential lies behind this segment, because they do not only dispose over their own money, but also indirectly influence their parents' purchases (Andó 2007). It further increases their marketing significance if we take it into consideration that the fashion innovators and trend leaders emerge from them, in addition, the life-long brand loyalty usually develops in this period as well

[4] Tari (2008) regards the group of Emos as a kind of alarm. These young people intend to express the disillusionment in the world, dropping out and the emotional nihilism of the desolate life with their typical appearance, the main motif of which is the hair covering one of the eyes. It refers to that this world is so horrible that it is enough to see it with one eye.

(Wolburg–Pokrywyczynski 2001). To sum up, it is justified to deal with this group's consumer characteristics.

We have already established about the youth that the development of consumer personality happens at this time, we can say that the young learn to behave as consumers. In this learning the environment has a great role, which, on the one hand, sets an example, on the other hand, influences and evaluates a consumer decision. Three reference groups play a decisive role in teenagers' decisions (Benedek 2001 pp.33):

- *The family*: Primary reference group with informal, normative and positive membership. A young person follows the model of his/her parents in connection with shopping as well. The parents evaluate their child's performance in this area as well, and they reward or punish in certain cases. In addition, the parents have an important informational role; they often give advice on shopping.
- *Peers:* This is a group which the youth has a more democratic relationship with than his/her parents. It is an informal, aspirational group, that is, they set an attractive example to be followed for the young. Since he/she longs for reinforcement, he/she tries to copy their behaviour, forming his/her self-image to them. Besides, sharing experience and giving advice also have a great role in this relationship.
- *Stars:* Not membership, symbolic aspirational group. The teenager does not have a chance to join the group, the members of which still mean a model to be followed. They are influences on his/her ideal self-image and ideal social self-image, however, it should be noted that they are rather relied on as information sources.

Consequently, the groups have a significant role in the shopping decisions of the young (Benedek 2001). "Belonging to a group (or desire for belonging) is expressed the consuming of certain products and brands, which is a carrier of important "messages" about the individual" (Tárkányi-Józsa 2006 pp.73). As a consequence, they express their belonging to a group and their personality with their consumption (Noble et al. 2009). However, it is essential that the products they purchase should be identifiable (for others as well). This can be the reason why brands play key role in the case of the young people (Sulyok 2003; Törőcsik 2006; Noble et al. 2009). The image carried by the brands influences the young consumers' both personal and social image (Törőcsik 2006; Kovács K. 2009), beyond that it refers to the reliability and high quality sought by the young (Szántó 2005).

While examining the consumer behaviour of the young we cannot evade mentioning their sceptical attitude to commercials. According to Cheung et al. this generation is a lot more resistant to commercials than any other previous generations (Cheung et al. 2008). Overall, it can be said about them that they know technique well; the media played a significant role in their socialisation, so they see quickly through the marketing communication tactics (Noble et al. 2009). They appreciate

what impresses them; it can have a great effect on them. On the other hand, they avoid easily what they do not appreciate; they turn it over, disable it, or ignore it (Andó 2007; Kollár 2007). Reaching this group means a serious challenge for advertising men. Johnson (2006) formulated five criteria related to the communication addressed to the young: it should provide enjoyment, be genuine, introduce new possibilities, be interactive and the product it popularises should be suitable for self-expression. Of course it is not guaranteed that the commercial impresses this quite critical audience.

On the basis of all this it can be stated that we can find several contradictions in the area of the consuming habits of the young. It is demonstrated in Figure 1. According to our point of view the basis of these contradictions is the heterogeneity of the young group. We can find reference to that in literature. According to Tárkányi and Józsa "the behaviour models found in the case of the young are varicoloured and it would be a mistake to treat them as a homogeneous group, taking only their age affiliation into consideration" (Tárkányi-Józsa 2006 pp. 79). Kollár (2007) also points to the same, who draws the attention to that recently more and more people in the profession have raised that the youth is not homogeneous. It is a question though that according to what characteristics it is possible (if it is at all) to classify them (Kollár 2007). Based on this we thought that the mentioned contradictions can be resolved if we examine the group of the young not in general but breaking it down into several smaller groups. The made an attempt at this in our primary research.

4. Primary research

We have showed in the foregoing that the consumption of the young has a serious economic and marketing significance; however, examining it is a quite complex task and often leads to contradictory results. According to our standpoint, the national (and presumably the international) research results (also) lead to contradictory results because of two factors:

- First of all, the sample choice makes it more difficult to compare the results, since – just like in the circle of the adult inhabitants, also in the case of the young – significant differences can be discovered in the consumption pattern according to the financial situation and the residence. In the case of the young, surveying the financial situation and the segmentation according to it is even more difficult compared to what is regular in case of adults. Concerning the residence it can be stated that a considerable part of the national research takes urban (mainly from the city) young people into the sample.
- Secondly, the effort to give general explanations can lead to confusing results. The young in itself is a particular target group, consequently we tend to treat them – as well as in common sense but in many cases in terms of science –

uniformly and generalised. However, it can lead to serious contradictions if we treat this quite heterogeneous age group as a homogeneous group.

Based on this, we set two objectives during developing our research. On the one hand that – concerning the reachable region by us – we choose representative sample, thus examining the consumer behaviour of the less researched youths from small towns and villages besides the better known – and more colourful – consuming habits typical of the young people from the city. On the other hand we aimed to study the young in breakdown by segments instead of – and occasionally besides – general explanations. Hereby we tried to get the answer to our central question whether we get closer to understanding the young people's consumption if we do not study them in general but we break them down in segments.

4.1. Research method and sample choice

In the course of our research we intended to formulate statements regarding the young living in the region, so choosing the quantitative method seemed to be appropriate, the questionnaire interview to be exact. There were five parts of the questionnaire, which examined the young people's behaviour as consumers from different aspects.

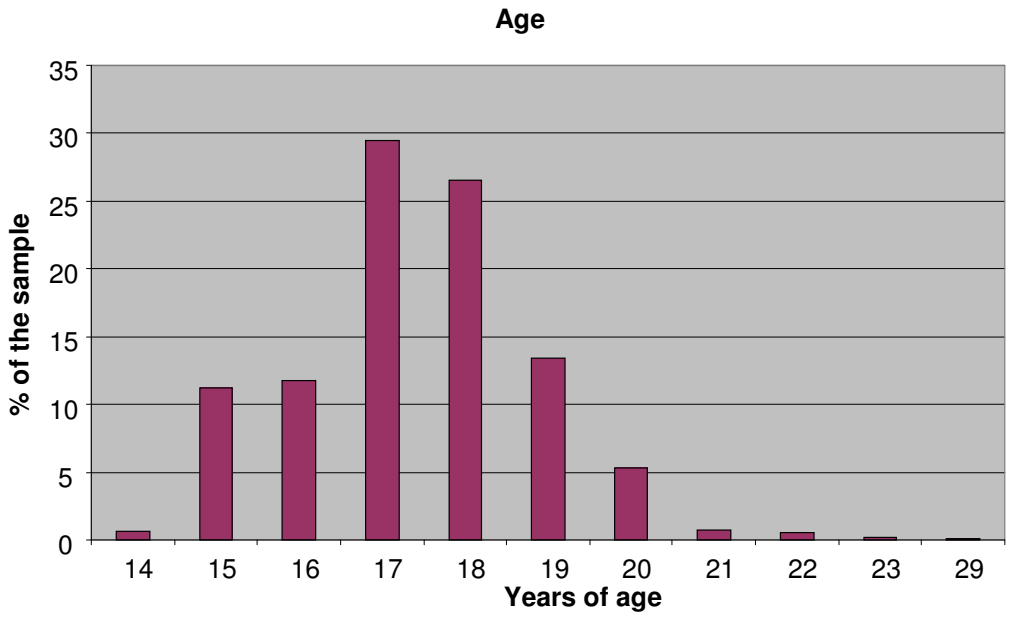
- In the first part we asked them about their shopping habits. Here they could give with multiple-choice where, what and from what source they shop in general.
- In the second part we studied their attitude to shopping and brands with Likert scale questions. Here we wanted to know the influence of the brand, the price and the quality factors on their consumption and the interaction of these factors.
- In the third part we examined their attitude to advertising also with the help of Likert scale questions. We were curious about the sympathy related to commercials and the manipulative associations linked to them.
- In the forth block we asked them about how they react if they are deceived and damaged in a purchase. Here the multiple-choice questions were mainly aimed at whether they try to take action in these cases and if yes, who they turn to.
- The last issue included taking down the sociodemographic data.

The questionnaire was filled out by 911 people in total. The main characteristics of the sample:

The 53,9% of those asked are women and 46,1% men. We aimed at people between 15 and 20 years of age (Figure 1). We find the age groups of 17 and 18 among the respondents for the most part – this age group gives the 53% of the total sample. The number of under 15 and over 20 is minimal, their proportion even jointly does not reach the 2,5% of the total case number. The dividing line

of the sample is almost exactly between the age of 17 and 18 since the proportion of under 18 is 53%, while the rate of majors is 47%.

Figure 1. Age distribution of the sample



Source: Own creation

- As we have already mentioned, we laid particular emphasis on examining not only the consuming habits of the young from cities. The 62 % of the respondents are town-dwellers, approximately every fifth of them (18%) lives in county town. The inhabitants of small towns and villages give 2/5 (38%) of the sample.
- The direct consequence of the distribution according to residence that the rate of grammar school students is lower than generally in studying the consumption of the young. The 11,4% of those who filled in the questionnaire attend vocational school and trade school, slightly more than its half attend secondary technical school (53,9%), while about a third attend grammar school (34,7%).
- Two thirds of the sample (65,6%) is given by those living in households of 3-4 person, a quarter (25,9%) live with more people, while 8,5% live in a household of 1-2 person.

*4.2. The general characteristics of the **frequency distribution** of answers*

Taking down and recording the data was carried out by the Sociology Department of the Faculty of Arts of the University of Szeged on the commission of the Southern Great Plain Regional Consumer Protection Association. The report they made analyses the distribution of the sample thoroughly, thus we mention only some main general characteristics in what follows. As we declared in our research objective we intend to examine the possibility of a more thorough and deeper inspection available by segmentation instead of general statements, thus these general characteristics are linked only indirectly to our research objective, so we describe them merely sketchily[5]:

- The big department stores and plazas are considered as most popular – the former is preferred by about the third of the young, the latter by the quarter. The fact that one of the two types of stores was named by 58,9% of the asked people shows the popularity of sale places with big ground-place well. It deserves attention that those living in small town and villages mentioned department stores as favourite shopping place in a significantly higher rate compared to town-dwellers, in addition, the boutiques and small shops also have greater popularity in their circles. At the same time, the plazas mean the favourite place to spend money rather for those who live in towns.
- If they wanted to buy more valuable things, the 2/5 of the asked people would ask the family for money, the third of them would save money and 17% would buy it at their own expense. The difference between the genders is notable: 24% of the men would plan to spend their own money, while the proportion of women is only 11,9%. Half of the girls (50,3%) would mainly ask their family for money for shopping, while less than a third (31,5%) of the boys would turn their parents.
- In making shopping decisions (according to the details in Table I.) the young people accept their close friends' and pals' opinion (30,5%) and their parents' advice (26,9%) most of all. Other family members and relatives have a key role in the choices of 6,5% of those who filled out the questionnaire, thus the opinion of close and more distant relatives together was the most important for the third of the participants (33,4%), gaining bigger emphasis to some extent than the friends' and pals' advice. As for the girls, the close friend's and pal's role can be considered as significantly bigger than in the case of boys.

[5] 'You are always wise customers" – research report of the Sociology Department of the Faculty of Arts in the University of Szeged, 2009

Table 1. Gender differences in who inluences the young during making a shopping decision

	Valid percent (men)	Valid percent (women)
An expert	5,3	0,2
Shop assistant	1,9	0,6
Other pal	1,9	1,2
Other person	0,7	0,6
I don't need anybody's opinion	23,2	19,8
Parent	27,1	26,6
Other family members	5,8	7,0
Partner	11,1	13,8
Close friend, pal	22,9	37,3

Source: You are always wise customers" – research report of the Sociology Department of the Faculty of Arts in the University of Szeged, 2009

- The survey shows that the young basically have a critical attitude to commercials. The majority of them think that they are manipulative, distorting, misleading and only good for manufacturers making money. The general evaluation of a value above 3 (measured on a scale from 1 to 5) was found only in the case of statements containing negative attitude towards commercials, so the five least likeable statements definitely have critical overtones. In accordance with this, the five least popular statements consist of positive assertions related to commercials. Consequently, it can be stated that the asked people are unequivocally "anti-commercial".
- The 98,4% of the respondents gave information about how they know most which the "trendy" brand is. The third of the young (33,1%) are informed based on commercials, while their similar proportion (31,1%) collect information about it from pals and friends. The fifth of the respondents (19%) picks up in the school what is "cool", while only 5,5% from home and the members of the family. So the role of non-home environment (pals, school) has crucial importance, together half of the asked (50,1%) learn form here what they should wear to be trendy.

4.3. Developing the segments within the sample

In the foregoing we generally described the young involved in the sample according to their consuming behaviour. However, our research objective was breaking them down into different segments instead of (and besides) a general description, and examining them along these segments. As the first step of developing the segments we classified with factor analysis, and we decreased the number of variables serving

as the base of segmentation. We did the factor analysis both on the variables of shopping attitudes and on the variables of attitudes related to commercials.

4.3.1. Segments based on the attitude to commercial

We ranked the variables related to commercial into four factors using principal component method. The KMO index and the Bartlett's spherical test both refer to that there is connection between the variables, furthermore, applying the factor analysis appropriate to analyse the data (KMO=0,814; Bartlett Test: χ^2 = 2580,911; degree of freedom: 66; Sig=0,000). We decreased the number of variables in the factor analysis until the weight of each element was over 0,5 in the final (rototilled) version; each variable correlates with only one certain component significantly; the explained variance ratio reached 63% cumulated level; and the received factors can be interpreted. The eigenvalue diagram also indicated four factors. We formed factors according to Table II. (we treated the variables normalised).

Table 2. Factors according to the attitude towards commercials

	Factor			
	1	2	3	4
Factor 1: Commercial love				
I like commercials	,799			
Most of the commercials are beautyful, it's nice wathing them	,832			
Most of the commercials are funny	,689			
Most of the commercials are amusing, even good for leisure	,662			
Factor 2: Commercial reality				
Most of the commercial statements are valid.		,772		
Most of the commercials shows reality.		,820		
Commercials bring the people's dream closer.		,647		
Factor 3: Commercial manipulaty				
Commercials mannipulate people for more consumption.			,744	
Most of the commercials are misleading, they con people.			,743	
Life is not like what commercials shows.			,779	
Factor 4. Commercial status				
Commercials shows what to strive for.				,630
Most of the commercials are made for wealthy people.				,847
Rotating method: Varimax, with Kaiser-normalisation, in 5 iteration steps..				

Source: Own creation

Afterwards we examined whether we could form – as homogeneous as possible – groups in the circle of the young involved in our sample according to the attitude to commercial. Besides the factors above we used the variables left out from the factors while developing the clusters. The hierarchic method resulted in too many clusters (15) both in terms of manageability and interpretability, so we worked with K-center method. We studied more cases, based on which the solution with 5 clusters seemed to be the most useable. In the followings we describe the clusters and their centres in Table III.

Table 3. Clusters and cluster centroids according to the attitude of the young towards commercials

	Cluster				
	1	2	3	4	5
Name	*Commercial hater*	*Consider commercial useful*	*Commercial adorer*	*Sceptic*	*Ignorer*
Numerosity	175	165	144	237	128
Commercial love factor	-,95749	,58946	1,01823	,07745	-,78646
Commercial reality factor	-,87087	,46719	,98987	-,12928	-,46802
Commercial manipulity f.	,91275	-,44201	-,55659	,45371	-,79161
Commercial status factor	-,46033	-,30296	1,15214	,11833	-,51934
Most of the commercials helps in consumption decisions	-,87912	,51435	1,03296	,03504	-,76244
Most of the commercials are untrustworthy	,76667	-,43607	-,30332	,45412	-,97819
Most of the commercials are violent	,53341	-,47111	,01187	,25587	-,70144
Commercials are essential	-,86428	,66835	,49691	,13827	-,57909
Commercials are only for the producers more interest	,63975	-,80369	-,11679	,42965	-,41199

Source: Own creation

The above values express a relative relation between the clusters and the variables. Describing the absolute relation was not our primary objective; however, we refer back to our earlier statement according to which the young generally had negative attitude to commercials. Nevertheless, an important statement is that behind this general antipathy we can find groups of young people who – with different

reasons but – like commercials. The main objective of our research is revealing exactly these kinds of – hiding behind the generality, typical of certain groups of the young – characteristics. Therefore, we can describe the clusters briefly according to the following:

- *Commercial haters:* These young people have expressly negative attitude to commercials. They dislike them, they think they are of no use, they cannot rely on them and with their help the manufacturers influence the consumers to swindle as much money as possible from them. In their opinion the world would be a better place without commercials.
- *Those who consider commercial useful:* They think the commercials are needed, mostly because we can get familiar with the market and the brands with their help. In their opinion the primary aim of the commercial is informing, so they can be regarded useful rather from the aspect of consumers than from the aspect of manufacturers.
- *Commercial adorers*: They expressly like commercials; they consider the presented things as examples to be followed. The commercials inform about the behaviour model they should follow.
- *Commercial sceptics*: They basically do not condemn commercials; they just have reservations about their manipulative character. They consider them as necessary evil, which are parts of our everyday life; however, if we rely on them too much, they can easily lead us wrong.
- *Commercial ignorers:* They do not deal with commercials, they ignore them. They dislike them, but they do not care about their manipulative character either, they consider them simply unnecessary.

4.3.2. Developing the clusters of marketing aspect

After classifying according to the attitude towards commercials we examined the characteristics of the young people's consumer behaviour with a more general segmentation having marketing aspect. For this clusterisation we used the factors formed form the variables related to shopping and the variables related to the financial situation of the family besides the earlier described factors related to commercials. We classified the variables related to shopping into six factors using principal component method (KMO=0,845; Bartlett Test: χ^2 = 3983,837; degree of freedom: 190; Sig=0,000). We decreased the number of variables in the factor analysis until the weight of each element was over 0,5 in the final (rototilled) version; each variable correlates with only one certain component significantly; the explained variance ratio reached 64% cumulated level; and the received factors can be interpreted. The eigenvalue diagram indicated seven factors but since only one variable belonged to the last factor, later on we treated it merely as a variable and we formed six factors:

- *Brand personality factor:* It refers to the relationship between the bought brand and the customer's personality. If its value is high the customer considers whether the given brand suits them and if yes, they stick to it.
- *Brand quality factor*: According to this purchasing brand-name products can be explained by their assumed higher quality.
- *Brand community factor*: According to this those people who buy the same brand think in a similar way, the identical consumption creates a kind of community relationship.
- *Appearance factor*: The appearance, the fashion, it refers to the significance of the visible consumption.
- *Considered budgeting factor*: It refers to the consumption determining role of the price.
- *Experimentation factor:* It shows the tendency to try out new products and new brands.

(Shopping in big stores factor: It is not necessary to treat it as a factor because it merely correlates with the "I like shopping in big stores" variable.)

After this we segmented the young based on the mentioned factors and the variables related to state. The hierarchic clusterisation turned out to be quite circumstantial (it referred to the existence of 25 clusters, which would have been anxious in terms of interpretation), so we chose K-center clusterisation. We examined the distance of two clusters also in the case of five, six and seven clusters, and finally we considered the six cluster solution as suitable (the later interpretability also played a role in the decision). We describe the main characteristics of the developed clusters in Table IV. according to the consumer behaviour of the six clusters which can be described complex.

Table 4. Complex clusters according to the consumer behavior of the young

	Cluster centroids							
	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6		
Name	*Self-value followers*	*The cool*	*Out-siders*	*Obeyers*	*Intellec-tuals*	*Day-dreamers*		
Numerosity	83	83	56	79	142	82	F-value	Szig
Family welfare	-1,2127	,77029	-,28565	,31680	,44649	-1,0182	109,457	,000
Recently upgrade of the family welfare	-,68503	,73303	-,42819	,52623	,07754	-,71293	45,063	,000
Able to save money.	-1,1450	,50646	-,39297	,59894	,55192	-1,0922	133,048	,000
Recently went on holiday abroad.	-,69098	,75839	-,63421	-,06464	,36131	-,62294	46,622	,000
Amount of pocket money	-,27885	,84811	-,39987	,00880	-,05859	-,42833	22,390	,000
Commercial love factor	-,67542	,27057	-,11599	,71666	-,43761	,80255	46,784	,000
Commercial reality factor	-,66651	,56396	-,18326	,66365	-,58277	,52914	50,095	,000
Commercial manipulity fact.	,82975	-,49796	-1,0793	-,20733	,52284	-,12459	52,688	,000
Commercial status factor	,06172	,38100	-,52843	,55133	-,61604	,59714	33,031	,000
Brand personality fact.	-,51605	,84867	-,84562	-,44628	,07228	,53654	46,938	,000
Brand quality factor	-,42115	,61960	-,39546	-,47743	,28256	,21015	22,206	,000
Brand community fact.	-,40865	,63694	-,93625	,02686	-,15721	,57781	34,488	,000
Appearance fac.	-,44126	,74732	-,65712	-,54999	,10045	,38538	29,592	,000
Considered budgeting factor	,21714	-,32520	,21901	,05587	,03009	,21849	3,799	,002
Experimentation factor	-,07270	,38714	-,09454	-,14097	,06132	,21440	3,556	,004
Distances of the cluster centroids								
Cluster 1		4,719	2,546	3,548	3,039	2,850		
Cluster 2	4,719		4,186	2,706	2,714	3,547		
Cluster 3	2,546	4,186		2,687	2,859	3,223		
Cluster 4	3,548	2,706	2,687		2,556	3,080		
Cluster 5	3,039	2,714	2,859	2,556		3,485		
Cluster 6	2,850	3,547	3,223	3,080	3,485			

Source: Own creation

In the followings the description of each segment can be found. We took other variables into consideration besides the variables used for developing the clusters.

- **Self-value followers:** Grammar school students from families with bad financial background living in small towns who spend little on appearance. They reject buying well-known and brand-name products because they think brands and the communication surrounding them is only good for manipulating the consumer. They are independent; they do not really listen to other people's opinion, especially not of their parents. Within the sample, the rate of smokers is the higher here. They expressly object to being influenced, they like going their own way. They would not change for anybody's sake. They are considered customers; they search for the price-value ratio adequacy.
- **The cool:** Vocational school and secondary technical school students from small towns, mainly boys. They have a relatively serious amount of own money which is due to the good financial situation of their family. They live for parties instead of culture and entertainment. The appearance is important for them, which they spend much on. They insist on well-known brands, they think it is important that a product should be fashionable and suit them. They typically shop in plazas, but they go to second hand shops for a brand name piece as well. The pals are important for them, besides commercials they can inform them what is trendy, they spend on keeping contact with them – Internet and phone – far more than the average. They consider themselves trendier than the other students in the school because, in contrast to them, they do not wear mass products. On the whole, they are typical innovators and opinion leaders.
- **Outsiders:** Poor trainees from villages who are isolated and they take part in the consumer society only tangentially. They do not ask their family for money but they save up themselves. They are not interested in brands and commercials at all, they do not care about what is trendy. They spend equally little on appearance, health care, phoning and Internet. On the whole, they can be described as considered customers, for whom one of the most important decision aspects is the price of the product and whether it is practical, while its quality is far less important. If they are cheated while shopping, they do not complain, they rather leave it at that. On the whole this is a group that tries to get on in life with their own meagre means.
- **Obeyers:** Secondary technical school students from well-to-do families, mainly girls. They accept their parents' opinion, they shop together with them, they live on their money, they do not have own savings. The rate of smokers is the lower here. There are several people who have a great effect on them (even their teachers and celebrities are among the influential people). They like commercials but they are interested in that what they buy should be at reasonable price instead of the brands. On the whole we can say that they try to identify with the girl role learnt in childhood as much as possible. They are susceptible to the behaviour and lifestyle

models given from outside, whether it is a „girl"-idea determined by the parent or the „girly"-role seen in the media.

- **Intellectuals:** Grammar school students from towns who depend on their parents significantly – although they try to decrease the dependence. Their own community and their friends (besides their parents) form their opinion, but they try to develop a particular, critical view and personal opinion. One part of this is that they have a critical attitude to commercials, they find them manipulative. During their shopping they follow their own head, they find it important that a product should be nice and high quality, which means that it suits them and well-tried. If a complaint takes place, they solve it themselves. On the whole, they have a critical way of thinking and less follow models. The consumption determines their personality to a lesser degree.
- **Daydreamers:** Vocational school students, mainly girls. They do not have own money, their parents support them, who has definitely bad financial possibilities. Their pals have a strong influence on them. Compared to the average, they like commercials very much, what they see there they consider as an example to follow, but it remains a dream because of their restricted possibilities. They find it important that what they buy should be of good quality and first of all it should look good, because the appearance is particularly important for them. It is interesting that they think a trendy brand has high quality but the cheapo brand is not of bad quality but a cheap mass product. They feel they have been cheated several times because they were made to pay more money (this can refer to that they wanted to get the wished products and brands at a lower price). On the whole, they are typical model followers, who would like to reach the wished idea – which is mainly determined by the Trendies – but for financial reasons they have less opportunities. Daydreaming and pretending remain.

5. Summary

In our study we examined the consuming habits of the young. Based on the professional literature we found several contradictions in connection with the consumption of this age group. With our quantitative examination we intended to reveal that to understand the young people's consumption first we have to accept that this is a very heterogeneous segment, thus treating it as a homogeneous segment is not adequate in terms of general description and marketing. Besides this main statement, we consider several literary statements justified based on our research:
The family has a serious influence on the young people's consumption. The parents have a key role both in providing the financial resources and in informing.The close, membership reference groups, friends and pals have a strong opinion-shaper role in the young's life. Those belonging to the given segment tried to adjust to the values represented by the community, which also appears in their consumption.

There are a great proportion of fashion innovators ("Trendies") among the young. For them buying brand-name clothes, the harmony between the personality and the consumption and keeping up with fashion are especially important. We find it quite an interesting statement that this is the group the members of which influence the behaviour of those belonging to other segments, setting a kind of example to be followed (we mean the "daydreamers" and "those who want to be suitable").

Besides these we reached such statements which supplement literary statements up till now, they may partly modify or refine them:
In our country, for the majority of the young the consumption does not have as important role in life as for the young in the western countries. In the case of certain segments ("Trendies", "Daydreamers") the consumption is very decisive, but the majority do not form their personality based on that. To interpret this result it is important to mention that the most frequently examined city youth (who the personality determination through consumption is rather typical of) made up only a small part of the sample.

Although the young people's attitude towards commercial turned out to be generally negative, we can state that it may extremely differ from each other by segments. Thus we found segments where the attitude to commercials is definitely positive.

On the basis of all this, our most important suggestion for practical experts to approach the group of the young in a segmented way. This is quite a heterogeneous age group, the members of which are now searching for their real consumer personality. This is an excellent opportunity for an occasional well-aimed communication to attract and charm them, and due to that, for the company to create a long-term relationship with their new customers. However, these young people have grown up already in the consumer society, they know the general marketing practices, and they are not fascinated by some striking clichés. The business that does not make an effort to examine this group thoroughly and broken into segments will not be able to gain their trust.

We have to mention as a barrier to our research that its statements can be generalised only in this particular region. While developing the clusters, the subjectivity of the researcher interpretation is large resulting from the characteristics of the method, thus the interpretation of the formed clusters bear the distortions originating from the personal interpretation of the people who carried out the research. To interpret the results more accurately it would be useful to conduct more comparative researches with the same questionnaire in other regions, with special regard to the capital.

References

Andó A. 2007: A reklámkerülő fiatalok elérése indoor eszközökkel. *Marketing & Menedzsment*, 41, 4-5, 79-90. p.

Benedek, J. 2001: A referenciacsoportok szerepe a tinédzserek fogyasztói döntéseiben. *Markering&Menedzsment,* 35, 4, 32-37. p.

Cheung, L. – Harker, D. – Harker, M. 2008: The state of the art of advertising from the consumers' perspective: a generational approach. *Marketing Review;* Summer2008, Vol. 8 Issue 2, 125-146. p.

Johnson, L. 2006: *Mind Your X's and Y's*, Free Press: New York.

Kazár, K. 2009: *Fiatal fogyasztók (?) A Szegedi Tudományegyetem jegyzettámogatási rendszerének elemzése.* (Szakdolgozat, Szegedi Tudományegyetem, Gazdaságtudományi Kar)

Kollár, Cs. 2007: A budapesti ifjúság fogyasztói csoportkultúrája az info-kommunikációs társadalomban és ennek marketingkommunikációs aspektusai. *Marketing&Menedzsment* 41, 1, 26-31. p.

Kovács, K. 2009: A divatterjedés és a divattermékek fogyasztását befolyásoló tényezők empirikus vizsgálata a hazai fiatalok körében. *Marketing&Menedzsment*, 43, 1, 62-71. p.

Noble, S. M. – Haytko, D. L. – Phillips, J. 2009: What drives college – age Generation Y consumers? *Journal of Business Research* 62 (2009) 617 – 628. p.

Smith, J. W. – Clurman, A. 2003: *Generációk, márkák, célcsoportok*. Geomédia Kiadó: Budapest.

Sulyok, A. 2003: *Az Y generáció elérésének titka.* (Szakdolgozat, Budapesti Gazdasági Főiskola, Külkereskedelmi Főiskolai Kar, Budapest)

Downloaded: http://elib.kkf.hu/edip/D_10719.pdf, 2009.07.11.

Szántó, Sz. 2005: *„Mondd marketing, van még esélyed?" Avagy betekintés a fiatalok fogyasztói magatartásába.* Budapesti Gazdasági Főiskola Külkereskedelmi Főiskolai Kar Szakmai Füzetek, 2005. 16. szám 102-111. oldal.

Downloaded: http://elib.kkf.hu/okt_publ/szf_16_13.pdf, 2009.09.26.

Tárkányi, E. – Józsa, L. 2006: A fogyasztó „második szocializációja". *Marketing&Menedzsment* 40, 5-6, 73-79. p.

Törőcsik, M. 2006: *Fogyasztói magatartás trendek*. Akadémia Kiadó: Budapest.

Tari, A. 2009: Y generáció. *Fordulópont*, 2008. (10. évf.) 3. (41). sz. 53-60. p.

Downloaded:http://www.ikerhold.hu/www.fordulopont.hu/FP-41_y-generacio.pdf, 2009.09.29.

Wolburg, J. M. – Pokrywczynski, J. 2001: A psychographic analysis of Generation Y college students. *Journal of Advertising Research*, 41, 5, 33–53. p.

Hypothetical models of food consumption behavior by the elderly

Sándor Illés[1] *– Katalin Végh*[2]

Aging is not only a challenge for food producers, but also provides a treasure trove of new opportunities for product development and innovations. We have designed a prototype model of food consumption of the elderly with the goal of calling the attention of producers and growers to the specific needs of this growing market segment. However, the food consumption habits of the elderly cannot be compacted into a single model, since this specific age group is not homogenous. By designing hypothetical models we are hoping to facilitate the further research of the issue. In the near future we want to test and try our general models in practice, involving the producers as well as the consumers. The models point at the specific needs of aging populations and their choice of food products based upon reasonable decisions and focusing on the future. Considerations of the future perspective are an important feature of this behavior, since it might significantly improve the quality of life and health awareness of the elderly, their decisions in the field of consumption, and ultimately their general health and welfare conditions.

Keywords: aging population, food consumption behavior model, longevity

1. Introduction

The *food consumption habits* represent a specific feature of the consumer's behavior in many respects. Firstly, they are directly or indirectly linked to the human existence; secondly, they result from historical, biological, social and cultural processes; and thirdly they represent one of *the most complex forms* of human beings. The rationality of eating habits is limited, since the information available to individuals and/or the public is not comprehensive: sometimes consumers have little information and sometimes they are flooded with too much. The limited ability of individuals to analyze and evaluate information, the reliability of information and the asymmetric information flow amongst the market actors just accentuate this fact. The specifics of decision-making in small communities make the overall mosaic even more colorful. The limited ability to absorb information (prior to, during and after consumption) and the consequent perception and observation (experience,

[1] Dr. habil. Sándor ILLÉS, senior researcher, HCSO Demographic Research Institute, Budapest
[2] Dr. Katalin VÉGH, associate professor, János Kodolányi University College, Székesfehérvár

knowledge, recollections etc.) lead to a distortion rather than a crystal clear rationality, and this distortion might even be further exacerbated by unconscious factors.

The differences between the consumption habits of the young, middle age and elderly generations are *undoubtedly clear* (Cseres-Gergely - Molnár 2008). Aging as a universal process has generated new requirements for food production and processing, as well as for the production of cooking tools. Nevertheless, marketing has continued to focus *primarily on the younger generation* and has seldom noticed the demands of *the elderly*, even though they represent a great market potential. Developed countries have taken notice of the aging society and have started to adapt their strategies to their new needs by introducing "senior-friendly" products to the goods and services market (Pettigrew 2005). These rapid changes have raised the question: to what extent will the goods aimed at the current seniors suffice the needs of *future seniors*. Expenditures in developed countries indicate that so far, foodstuffs have represented a major share in the overall consumption, but we can realize signs of a decreasing trend. We are living in a time of expanding goods and services not linked directly to simple subsistence. In general we can say that a part of income spent on foodstuffs (with some exceptions) has been decreasing and we are facing a quantitative saturation (Keszthelyiné 2004, Lehota 2004). The desire to preserve good health is a new phenomenon of growing importance. We can presume that the Hungarian society will also start paying more attention to the quality and origin of the goods it buys.

2. Time related aspects of food consumption by the active elderly

The conceptual scheme in this chapter represents the time-related aspects of *food consumption by the active elderly* and the long-term changes in the habits of this same group in three different time periods: in the past, at present and in the future. The Chart 1 also includes *the general factors* we consider relevant in this aspect.

The most *significant determinant* of food consumption in the *past* was the strong influence of experience and standards. Due to lack of *information, experiences and traditions* were inherited from fathers/mothers by sons/daughters. The under-developed delivery and trading methods were not able to ensure unlimited access to food products originating from different geographical areas. Many food products were available only seasonally, since their shelf life was limited. The purchasing methods were also limited: most often the consumers bought products directly from the producers. Before the development of a global economy, there were only *local and/or regional markets*. However, the past factors were also an advantage for many reasons, e.g. buying directly from the producer in an open-air market made the buyers feel safe about the origin and the freshness of the product. In the past, consumers strived for extending the shelf life of products and in order to

achieve this they used different procedures and storage chambers, according to actual experience and possibilities.

The *modern era* can be characterized by an unlimited flow of *goods and information.* The changes that have taken place in the food market in recent decades, as well as those to come in the near future, are resulting from the following processes: aging population, increase in the number of single-member households, improved health care, changed employment structure, rapid development of technology, increase in international traveling and the increasingly important role of the media.

Since the start of *first demographic transition* (i.e. the last quarter of the 19th century), life expectancy has increased significantly in Hungary. This process can be associated with the changed lifestyle and improved standard of living. In addition, the population is well informed by *medical scientists* about different factors causing diseases, and this logically has lead to a change in our dietary habits (e.g. massive increase in allergies, sensitivity to foodstuffs, high mortality rates due to cardio-vascular diseases, gastroenterological disorders etc.). In many cases, however, one of the consequences of the efforts to extend the shelf life of products is the excessive use of preservatives in food products, which may cause allergic symptoms. One of the distinctive features of the present time is the unlimited food supply. The seasonality (the original indicator of the rhythm of life over the year) is slowly fading away. This is due to exporting and importing, and also to the fact, that irreconcilable distances no longer exist (Simai 2007, Nemes Nagy 2009). This is partly a result of increased *international traveling* (Rátz et al. 2008), and partly due to the expansion of food store chains, which offer a broad product selection throughout the year.

Our objective is to outline a possible scenario of the future based upon present considerations (Nováky 2006). In all likelihood, the change in the structure of consumption will continue. Focus on the near *future* (as a key factor of behavior) will be a *typical feature of elderly consumers.* This can have a very serious impact on their health awareness, health-related decisions and the development of their health conditions. Information on food and nutrition based on the most recent facts might become an important sector of mass media activities, since once provided with the necessary information one could actively contribute in the preservation of his/her own health.

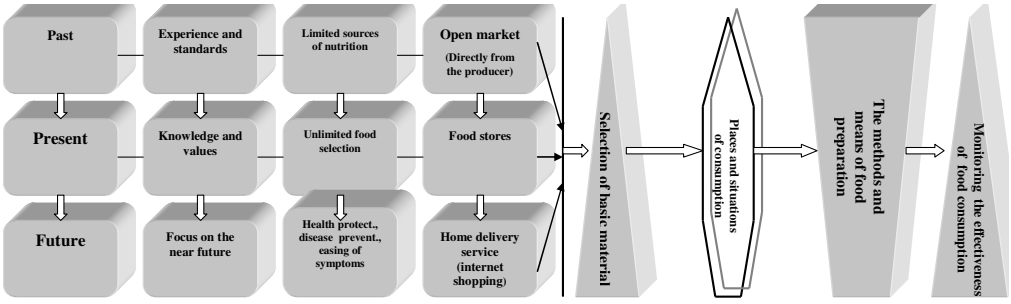
In general we can state that a new wave of purchasing and consuming habits is unfolding and influencing the behavior of the middle-age generation, and in time, it will integrate into the everyday habits of the elderly consumers. The *key characteristics* of this new wave are the following:

- Increasing popularity of natural products (“bio”);
- Increasing demand for functional food products;
- Increasing interest in the healthy life style (i.e. physical activity, diet…);
- Increasing sale of “convenient” products with short preparation time;

- Lessening of time spent shopping;
- Possible increase in the time spent preparing food;
- Preferring "at one place, once a week " shopping;
- Increase in eating out options.

The last phase of the conceptual scheme depicted in Chart 1 indicates "usefulness" as one of the aspects of food consumption by the elderly. In the past, this factor was not the focus of attention, when the goal was mere nutritional survival. However, in the present and in the future, the improvement of the quality of life will become increasingly more important. The priority for the consumer will be to rank the product characteristics by usefulness and effectiveness. These can be determined upon the previous information and experience the buyer had available. The key characteristics of food products are specifically important, including taste, freshness, healthiness and also the price. In regards to the uncertainty factors, we need to understand, that it is not possible to define elderly consumers in general, because this age group is highly heterogeneous. In addition, the mental state, income situation, health condition of a person varies and, depending on these factors, one may often reconsider the usefulness and effectiveness of food products. Therefore if we want to maximize the value of a product, we need to carefully monitor the changes in the consumption habits of this fragmented target group, as well as the characteristics of the products intended for it.

Chart 1. Time-related aspects of food consumption by the active elderly



Source: Designed by the authors

3. The future oriented model of quality and healthy food consumption behavior by the elderly

The model represented in Chart 2 has *three fundamental starting points:*

- Economic and societal (cultural) factors;
- Individual factors;
- Food product factors.

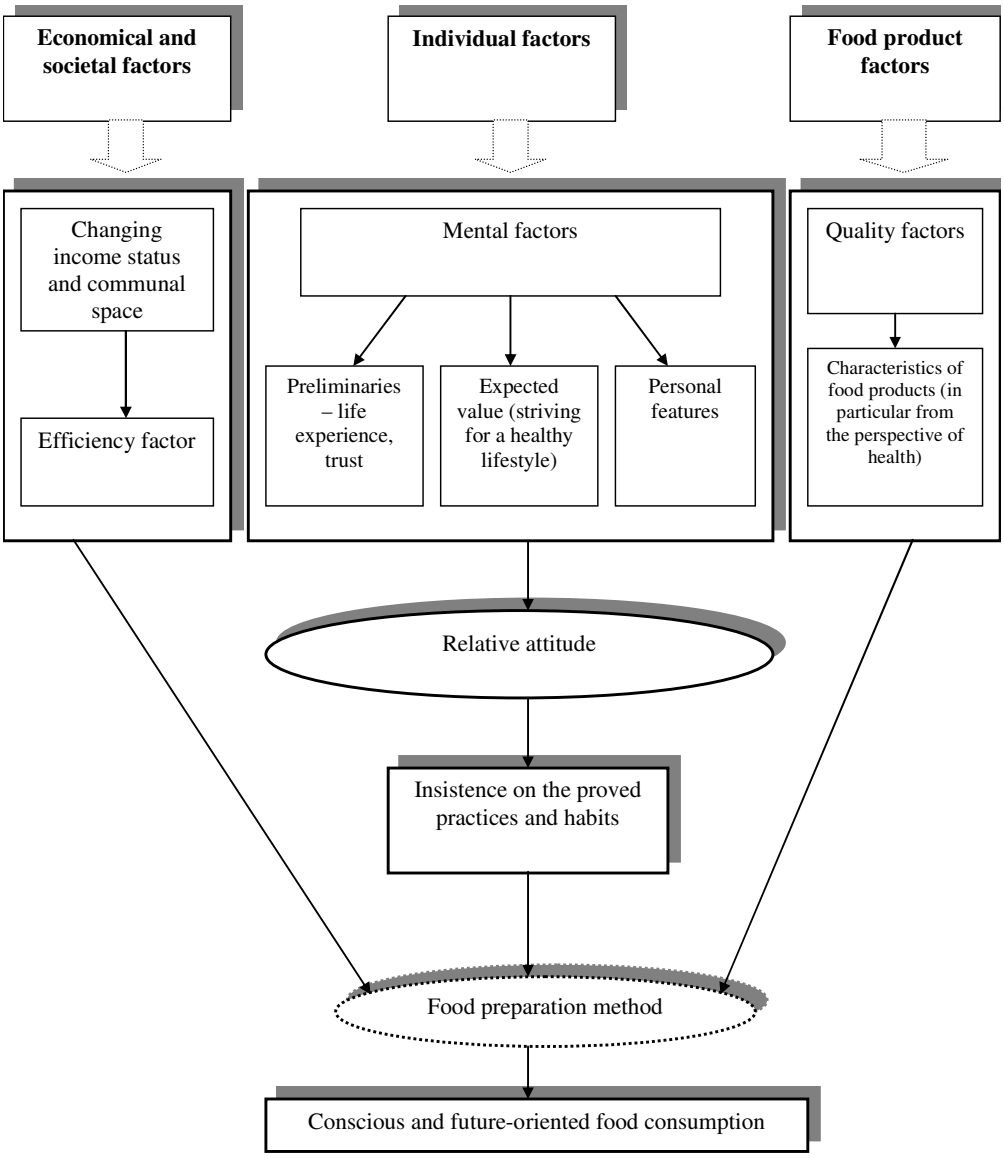
The model has been elaborated based upon the classic Pilgrim model of *food consumption behavior* (1957). The fundamental determinant of this model is perception. The model starts from the mental and chemical characteristics of a food product influencing the physiological needs. While shopping and eating, one is striving for a balance between the perceived reality and objective reality (cognitive dissonance). This indicates that there is often a difference between what we think in relation to food consumption and what we actually do.

Analyzing the *economic and societal factors*, it is important to take into consideration the changing income situation of the elderly and their relationships with spouses and any household members. Given this changing situation, the elderly have to be efficient and balance consumption with income.

The *health condition* of the Hungarian population had alarmingly worsened before the change of the political regime (Gaál 1998). Interestingly, since the mid-1990's, there has been an unexpected improvement in the Hungarian mortality rates (Józan 2008), which has stabilized and indicates the improvement in the general health state of the population. The accelerating speed of life also requires education about what constitutes a healthy lifestyle. To a great extent, maintaining good health results from conscious efforts, since a person is not only a passive sufferer of outer conditions that make him/her feel good/bad, but also is central in creating these conditions.

The *individual factors* are evidently mental factors. We defined three subcategories in our research: the specific cognitive factors, perception factors, and the influencing factors including motives. The fundamentals of the specific *cognitive factors* are the importance of health (relative importance), external health control, and the perceived health condition. The mental factors are associated with relative attitudes towards a product, i.e. they indicate the relation of the consumer to the food product (Hofmeister-Tóth 2006, Törőcsik 2006). Since the shaping of *attitudes and how* they can be changed precede knowledge and learning, marketing can play an important role. The effective use of communication along with personal habits rooted in a cultural context can draw attention to the importance of the methods of preparing food. The information flow facilitates food consumption in line with healthy lifestyle standards and considerations of rational values (Szántó 1998).

Chart 2. The future-oriented model of quality and healthy food consumption by the elderly



Source: Designed by the authors

4. Developments and process expected in the near future

The f*undamental factors relating* to food products include food quality factors and food product characteristics. When designing their quality model, Csete - Láng (1999) tried to capture all distinctive dimensions of quality. According to their interpretation, quality is the result of a complex and multi-faceted process with various different aspects, including biological, technical, technological, health, consumption, marketing, nutritional, environmental, food safety etc. In a concise wording: Quality is the suitability for the purpose.

It is increasingly characteristic for the category of consumers with a higher level of health awareness to draw back from mass-produced goods and to demand products with a higher added value and special quality meeting their specific dietary needs. When we talk about quality, firstly we refer to advantages of nourishment, and secondly to the natural origin, value for the consumer and reasonable price – briefly the marketability (Lakner–Sarudi 2004). Food producers develop f*unctional food products and modify their composition* - some ingredients are decreased, others are increased. Such food products can positively influence a number of physiological functions, which will improve health and general wellbeing and decrease in the risk of certain diseases (Szakály–Berke 2004). In addition to nutrition, the purpose of functional food products is to improve immunity, prevent certain diseases (e.g. high blood pressure, diabetes), facilitate the recovery from certain diseases, maintain good physical and mental condition and modify the process of aging.

Evidence of the aforementioned interrelations is currently available. For example it is generally known that fruits and vegetables of different color contain the most antioxidants, which improve the *immunity system* and *life expectancy*, but they do not slow down the aging process. A strict diet (1,500 calories a day) consisting of mainly fruits and vegetables not only decreasing the body temperature and eliminates the feeling of hunger, but also slows down aging (László–Falus 2002). The aforementioned examples indicate the specific needs of the elderly and help us to arrive at the shaping of a conscious and future-oriented choice of food products. Orientation to the future (as a behavioral element) is an important factor (Hideg 2007), because it (as a mental attitude) can have extremely serious implications on the health awareness of the elderly, their health related decisions and thus also on their *progressive health condition.*

References

Cseres-Gergely, Zs. – Molnár, Gy. 2008: Háztartási fogyasztói magatartás és jólét Magyarországon a rendszerváltás után. *Közgazdasági Szemle,* 55/2, 107–135. p.

Csete, L. –Láng, I. 1999: Minőség és agrárstratégia. *Magyar Tudományos Akadémia,* Budapest, 15–38. p.

Gaál, B. 1998: A magyar háztartások élelmiszerfogyasztási szokásai. *Marketing&Menedzsment,* 50/3, 71–75. p.

Hideg, É. 2007: A forecasttól a foresightig. *Magyar Tudomány,* 168/9, 1167–1170. p.

Hofmeister-Tóth, Á. 2006: *Fogyasztói magatartás.* AULA Kiadó, Budapest.

Józan, P. 2008: *Válság és megújulás a második világháború utáni epidemiológiai fejlődésben Magyarországon.* MTA Társadalomkutató Központ, Budapest.

Keszthelyiné Rédei, M. 2004: Az időskorú háztartások fogyasztási színvonala és szerkezete. In: *Időskorúak Magyarországon*. KSH, Budapest, 144–174. p.

Lakner, Z. – Sarudi, Cs. 2004: Ways and deadlocks in the strategic development of the Hungarian food chain. *Gazdálkodás,* 49/8, 48–57. p.

László, V. – Falusi, A. 2002: Az öregedés sejttani és genetikai alapjai. *Magyar Tudomány,* 58/4, 406–411. p.

Nemes Nagy, J. 2009: *Terek, helyek, régiók. A regionális tudomány alapjai.* Akadémiai Kiadó, Budapest.

Nováky, E. 2006: Jövőkutatás és felelősség. *Magyar Tudomány,* 167/9, 1090–1098. p.

Pettigrew, A. M. 2005: Segment with segments: Younger and Older Seniors Expectations of Financial Planning Service. *Journal of Research for Consumer*, Issue 7, 24–34. p.

Pilgrim, F. J. 1957: The Component of Food Acceptance and Their Measurement. *American Journal of Clinical Nutriton,* 5/2, 171–175. p.

Rátz, T. – Smith, M. – Michalkó, G. 2008: New places in old spaces: mapping tourism and recreation in Budapest. *Tourism Geographies,* 10/4, 429–451. p.

Simai, M. 2007: *A világ a XXI. század forgatagában: útelágazások és útvesztők.* Akadémiai Kiadó, Budapest.

Szakály, Z. – Berke, Sz. 2004: A táplálkozás, a minőség és a marketing kapcsolata élelmiszereknél. In Berács, J. – Lehota, J. – Piskóta, I. – Rekettye, G.(szerk.): *Marketingelmélet a gyakorlatban.* KJK Kerszöv, Budapest, 319–335. p.

Szántó, Z. 1998: A racionális döntések elméletén nyugvó társadalomtudomány. In Csontos László (vál.): *A racionális döntések elmélete.* Osiris-Láthatatlan Kollégium, Budapest, 7–24. p.

Törőcsik, M. 2006: *Vásárlói magatartás.* Akadémia Kiadó, Budapest.

Using Web Technologies and Online Databases to Develop Research Methodology for Determining Level of Integration of Effective Sales Management and Marketing from a relational approach point of view

Florin Mihoc[1] *– Adela Laura Popa*[2] *- Cristian Hnatiuc*[3]

In today's complex environment, every organisation has to cope with a multi-facets and dynamic environment. In order to maintain effective relationships with its stakeholders and develop long term relationship, organisations are to better organise, store and value their key data about those. These relationships can be nurtured and capitalised maximising customer lifetime value through professional sales management that uses relational databases. Most of local SMEs companies have a vague notion of marketing and operate sales as a total distinct activity. Traditionally sales is arguing that marketing is only interested in sophisticated budgeting, new product development, market findings and setting the prices too high, while short term sales are neglected. On the other hand marketing points out that sales department tend to have a myopia, being focused only on short-term goals, neglecting market signals and even not seizing long term advantages. Too often, these realities lead to poor coordination, reduced control, higher costs, lengthens sales cycles and increase total costs of sales. Developing an integrative model of work between marketing and sales departments and teams improves overall company performance by sharing their performance metrics, and also results in reduced internal misuse of energies and focus, enabling former combatants to aim towards same goals. The aim of this paper, which represents a work in progress, is to develop a research model based on web technologies and online databases for determining correlations between relationship marketing approach and effective sales management withing SMEs located in western region of Romania.

Keywords: relationship marketing, sales management, web technologies, databases

[1] Florin Mihoc, PhD, associate lecturer, Emanuel University of Oradea, Griffiths Management School (Romania)
[2] Adela Laura Popa, PhD, assistant lecturer, University of Oradea, Faculty of Economic Sciences (Romania)
[3] Cristian Hnatiuc, PhD, associate lecturer, Emanuel University of Oradea, Griffiths Management School (Romania)

2. Introduction

The nature of business and the role of salespeople have changed a lot in the last years. If in the 1930-1960 the main objective of salespeople was making the sale, in the 1960-1990 the focus shifted into satisfying customer needs. Nowdays, the main goal is to build strong, long-term relationships. The role of the salespeople also changed, from a mear provider, persuader and problem solver to a value creator. In meeting the everchanging demands of customers, the company now focuses on creating new alternatives, matching buyer needs with it's own capabilities. Gummesson (2002) argues that companies understand the importance of relationships and relationship marketing, but they do not clearly grasp how to implement a long-term customer relationship strategy. Traditionally, the sales department argues that marketing is only interested in developing sophisticated marketing programs, elaborate marketing budgets, coordinate new product development, integrate the marketing findings and setting the prices too high.

The premise is that the implementation of a good relationship marketing will result in an increased revenue and a reduction of costs, resulting from the retention of existing customers. Additionally, the view is that relationships will lead to better communication and coordination between the business units or different departments involved in the relationships, resulting in lower costs to serve the customers. Since it is more efficient to keep existing customers than to invest in getting a new customer, the premise seems to be correct, but unfortunately this has not always proved to be the case. In some instances, the relationships between consumers and suppliers have deteriorated from direct implementation of relationships strategies and relationship marketing (Grayson et. al. 1999). This result has been obtain from a lack of knowledge and understanding in how to design, implement and conduct relationship marketing programs.

Gummesson (2002) states that marketing managers and, in the end, all of the company's employees need a paradigm shift. In his paradigm, everyone in the company is involved in marketing. He introduces the term of part-time marketers, where they are solicited to supplement full-time marketers. Thus, the company becomes focused on the strategy of integrating different business units and departments that work together to obtain the same business goals. However, common knowledge opinates that main responsibility and activities for marketing and sales are aligned to their respective departments, and as such the higher degree of coordination, control and in the last integration of the two the more effective the outcome.

Successful implementation of sales management and relationship marketing will nonetheless lead to better communication and information sharing, will improve the company's overall performance and will enable the business units to achieve and exceed the company's goals.

- **Relationship marketing**

Marketing of the 21st century is significantly shaped by its relationship orientation. Without altering core and aim, marketing is refocusing its efforts of increasing organizational performance through the development of long-term relationships with all its partners (suppliers, customers, other stakeholders), (Pop et al. 2009). Such a mutation, takes place on the basis of shift from transactional marketing to relational marketing (Pop–Petrescu 2008), the latter requiring a new approach to business relationships among all the partners mentioned above.

Relationship Marketing (RM) represents a relative new approach and a marketing contemporary focus setting a new way of thinking and action of marketing based on *the existence of a network of connections* between all parties involved in the exchange of products, services, ideas and rights, not limited to producer, distributor or consumer. The prestigious American Marketing Association (AMA) defines RM as "Marketing with the conscious aim to develop and manage long-term and/or trusting relationships with customers, distributors, suppliers, or other parties in the marketing environment" (AMA 2009). RM strategy is based on initiating, nurturing and valuing relationships between supplier and clients mainly, but not only as it is a wider concept taking into consideration all stakeholders. Also, RM aims to develop customer loyalty developing marketing mixes in order to achieve this goal. Core philosophy of RM states simply that a key customer has to receive continual attention (Futrell 2002). American marketing school defining RM as the process of customer loyalty development (Futrell 2000), with main objective of identifying long term partners among customers with which to develop long term and profitable relationships rather than short term transactions.

- **Sales Management**

According to recent literature on Sales Management (SM) there is a gap between theory and practice of sales, with the later being more advanced. (Jones et al. 2005), in his own words: "much of our knowledge rests on models and assumptions that were advanced in past decades and that may need revision in light of rapidly evolving demands of the marketplace".

Some of the most recent definitions and reviews of the SM concept describe it as a process and a discipline. One of the most comprehensive and relevant definitions and content reviews is one formulated by the prestigious American Marketing Association (AMA 2009) which states that SM represents "the planning, direction, and control of the personal selling activities of a business unit, including recruiting, selecting, training, equipping, assigning, routing, supervising, paying, and motivating as these tasks apply to the sales force." SM consists of three major interdependent processes: formulation of the strategic sales objectives and program; implemen-

tation of the program; and coordination, control and evaluation of personal selling and overall sales performance.

Key aspects worth underlining when analyzing and reviewing as a whole the concept of SM and its current macro environment with all dramatic changes and challenges occurring at a very fast rate further pointing some insights and directions regarding research, are as follows, (Geiger et al. 2009):

- Rapidly changing and highly complex business environments are affecting the selling function; sales organizations have to adapt at an accelerated pace to challenges, some of which research may currently be lagging behind sales practice.
- Nowadays sales is being viewed as a strategic value-creating function in many organizations and often attracts high-level managerial involvement.
- More than ever before, sales forces are held accountable for their actions in both ethical and monetary terms; profitability is an important factor for contemporary sales management, yet very little research attention has in the past been devoted to budgetary and other financial sales issues.
- Sales practices and policies are in the present stage more often than not integrative in nature, bridging both across internal departments and between a number of customer-facing channels such as the internet or telemarketing. Sales research however has traditionally isolated the sales function at an individual level.

- Informatic applications. Web techonolgies

Data technology applications (i.e. CRM software) provide with the integration possibility of all the data available regarding every customer. This information might be available at different levels of organization and are gathered through different communication channels, with every customer interaction experience in order to achieve a holistic perception of the customer which is extremely valuable for the organization (Bălan 2007).

Any application that aims to achieve effective customer relationship management (CRM) is relying on advanced databases containing of such inputs which further permit a three levels approach for every stage of CRM, which are: strategic, operational and analytical (Bălan 2007). One of the most important informations the CRM system can provide is to determine the most profitable customers and what are the common characteristics these share. But, first of all in order to offer appropriate and essential information about the customers, a CRM is based on an effective marketing database. (Popa–Sasu 2008)

In order to reach marketing objectives, organisations are to know their customers. In order to know their customers, they are to gather information which have to be organized in a database fashion. This whole process allows organizations to practice database marketing (Kotler et al. 2006). Nonetheless databases are not a scope themselves, they are rather instruments that facilitate reaching primary marketing objectives. This aim is operationalised through a DataBase System – DBS,

represents an assembling of interrelated elements that contribute to realization and usage of a database application.

Marketing decisions, both related to market and to marketing mix are founded on marketing research and data analysis, whether data are obtained from external or internal sources. Data storage in databases constructs that are developed according to organization demands represents a solid foundation for achieving decision related marketing objectives such as: awareness and knowledge of customer needs, identification of potential customers (prospects), identification of ideal customer vis a vis different criteria (customer value, sales volume, customer profitability, etc.), customer loyalty and customer relationship analysis through surveys that identify loyal customer profiles, characteristics of these ones, relationship evolution over a time period.

As two well know professors underlined in Marketing Management (Kotler–Keller 2008), through complex data mining marketing researchers are able to extract out of total information amount, useful information related to: individuals, patterns, and different customer segments. Data mining requires advanced statistical and mathematical analysis techniques such as cluster analysis, predictive data formation and neuronal analysis.

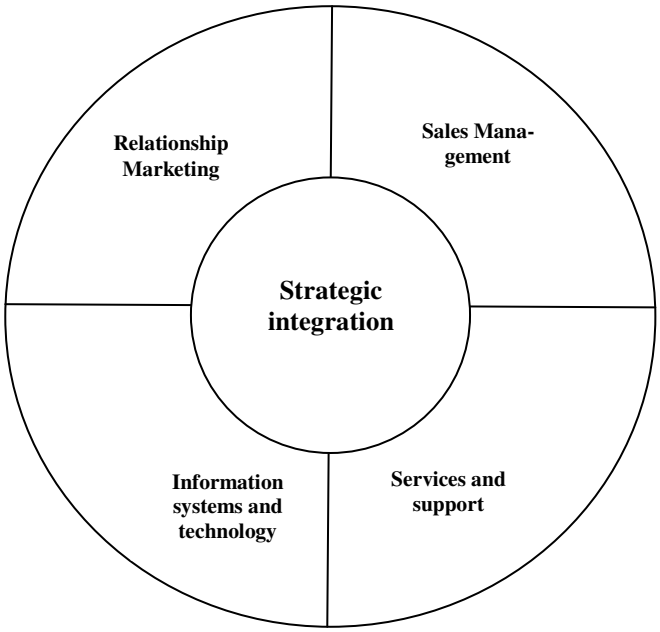
Also, the research in marketing can benefit from the availability of Web technologies. According to Gay and colleagues (Gay et al. 2007), marketing research generally involves the compromise of three factors: the cost, the speed and the accuracy. Using Web technologies when appropriate, may have a positive impact on all these three factors mentioned before. The required time for research is significantly shorter, the research is more price convenient, there are no geographical borders, and the possibility to store data directly in online databases – are just a few of the advantages provided by web technologies. Even if we reffer to the exploratory research or to the descriptive research, researcher must consider web technologies as an alternative to the classical reasearch tools. Findings of McDonald and Adam show that online data collection offers some advantages over postal collection from an efficiency point of view. (McDonald et al. 2003).

3. An integrative model proposed by literature

Relationship marketing and sales management have been studied for a long time as distinct functions of a company. Information systems and the use of technology in better understanding the customers need and wants, anticipating future trends and coming forth to meeting those needs, has also been studied. Services and support are also used as a successful tool of response to customers' complaints and increase the trust between parties involved in the relationship.

The premise of this integrative conceptual model is that by strategically integrating all these functions the overall effectiveness and performance of the company will synergistically evolve, as exhibited in Figure 1.

Figure 1. Integrative model



*Source***:** after Ngai, E.W.T., (2005), Customer relationship management research, Marketing Intelligence and Planning, 23, 582-605

2.1. The Evolution of Marketing and Sales Integration

In the last century, the marketing concept and sales management have changed due to the changes in the marketplace. There are four main orientations that can be identified in this evolution: production orientation, sales orientation, marketing era and partnering era.

2.2. Production orientation

Prior to 1930 and after the World War II the demand for products exceeded the supply. The business world equated marketing with selling. Under this business perspective, the key into achieving profitability was greater sales volume and the marketing main responsibility was to sell what the company produced. Competition was limited, customers had no alternative or competing products to chose from. Compa-

nies had little concern for buyers' needs and developing new products in order to satisfy those needs. The role of production oriented marketing and sales management was dominated by taking orders and promoting existing products.

2.3 Sales orientation

After the initial production orientation, the marketing concept evolved and the sellers market shifted to a buyers market. There were not enough customers to buy all the products that were manufactured and the competition among companies increased significantly. The role of marketing shifted into creating and increasing demand to existing products and engaging into aggressive selling techniques. The customer was presented to the product and convinced to buy it even if he didn't want or need the product.

2.4. Marketing Era

As the economy matured into a consumer society, the post war conditions of scarcity were replaced by increased competition, an abundance of manufacturers and brands, a more increasing customer influence over companies. The marketing concept was introduced as a response to these factors and as an influencer. The marketing concept is a business philosophy emphasizing that the key to business success is to satisfy customers needs (Webster 1988). According to this philosophy, all of the company's employees, not just marketers, need to be involved in marketing activities and toward satisfying customer needs. Research and development departments designed products customers want, not just products that were challenging to design. Production started to build products that were needed on the market, not just easy and cost effective to produce. Marketing departments engaged in activities of identifying customers' needs and solving customers' problems. The aggressive sales orientations as volume, price and promotional orientations were seen as less profitable than focusing on the needs of a particular set of customers. The view was that the company must service a selection of customers, specifically those that were the most profitable.

Selling became more complicated, due to increased market and customer complexity. Salespeople developed a greater product knowledge, better communication tools and improved their usage of new technology advantages in reaching customers. For example, salespeople made use of new computing technologies: databases were employed and populated with the use of telemarketing, increased complexity of reporting offered new opportunities and niches and by the turn of century, the internet started to offer new tools in marketing and sales management.

2.5. Partnering Era

By early 1990's both companies and customers recognized the strategic advantages that can be obtained by working together. Rather than treating each party as a one time customer / seller, long term relationships offered an increased synergy effect. Manufacturers developed close, long term relationships with a selection of suppliers, and by freely exchanging information the suppliers could develop tailored and customized products to the manufacturer.

Marketing became oriented towards value creation. Value represents the difference between the benefits that can be obtained from a specific product or service, and the respective cost of that product. A benefit represents the subjective perception of the customer and how that product improves his or her current situation. If the perceived benefit is bigger than that of competition, than the total value of the customer will increase. Through a long term relationship, customers can offer to the company important information regarding perceived value. Thus, marketing department's responsibility is to closely work with customers to develop solutions that enhance the profits of both. In this capacity, marketing has two roles. First, it must develop a deep understanding of the customers needs and convince him that the company has the capabilities to satisfy those needs. Secondly, it must be sufficiently influential in order to strategically integrate all of the company's functions in order to create a customized offer to the customer's needs.

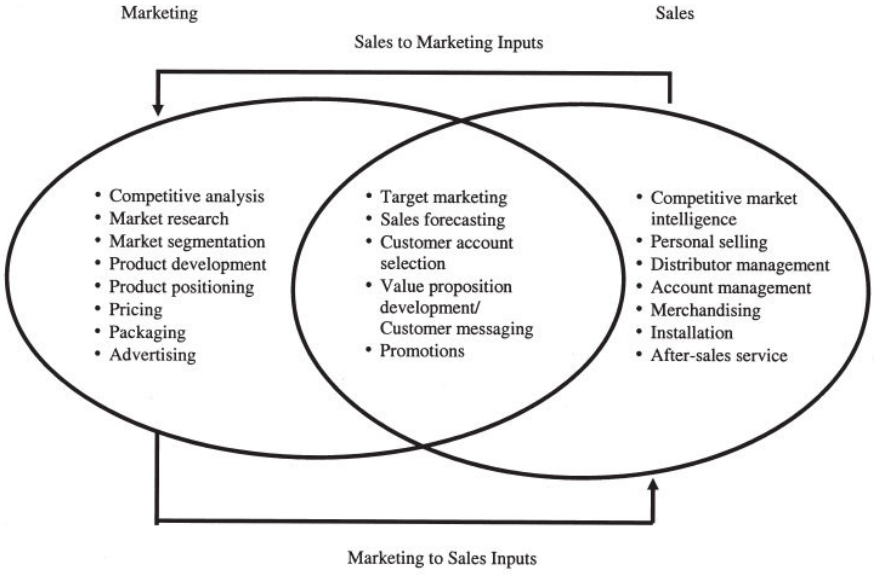
2.6. Integration of the RM and SM

In order to define integration of RM and SM we can start from a very early definition developed by Lawrence and Lorsch (1969) who noted that integration is "the quality of the state of collaboration that exists among departments that are required to achieve unity of effort by the demands of the environment" (Lawrence–Lorsch 1969). Whilst a more recent work set the broad perspective, yet still vague general notion, that of Song, Xie, and Dyer (2000) who opinate that integration is involvement, information, quality, and harmony. We agree a more recent definition worked by a group of professors who went into deeper insight and more specific terms. Rouziès and colleagues (2005) which view that sales–marketing integration as the extent to which activities carried out by the two functions are supportive of each other. They also underline that in order for the two functions to be supportive for the other there are two prerequisites: consistency and congruency need with each other and coordination need for the timing of the activities between the two departments (i.e the timing of a salesperson's meetings to a customer to introduce a new product or service must be aligned with marketing's launch of the advertising campaign for that item).

It is worth underlying that sales–marketing integration is a dynamic process in which the two functional areas create more value for their firms by working together than they would create by working individually (Rouziès et al. 2005).

In order to illustrate this need for integration between the two key functions we will use The Venn Diagram, in Figure 1. The activities in left circle are primarily marketing function tasks, valuing input from sales, while the ones from the right are executed by sales using input data from marketing. At the intersection between the two are activities that can be performed only through coordinated effort between sales and marketing.

Figure 2. Example of Sales and Marketing Tasks Integration



Source: Zoltners (2004)

3. Literature review considerations. Proposed model

Wotruba (1991) suggests the nature of personal selling, like marketing, has evolved through the four eras: production, sales, marketing, and partnering. Main aim of the sales force in the partnering era, which is the interest of this paper is to develop and

maintain relationship exchanges under overall objective of the salespeople to develop long term relationship with the customers

We propose the following paradigm shift as the position of this paper for the partnering era and role of the sales organization developed by Weitz and Bradford (1999) who argue that there are three major differences between the partnering role and the other roles: (1) the focus of interpersonal communications - managing conflict versus influencing purchase decisions; (2) the salesperson's objective - building and maintaining the relationship with the customer versus maximizing short-term sales; and (3) the unit of analysis-- the sales team versus the individual salesperson.

Although effective selling and sales management are often critical to marketing success, and vice versa it is relatively more facile to evaluate the first, as results of sales are readily measurable, while effectiveness of marketing activities (marketing strategies, advertising campaigns and promotional programs) are often more ambiguous.

In light of this relationship-oriented approach (Weitz–Bradford 1999), many firms with salespeople in partnering roles are taking into the account additionally to the traditional output measures of sales effectiveness such as sales, sales to quota, and gross margin, measures of customer satisfaction and even using these satisfaction measures as part of the incentive compensation for their sales force. (Lambe et al. 1997).

Effectiveness appraisal in the partnering role is based on knowledge of what the seller can do and what the buyer will want to do in the future that is why furthermore collecting, storing and valuing key data about customers and developing customer base knowledge in a very dynamic context of current market through web technologies and online databases is vital for overall company effectiveness.

In order to attain integration some companies use so called integrators, assigning people tasks roles. Typically such integrators have the responsibility for improving the interactions between the two functional areas, however they do not have formal authority to involve in decision making related to specifics. If we are referring to the whole sales and marketing integration effectiveness, literature reveals some of the factors that enhance the effectiveness of integrators are (Rouizés et al. 2005):

- maintaining a balanced orientation that recognizes the perspectives and goals of the sales and marketing functions
- providing integrators with the information, responsibility, and conflict management skills necessary for resolving conflicts that develop between sales and marketing
- using unique knowledge and skills as a basis for influence rather than formal authority

4. Methodology

However, not much research has been conducted on such topics as the integration of Sales Management (SM) and Relationship Marketing (RM) with the use of electronic applications such as web technologies or online database, and how to make this construct functionable and superior to merely distinct and individual aim and action of the two functions.

These observations within literature review section lead to the key hyphotesis that the greater level of sales-marketing integration the higher the business performance of the SMEs from western area of Romania. This primary hyphotesis is at the very core of this present paper.

First step of the model is to demonstrate there is a correlation between level of integration between relationship marketing and sales management within SMEs in the western region of Romania. Afterwards, we will measure the existing relationship level and group the analyzed companies in four clusters, based on the definitions of each of these categories as presented by Philip Kotler, Neil Rackham, and Suj Krishnaswamy (2006), as follows:

- **Undefined -** The first category refers at companies where the relationship between marketing and sales is undefined. In such a situation, both Marketing and Sales coexist mostly as independent entities and are preoccupied with their own activities, plans;
- **Defined -** Second type of relationship illustrate companies where SM and RM are characterized by a defined relationship. In this case, units are merely tolerating each other; there are processes and rules created in order to prevent conflicts. Also, under defined relationship they start to build a common language and, for some major events as trade shows and customer conferences, they work together;
- **Aligned -** There is a third type, where the relationship between Marketing and Sales is best described as aligned and includes companies where there is a specific border between the two groups, however this is flexible. They understand each others role;
- **Integrated -** The fourth category will consist of companies where the relationship between the two is integrated. Integration characterizes companies where Marketing and Sales work together; the boundaries between them are vague, they share structures and processes, develop and implement share metrics.

In order to explore the existing situation and to demonstrate the hypothesis that there is a strong relationship between effective Sales Management and Relationship Marketing, we adopted a quantitative research designed in part confirmatory – in order

to demonstrate the hypothesis –, and in part exploratory – so as to analyze the situation for the SMEs in western region of Romania.

4.1. Research instrument

In order to collect the necessary data, we will use instrument for research the questionnaire adapted after the one proposed by Philip Kotler, Neil Rackham, and Suj Krishnaswamy (2006). Questionnaire is made of twenty statements, each of them being rated on a scale from 1 to 5, where 1 means strongly disagree, 2 – disagree, 3 – neither, 4 – agree and 5 - strongly agree. The statements in the questionnaire are as follows:

- Our sales figures are commonly in a very close range to the sales forecast.
- If the situation and overall results are negative neither function criticizes the other.
- Marketing people with specific roles in relationship development often conduct meetings with key customers and prospects during the sales process together with sales reps.
- Marketing function (relational) always solicits participation from Sales in drafting the marketing plan and overall relational focused strategies and tactics.
- Sales force believes the collateral supplied by Marketing is a valuable tool to help them improve sales results.
- The sales force willingly cooperates in supplying feedback requested by Marketing, and there is a continual information sharing among the two functions related successful and unsuccessful customer experiences. (i.e re-acquisition rate, cross-buying, revenue generated from refferals).
- There is a great deal of common language within the company between Sales and Marketing and both compartments understand how can contribute to create customer value.
- The heads of Sales and Marketing regularly confer about upstream (strategic) issues such as idea generation, market sensing, and product development strategy.
- Sales and Marketing work closely together to define segment buying behavior and are integrated in serving the needs and achieving customer loyalty within our target markets.
- When Sales and Marketing conduct meetings, they do not need to spend much time on dispute resolution and crisis.
- The heads of Sales and Marketing collaborate on business planning for products and services that will not be launched for two or more years.
- We discuss and use common metrics for determining the success of Sales and Marketing both quantitative and qualitative. (i.e customer retention rate, customer satisfaction-dissatisfaction, customer lifetime value).

- Marketing actively participates in defining and executing the sales strategy for individual key accounts, acting as relational managers.
- Sales and Marketing manage their activities using jointly developed business guidelines, processes, or pipelines that span the business chain – from initial market research to customer service and long-term partnerships.
- Marketing makes a significant contribution to analyzing data from the sales funnel and using those data to improve the predictability and effectiveness of the funnel.
- Sales and Marketing share a strong "We rise or fall together" culture.
- Sales and Marketing report to a single Chief Sales and Relationship Marketing, or equivalent C-level executive.
- There's significant interchange of people between Sales and Marketing.
- Sales and Marketing jointly develop and deploy training programs, team building, different events, and learning opportunities for their respective staffs.
- Sales and Marketing actively participate in the preparation and presentation of each other's plans to top executives.

Valuing the same scoring methodology as originally developed by Philip Kotler, Neil Rackham, and Suj Krishnaswamy in Harvard Business Review (2006) we will include in the Undefined relationship category the companies scoring between 20 and 39, in the Defined relationship category the companies scoring between 40-59, in the Aligned relationship category the companies scoring between 60-79 and within the Integrated relationship category the companies scoring between 80-100.

The population elements will consists of the SMEs companies in western region of Romania, more specific those registered in the area of seven counties (Maramures, Satu Mare, Zalau, Cluj, Bihor, Arad and Timis). Moreover, these organizations will be considered suitable only if they have a distinct marketing department included in their organization structure with an orientation and practice of specific relationship marketing policies and practices. In order to clarify the SME concept, we will further render to the definition for SMEs as given by the European Commission in 2006, which is in vigor in Romania as member state as well, it states the following: "an enterprise is considered to be an SME if it employs no more than 250 persons and it has an annual turnover not exceeding EUR 50 million and/or an annual balance sheet total not exceeding EUR 43 million" (European Commission, 2006). In order to be able to efficient collect and analyze data, the questionnaire will be published on a web-site and the link will be sent via email to the management of the selected companies, who are in the position to evaluate and provide accurate info and scale the statements. After feedback will be transmitted by the companies the data will be automatically collected in an online database, which is specially designed to store them in order to provide the most appropriate way to be further proc-

essed. Prior to processing the data will be imported in specialized statistical software. Using statistical and mathematical analysis techniques such as correlation indicators, analysis of cross-classifications, cluster analysis or basic statistics the necessary data and indicators will be generated to eventually demonstrate the hypothesis that the relationship marketing vision is correlated with effective sales management and will allow us to unveil the image of the situation regarding the relationship between RM and SM for Western Romanian SME companies.

5. Conclusions - Limitations, managerial implications & research avenues

Main aim of this work was to develop a methodology that uses web technologies and online databases in order to determine integration of RM and SM in the range of SMEs companies in western region of Romania.

This whole enterprise is constituted as a tree phased process: diagnosis, development and design. In the first phase of diagnose we sought to determine the type of relationship between RM and SM, beginning with the premise that there is a direct and integrative relationship that results in superior performance for the sales and marketing. This methodology using online web tools and databases with some advanced statistical algorithms will be able to cluster the organizations in search according to the four types of functionalities existing between their sales and marketing, as described in the respective section: undefined, defined, aligned and integrated.

Secondly, development phase will seek to identify ways and strategies, methodologies that reveal integrators and to set the workable interdependencies between the two functions. Such shared metrics for instance are establishing a sales goal to which both teams commit. Important aspects define key sales metrics—such as number of new customers and closings—for salespeople. Some of the most common measures are percent of sales quota achieved, number of new customers, number of sales closings, average gross profit per customer, and sales expense to total sales. Also here, some of the common impediments to integrations, costs and drawbacks for integration will be considered.

Design's purpose is to value those integrative, devices, mechanisms and tools identified within previous two stages and to validate them through scientific methodology before empowering managerial staff with them.

There are definitely limitations to present survey as it represents only the first stage in what is intended as a multi-phased research, with this one being just the first step. There are for sure facets of the integration of the two functions which are not yet to be taken into account. For instance, present paper does not take into consideration costs associated with the integration and perils of excessive integration (i.e.

costs of using additional people-integrators, costs and efforts resulting as a consequence of excess of communication).

As for implications for managerial practice of the paper, executives from both sales and marketing will gain a wealth of understanding and insight of how sales and marketing can be aligned to each other, tuned to jointly aim towards shared goals, evaluate themselves by shared metrics and conduct planning and executing activities in common. Managers from surveyed SMEs and not only will have at disposal a relevant set of indicators, variables. As development phase of the research project aims developing an integrative model of work between marketing and sales departments and teams improves overall company performance.

Future research avenues for the next phases of the work will aim to measure intensity of the relationship of the two functions and more specifically to identify key complementary activities from the two areas and determine how they influence each other and to select a distinct set of integrators, factors that facilitate and promote integration process. Starting from some of the limitations of this paper in the future will look at the conditions affecting the trade-off between these costs and benefits of integrations, which are likely to moderate the integration–performance relationships. Will identify those moderators and elaborate on their content, influence and effective usage. Another important matter which will be further analyzed is that of ways and mechanisms for improving sales and marketing integrations and some of the literature proposed solutions on this topic.

We believe that synergies attained through RM and SM integration that enables SMEs to achieve superior performances while resulting in reduced internal misuse of energies and focus achieves a higher relational objective, that of helping former combatants (sales and marketing) to become allies towards same goals. One of the most valuable aspects of this present paper is represented by the fact that is a premiere in national literature, according to our knowledge no such survey on the topic has been conducted at a local, regional or national level for SMEs throughout Romania.

References

Aberdeen Group 2002: Bridging the Divide: Process, Technology, and the Marketing/Sales Interface. *Marketing View Point,* 15, 10, 1–12. p.

Bălan, C. 2007: Managementul relaţiilor cu clienţii: valenţe strategice, operaţionale şi analitice. *Revista de Marketing Online,* Vol. 1, Issue 3, 6-17 p.

Futrell, C.M. 2000: ABC's of Relationship Selling, McGrawHill, New York.

Futrell, C. M. 2002: Fundamentals of selling, customers for life. McGraw Hill, Columbus OH, New York.

Gay, R. – Charlesworth, A. – Esen, R. 2009: Online Marketing. A customer-led approach. Oxford University Press, Oxford.

Geiger, S. – Guenzi, P. 2009: The sales function in the twenty-first century: where are we and where do we go from here? *European Journal of Marketing*, Vol. 43 Issue 7/8, 873-889. p.

Grayson, K. – Ambler, T. 1999: The dark side of long term relationship in marketing services. *Journal of Marketing Research*, 36, 1, 132-141. p.

Gummesson, E. 2002: Relationship Marketing in the new economy. *Journal of Relationship Marketing*, 1, 1, 37-58. p.

Homburg, C. – Jensen, O. – Krohmer H. 2008: Configurations of Marketing and Sales: A Taxonomy. *Journal of Marketing* Vol. 72, 2, 133–154. p.

Jones, E. –Brown, S. P. – Zoltners, A. – Weitz B. A. 2005: The Changing Environment of Selling and Sales Management. *Journal of Personal Selling & Sales Management,* Vol. 25, Issue 2, 105–111. p.

Kotler, P. - Keller, K. 2008: Management Marketing. V-th edition, Editura Teora, Bucharest.

Kotler, P. - Rackham, N. –Krishnaswamy, S. 2006: Ending the War Between Sales and Marketing. *Harvard Business Review,* 84, 7/8, 68-78. p.

Lambe, C. J., - Spekman, R. 1997: National Account Management: Large Account Selling or Buyer-Seller Alliance, *Journal of Personal Selling and Sales Management*, 17, Fall, 61-74p.

Lawrence, P. R. - Lorsch, J. W. 1969: *Developing Organizations: Diagnosis and Action*. Addison-Wesley Publishing Company, Upper Saddle River.

McDonald, H. – Adam, S. 2003: A comparison of online and postal data collection methods in marketing research. *Marketing Intelligence & Planning*, 21, 2, 85-95. p.

Ngai, E.W.T. 2005: Customer relationship management research. *Marketing Intelligence and Planning*, 23, 6, 582-605. p.

Pop, N. Al. - Petrescu, E-C 2008: Marketing et gestion de la relation client. Editura Uranus, Bucharest.

Pop, N. P. - Mihoc, F. – Fotea, I. S. – Pop L. N. 2009: A Holistic Approach of Relationship Marketing in Launching Luxury New Products. Case Study: Research of the Demand for Housing in Residential Complexes in Oradea, *The Annuals of University of Oradea – Economic Science*, University of Oradea Publishing House.

Popa, A. L. - Sasu, D.V. 2008: Marketing Databases – The Shortest Way to Client's Heart. *The Annuals of University of Oradea – Economic Science*, University of Oradea Publishing House, Tom XVII 2008.

Rouziès, D. - Anderson, E. – Kohli, A. K. –Michaels, R. E. –Weitz, B. A. –. Zoltners, A. A. 2005: Sales And Marketing Integration: A Proposed Framework. *Journal of Personal Selling & Sales Management,* Vol. 25, Issue 2, 113–122. p.

Song, M. X. - Xie, J. –Dyer B. 2000: Antecedents and Consequences of Marketing Managers' Conflict- Handling Behaviors. *Journal of Marketing,* 64, 1, 50–66. p.

Webster, E. F. 1988: The Rediscovery of the Marketing Concept. *Business Horizons*, 31, 3, 29-39. p.

Weitz, B. A. - Bradford, K. D. 1999: Personal selling and sales management: a relationship marketing perspective. *Journal of the Academy of Marketing Science*, Vol. 27, Issue 2, 241-54. p.

Wotruba, T. 1991: The Evolution of Personal Selling? *Journal of Personal Selling and Sales Management*, 11, 3, 1-12. p.

Zoltners, A. 2004: Sales and Marketing Interface, paper presented at the Sales Force Summit, May, University of Houston, Texas.

European Commission - Enterprise and Industry, http://ec.europa.eu/enterprise/policies/sme/facts-figures-analysis/sme-definition/index_en.htm

Cross-cultural Fairness and Equity Research in Inter-organizational Relationships and Applied Statistical Methods for its Qualitative Measurement

Eszter Megyeri[1]

Research on Fairness and Equity linked closely to Justice has academic roots in multiple disciplines such as social psychology or philosophy. In the past few years an emphasized focus was placed on fairness among entities participating in business transactions with the objective of identifying the impact of human and business ethical attitude on business interactions. Measuring fairness by itself has its challenges due to its subjective nature. It is assumed that a business relationship, which is considered fair, is more balanced and stable. As a part of broader research on this topic, it is essential to understand what methods and techniques were used previously to evaluate fairness and equity when comparing findings in a cross-cultural setting. My objective in this paper is to summarize and evaluate the statistical methods used by the international research community to analyze and interpret fairness and equity in Inter-organizational Relationships.

Keywords: fairness, inter-organizational relationships, applied statistics

1. Introduction

The research on Fairness and Equity often referred to as Justice has its academic roots in many fields of social sciences. This paper intends to review some fundamental approaches to the measurement of Fairness and Equity. The objective is to compare statistical methodologies used in the current fairness research in the field of management specifically applied when a model is set in a cross-cultural environment. In the first part, I will give an overview on fairness research by taking a walk through the different disciplines' approach to fairness. This introduction will lead the way to the managerial interpretation and research of fairness, specifically focusing on the interorganizational application. The second part of the paper will show specific statistical methods with examples from empirical studies on fairness.

[1] Megyeri Eszter, PhD student at the University of Szeged, Doctoral School in Economics (Szeged)

2. Fairness and Equity Literature Overview

Fairness has an extensive academic literature. Navigating through the different interpretations of fairness is a challenging task. The question of fairness was embraced by many great thinkers of human history. Discussions of fairness and justice are illustrated in Aristotle‘s teaching. "Equals should be treated equally and unequals unequally" (Aristotle 350BC, Book VIII.). Aristotle pointed the attention to horizontal and vertical equity where equals and unequals are defined according to principles such as income or wealth levels. In this basic approach, the question is how fair an action or policy is to treat everyone in the same way unless the individuals are involved differently in such a manner that makes it relevant to the situation to be handled in a differentiated way. *Political philosophy* continuously emphasizes the importance of applying ethical aspects when defining fairness which is linked and related to defining justice as well. John Rawls' work on "A Theory of Justice" (Rawls 1971), influenced justice and fairness debates in the last decades, assumes that justice can only be developed under conditions which are fair to all involved parties.

Fairness is not only a central theme of philosophy. It's a basic element of *economics research* as well in terms of viewing the different ways of wealth or income distribution. Looking specifically at welfare economics, which focuses on maximizing social welfare, distribution must be Pareto efficient where no individuals or groups of individuals can be better off without making anyone else worse off. However, this does not help the choice of set of ethical values guiding the differentiation of individual preferences which are aggregated in the total social welfare. To reach a fair distribution, the division of goods must be *Pareto efficient* and *equitable* which means that every person's subjective valuation of their own share of basket of goods must be the same (Varian 2008). A *fair division* assumes equitable, envy-free share of goods in a Pareto efficient balance. „Because this type of policy is almost never possible, economists have been forced to fall back on the concept of potential Pareto improvements, for instance, in cost-benefit analysis. This is where winners gain more than losers lose and therefore, potentially, are able to make compensation so that no one loses. Compensation schemes are very difficult to design, however, because it is so hard to identify the winners and losers." (Wilber 1998, 5). As a result, ethical guidelines used are in line with the utilitarianism approach providing "the greatest good for the greatest number" (Wilber 1998, 5). Ethics can influence economics in the following channels. Firstly it shapes the way economics theories are developed and applied. This is followed by the economic actors and their ethical values which can impact and shape their own actions and behaviours. Thirdly, institutions and policies related to economics can fall under ethical evaluation as well. (Wilber 1998).

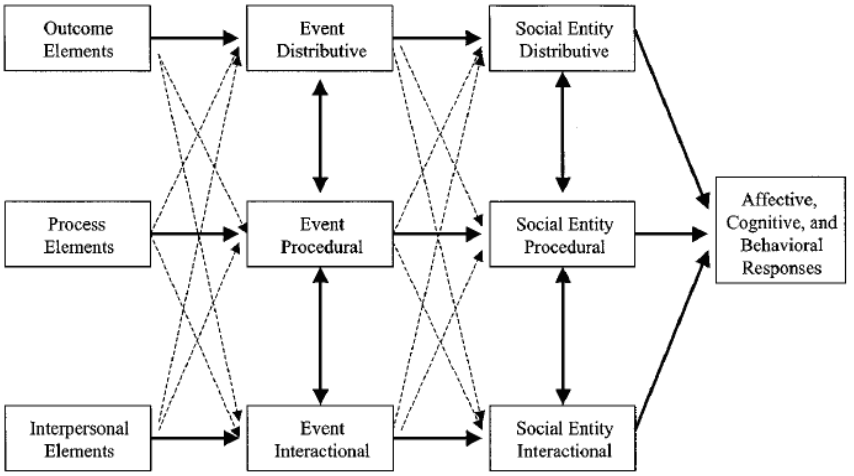
Social Psychology has a number of researches in the twentieth century theorizing on fairness under the Organizational Justice studies. In the 1960s Homan put down the foundation of *social exchange theory* in which he proposes that an indi-

viduals' objective is maximizing gains while minimizing cost based on which they evaluate relationships (Homans 1961). This work inspired Adam's well-known performance based *Equity Theory* in which he suggests that people compare the ratio of inputs and outputs to determine the equitable state driving their satisfaction. This concept is applied in two ways. An individual looks at his or her own rewards and contribution ratio and adjusts either his or her inputs or outputs if unsatisfied. Secondarily, the output-input ratio of an individual is compared to others output-input ratio as well within an organization. If ratios are unequal, adjustments are made to rebalance them (Adams 1965). Adam's theory triggered the examination of the so called *Distributive Justice* research in social psychology during the 60s and 70s. Distributive Justice aims to determine the mechanisms driving the way of final allocation goods in a society. In the 70s the debate among sociologists continued with examining the way how a decision is made on distribution by developing the *Procedural Justice* approach stating that the distribution of outcomes cannot be fair without a fair procedure of making such a decision. Thibaut and Walker emphasized this point based on simulated trials, in which the chance of self-representation in the process positively influenced the capability of accepting a final decision even if it was a negative outcome for the individual (Greenberg 1990). Procedural justice deals with the aspect of an individual's fairness perception of the outcome allocation process by focusing on the formal procedures used for making the decision. Among the procedural justice models we find the so called Group Value, later renamed as Relational model. This approach states that besides Distributive and Procedural justice, feedbacks and responses from the group, especially the ones coming from the authority, play an important role when developing fairness perception (Tyler-Lind 1992). The authority is assessed through three main characteristics which are neutrality in decision making, trustworthiness representing the unbiased attitude and honour for the employees which includes honouring the employees' rights as well. The empirical testing of the Group Value model in a cross-cultural comparison showed rather similarities than differences among different cultures (Cropanzano et al. 2001). A third dimension of the organizational justice research is the *Interactional Justice*, which "refers to the perceived fairness of the enactment or implementation of procedures and has two sub-facets. Interpersonal justice captures the sincerity and respectfulness of authority communication, while informational justice concerns the use of honest and adequate explanations for decisions" (Colquitt et al. 2006, 111).

Integrative Justice theory targets to embrace all the three (distributive, procedural and interactional) aspects. Their main integrative models are to be presented here. Folger's *Referent Cognitions Theory* „maintains that an unfair judgment will result from a situation where an individual believes a more favourable outcome would have resulted from an alternative procedure that should have been used. Thus, the referent in this model refers to the awareness of procedural alternatives that

would lead to a more favourable outcome" (Cropanzano et al. 2001, 168). The contribution of the theory was that it defined the conditions needed to hold others accountable for an unfair treatment. The criticism of the theory was that it did not address the process by which accountability judgments are made. (Cropanzano et al. 2001). Based on the limitation of the theory, Folger modified and adjusted his theory referring to it as fairness theory in which he suggests that distinct judgments must be made to determine if a given situation is fair. „These judgments contrast the negativity of the situation, the actions of the target, and the moral conduct employed with counter factual scenarios of what would, could, and should have taken place" (Cropanzano et al. 2001, 168). Soon after the development of the fairness theory, a new model, the *Fairness Heuristic Theory* was introduced which synthesized several previous models. Its basic approach was that individuals often face situations in which they must concede to authority, which gives an opportunity to be exploited. It is being referred to as the fundamental social dilemma (Lind 2001, Cropanzano et al. 2001). "As a result of the possibility of being exploited and having one's identity threatened, individuals are often uncertain about their relationships with the authority. This uncertainty leads an individual to ask questions such as whether the authority can be trusted, if the authority will treat him or her in a nonbiased manner, and if the authority will look at him or her as a legitimate member of the society, organization, or work group" (Cropanzano et al. 2001, 168). However, to evaluate these situations correctly, accurate and adequate information is often unavailable which makes the individual rely on heuristics to guide their subsequent behaviours, responses and action (Lind 2001). "*Uncertainty management* theory recognizes that many aspects of work and family life may contain uncertainty. According to the theory, fairness can remove trust-related uncertainty and mitigate the discomfort associated with other forms of uncertainty—even if they have nothing to do with authorities. The authors summarized this key tenet as follows (Colquitt et al. 2006, 113): "What appears to be happening is that people use fairness to manage their reactions to uncertainty, finding comfort in related or even unrelated fair experiences" (Lind-Van den Bos 2002, 216).

Table 1. An integrative model of organizational justice



Source: Cropanzano at al. 2001, 191.

From the early 2000s, an emphasized focus was placed on fairness among entities participating in *business transactions.* There were a number of studies in game theory and economics have investigated in how far different normative theories are applied by people in their subjective judgments (Fortin-Fellenz 2008). These studies were reviewed by Konow who clustered the different distributive fairness norms into three main theoretical streams (Konow 2003):

1. Equality and need,
2. Utilitarianism and welfare economics,
3. Equity and desert

"Fairness views are best explained by an integrated approach that acknowledges the influence of the three principles of justice, whereby the weight on each is determined by the context" (Konow 2003, 1190).

When looking at the *management application of fairness theories*, we find that many areas of management are impacted by the findings of Social Psychology. Areas of Human Resource Management have specifically built a number of practical applications in hiring, performance management, compensation and benefit management in order to build a genuine and adherent business culture (Brockner 2006). In management, there are further splits in the investigation. The first is the intra-organization application of fairness, which will not be detailed here. The second area focuses on examining and measuring fairness among business entities in the inter-organizational context. One of the first empirical studies on this topic was done to investigate reactions to perceived inequity in inter-organizational relationships in two countries, namely the US and the Netherlands. These studies were focusing on

automobile deals' perceived inequity and attitudes towards their automobile (OR: car producer) supplier partner (Kumar et al. 2003). Sheer's, Kumar's and Steenkamp's work also lead to set increasing statistical challenges to measure fairness in normative manner while researching it in an intercultural and cross-national comparison.

3. Measuring fairness in a cross-cultural context

One of the first findings from Sheer-Kumar-Steenkamp's article was that the cross cultural comparison requires a special attention before moving on to analysing any of the data collected from a cross national environment (Kumar et al. 2003). How can we make the data and the analysis free from cross-cultural effects? In the next paragraphs we are going to review scaling and invariance issues specifically applicable for cross-national research.

3.1. Scales

The scaling method requires additional attention. To measure social attitudes the Semantic Differential Scale (SDS) is the recommended method (Malhotra 2008). The SDS is a scaling tool first developed in the 1950s (Osgood et al. 1957). The scale is usually a seven-point, bipolar rating scale using opposites of adjectives. Some studies have used five or six-point scales as well. "The SDS has been used extensively in language attitude studies as a means of measuring subjects' attitudes towards various languages, dialects, accents, as well as the speakers of different varieties." (Al-Hindawe 1996, 1). A feature such as fairness, for example, would be represented by the semantic differential scale in the following form, which can be presented with or without the numeric scaling.

unfair 1 2 3 4 5 6 7 fair
unfair ○ ○ ○ ○ ○ ○ ○ fair

The advantage of the SDS is that it is relatively easy to implement it from the communication and the procedural point of view. One of the widely-used SDS is the Likert scale where the respondents evaluate a statement or situation based on pre-set subjective or objective criteria, which are called the Likert items. "A Likert item is composed of a stem (word, phrase, or sentence) followed by an endorsement scale running from strongly disagree to strongly agree" (Colton-Covert 2007, 159). The Likert scale is often used in cross-national researches.

3.2. Measuring Invariance - Multigroup Hierarchical Confirmatory Factor Analysis

When collecting data in a cross-cultural environment, the data might not be directly comparable due to a number of potential variations and interpretation of the questions or the answers. Marketing research defines the following invariance requirements to address the variance issues in a cross-national research (Malhotra 2008):

- Definition invariance
- Concept invariance
- Functionality invariance
- Category invariance
- Invariance in operationalisation
- Metric invariance
- Language invariance

There are a number of ways addressing the requirement of invariance. Steenamp is among the leading statisticians developing the consumer research framework for cross-national research and comparison (Steenkamp-Baumgartner 1998). In the early 90s, the Multigroup Hierarchical Confirmatory Factor Analysis (MGHCFA) was advised to be used as the suitable method for testing model invariance (Steenkamp-Baumgartner 1998).

Table 2. An example of the MGHCFA

MODEL COMPARISONS FOR ETHNOCENTRISM DATA

	χ^2 value	*df*	RMSEA	CAIC	CFI	TLI
Calibration data:						
Equality of Σ^g and μ^g	1,853.11	130	.0992	2,396.14	n.a.	n.a.
Equality of Σ^g	1,137.60	110	.0774	1,847.72	.922	.905
Equality of μ^g	643.80	20	.1410	2,105.81	n.a.	n.a.
Configural invariance	936.09	105	.0712	1,687.99	.937	.919
Full metric invariance	1,078.45	123	.0706	1,679.96	.928	.921
Final partial metric invariance	956.52	119	.0672	1,591.45	.937	.928
Initial partial scalar invariance	1,231.85	133	.0728	1,749.82	.918	.917
Final partial scalar invariance	1,024.36	130	.0664	1,567.40	.932	.930
Full factor variance invariance	1,050.42	132	.0668	1,576.75	.931	.929
Final partial factor variance invariance	1,025.18	131	.0661	1,559.86	.932	.930
Initial partial error variance invariance	1,551.86	147	.0783	1,942.29	.894	.903
Final partial error variance invariance	1,098.95	142	.0657	1,541.73	.928	.931
Validation data:						
Configural invariance	952.42	105	.0721	1,703.79	.934	.915
Full metric invariance	1,058.15	123	.0700	1,659.25	.927	.920
Final partial metric invariance	983.49	119	.0684	1,617.98	.932	.923
Initial partial scalar invariance	1,257.36	133	.0738	1,774.97	.912	.913
Final partial scalar invariance	1,035.03	130	.0670	1,577.69	.929	.926
Full factor variance invariance	1,091.51	132	.0685	1,624.02	.925	.923
Final partial factor variance invariance	1,046.15	131	.0671	1,580.46	.928	.926
Initial partial error variance invariance	1,498.81	147	.0770	1,899.55	.894	.903
Final partial error variance invariance	1,152.70	142	.0677	1,595.17	.921	.925

Source: Steenkamp, Baumgartner, 1998, p159

Comment: In this model Observed response variable ( ) for a respondent (i), in a country (g) for a specific item (k) composed by latent construct ( ) for i respondent in country g plus the slope of regression of on the value ( ) of latent construct of representative i in country g, used as factor loading plus the expected value of the observed response ( ) when latent construct for i respondent in country g equals 0 plus the error of the measurement ( ). By further following Steenkamp's and Baumgartner's method, the following invariance can be calculated for configural invariance, metric

invariance, scalar invariance, factor covariance invariance, factor variance invariance and error variance invariance.

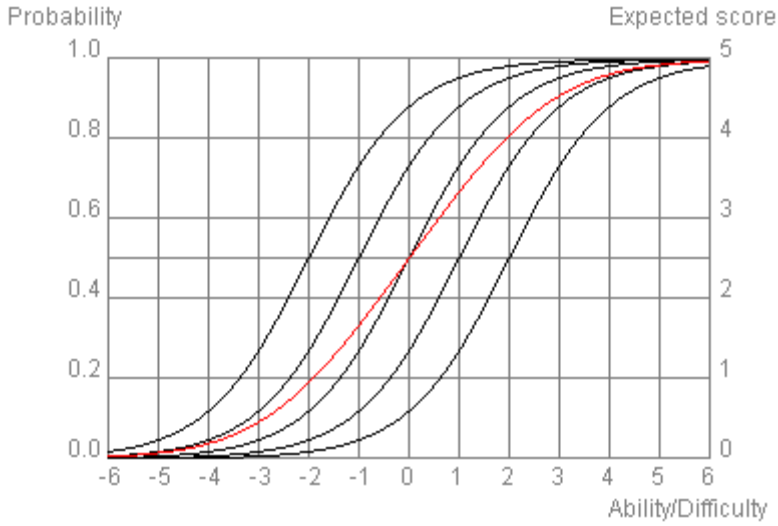
For the model fit, the following measures are recommended to use by Steenkamp and Baumgartner: RMSEA=root mean square error approximation, CAIC = consistent Akaike information criterion, CFI = comparative fit index, TLI = Tucker-Lewis fit index.

A number of critics were made on the Confirmatory Factor Analysis (CFA) applied. It requires at least partial invariance for a minimum of 2 items which is considered especially an issue when measurement construct consists of only a few items (Steenkamp-Baumgartner 1998). When measurement invariance is not satisfied, sub-groups of countries can be compared. However, that usually does not satisfy the researchers' objectives. In response to criticism and based on findings on the empirical application of the CFA, a new approach was proposed by the previously CFA supporter Jan-Benedict Steenkamp and his co-writer De Jong Martijn in 2007.

3.3. Relaxing Measurement Invariance – Item Response Theory

Steenkamp and Martijn recommended to relax measurement invariance in cross-national consumer research and proposed a hierarchical IRT Model instead. (Martijn-Steenkamp 2007). IRT stands for *Item Response Theory*, which is a relatively new model. IRT is a test of measurement equivalence across experimental groups, where groups are expected to show mean differences based on their latent trait or personal parameter on the attribute being measured (Partchev 2004). IRT originates the probability of each response as a function of the latent trait and some item parameters. The Item response function (IRF) describes the probability of a response to the item as a function of a person or ability parameter. It has two main families dependent on the number of traits examined. One is the *unidimensional*, which involves the analysis assuming a single trait. The *multidimensional* one includes multiple traits or multiple personal parameters. IRT model categories are based on the number of scored responses, which can be *dichotomous*, where scored only as correct/incorrect responses or *polytomous* outcomes, where each response has a different score value, for example a Likert-scale (Partchev 2004). The below chart shows the item response functions of the one-parameter logistic model for five items indicated by the black solid lines. In this example the probability of a test score is targeted to be estimated by the given ability of the responder examining five questions as items. The five item response functions estimates the probability that a person of a given ability will give a correct response to the corresponding item. The test response function in red shows very much the same, but for the test as a whole for any ability predicting predicts the expected test score.

Table 3. Item response and test functions of the one parameter logistic model



Source: Partchev, I. 2004: A visual guide to Item Response Theory.

In the IRT model, in a more complex model than above can provide sets of items for different ability or personal traits that can be linked to the observed nations or countries in a cross-cultural comparison. By that different response profiles can be developed to each responder group. The IRT model will not set the need for invariance but provides a base set or profile of the responders.

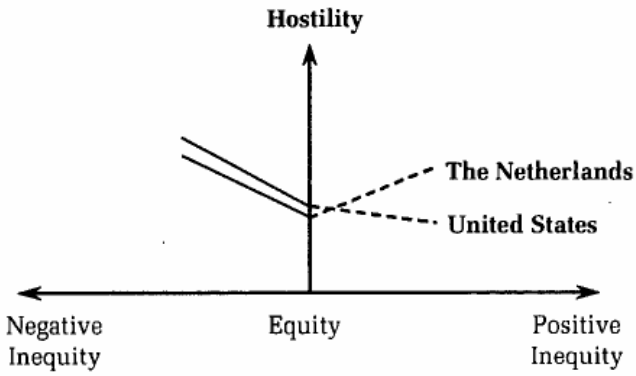
4. Fairness and Equity Spline Regression

The third part focused on the data calibration from the cross-cultural comparison point of view. In the fourth part, a special type of regression analysis will be reviewed, namely the Spline regression which is suitable to display bipolar attributes. At first, the Spline Regression will be reviewed from the methodology point of view, after which an example will be presented. "Spline regression is a type of regression in which different linear slopes are estimated for different ranges of the independent variables. The ranges have endpoints. They are called knots. Splines are lines or curves. They are usually required to be continuous and smooth. Uni-variate polynomial splines are piecewise polynomials. They show one variable of some degree d with function values and first d-1 derivatives that agree at the points where they join." (Hurley et al. 2006, 2). The joining points transitioning from one to the next, are called break points, interior- or simple knots. Knots give the curve freedom to bend and more closely follow the data. Splines with few knots are usually smoother

than splines with many knots. The increasing the number of knots usually increases the fit of the spline's function to the data (Hansen-Kooperberg 2002). There are many types of splines such as the Polynomial Spline, Periodic Spline, Spline on the sphere.

Sheer, Kumar, Steenkamp used spline regression to compare perceived inequity in US and Dutch inter-organizational relationships by measuring responses to positive and negative inequities between car deals and their suppliers in the two countries (Kumar et al. 2003). Positive and negative equity was defined as independent variables. Hostility, trust, relationship continuity and guilt were the dependent variables. The spline regression analysis provided a well interpretable outcome, of which a sample is provided below. The finding was that when experiencing negative inequity (referring back to Adam's equity theory, when the dealer's output/input ratio was smaller than the supplier's), auto dealers both in the US and the Netherlands responded by increasing hostility.

*Table 4.*Graphical Depiction of Effects of Negative and Positive Inequity-Hostility



Source: Reactions to perceived inequity in US and Dutch inter-organizational relationships, Sheer, Kumar, Steenkamp, Academy of Management Journal. 2003.

However, when experiencing positive inequity (when the dealer's output/input ratio was greater than the supplier's), the impact on the dealer's hostility toward the suppliers were different. US dealers showed decreasing hostility. On the contrary, Dutch dealers showed increasing hostility while their net gain increased. The referred article further analyzed the similarities and differences of responses impacting other attributes of the relationship using spline regression analysis.

5. Conclusion

Fairness, Equity, in some disciplines referred to as Justice Studies is a subjective matter. A number of different disciplines intended and continue to build simple and complex models to explain fairness among people or groups in a society, in a country within or among business entities. In the first part of the paper the base approach of fairness from different disciplines' point of view was reviewed. The second part focused on measuring fairness in an inter-cultural and inter-organizational setting. Within consumer behaviour research methodology, statistics plays a key role in putting quantitative measures behind the research settings. The most frequently used scaling methods were reviewed which are advised when comparing cross-cultural data. Two methods, CFA and IRT were reviewed as a tool to test the data to be free from cross-cultural effects. In the last part, spline regression was reviewed as a tool to analyze and graphically display bipolar data sets.

References

Adams, J. S. 1965: Inequity in social exchange. In Berkowitz, L. (ed), *Advances in Experimental Social Psychology*. Vol. 2, Academic Press, New York, 267-299.

Al-Hindawe, J. 1996: Considerations when constructing a semantic differential scale. *La Trobe Papers in Linguistics* Vol 9, Available at http://www.latrobe.edu.au/linguistics/LaTrobePapersinLinguistics/Vol%2009/Contents.htm [downloaded 19th December 2009]

Aristotle, 350 BC*: Nicomachean Ethics.* translated by W. D. Ross. Internet Classic Archives, Available at: http://classics.mit.edu/Aristotle/nicomachaen.8.viii.html [downloaded 19th December 2009]

Brockner, J. 2006: Why It's So Hard to Be Fair. *Harvard Business Review*. March, 122-129

Colquitt, J.A.-Scott, B. A.-Judge, T. A.-Shaw, J. C. 2006: Justice and personality: Using integrative theories to derive moderators of justice effects. *Organizational Behavior and Human Decision Processes*. Vol. 100, 110-127.

Colton, D.-Covert, R. W. 2007: *Designing and constructing instruments for social research and evaluation*, Jossey-Bass, San Francisco 159.

Cropanzano, R.-Byrne, Z. S.-Bobocel, D. R.-Rupp, D. E. 2001: Moral virtues, fairness heuristics, social entities, and other denizens of organizational justice. *Journal of Vocational Behavior*. Vol. 58, 164-209.

Fortin, M.-Fellenz, M. R. 2008: Hypocrisies of fairness: Towards a more reflexive ethical base in organizational justice research and practice. *Journal of Business Ethics*. Vol. 78(3) 415- 433

Greenberg J. 1990: Organizational Justice: Yesterday, Today, and Tomorrow, *Journal of Management*. Vol. 16, No. 2. 399-432.

Hansen, M.-Kooperburg, C. 2002: Spline Adaptation in Extended Linear Models. *Statistic Science* Vol. 17.1, 1-51.

Homans G. C. 1961: *Social behavior: Its elementary forms.* Harcourt Brace & World, New York Available at: http://books.google.hu/books?id=pqA9AAAAIAAJ&dq=Social+Behavior:+Its+Elementary+Forms&printsec=frontcover&source=bl&ots=a56gfyWNfP&sig=lCdBCC56wi7FFfmUJZgItu1lee0&hl=hu&ei=_iIhSuuNN9DdsgaqtLi3Bg&sa=X&oi=book_result&ct=result&resnum=1#PPP13,M1 [downloaded 19th December 2009]

Johnston, J. 1984: *Econometric methods.* (3rd ed.) McGraw-Hill. New York

Hurley, D.- Hussey , J.- McKeown, R.-Addy, C. 2006: An Evaluation of Splines in Linear Regression. *Paper 147-31* SAS Institute Available at: www2.sas.com/proceedings/sugi31/147-31.pdf [downloaded 19th December 2009]

Konow, J.: 2003: Which Is the Fairest One of All? A Positive Analysis of Justice Theories. *Journal of Economic Literature* Vol. 40, 1188–1239.

Kumar, N.-Scheer, L. K.-Steenkamp, J. B. E. M. 1998: Interdependence, punitive capability, and the reciprocation of punitive action in channel relationships. *Journal of Marketing Research.* Vol. 35: 225-235.

Lind, E. A. 2001: Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. In Greenberg, J-Cropanzano, R. (eds), *Advances in organizational justice.* Stanford University Press, Stanford 56–88.

Lind, E. A.-Van den Bos, K. 2002: When fairness works: Toward a general theory of uncertainty management. In Staw, B. M.-Kramer, R. M. (eds), *Research in organizational behaviour.* Vol. 24 181–223.

Malhotra, N. K. 2008: *Marketingkutatás.* Akadémiai Kiadó Zrt. Budapest

Martijn, D.-J. Steenkamp, J. B. E. M.-Fox, J. P. 2007: Relaxing Measurement Invariance in Cross-National Consumer Research Using a Hierarchical IRT Model. *Journal of Consumer Research*. Vol. 34 August, 260-278.

Osgood, C.E.-Suci, G.J.-Tannenbaum, P.H. 1957: *The measurement of meaning.* Urbana:University of Illinois Press, Champaign

Partchev, I. 2004: *A visual guide to Item Response Theory*. Available at: http://www.metheval.uni-jena.de/irt/VisualIRT.pdf [downloaded 19th December 2009]

Rawls, J. (1971). A Theory of Justice. Harvard University Press, Cambridge, MA. Available at: http://www.google.com/books?hl=hu&lr=&id=kvpby7HtAe0C&oi=fnd&pg=PR11&dq=Rawls+A+Theory+of+Justice&ots=tgesv6ze86&sig=MbW1JdVJ3T8t9z0058iL-HYgkKQ#v=onepage&q&f=false [dowloaded 19th December 2009]

Scheer, L.-Kumar, N.-Steenkamp J. B. E. M. 2003: Reactions to Perceived Inequity in US and Dutch Interorganizational Relationships. *Academy of Management Journal* Vol. 46.3, 303-316.

Steenkamp, J.B.E.M. Baumgartner, H. 1998: Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research* Vol. 25.1, 78-90.

Tyler, T. R.-Lind, E. A. 1992: A relational model of authority in groups. In Zanna, M. (ed), *Advances in experimental social psychology*. San Diego: Academic Press, San Diego, Vol. 25, 115–191.

Varian, H. R. 2008: *Mikroökonómia középfokon*. Akadémiai Kiadó Zrt., Budapest

Wilber, C. K. 1998: Economics and Ethics. In Davis J.B.-Hands D.W.-Mäki U. (eds.), *The Handbook of Economic Methodology*. Edward Elgar, Cheltenham. Available at: http://www.nd.edu/~cwilber/pub/recent/ethichbk.html [downloaded 19th December 2009]

Social networks as a support of marketing for nonprofit organizations

Drago Ružić[1] *–Antun Biloš*[2]

Social network web sites or social networks in short, became a surprisingly interesting and popular subject, especially since the mass acceptance of this specific form of online communication and interaction. Social networks attract the attention of both experts and scientific researchers largely because of the numerous possibilities of using the specific form of networking provided on almost every type of widely accepted social networks. Social networking is one of the key components of Web 2.0, which is characterized by openness, freedom and collective intelligence. However, the most important characteristic of Web 2.0 is the exceptional interactivity between users. All the specific aspects of nonprofit organizations and especially of the education sector are apparent in the online environment as well. Furthermore, enhanced online communication systems enable a high level of usage of the aforementioned aspects for accomplishing the organizations' goals. This paper deepens the research carried out during the spring of 2009 on the similar subject of general usage of social networks amongst its users. The paper starts with a brief overview of the historical development of social networks in order to show how these communities have changed and developed over time. Current characteristics of social network web sites are determined and described in order to give a full picture of their state and marketing possibilities as well as the implications they have on future development of Internet based interactions. In addition, there is a special comment about the value of social networks and the way of determining their value and specific types of networks that are present in the global Internet space today. This paper also gives a general picture of the state of social networks and their users in the Croatian Internet scene with estimates of further development in the near future. The practical example of Internet presence within the social network environment takes the central place in the paper. An overview of the postgraduate study profile, on the currently most popular social network Facebook, is given. In addition, the advertising campaign of the aforementioned study within the same social network, as well as its overall efficiency is accentuated and commented. At the very end, the current marketing value and the reach of social networks concerning the marketing for nonprofit organizations is summed up and directions in which the future research of this subject will go are given.

Keywords: social networks, Web 2.0, Internet, online advertising, education

[1] Drago Ružić, PhD student, senior lecturer, University of J.J.Strossmayer Osijek, Faculty of Economics in Osijek (Croatia)
[2] Antun Biloš, M.Econ., assistant, University of J.J.Strossmayer Osijek, Faculty of Economics in Osijek (Croatia)

1. Introduction

One of the undeniable innovations of Web 2.0 is the, today unavoidable, social network. All users who are up to date with the recent developments in the omnipresent Internet space are well acquainted with this and are probably participants themselves in some of the existing social networks. Over 300 different, active social networks already exist today. Each has its own specific characteristics, beginning with the ideas of connecting, the dominant target group, the possibility of interaction and visual identity.

Parallel with the most important, general (horizontal) social networks like Facebook, Myspace, LinkedIn etc., less specific (vertical) social networks of various marketing niches, which do not stand out with its size and the number of users but with the homogeneity of their user base and their interests are described. Further, a general overview of the state of social networks and their users on the Croatian Internet scene is given, with the prediction of future development in the near future. Since we are observing online social networks from a marketing perspective, their marketing potential, which mostly refers to integrated marketing communication and to all the effects that can be achieved by appropriately using social networking, is accentuated and commented. Here lies the basic problem of the marketing approach to social networks because it is not easy to determine the cause and effect relationships of the success of an individual online marketing (or advertising) campaign. Numerous skeptical opinions concerning the efficacy of Internet advertising speak in favor of that. However, a large number of examples which are examined in this work illustrate best all the restriction, but also the great possibilities of Internet integrated marketing communication.

2. Social networks

We can examine social networks even if we exclude them from their digital surrounding. A social network is a sociological structure which describes relationships between individuals. We all belong to one great social network, but we also belong to those social networks which are smaller in scope and which originate from our families, friends, work places, interests, hobbies etc. The term social network itself has roots in the early sociometric theories.[3]

There are numerous ways in which users may use a social network; starting from creating an own profile and adding personal facts, areas of interest and preferences, searching for old friends and creating new, exchanging music, photos and videos and a plethora of related activities (which best describes the type and the

[3] International Sociometry Training Network (2009), (accessed March 15, 2009), [available at http://www.sociometry.net/].

complexity of the applications). It is also possible to become a member of different interest groups connected with e.g. favorite TV shows, sports or political parties. More than 300 different social networks already exist today.[4]

In order to determine the popularity of social networks as precise as possible, we first must take interest in the number of users each of the social networks has. The number of overall registered users exists as an indicator for a longer period of time, but is rarely used today as relevant data because of its inaccuracy. There are more significant indicators which describe the popularity of an individual social network. Here we should name two key indicators: the number of active users (as a part of the total number of registered users) and, the indicator that analysts and analytic agencies use more often in their analysis: the unique visitor number[5].

Two social networks which definitely stand out from the others concerning popularity are Facebook and Myspace. According to data from the end of 2008[6], Facebook is the leading world online social network with 222 million unique visitors in December of the same year. From the company, they state with pride that during January 2009 they have exceeded the number of 175 million active users[7] and in March of the same year that number exceeded the 200 million mark.[8]

In second place is Myspace with 125 million unique visitors. This is peculiar because Myspace was the leader in this field for a longer period of time until Facebook had caught up with it in March 2008 and later took the lead. [9]

Facebook recorded almost a record growth from 2006 (as far as the comScore research goes) to this day (e.g. the growth of 270% in the period from June 2006 to June2007 or the growth of 154% in the period between 2007 and June 2008)[10]. The information of the company Alexa, which deals with the tracking of the number of visitors of the leading world web sites, is also interesting. It says that 7 of the 20

[4] Wikipedia (2009), «List of social networking websites», (accsessed March 14, 2009), [available at http://en.wikipedia.org/wiki/List_of_social_networking_websites].
[5] Webopedia Computer Dictionary (2009), «Unique visitor», (accsessed March 14, 2009), [available at http://www.webopedia.com/TERM/U/unique_visitor.html].
[6] TechCrunch (2009), «Facebook Now Nearly Twice The Size Of MySpace Worldwide», (accessed March 8, 2009), [available at http://www.techcrunch.com/2009/01/22/facebook-now-nearly-twice-the-size-of-myspace-worldwide/].
[7] Facebook (2009), «Statistics», (accessed March 9, 2009), [available at http://www.facebook.com/press/info.php?statistics].
[8] Facebook (2009), «Statistics», (accessed March 15, 2009), [available at http://www.facebook.com/press/info.php?statistics].
[9] TechTree (2008), «Facebook: Largest, Fastest Growing Social Network», (accessed March 8, 2009), [available at http://www.techtree.com/India/News/Facebook_Largest_Fastest_Growing_Social_Network/551-92134-643.html]
[10] TechTree (2008), «Facebook: Largest, Fastest Growing Social Network», (accessed March 8, 2009), [available at http://www.techtree.com/India/News/Facebook_Largest_Fastest_Growing_Social_Network/551-92134-643.html]

worlds most visited web sites are social networks (or web sites like YouTube or Hi5 which for the most part consist of the components of social networking) (Roos 2007).

With the growth of active users of social networks, the demographic picture significantly changes. Although great popularity was first achieved with younger groups of Internet users, today social networks increasingly attract a more mature population. For example, until august 2006 more than two thirds were older than 25 with more than 40% of them being in the age group from 35-54 years (Roos 2007). Facebook started as a social network exclusively for Harvard students in order to later expand to all Internet users. Today, more than half of Facebook users are not students.[11]

An interesting and very practical point of view about estimating the value of social networking was published on the futuria.hr blog (Kovač 2007). The most significant is the economic background of every online social network. Every network is created with the goal of attracting as large a number of users as possible in order to implement the advertising model on the basis of which online networks earn their profit. In other words, a social network has to rent advertising space if it wants to be profitable. In order to achieve that it has to create an added value that will attract users. When visitors come (and become users) it is almost certain that advertisers will come to.

The social network has the task to continually take care of the benefits for the users and advertisers at the same time. The concept is founded on the balance with which "[…] the network will achieve […] "adequate" […] income and, on the other side, have advertisements which will not be to intrusive and tiring for the visitors, but will be relevant and connected with their preferences so that they will practically offer them useful information." (Kovač 2007). That balance is one of the key elements of the Google advertising model which is for now showing to be extremely successful.

3. Types of online social networks

It was already mentioned that over 300 active online social networks exist currently. But, they do not promote the same forms and types of activity. It is, because of these different ways of connecting and the basic line that connects one social network, possible to place most of them in a few general categories (Roos 2007).

- **Community Social-networking Sites**

In this broad category we place every site for social network connecting which purpose is to enable users to expand their social network through existing contacts and

[11] Facebook (2009), «Statistics», (accessed March 14, 2009), [available at http://www.facebook.com/press/info.php?factsheet#/press/info.php?statistics].

searching for new contacts. Some online social networks aim for a very specific demographic group. For example, LinkedIn has established itself as a social network for business professionals, Black Planet and Asian Avenue are directed towards racial and ethnic groups. Other sites are directed towards the global online community within which smaller groups are created according to interests or demographic characteristics (Roos 2007).

- **Media sharing sites**

The basic idea of these networks is to send your own graphic, audio or video recordings and make it possible for them to become publicly accessible through a social network. In other words, it is not necessary to become a user in order to watch the media, but it is necessary to be one in order to upload media to the server. The most popular among the sites for sharing video contents is YouTube and Flickr stands out as the most popular site for the exchange of graphic material (photographs, pictures).

- **Music Social Networks**

According to the definition they are sites for the exchange of multimedia contents, music social networks can, according to some experts, be treated as a separate group. This is mostly because of the fact that they are quite popular among users of social networks. Last.fm is a great example of such social networks. Users can search for and listen to songs of their favorite artists and groups, comment, rank them and create their own playlists.

- **Blogging Social Networks**

These are sites for social networking which gather individuals and their blogs. Instead of a usual profile, every user writes and edits its own blog which has a certain theme. Sites like Live journal or Blogger enable their users to search for and read through interesting blogs (Roos 2007).

Since social networks are so many, the above categorization did not take into consideration all of them, but it did cover the most significant and popular groups. None the less, regardless of the type, specific target group and aims of an individual social network, such communities have already become dominant places for the gathering of Internet users. This trend will probably continue to develop in the future.

The development of social networks in the worldwide Internet space had to stimulate the development of the domestic web scene. Some networks have appeared and disappeared but a few remained and still exist today. In 2007 Tulumarka was started as the first social network about night life and later that year Trosjed was introduced to the Croatian Internet user, "the first Croatian social network based exclusively on Microsoft technologies".[12] Additionally, the currently available Croatian social networking sites are smajli.hr, linker.hr and Mojnet.

[12] Ict.hr (2008), «Domaće društvene mreže», (accessed March 21, 2009), [available at http://www.ict.hr/date/2008/02/15/].

The current situation on the Croatian social networking scene is best illustrated by the fact that the established worldwide social networks are much more significant than the domestic ones and that they have a larger number of Croatian users.[13] Although the data somewhat varies from source to source[14], we can point out the data from the website AllFacebook[15] which exclusively deals with statistic data concerning that popular social network. According to that source, on March 15th Facebook had 576 540 active users from Croatia, which is a respectable number. This implicates that Facebook is not just the most visited social network, but the most visited website if we observe the Croatian Internet users.

4. Facebook

It was mentioned that Facebook is currently the most popular online social network among Internet users. In its beginning Facebook was not accessible to everyone, but was exclusively directed towards the college population of certain American universities. The whole story started in 2004 when Mark Zuckerberg, Dustin Moskovitz and Dustin Hughe, all three Harvard students, started thefacebook.com.[16] The basic idea was to link students of that university and creates a web location on which the students will be able to communicate, publish and exchange various digital contents. By the year 2005 students from 800 American university networks were able to become users of the application which then had over 5 million active users. In August of the same year the network started working under the new name Facebook. Today, every Internet user, if he is over 13 (that is the only restriction) and has a valid e-mail address, can become a user of Facebook. According to Facebook's information, around 70% of its users come from outside the US.[17]

- **Facebook profiles**

A user account on Facebook is created in a very simple fashion by entering the necessary data in the application form. After the registration Facebook will generate a user profile. This profile is a web page that all other users can find. When a user finds a familiar person he sends a friendship request and will be able to view the entire profile only when that person approves the request.

Every user profile looks basically the same. The structure is the same and the contents within the page are distributed in the same way. However, users can modify

[13] Ict.hr (2008), «Domaće društvene mreže», (accessed March 21, 2009), [available at http://www.ict.hr/date/2008/02/15/].

[14] Manjgura (2009), (accessed March 21, 2009) [available at http://manjgura.hr/tag/drustvene-mreze/].

[15] AllFacebook (2009), «Facebook Demographic Statistics», (accessed March 21, 2009), [available at http://www.allfacebook.com/facebook-demographic-statistics/ ().

[16] Wikipedia (2009), «Facebook», (accessed March 22, 2009), [available at http://en.wikipedia.org/wiki/Facebook].

[17]Facebook (2009), «Statistics», (accessed March 15, 2009), [available at http://www.facebook.com/press/info.php?statistics].

their profile by inserting numerous contents and change the basic appearance of their profile in this way. The Facebook team also continually works on improving the layout of the profile and other pages. The visual identity stays basically the same and the biggest changes are concerning the arrangement of individual elements within the page. In the beginning these changes come upon open protest from the users but end with the user's adaptation.

Along with the mentioned personal profiles, Facebook enables the creating of different profiles which are not intended for private individuals. Companies, artists, musical groups and public figures can create their own special user profiles using Facebook Pages. These profiles look very similar to the usual profiles, but differ in the fact that friendship is not necessary for viewing them. That means that anyone can view them even if he is not a registered Facebook user. A user cannot become a friend with such a profile but can become a fan and connect in that way with what is represented through the profile (Strickland 2007).

- **Facebook Groups**

The Facebook platform enables the creating of groups which connect users with the same interests. One company can create its own specific profile, but it can also create its own group. The group enables the communication with all members of the group through messages which can be treated as some kind of internal e-mail service. However, we should point out that a user is not automatically a member of the group if he/she is a fan of the company, i.e. of the Facebook profile of the company. One does not condition the other. A group can also be marked as private and allow membership only to a certain population of users. Facebook groups can only be searched through the Facebook system while Facebook Pages profiles can be found with tools for the searching of Internet resources (dominantly search engines).

Philip Robertson, the director of marketing communications of the company ooVoo (applications for video face to face communication), wanted to incorporate the application within social networking immediately after its launch in mid 2007. Since Facebook soon became a place in which people are connected online, at the same time the company started searching for new ways of marketing and communicating with its users. Robertson started a public Facebook group as a way of communicating with existing ooVoo users and in order to create a larger population of users of the service simply because the existing users promoted the application to their friends. Robertson pointed out that Facebook groups were a very useful tool for marketers mainly because of the transfer value they had (Meerman Scott 2009). On a certain level members of a group work for the group and promote it with their activities.

- **Advertising on Facebook**

Along with the above mentioned Facebook Pages profiles and Groups there are a few other tools which marketers can manage in order to achieve certain marketing goals. We should mainly point out Facebook applications, Social Ads, Facebook Polls and Facebook Beacon (Alba 2009).

Social ads can be placed on two locations, as a banner which is shown on user profile pages on the right hand side or as a banner which is shown as sponsored content in the news section that the user can see. Both options offer an approach to very precisely segmented user groups according to demographic data they themselves have submitted (age, location, interests, university, college, work place etc.). Both types of advertisements can be bought on the basis of cost per one thousand impressions (CPM[18]) or on the basis of the number of clicks a banner has generated (CPC – cost per click). Detailed analytics come with each type of advertising which help to determine the success of an advertising campaign, achieve a more precise segmentation and lead to corrective activities.[19]

Facebook Polls use identical segmentation as Social ads and enable the polling of a chosen segment, i.e. of those users which suite the set criteria. It is possible to get the poll results in real time and they are segmented according to the age and sex of the individuals questioned. But the most significant aspect of Facebook is the applications.

- **Facebook applications**

The basic Facebook applications include adding pictures and videos, creating groups and events, taking part in the marketplace (classifieds), publishing information about external contents (e.g. other website addresses) and the sending of virtual gifts (Strickland 2007). These applications have been developed by Facebook and are at disposal to all users. There are also hundreds of applications which have not been developed by Facebook but by the users themselves. Those applications are called third-party applications. Many websites, but mostly social networks, give their users a limited access to the application programming interface (API) which enables them the development of applications that are executed on the website platform. Authors of the applications aim for entertainment value and for a new level of interactivity which will motivate the user to visit or/and use it often.

A very interesting example is the "Životinjsko carstvo" application which was developed by the Croatian company Drap for Kraš. The users of this application get daily new stickers from the Životinjsko carsto sticker album. Among those are also duplicates which can be traded with friends (if they use this application). The idea is, as with the offline version of the Životinjsko carstvo sticker album, to educate the user[20] about animals and motivate them to complete the album by collecting stickers. Kraš rewards those users with their gift packages. The important thing to point out here is that the user gets new stickers on a daily basis only if he opens the application which necessarily means visiting the application often. According to informa-

[18] The letter M in the acronym stands for the word mille which is the Latin word for one thousand.
[19] Facebook (2009), «Advertising», (accessed March 04, 2009), [available at http://www.facebook.com/advertising/].
[20] Usually part of younger age groups, although the rule has exceptions, especially within the boundaries of a Facebook application.

tion on the description page[21], the application had more than 120 000 active users in March of this year.

The use of all the advancements of Web 2.0 helped with a strong interactivity is definitely becoming an everyday occurrence in the Internet space. Similar applications are developed for Podravka, mostly connected with their Čokolino (Ivezić 2008). Along with Kraš and Podravka, the second largest Croatian telecom operator Iskon has also created an interesting Facebook application – "Zapali s ekipom u London" (Visit London with your crew). The idea is to reward the most original photographs and comments (and the platform for that is Facebook) with a trip to London.[22] According to Facebook's information, more than 70% of their users actively use applications (according to data from April 2009). This report talks about 52 000 applications which are currently available to Facebook users.[23]

- **Mobile Facebook**

Using a mobile phone a user can send text messages, pictures and videos to his/hers Facebook profile. Text messages are treated as classic SMS (short message service[24]), for sending pictures and videos the multimedia format MMS (multimedia messaging service[25]) must be used (Strickland 2007). The user must send all these formats and types of messages to a specific e-mail address in order for the message to appear on his profile.

The popularity of mobile Facebook is also showing strong growth. In April of this year the number of active users of mobile Facebook (i.e. the users which accesses Facebook via mobile phone) has exceeded 30 million.[26] An interesting fact is that "mobile users" are 50% more active than the users that access Facebook through a computer. The whole situation has reflected on mobile operators all over the globe. More than 150 of them spread over 50 countries are currently working on developing and promoting mobile Facebook.[27] Part of the popularity can certainly be attributed to the most popular electronic device of 2008 – iPhone.

[21] Facebook (2009), (accessed March 29, 2009), [available at http://www.facebook.com/apps/application.php?id=21049955977&b&ref=pd_r]
[22] Iskon (2008), «Zapali s ekipom u London», (accessed March 29, 2009), [available at http://www.iskon.hr/za_kucu/zapali_s_ekipom_u_london].
[23] Facebook (2009), «Statistics», (accessed April 15, 2009), [available at http://www.facebook.com/press/info.php?statistics].
[24] Wikipedia (2009), «Short message service», (accessed April 04, 2009), [available at http://en.wikipedia.org/wiki/Short_message_service].
[25] Wikipedia (2009), «Multimedia message service», (accessed April 04, 2009), [available at http://en.wikipedia.org/wiki/Short_message_service].
[26] Facebook (2009), «Statistics», (accessed April 15, 2009), [available at http://www.facebook.com/press/info.php?statistics].
[27] Facebook (2009), «Statistics», (accessed April 15, 2009), [available at http://www.facebook.com/press/info.php?statistics].

Newer mobile devices enable Internet surfing. Because of that Facebook has created a modified version of its site which is adapted for mobile device usage. Websites can be accessed on m.facebook.com.

- **W – Marketing**

Widget marketing (and widget advertising) is a completely new approach to Internet marketing which takes place on social network platforms. Widget is generally defined as a transferable piece of programming code which can be installed and executed in any separate HTML page and does not require separate compiling on the end user's computer.[28] From a marketing point of view, this term can be explained as an interactive application in a social network that is on the one hand interesting to use, and on the other hand business goal oriented (Bhargava 2006). Those goals are mostly connected to advertising of companies, products or services or improving brand awareness. Companies that have tried this electronic marketing technology are widely known names: Electronic Arts, Viacom, Sony, Gap, Hewlett-Packard and others.

Expert advertising teams dealing with Facebook have come up with this kind of advertising mostly because classic banner advertising did not bring a large enough number of clicks to advertisers. Facebook users have always seen advertisements as a nuisance which sometimes irritated them while they were socializing with other users (the same applies to other social networks). The key to success is hidden in "widget advertisements" and their adequate use. Companies need to create content in which the community can play the main role, in which user interaction dominates and which can be integrated with other tools. The results of this will be visible in a longer period of time if the procedures for measuring their success change. Traditional techniques of advertising and measuring of successfulness will not be of use here (Owyang 2008).

Parking wars is a Facebook application which was developed by the marketing team of the company A&E Television Network in order to promote a reality show about a group of Philadelphia Parking Authority employees. Although the idea of a reality show about employees of a company that deals with maintenance and collecting parking fees does not sound very interesting, using the Facebook application, the company achieved outstanding success (Jurman 2009). The application (widget) enables the users to "transform" their profiles into parking spaces and to park their virtual car on other user profiles. The fun in all this is that the user is trying to catch "illegally parked" users on his profile and charge them a fine while at the same time avoiding being caught and fined on other user's profiles. The result of this technique is fascinating. In the first two months of existence 200 000 users installed this application on their profiles and the number of visits to the before mentioned reality shows website soared to 45 million (Jurman 2009). The success of this

[28] Wikipedia (2009), «Web widget», (accessed April 17, 2009), [available at http://en.wikipedia.org/wiki/Web_widget].

application results from its interactivity which did not necessary have a strong link with what it was promoting, but was interesting enough to attract a large number of users. The large popularity of the application was indirectly transferred to the reality shows website.

- **The advertising potential on Facebook**

Advertising on Facebook is today used by numerous companies from very different areas of business, but organizations and individuals also use it. A relatively large number of those who are criticizing banner Facebook advertising point out the fact that a large number of impressions results in a very small number of clicks (in most cases comments refer to business subject experiences). However, it is important to emphasize that experiences vary and mostly depend on the manner in which the banner was created and on the target group towards which it was directed (O'Neill 2007). Because we are speaking of a rather new advertising trend, a large number of individuals who work in marketing have agreed that every advertising campaign should be tested and refined during the time it is public and that one should learn of his own experience. The other key issue of this discussion is the cost of banner advertising. On average, Internet advertising has by far the lowest cost compared to traditional techniques. Banner advertising on Facebook is, concerning expenses, very acceptable especially because the advertiser can decide himself how much he is willing to pay for a thousand impressions or one click (the smallest amount is a symbolic 1 cent) and what his weekly advertising campaign budget should be.

5. Facebook in the marketing of non-profit organizations

One of the most frequently used specific functions which the social network Facebook provides for its users are the special profiles called Facebook Pages[29] which are intended for a special subject group. This group consists of companies, products, brands, organizations, bands, artists or public figures. It is apparent that we are not dealing with private users, but with special subjects who want utilize this, still relatively new, way of communication. The subjects create these profiles free of charge. Of course, the only thing that is necessary is that the author is a Facebook user. The process of creating a profile is similar to creating a regular personal Facebook profile and takes up very little time. The type and amount of information which will be placed on this specialized profile is determined by the author. This will, of course, vary depending on the theme of the profile i.e. the subject that is represented. There is a number of tools which help you place the information. The possibility of adding general information about the subject, news or other textual information, photo galleries, videos and links is at the disposal of the user. In addition, it is possible to cre-

[29] Facebook, (2009), „Create a page", (accessed April 15, 2009), [available at http://www.facebook.com/pages/create.php].

ate events connected with the subject. As with other web sites, visitors (or users) need to be continually attracted with quality content which is regularly updated and with newly created content (Borges 2009). Facebook users can confirm their interest in the subject by becoming fans of a certain Facebook Pages profile (Belwani 2009). Creating a profile is very easy, but the complexity of maintenance and usage of the right tools becomes apparent very soon. Numerous experts give very precise advice on how to correctly use a social network and the media of social networking and how to contribute to the goals of the organization (Meerman Scott 2009). Belwani focuses on Facebook profiles and their fans (Belwani 2009):

1. Connecting with other platforms – the connecting of multiple social networking platforms through your own web site helps to direct users through the network;
2. Creating of resources (additional content) – adding additional content which is not directly linked to the primary activity of the company helps to broaden the user base;
3. Improvement of sales through participation – sales improvement activities (coupons, free delivery, weekly sales) significantly broaden the user base and improve popularity;
4. Support of existing profiles – companies need to encourage users to be proactive in creating official profiles of the subject and build independent brand promoters.
5. Precise segmentation – all segments are not equally represented on social networks. Therefore, it is necessary to choose the message precisely with the knowledge to whom you communicate.

These pieces of advice are aimed at business subjects, although they can be used when dealing with non-profit organizations and all those who use social networks for non-personal purposes. The interactivity is of utmost importance and the social networks and the media of social networking contribute to it, because of its two-way communication, in ways that are unprecedented to date (Evans 2008).

Postgraduate study Marketing of significant areas on the social network Facebook

Soon after creating a classical web site of the postgraduate study Marketing of significant areas[30] at the Faculty of Economy in Osijek, a need has occurred to bring the course program closer to existing and new students in an environment that is familiar to them according to the demographic segment they belong to. This conclusion arises from the fact that this is a relatively young population that has good computer skills and for which internet communication is part of the everyday routine (Bonfield 2009). It is easy to conclude that social networks are a good platform for building social relations and bringing the users together (Hargadon 2009).

[30] More about the subject on http://www.mpp.efos.hr/

Marketing of significant areas Facebook profile[31] started its existence at the beginning of April 2009. All important information about the course has been added to the profile, special notes were created with all the basic questions and answers which a potential (even an existing) student could have. In addition, photo galleries of recent course events with appropriate comments were added. Part of the content was taken from the web site of the study through which visitors were directed to the Facebook profile.

With the creation of this profile, all important events and information were published on both the web site and Facebook profile and links were created which connected them between the two pages. Be it a visiting professor or a marriage of a colleague, all the interesting information found their place on the 2 mentioned web sites. By using the advantages of the individual web sites we have achieved a synergy of the two. The classic web site has no limitations when creating content and choosing the form in which it is published. It also enables the protection of certain data (creating groups and content which are unavailable to the public). Of course, the classic web site is appropriate for those who are not part of Facebook and is a necessity today.

On the other hand, the Facebook Pages profile has proven itself as a great location for publishing photo galleries with the special function of connecting photos with Facebook users, their mutual commenting and their content suggestions. Facebook profiles enable the private users (and fans) to create their own part of the content and in this way improve the level of interactivity and social networking. Finally, it is important to point out that the Facebook profile of the study ranks high with search engines which is indisputably of great advantage for the course and is gained through a profile that is free of charge. Most users come to the web site through these search engines as a dominant tool for searching the Internet.[32] Because of this the position according to the key words in the data bases of search engines is treated as one of the most important indicators of success of web sites.

At the moment of writing of this paper the Facebook profile of Marketing of significant areas had 40 fans, which is a very small but homogenous group. However, we should emphasize the fact that the course has only two generations of students and that this has undoubtedly affected the relatively small number of fans, since fans of this profile are mostly students of the course. Of course, with proper management of the profile and with special emphasis on the quality of the content, the goal of expanding to other users and improving the level of popularity and awareness of the broader public through these activities should be achieved.

[31] More about the subject on http://www.facebook.com/pages/Osijek-Croatia/PSS-Marketing-posebnih-podrucja/130184715091

[32] CyberMark Inc., (2009), „Why is search engine optimization important?", (accessed November 5, 2009), [available at http://www.cybermarkintl.com/why-optimization-important.html/]

6. Example of Facebook advertising

It was stated earlier that almost every social network, primarily those directed at online communities, make the most of their profit by renting their own advertising space within their web-sites. Facebook, as the most significant social network according to the number of unique monthly visitors, is not the exception. Despite that, it was mentioned that the return of investment in Facebook advertising is a subject open to discussion, since it is very difficult to precisely measure the efficiency and, very often, we lack the precise tools which can reliably measure that conversion. An apparent fact is that the very efficiency of advertising will vary primarily depending on the goals of advertising, rather than the object of advertising i.e. the message that is communicated to the target group, but also on the way the target group is chosen and the advertisement is designed. Speaking for this is the following fact: a number of experts in internet advertising and social media have agreed that products or industries exist which are grateful subjects, such as dating, weight loss (in other words, advertisements in these areas generate a large amount of clicks, but also a large level of conversion) (Digital Marketing World online conference 2009). On the other side of the medal, there are products or industries where advertising will not achieve the desired effect not even remotely as some other types of advertising would. The most frequent examples for this are connected to the agriculture industry, although a few exceptions can be found here.

With the goal of shedding light on the dilemma of how efficient Facebook advertising is, the advertising of the postgraduate specialist study Marketing at the Faculty of economy in Osijek was conducted. It is important to highlight the fact that the target group for this study is highly represented in the overall population of Croatian internet users (a young to mature population, between the ages of 20 to 40, with a completed graduate study)[33] which is of great importance when considering online advertising and the possible options that come with it. From the few possible options of Facebook advertising the most common one was chosen – banner advertising. This choice was the result of the consideration of practical factors; banner advertisements can be created easily and fast with almost no additional financial expenses, in contrast to e.g. developing special Facebook applications. An interesting fact is that some universities (e.g. Stanford) have started advanced courses for creating Facebook apps as early as 2007.[34]

In the fall of 2009, the managing staff of the Marketing of significant areas study prepared for the enrollment of a third generation of students and the goal of Facebook advertising was to attract new students. The secondary goal was to collect primary information which would result from the conducting of the campaign with

[33] GfK Research, (2009), „Internet – usporedba: Hrvatska i zemlje centralne i istočne Europe", (accessed July 23, 2009), [available at http://www.gfk.hr/press1/internet4.htm]
[34] Virtual Marketing Blog, (2009), „How to make Facebook applications", accessed July 23, 2009), [available at http://www.virtualmarketingblog.com/how-to-make-facebook-applications/]

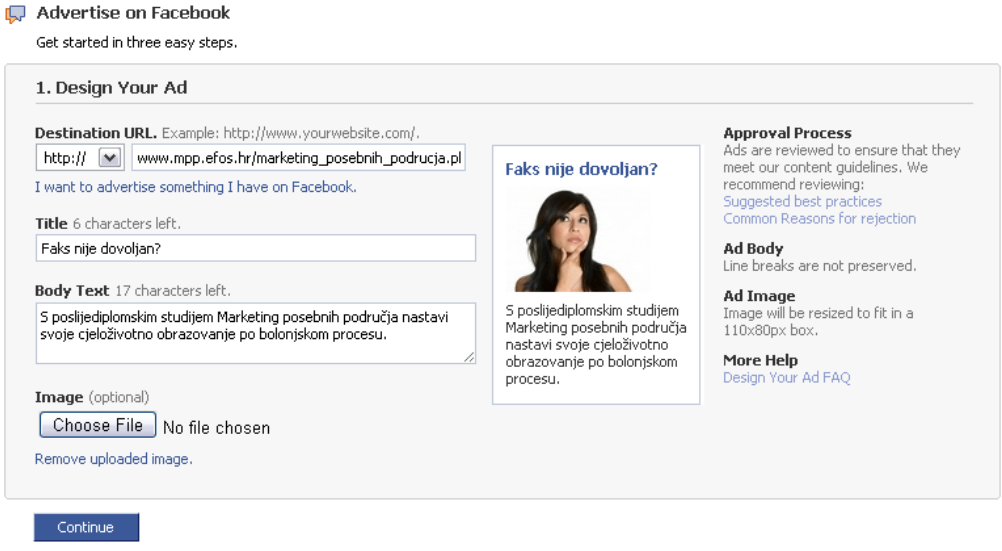
the final goal of testing the success of this type of advertising. A two week (fourteen days) frame was set, from the 15. – 28. of July. This course has its own Facebook profile but it was decided that the placed ads would lead to the, for this occasion specially created, page within the web site of the course (landing page). The ad led to the web site of the course since it enabled more freedom in creating and visually presenting the content of the page and the possibility of an online application for the course.

The Facebook advertising system enables a very simple and fast creation of the desired ad in 3 steps:

1. Visual layout creation
2. Target group determination
3. Creating of the campaign i.e. determining the prices

In order to determine the different options for the advertised text in the ads, two Facebook ads in two parallel advertising campaigns were created. Both ads had identical titles: „College wasn't enough?“, alluding at the need to continue education after college graduation i.e. the trend of whole-life education. The ads were different in the images that were displayed and the textual part. The first ad contained the image of a young woman wondering about something which is consistent with the title of the message. The procedure by which you create the ad is shown in Figure 1.

Figure 1. Step one: creating the visual layout of the Facebook ad



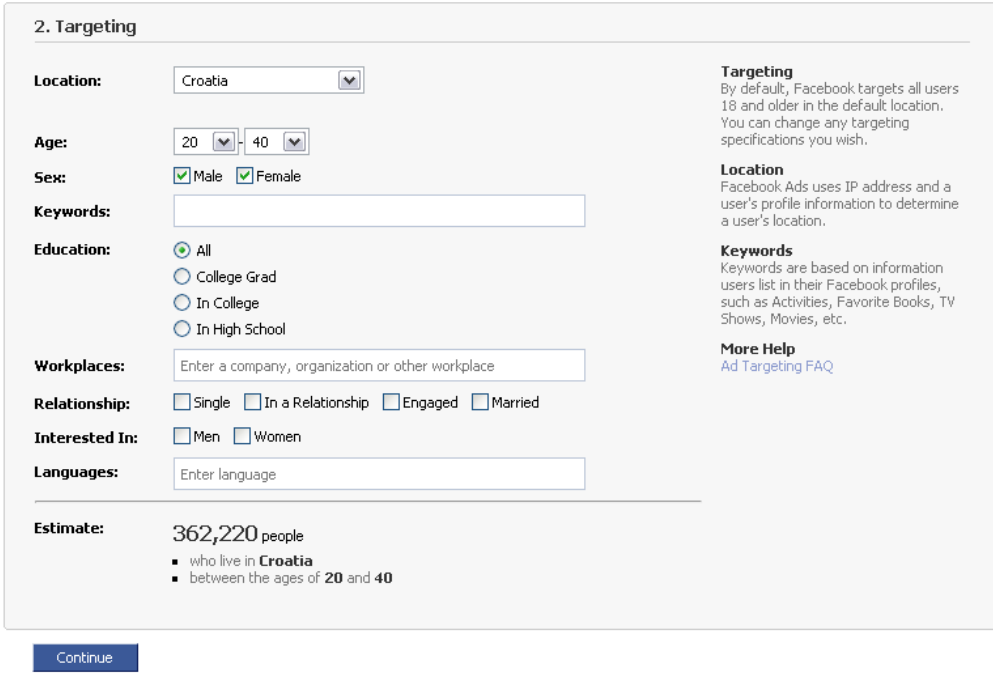
Source: Facebook

The second ad had the image of a young gentleman with a similar facial expression. This approach was chosen because the faces of people attract more attention then, for example, just the logo of the course or other font types. It is also apparent that the users of Facebook are exposed to new ads on their profiles and it is a demanding task to attract their attention. Of course, the images of a young woman and man are not in contrast with the goal of the ad and do not constitute false advertising since the potential students of the course are mostly young people (relevancy of the images is important since every ad has to pass testing from the Facebook staff which is called the approval process). The bodies of both ads were very similar, only being different in that he first contained the information that the course was created in concordance with the Bologna process and the second contained the title which is awarded when graduating (univ. spec. oec.). Both pieces of information are important to the target group, but because of the limited amount of signs in the textual part of the ad they were distributed in two ads.

The second step (Figure 2) in the process of Facebook advertising refers to determining the target group. Since Facebook has precise information about its users, it enables an accurate segmentation and the precise determining to whom the ads will be placed. More developed markets have more developed segmentation, but in Croatia it is also possible to achieve a high level of target group determination.
Besides determining the country, sex, age and education segmentation is conducted. It is also possible to determine the marital status of the desired users and the language which they speak. It is also possible to choose the search of users according to keywords, although it has proven that this function is not at an adequate level when having Croatian internet users in mind (it is assumed that the language barrier is a big problem for this type of filter). For the needs of this campaign it was decided to segment the users exclusively according to location[35] and age; i.e. users from Croatia between the ages of 20 and 40 of all educational levels. After the filtration of the overall population of users from Croatia, Facebook projects with how large a segment it will be communicated i.e. how many users exist in the chosen target group. In this example, there were 362 200 users who constitute the target group according to the given parameters.

[35] It is important to point out that more developed economies have a more precise geographical segmentation in the sense that besides choosing the country, one can choose a more narrow territorial unit, an individual city or place. It can be estimated that with an increasing number of Internet users this will also be possible in the croatian market.

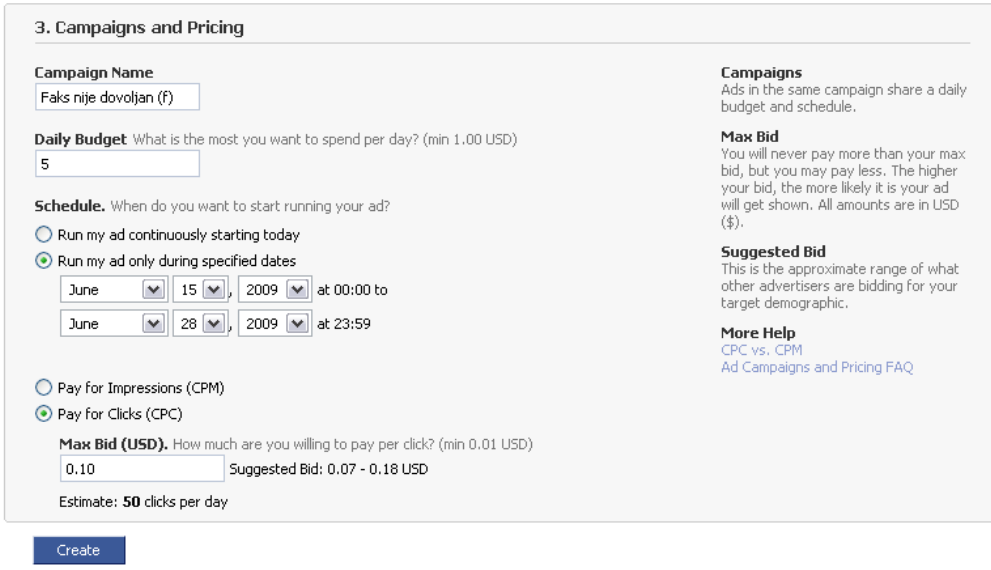
Figure 2. Step two: determining the target group



Source: Facebook

In the third step (*Figure 3*) the process of creating the campaign is finalized; a name is assigned to the campaign, the time frame of advertising is determined and the financial parameters are defined. There are two expenditure levels which are determined. The first is the daily budget of the individual ad which is the maximum that can be spent on one ad in one day of advertising. On the second level it is possible to choose the manner in which the expenses are calculated; the payment of one thousand displays (CPM –cost per mille) or payment per click on the ad (CPC cost per click). Since the displaying of an ad is not the most adequate indicator because of the aforementioned large exposure to ads, the majority of advertisers decides for the payment per click method which was also the case with this campaign. The level of 5 American dollars was taken for both ads and the cost per click was set at 0.10 dollars and 0.08 dollars respectively. The Facebook system suggested the expense frame of the individual click after the parameters were given and because of this these expense levels were chosen (the system suggests the expenditure class, and in this example the middle values of these classes were taken). We must emphasize the fact that the lowest daily budget is 1 dollar and the cost per click is 0.01 dollar. Undoubtedly, the levels would mean a smaller amount of displays i.e. a more poorly positioned ad.

Figure 3. Step three: finalizing the process of creation

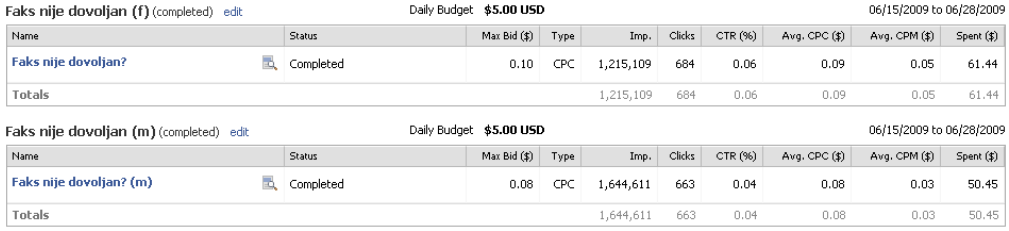


Source: Facebook

After creating the advertising campaign, the Facebook system has to check the suggested ad. Within the system itself there is a guide on how to correctly create an ad and which are the most frequent mistakes when creating individual ads or campaigns. A few hours were necessary in order for the ads to pass the testing and for the campaign to start.

The two week advertising campaign achieved the following results (shown in *Figure 4*).The first ad „College wasn't enough? (f)“ was displayed 1 215 109 and 684 clicks were generated. The overall cost was 61.44 dollars with an average cost per click of 0.09 dollars. The second ad „College wasn't enough? (m)“ was displayed 1 644 611 times which generated 663 clicks. The overall cost of this ad was 50.45 dollars with an average cost per click of 0.08 dollars (which was the maximum amount set for this ad). It is apparent that the first ad generated more clicks from a lesser amount of displays with a larger overall cost. Of course, one should bear in mind that a higher cost per click (0.10 dollars) was set for this ad. However, it is very important to point out that the limited time period of two weeks was not enough for a precise determination of the reason for the larger amount of clicks of the first ad which means that we cannot, a priori, conclude that the reason for that was the young lady in the ad image or the information that the course is in concordance with the Bologna process.

Figure 4. Statistical data of the Facebook advertising campaign



Faks nije dovoljan (f) (completed) edit — Daily Budget $5.00 USD — 06/15/2009 to 06/28/2009

Name	Status	Max Bid ($)	Type	Imp.	Clicks	CTR (%)	Avg. CPC ($)	Avg. CPM ($)	Spent ($)
Faks nije dovoljan?	Completed	0.10	CPC	1,215,109	684	0.06	0.09	0.05	61.44
Totals				1,215,109	684	0.06	0.09	0.05	61.44

Faks nije dovoljan (m) (completed) edit — Daily Budget $5.00 USD — 06/15/2009 to 06/28/2009

Name	Status	Max Bid ($)	Type	Imp.	Clicks	CTR (%)	Avg. CPC ($)	Avg. CPM ($)	Spent ($)
Faks nije dovoljan? (m)	Completed	0.08	CPC	1,644,611	663	0.04	0.08	0.03	50.45
Totals				1,644,611	663	0.04	0.08	0.03	50.45

Source: Facebook

The provided numbers of ad clicks are still not an adequate indicator of the success of this advertising campaign since the individual click and the possible viewing of the page of the whole web site does not mean that the user will actually apply for the course. In order for this conversion to be calculated more precisely an online application form was used as an exact indicator for the number of applicants. In addition, the analytical tool Google Analytics was used to observe the number of visitors to the course web site. Google Analytics confirmed a significantly larger number of visits of the web site during the time of advertising and the significant majority of those visitors came from Facebook. A relatively large number of Facebook visitors generated a satisfactory amount of 3 i.e. 4 applications (one applicant showed interest for the course earlier but the Facebook advertising accelerated her process of application) for the new generation of the study Marketing of significant areas.

From all that was said above we can conclude that in this case Facebook advertising was extremely efficient since a relatively satisfactory amount of applications was generated in a short period of time and with relatively small expenses. This research will undoubtedly be more complete when other methods of internet advertising as well as traditional advertising in the offline world are tested.

7. Conclusion

The term social network exists for more than half a century, but the fast development of social network websites like Facebook, Myspace, LinkedIn and others have transformed the old sociometric term into one of the most significant and most used Internet terms in recent time. Social network websites are on a global level today the most visited web locations with very active users who spend much time in online socializing. But online social networks with their advantages make a very successful platform for a wide array of business activity. Two social networks that definitely stand as the globally most popular are Facebook and Myspace and since April 2008 Facebook took over the dominant role on that market. With the growth of the number of active social network users, the demographic setup of the users significantly changes. Although great popularity was first achieved with younger groups of Inter-

net users, today social networks attract a more mature population which is of great importance for usability of these networks for business purposes.

Facebook is definitely an interesting platform for marketing activities on all levels: this platform constantly develops and finds new ways for the users to attain useful information and not just irritating ads. On the other hand, the goal is to enable the advertisers a very segmented approach which can achieve the best advertising results and achieve the best level of conversion, i.e. indirectly speaking a return on invested funds. Although a certain level of justified skepticism exists towards this type of promotional activity, there are enough examples that speak for the fact that it is the future of advertising.

The fact is that Internet advertising is an area that rapidly changes and creates completely new techniques and forms adjusting according to those at which they are aimed. The situation is interesting if we observe social networks which almost perfectly know their users. From all the above mentioned a great marketing potential can be seen, but also the necessity to develop and adjust to users and advertisers. All this speaks for the thesis that we are in a time of increasingly direct marketing in which the target group is not part of the mass but many individuals with which we communicate almost directly.

References

Alba, J. 2009: Facebook Success Stories: How 12 companies are leveraging the social network to advance their marketing objectives. MarketingProfs Case studies collection, (accessed March 04, 2009), [available at http://www.marketingprofs.com/newprem/library/item.asp?id=2610].

AllFacebook 2009: Facebook Demographic Statistics. (accessed March 21, 2009), [available at http://www.allfacebook.com/facebook-demographic-statistics/].

Benzie, R. 2007: Facebook banned for Ontario staffers, TheStar.com, (accessed March 22, 2009), [available at http://www.thestar.com/News/article/210014].

Belwani, S. 2009: 5 Elements of a Successful Facebook Fan Page. Mashable, (accessed November 5, 2009), [available at http://mashable.com/2009/03/30/successful-facebook-fan-page/]

Bhargava, R. 2008: Personality not included. McGraw-Hill, New York.

Bonfield, B. 2009: Should Your Organization Use Social Networking Sites? Idealware, (accessed November 6, 2009), [available at http://www.idealware.org/articles/should_you_social_network.php/]

Borges, B. 2009: Marketing 2.0. Wheatmark, Tucson, Arizona, 198-199. p.

ComScore 2007: Social Networking Goeas Global. (accessed March 8, 2009), [available at http://www.comscore.com/press/release.asp?press=1555].

CyberMark Inc. 2009: Why is search engine optimization important? (accessed November 5, 2009), [available at http://www.cybermarkintl.com/why-optimization-important.html/]

Digital Marketing World online conference 2009: Wining Against the Odds. April 1, 2009, (accessed July 23, 2009), [available at http://www.marketingprofs.com/events/6/conference/]

Evans, D. 2008: Social Media Marketing. Wiley Publishing Inc., Indianapolis, Indiana.

Facebook 2009: Advertising. (accessed March 04, 2009), [available at http://www.facebook.com/advertising/].

Facebook 2009: Statistics. (accessed April 15, 2009), [available at http://www.facebook.com/press/info.php?statistics].

Facebook 2009: Create a page. (accessed April 15, 2009), [available at http://www.facebook.com/pages/create.php].

GfK Research 2009: Internet – usporedba: Hrvatska i zemlje centralne i istočne Europe. (accessed July 23, 2009), [available at http://www.gfk.hr/press1/internet4.htm].

Hargadon, S. 2009: Thoughts on Social Networking in Education. (accessed November 6, 2009), [available at http://www.stevehargadon.com/2008/11/thoughts-on-social-networking-in.html].

Ict.hr 2008: Domaće društvene mreže. (accessed March 21, 2009), [available at http://www.ict.hr/date/2008/02/15/].

International Sociometry Training Network 2009: (accessed March 15, 2009), [available at http://www.sociometry.net/].

Iskon 2008: Zapali s ekipom u London. (accessed March 29, 2009), [available at http://www.iskon.hr/za_kucu/zapali_s_ekipom_u_london].

Ivezić, B. 2008: Domaće tvrtke krenule u osvajanje kupaca preko Facebook-a Poslovni dnevnik. (accessed March 29, 2009), [available at http://www.poslovni.hr/99268.aspx].

Jurman, H. 2009: Eureka internetskog oglašavanja. *PC Chip*, 166, 132-133. p.

Kovač, T. 2007: Kako procjeniti vrijednost društvene mreže? futuria.hr, (accessed March 15, 2009), [available at http://futuria.hr/2007/10/29/kako-procijeniti-vrijednost-drustvene-mreze/].

Manjgura 2009, (accessed March 21, 2009) [available at http://manjgura.hr/tag/drustvene-mreze/].

Meerman Scott, D. 2009: The New Rules of Marketing and PR. John Wiley & Sons Inc., Hoboken, New Jersey.

Meerman Scott, D. 2009: World Wide Rave. John Wiley & Sons Inc., Hoboken, New Jersey.

O'Neill, N. 2007: Is Facebook Advertising Effective? AllFacebook, (accessed April 22, 2009), [available at http://www.allfacebook.com/2007/11/is-facebook-advertising-effective/].

Owyang, W. 2008: What's Facebook's New. Engagement Advertising. Means to Brands. (accessed April 20, 2009), [available at http://www.web-strategist.com/blog/2008/08/21/facebooks-engagement-ads/].

ROI Factory 2009: Facebook Advertising Analysis – Part Two. (accessed April 22, 2009), [available at http://roifactory.com/index.php?option=com_content&view=article&id=22:facebook-advertising-analysis-part-two&catid=3:thedogfoodchronicles].

Roos, D. 2007: How social networks work, How stuff works. (accessed March 7, 2009), [available at http://communication.howstuffworks.com/how-social-networks-work.htm].

Ružić, D. 2003: e-Marketing. Ekonomski fakultet u Osijeku, Osijek.

Strickland, J. 2007: How Facebook works, How stuff works. (accessed March 22, 2009), [available at http://computer.howstuffworks.com/facebook1.htm].

Strickland, J. 2007: How Facebook works, How stuff works. (accessed April 03, 2009), [available at http://computer.howstuffworks.com/facebook4.htm].

TechCrunch 2009:Facebook Now Nearly Twice The Size Of MySpace Worldwide. (accessed March 8, 2009), [available at)http://www.techcrunch.com/2009/01/22/facebook-now-nearly-twice-the-size-of-myspace-worldwide/].

TechTree 2008: Facebook: Largest, Fastest Growing Social Network. (accessed March 8, 2009), [available at http://www.techtree.com/India/News/Facebook_Largest_Fastest_Growing_Social_Network/551-92134-643.html]

The Economist 2008: Red line that cannot be crossed. (accessed March 22, 2009), [available at http://www.economist.com/world/mideast-africa/displaystory.cfm?story_id=11792330].

Virtual Marketing Blog 2009: How to make Facebook applications. (accessed July 23, 2009), [available at http://www.virtualmarketingblog.com/how-to-make-facebook-applications/].

Webopedia Computer Dictionary 2009: Unique visitor. (accessed March 14, 2009), [available at http://www.webopedia.com/TERM/U/unique_visitor.html].

Wikipedia 2009: Facebook. (accessed March 22, 2009), [available at http://en.wikipedia.org/wiki/Facebook].

Wikipedia 2009: List of social networking websites. (accsessed March 14, 2009), [available at http://en.wikipedia.org/wiki/List_of_social_networking_websites].

Wikipedia 2009: Multimedia message service. (accessed April 04, 2009), [available at http://en.wikipedia.org/wiki/Short_message_service].

Wikipedia 2009: Social bookmarking. (accessed March 14, 2009), [available at http://hr.wikipedia.org/wiki/Social_bookmarking].
Wikipedia 2009: Short message service. (accessed April 04, 2009), [available at http://en.wikipedia.org/wiki/Short_message_service].
Wikipedia 2009: Web widget. (accessed April 17, 2009), [available at http://en.wikipedia.org/wiki/Web_widget].

Thoughts on the possibilities of empiric study of organisational culture and competences

Málovics János[1] *– Málovics Éva*[2]

In the special literature we can often read about the interaction of organisational culture and competences. Culture and competences are considered important parts of the conditions of organisational success by the management literature. In our opinion the mechanism of the impact between culture and competences has not been discovered or empirically completely proven as of yet. The aim of our research was the completion of the above mentioned gap, therefore we researched this topic in a pilot study, the results of which we present in this article.

1. Introduction

In professional literature you can often find thoughts about organisational culture having cardinal importance on the competences of organizations and their members. These factors define the successful survival of organisations, the perceived quality of service, and the satisfaction with this service. The empirical analysis of this coherence has not been performed yet and its theoretical cogency is also disputable.

Neither in the literature of culture or in the literature of competence has any consensus evolved about a definition either, however, the practical application of these two areas are being carried out without any agreement in the scientific part.

The leaders of organizations, practitioners, consultants and researchers are greatly interested both in the measurement and the modification of organisational culture, and in the development of competencies as well. Thus, the research about connections among the topics mentioned in the title is required from both theoretical and practical sides.

[1] János Málovics, Dél-alföldi Regionális Munkaügyi Központ (Szeged)
[2] Éva Málovics, University of Szeged, Faculty of Economics and Business Administration, Faculty of Business Studies (Szeged)

2. About organisational culture

Organisation researchers took over the concept of culture from anthropology and they have not succeeded to frame a consensus concerning its meaning and conceptualization as of yet. In this study we do not deal with the numerous interpretations and conceptualisations of culture, from our point of view, the integration approach, used commonly in business sciences, is important. Therefore, we will continue with a brief exposition of it.

The most complex concept of organisational culture in integration approach was formed by E. H. Schein, who defined culture as:
"(…) a pattern of basic assumptions – invented, discovered, or developed by a given group as it learns to cope with its problems of external adaption and internal integration – that has worked well enough to be considered valid and, therefore, to be taught to new members as the correct way to perceive, think, and feel in relation to those problems" (Schein 1985.)

Conclusions that can be drawn from the definitions: culture determines how members of an organisation perceive reality, how they think of it, and their feelings in connection with it.

According to this approach organisational culture is the source of consistency inside of an organisation and it also helps members to explain the inner characteristics and outer surroundings of the organisation homogeneously. Its main representatives are E. H. Schein, G. Hofstede, Deal and Kennedy. According to these approaches organisational culture encourages members to agree on what, why and how must be done, and it also helps to form the understanding and identity of the organisation. The culture works like an integrating mechanism that keeps the numerous variations of individual behaviours within bounds. The common part of these numerous definitions is that the culture is considered to be a system of assumptions, values and beliefs that are shared by the members of the organisation. Thus culture works like a certain kind of a social bonding material that holds theoretically discordant groups and individuals together since it restricts the enforcement of individual ambitions and interests. How narrow or wide the restrictions are actually depends on the characteristics of the culture.

According to this concept, the parts of culture are consistent, therefore culture is a source of harmony to a certain degree. The basis of this harmony is provided by the common norms and values. This approach considers it the role of the management to create and form culture. Empirical investigations often examine the values deemed important by the management, and consider these to be deterministic.

Representatives of this approach agree that culture is made up of different levels, the elements of a certain culture are integrated by premises and fundamental assumptions. Some of them are easily recognisable, they appear "on the surface" and can be studied. Others are hidden more deep in the culture, they can not be easily

identified, however, they are needed to understand the substance of culture. (Schein 1985). G. Hofstede's (1991) researches proved, that organisational culture can not be explained all by itself but only as a part of several force fields. Within the organisation, a leading role is played by the values and norms deemed important by the management. In other words, leadership culture plays a defining role, however, the organisation itself has reflexion on it.

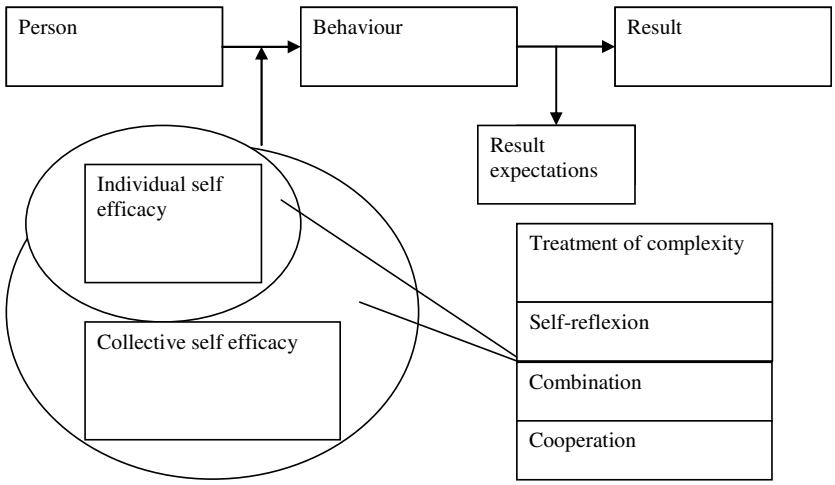
Deal and Kennedy (1982) emphasised the importance of strong cultures, stating that „culture is power". They consider culture the most important factor in the formation of commitment and enthusiasm. According to them managers have a task to plan and mediate the culture to the members of the organisation. According to this approach, the primary question is how managers can generate a feeling in their people that they are working on something rational and useful. In the process of generation, a key role is paid by attention paid to employees and the attainment of common values and norms.

3. The connection between culture and competences

Special literature has extended the notion of competence to several levels of the organisation, and there is also vivid discussion about it among the researchers. The topic of individual competences have been important for the management for a long time now, because they have undoubted impact on the performance of organisational members. The concept of competence was extended to social systems as well, and for example literature mentions organisational competences as the main factor of competitiveness. The subjects of organisational knowledge are of course organisational members and their competences, and the organisational culture (which includes norms, values, myths, stories). Several authors consider values, norms and shared mental models to be competences. (Schreyögg 2003)

In our study we would like to shed light on the reaction between organisational culture and competences. In order to achieve this, as a first step we applied a questionnaire of the German special literature, created by Uta Wilkens (2006) for the measurement of individual and collective competences. U. Wilkens and her partners operationalised the dimensions of the social cognitive theory according to figure 1:

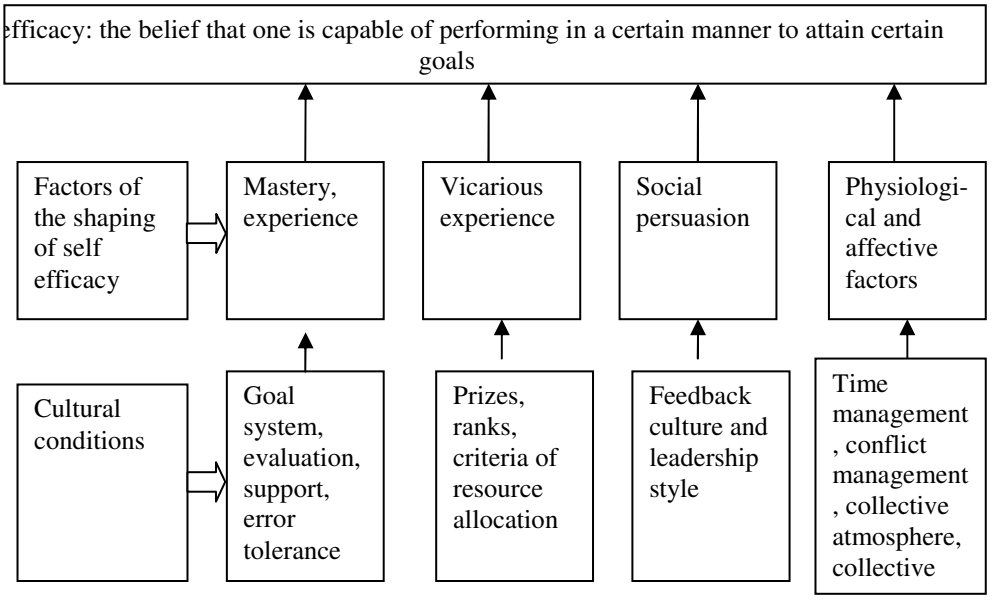
Figure 1. The dimensions of competence based on the social cognitive theory



Source: Wilkens et al. 2006.

In order to better grasp the reactions between organisational culture and competences, we applied A. Bandura's (2001) model, which has been well-confirmed by experiments. This model also contains the factors for improving individual and collective competences. Bandura assumes a strict connection between individual and collective self efficacy. The investigated relationship is shown in figure 2:

Figure 2. The effect-mechanisms of organisational culture that have an impact on efficacy



Source: Own creation based on Wilkens (2004)

The corresponding literature specifies several effects that the organisational culture has on organization members. We accept the idea that is rather widespread in German special literature which states that the efficacy described and experimentally well-confirmed by Bandura is a key competence - the high status of which is the base of individual competence. According to the connections described in the figure above, Bandura's well-confirmed efficacy-improving factors are derived from certain cultural dimensions.

The effect that has the most impact on efficacy should be those mastery experiences that we acquire at the workplace on a daily basis, the kind of which depend on the organisational culture. In the world of work, the base of mastery experiences is provided by the setting of obtainable, yet challenging goals; that is, if a compromise with leaders concerning goals and deadlines is common in the organization, and there is also a constructive feedback on the acquiring of these goals, and furthermore, creative errors are not sanctioned, then that presumably improves members' efficacy.

The other factors with a great impact on efficacy are vicarious learning, model application, and the experiences of social comparison, that is learning by observing others' behaviour and the corresponding consequences. A central category of learning by using a model is substitutional confirmation. This means that a

person, who has just observed that another person, who is in some way subject to the model, had been rewarded, will also learn that proper behaviour. (Bednorz-Schuster 2006.) This experimental result naturally raises the question that which model is to be adopted. In order to answer this, several studies have been carried out. Out of the assumptions summed up by Secord and Backman (1981), the following are important for our topic:

1. We select that particular person as a model, whose behaviour receives substitutional confirmation.
2. A person is chosen as a model because he himself hands out and checks rewards.
3. A person is chosen as a model because his social roles are deemed similar by the imitators to their own.

Bandura (1970) does not limit model learning only to cases where the observation of a particular behaviour actually occurs, but instead extends it to those learning forms where the individual meets behaviour samples in a symbolic way. "A great part of a culture's behaviour samples are supplied by learning from a model, which gains significant stability during adulthood. Naturally, learning by confirmation and by a model both play an important part in this process. (...) We do not only gain behaviour forms from the model, but norms and samples as well, which control our behaviour." (Bednorz-Schuster, 2006.)

The effect that discrepant practice has on model imitation was also examined. Those models that allowed themselves much freedom when it came to keeping themselves to norms, while still setting a rather high level of performance for others to follow, were barely imitated. An interesting experiment concerning model learning was carried out with nurses. (Bednorz-Schuster 2006.) The question of the study was if it is possible to enlengthen the interaction between nurses and patients by model behaviour.

In the psychiatric institute where the research has taken place, social interaction classes were introduced with the participation of both nurses and patients. During these classes table tennis and Monopoly could be played. Model behaviour was first caused by the presence of the two psychologists of the institute: 40% of the staff took part in the game. In case of the presence of the head nurse, this amount has risen to 70%. The same ratio was observed six weeks later. However, when no model took part in the games, the ratio dropped to 3%. Naturally we can clearly see that learning by model imitation is mixed with other factors. Participation while the head nurse is present could also be motivated by the fact that the nurses' rewards depend on the head nurse. In conclusion, if the behaviour of the leaders is considered to be a model by the employees, that has an effect on their efficacy.

The third influencing factor is verbal information sharing, which in our case is feedback culture within an organization, considered to be a part of the leading style. If regular and constructive feedback are present in the leading style, that has a positive effect on employees' efficacy.

The fourth influencing factor is the person's physical and emotional status, which, in our opinion, is in close relationship with collective atmosphere, time- and conflict-management within a work-organization.

4. Results of empirical researches

In order to measure the connection mentioned above we used the appropriate parts from two questionnaires. One questionnaire was created by U. Wilkens to measure individual competences, and the other one was our very own created on the basis of an SHL idea to measure culture developing competencies. Measuring instruments were tested at two faculties of the university with online questioning. The questionnaire was completed by 32 people, 12 men and 21 women.

The Cronbach - Alpha values of the items used int he questionaaire, which measure culture dimensions were between 0,7 and 0,9, thus their reliability can be considered suitable. The first part of the other instrument used for measurement was made up of the items already tested by R. Schwarzer in an international panel and were considered to be valid. These were used to measure individual self efficacy. The Cronbach – Alpha values of the items of the dimensions worked out by Wilkens turned out to be too low, which indicates reliability problems. The authors have experienced similar problems using the German sample. Results can be found in table 1.

Table 1. Cultural factors developing competencies and competence expectations

Faculties						
	1st Faculty		2nd Faculty		Total	
Cultural Factures	Mean	Deviation	Mean	Deviation	Mean	Deviation
Endowment	3,6	0,6	3,6	0,8	3,6	0,8
Delegation	3,5	0,7	3,8	0,8	3,7	0,7
Targeting	3,2	0,6	3,5	0,6	3,4	0,7
Aim access	3,4	0,6	3,5	1,0	3,4	0,9
Mentor-models	3,4	0,8	3,6	0,9	3,5	0,8
Feedback - culture	2,6	0,9	3,3	1,0	3,0	0,9
Group climate	3,4.	0,6	3,5	1,0	3,5	0,8
Competence expectations						
Self efficacy	3,0	0,5	3,0	0,6	3,0	0,6
Handling Complexity	2,8	0,4	2,7	0,6	2,7	0,5
Reflexivity	2,5	0,6	2,7	0,6	2,6	0,6
Combination	3,3	0,5	3,1	0,7	3,2	0,6
Cooperation	3,0	0,5	3,1	0,6	3,0	0,6

Source: own creation

Significant difference between the faculties can not be found either at the area of culture or competence dimensions. On the basis of the results it is presumable of the discriminancy analysis that there is difference between the faculties in performance-feedback. We have not found correlation between competencies and culture developing competencies.

5. Summary

In this study we attempted to exemine the competence developing impact of organisational culture in the areas of both theory and practice. Both categories are complex, disputable, and uneasy to measure – in fact, some authors consider them to be immeasurable. The development of competences happens as a result of reactions between the individual, the organisation, and society. In spite of this in special literature one can find the opinion that competencies of the members of the organisation can be developed. This is one of the managers' tasks and one of the

most effective instruments to achieve this is culture. In this study we have showcased the first steps of our research in this field as we looked for and tested theoretical models and research instruments. The question of the research can not be answered by the adoption of a sample of such size, however, it was realised that we have to change our instruments of measurement and we now plan to complete our following research using qualitative research methods.

References

Atkinson, R.S. 1998*:* Pszichológia. Osiris Kiadó, Budapest.

Baitsch, C. 1996: Kompetenz von Individuen, Gruppen und Organisationen - Psychologische Überlegungen zu einem Analyse- und Bewertungskonzept. In: Denisow, K. - Fricke, W. - stieler - Lorenz, B. (Hrsg.): *Partizipation und Produktivität – Zukunft der arbeit 5*, Bonn, 102-112. p.

Bandura, A. 2001: Social Cognitive Theory. An Agentic Perspective. *Annual Review Psychology*, 52, February, 1-26. p.

Bednorz, P. – Schuster, M. 2006: Bevezetés a tanulás lélektanába. Medicina Könyvkiadó Rt. Budapest.

Davies, B. – Allestair, P. – Warr, P. 1994: *CORPORATE CULTURE QUESTIONNAIRE,* Manual and User's Guide. SHL Hungary Kft.

Deal, T. - Kennedy, A. 1982: Corporate Cultures: The Rites and Rituals of Corporate Life. Rading, MA., Addison-Wesley.

Erpenbeck, J. (Hrsg.) 2004: Handbuch Kompetenzmessung in Unternehmen. Schäffer-Poeschel, Stuttgart.

Hofstede, G. 1991: *Cultures and Organisations: Software of the Mind.* McGraw-Hill, London.

Klieme, E. 2004: Was sind Kompetenzen und wie lassen sie sich messen? *Pedagogik 56*, 6, 10-13. p.

Secord, P. F. – Backman, C. W. 1972: Szociálpszichológia. Kossuth Könyvkiadó-Mezőgazdasági Könyvkiadó, Budapest.

Pawlowsky, P. - Menzel, D. - Wilkens, U. 2004: Wissens- und Kompetenzdiagnostik in Organisationen. In: Erpenbeck, J. (Hrsg.) 2003: *Handbuch Kompetenzmessung in Unternehmen*. Schäffer-Poeschel, Stuttgart.

Schein, E. H. 1985: *Organisational Culture and Leadership.* Jossey-Bass Publishers, San Francisco - Washington – London.

Schreyögg, G. - Kliesch, M. 2003: Rahmenbedingungen für die Entwicklung Organisationaler Kompetenz. QUEM-Materialen, 48, Berlin.

Schwarzer, R. 2004: Psychologie der Gesundheitsverhaltens (3. ed.). Hogrefe, Göttingen.

Wilkens, U. 2004: Von der individuellen zur kollektiven Kompetenz? *Herbstworkshop der Kommission Personal*. 24/25. September 2004, Konstanz.

Wilkens, U. - Keller, H. - Schmette, M. 2006: Wirkungsbeziehungen zwischen Ebenen individueller und kollektiver Kompetenz. In von Schreyögg, G. – Conrad, P. (szerk.): *Managementforschung Band 16: Management von Kompetenz*. Gabler Verlag,Wiesbaden, 121-161. p.

How Can Gold Collar Workers Become Gold Medalists? The Most Important Factors of the Successful Consulting Policy

Harazin Piroska[1] *– Pádár Katalin*[2]

Due to the importance and specialty of the sector of services, there are many questions about this sector which should be answered so that it can be further developed. In our work, we made an attempt to define to what extent is the policy of consultants in accordance with the expectations of their present and potential clients, as the knowledge and satisfaction of these expectations are the most important factors of the successful consulting policy.

Having reviewed the theoretical background, we could outline a wide spectrum of the peculiarities of services and consultancy. Learning on these secondary information, we built up our primer research in which we collected qualitative (deep-interviews) and quantitative (questionnaires, conjoint analysis) information on the three main groups of actors of consultancy: consulting firms, present clients, potential clients (in our work: university students). Finally, with the analysis of the collected information and data, we were able to determine how gold collar workers can become gold medalists…

Keywords: gold collar workers, consultancy, consulting policy, expectations of present and potential customers, interview, questionnaire, conjoint analysis

1. Introduction

In these days in Hungary, especially in the west part of the country, more than 50 percent of the employees work in the sector of services. However, as previously this sector has received less attention because the industrial and agricultural sectors were of higher importance, there are still many questions to be answered.

Compared to products, services are more difficult to handle and deal with, therefore the problems in this sector are also difficult to solve. This is also true for the relatively new field of consulting services, especially for the essential but hardly understood field of financial services.

In our work, we tried to define to what extent is the policy of consultants in accordance with the expectations of their present and potential clients, as the knowl-

[1] Harazin Piroska, M.Sc. in Engineering and Management, PhD-Student, Budapest University of Technology and Economics, Department of Environmental Economics
[2] Pádár Katalin, M.Sc. in Engineering and Management

edge and satisfaction of these expectations are the most important factors of the successful consulting policy.

Having reviewed international and Hungarian literature, respectively previous case studies and projects, we could outline a wide spectrum of the peculiarities of services and consultancy. Learning on these previously mentioned pieces of secondary information, we built up our primer research in which we collected qualitative and quantitative information on the three main groups of actors of consultancy: consulting firms, present clients and potential clients (in our work: university students).

We asked the consulting companies by using deep-dive interviews (11 interviews), examined the present (77 persons) and potential (120 persons) clients with the help of questionnaires, and with the potential clients we also implemented a conjoint analysis (354 participants). A year later we continued the conjoint analysis (creating a time-series), adding an extra sample of 193 people to it, thus extending the whole sample to 547 persons this way.

As we analyzed the market of consultancy from the viewpoint of all the actors, we had an opportunity to define the relationship between the policy of consulting firms and the expectations of customers. Finally, with the analysis of the collected information and data, we were able to determine how gold collar workers can become gold medalists.

2. Theoretical overview

Kelley (1985) regards the gold collar workers as the third group of the employees, beyond white and blue collar workers. The gold collar workers, who promote the use and transmission of knowledge in the various fields of life, nowadays also have to face difficulties. “Intelligent, independent, and innovative, these employees are incredibly valuable. They are lawyers and computer programmers, stock analysts and community planners, editors and engineers. They are as distinct from their less skilled white-collar counterparts – bank tellers, bookkeepers, clerks, and other business functionaries – as they are from blue-collar laborers.” Kelley (1985)

The goal of our research was to find out to what extent is the policy of consultants in accordance with the expectations of the current and potential customers, as the knowledge and satisfaction of these expectations are the most important factors of the successful consulting policy.

Based on our review of the international and Hungarian literature, respectively previous case studies and projects, we could outline a wide spectrum of the peculiarities of services and consultancy.

By the twentieth and twenty-first centuries, the sector of services has become more important and the number of employees in the sector has also increased. This fact is supported by various statistical data, for example that in the developed coun-

tries roughly 60 percent of the GDP comes from this sector and this figure is 50 percent even in the developing countries (Aharoni 1993). This is an evidence of why it is important to examine this sector also answering many questions and solving difficulties.

The first and most important one of these questions is how the expression of services can be explicable. According to Kotler et. al (2004) "a service is any act or performance that one party can offer to another that is essentially intangible and does not result in the ownership of anything. Its production may or may not be tied to a physical product." Even though many different definitions are available, there are some features of services which are common in the different perspectives. These are the following ones (according to Grönroos 2001, Veres 1998): heterogenity, intangibility, perishability or inability to inventory, and inseparability. These features are all also true for the expertise-based service of consultancy.

There is also no exact definition for the concept of consultancy, but this is one of the purest services, which is used in various fields of life. According to Biech (2006) consultancy is a process where a person or a company can help other people or companies in reaching results and success. Comments frequently show themselves regarding the concept which lack the realization of consultancy's importance and are merely trite phrases due to the fact that this is one of the fields with the least amount of research work about and it also lacks authentic publications on the topic yet. Canback (1998, 1999)

The consultancy process starts with the appearance of a problem on the side of the client/customer which is followed by the contact making of the parties. The relationship between the consultant and the customer is permanent. This had been a one-way relationship but as the economic environment has changed it became a two-way relationship where the most peculiar factors are communication and trust. Without the close relationship these services cannot be supplied. Consequently, the consultants have to win new customers and also have to maintain a good relationship with them just as with their old clients. The features of the close relationship are trust, availability and customer-knowledge (Nádor 2005).

Due to the special features of services, the related marketing activity is also a really interesting question. According to Veres (1998) the principle of HIPI expresses the main features of the marketing regarding services: heterogenity → H, intangibility → I, perishability → P, inseparability → I. In the case of services, the marketing mix is completed by three other factors, which result in the 7Ps of the services: Product, Price, Place, Promotion, People, Physical evidence and Process (Veres 1998 and Grönroos 2001). As it was mentioned before, consultancy is a special type of services, thus the marketing features of services also apply to it. According to Gyöngyösy (2000) and Szabó (1993), the main marketing activities of the consulting sector are the continuous seek for clients, conviction, service and retention. Nádor (2001) says that the relationship should be in the center of the marketing activity and the key factor is the exact definition of the target audience. Reputation,

image and the available references are the most important factors in international companies' marketing strategy and for the Hungarian firms and clients there is also emphasis on personal relationships and relationship management. Gyöngyösy (2000) also mentions that in order to reach success the consultants have to make a continuous, well-planned marketing strategy and Kubr (1993) claims that this strategy can only be successful if the demand of the clients is in the center. Despite this fact, real-life practice shows that the demand of the customer is not in the center of the marketing strategy, but reputation, fame and image are the mostly communicated elements of consulting firms. This strategy can inspire the clients to choose their consultant by right of fame, reputation instead of their actual demand.

Summarizing the specialties of the consultancy market, Poór (2000) cites Kéri (1971 and Ladó (1986), stating that the evolution of consultancy as a vocation was not an unbroken process in Hungary. As of today, concerning the number of consultancy firms and consultants, the Hungarian market has approached the respective figures of developed countries – there are 3000 consultants for 10 million inhabitants with a number of consultancy firms over 300. (Jakó 2004) Having a look at today's economic situation, one should not exclude the possibility that the consultancy market will grow narrower due to restrictions. Even though these restrictions might also strengthen the demand for consultancy, as amidst less venture-friendly conditions the role of efficiency-improving actions will be valorized. (Kiss-Kornai 2007) According to a FEACO (Fédération Européenne des Associations de Conseils en Organisation – European Federation of Management Consultancies Associations) study, which presents the financial results of 2005 and the guidelines for 2006 of the European leadership consultancy sector, the consultancy market showed growth in 2005 and forecasts for 2006 are also positive. In 2005 the revenue of the sector (50000 firms, half a million employees) was €61.6 billion. In respect of 2004, revenues showed a 14% growth rate. This growth can be the result of a general smartening process throughout the whole consultancy market and/or that of a new analyzing method, which regroups the leadership consultancy sector. Using the genuine grouping, the growth rate is still 8 %. (FEACO Survey of the European Management Consultancy Market (2005-2006))

3. Hypothesis

With the help of the theoretical overview we defined our hypothesis, and our main aim was to confirm this hypothesis with our primary research. In our hypothesis we assert that reputation (fame) is not the primary factor when the clients choose consultants. Having defined our hypothesis, we state that the expertise of the consultant is the product of the consulting service, therefore this is the first, essential factor when the clients choose consultants. Accordingly we do not want to prove this fact,

we focus on finding further aspects which are able to influence the clients' decision making.

4. The Structure of the Empirical Research

To confirm our hypothesis, we designed and implemented an empirical research where we tried to examine the opinion of all of the actors: consulting companies, present clients and potential clients (university students). We asked the consulting companies by using deep-dive interviews (11 interviews), examined the present (77 persons) and potential (120 persons) clients with the help of questionnaires, and with potential customers we also implemented a conjoint analysis (354 participants). In a later phase of the research, a year later, we continued the conjoint analysis (creating a time-series), adding an extra sample of 193 people to it, thus extending the whole sample to 547 persons this way.

It is important to notice that the sample of our research was rather small and we never endeavored to representativeness, though it can serve as a base of a more representative research (like a pilot study).

4.1. Consultants

To reach our aim, first of all we examined the consultants, the representatives of consulting companies, who actually provide the service. We collected information from them by using deep-dive interviews (qualitative data), and we prepared 4 longer and 7 shorter interviews. . In the case of the 4 longer interviews we asked 2 of the biggest financial consultants in the Hungarian consultancy market, and we also asked 2 auditors (accountants), who were also financial consultants. The 7 shorter interviews were prepared during a workshop. The representatives of four human resource consulting companies, two information technology consulting companies and one management consulting company were asked by interviews.

For the deep-dive, revealing interviews we used a list of questions which had derived from our former secondary research. Therefore these were well-structured interviews, but as Kvale (2005) said interviews allow interviewees to show a face of their own 'lived reality' regarding the topic throughout the flow of the conversation. During our research we have spent quite some time with the interviewees and this way we also had the chance to receive accommodating and friendly answers also to those questions which emerged during the flows of the interviews.

An important question from the viewpoint of our hypothesis applied to the opinion of consultants on the marketing strategy their company uses. Every representative mentioned the references which can help clients to find and choose consultants. But no one can say that this is a marketing strategy… They also listed traditional tools of advertising (i.e. newspapers, flyers) but it was also mentioned that

these solutions are not effective. Consultants think that creating relations in person is the best possible marketing activity.

With help of the interviews we had an opportunity to get to know numerous things from consultants. One of these was about the factors of success. According to the consultants, cooperation and appropriate communication with the clients, availability, flexibility, quality of service and professional knowledge are the success-factors. Leaning on this result, we thought that all these can be the critical factors which can outrun the reputation in rank.

The most important question regarding our hypothesis asked about the opinion of consultants on the claims, requirements of clients. As a result of the interviews, the aspects which can influence the decision of clients are professional knowledge, expertise; customer policy; trust; availability; customer-oriented behavior; close consultant-client relationship. Reputation was also mentioned by the consultants but with lower importance.

Subsequently, according to the consultants, reputation is not the base of decision making for customers; thus, as a summary, we can make the conclusion that we are able to prove the hypothesis.

4.2. Present Clients

As it was mentioned already, we asked the present clients with the help of a questionnaire. 77 questionnaires were filled by present clients, who have already used consulting services. We believe that their experience enables them to give relevant and useful answers about the right and maybe the possibly best consulting service.

Hoffman et al. (2000) point out that while using a questionnaire one can size up the respondent's opinion on and knowledge on a given question or topic together with his future goals and intentions; and besides all these the query provides objective data which are not or hardly accessible anyhow else. Accordingly, we have chosen to use this method and used questionnaires. We have tried to integrate all our questions into the questionnaire in a way that promised the most valuable and useful answers. After several pilot versions we have reached the final one, ending up completely agreeing with Scipione (1994) who claims that the design of a questionnaire is more of art than science.

Our final questionnaire basically consisted of closed-ended questions but in order to get to this final version we had also used two previous questionnaires which mainly contained open-ended questions. We used these open-ended questionnaires to verify the actual importance of those factors which we considered to be the relevant ones. Furthermore, we also wanted to collect additional important factors.

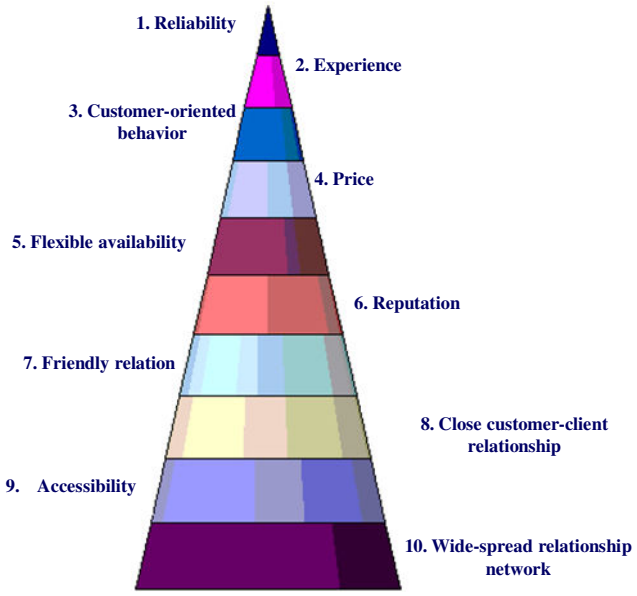
It is important to mention some data about the clients who filled the questionnaires. 55 percent of the clients were individuals, and the 12 percent of the rest represented a huge company (more than 1000 employees) and 49 percent represented a small company (1-20 employees). 55 percent of the sample has used the service for

more than 5 years. We were also interested in what kind of consulting services are used by these clients. Financial, insurance-related and taxation consultancy appeared to be the most popular services. We also asked the clients how and from where they can get information about the consulting services. The answers show that according to the present clients, reference is the best way to choose a consulting company (the second most often mentioned way was online advertisements).

From the viewpoint of our hypotheses, the first question we would like to highlight was an open-ended one where we were interested in the claims, requirements that come from the clients' side. There were many different answers (25 different features), but the most popular one was professional knowledge, expertise which was mentioned 43 times and this high number proves our supposition that the expertise of the consultant is the first, essential factor when the clients choose consultants. Professional knowledge was followed by reliability and preciseness. Reputation was also mentioned by the clients as a claim, but it was not over-emphasized, only 5 out of 77 clients mentioned it as an important requirement in decision making. (It is interesting to mention that reputation was mentioned only 5 times which is much lower than the 43 times how often professional knowledge was mentioned, and this is not a negligible fact if we want to prove our hypothesis.) Therefore with this question we managed to prove our hypothesis and we have also found new factors which are important for clients, and which can be the determining factors of the successful consulting policy.

For a more exact answer in connection with the most important factors, a further, really important question was a close-ended one. Here we asked the clients to make a gradiation, rank between the different, pre-defined factors (experience, reliability, price, customer-oriented behavior, reputation, widespread relationship network, friendly relation with the client, close consultant-client relationship, flexible availability, accessibility), which influence their decision making. The results of this question are important, as it will serve as a base for a later comparison between the opinions of present and potential clients. The outcome of the ranking can be seen on Figure 1, where reliability is in the first place and then comes experience. Customer-oriented behavior was ranked third. Reputation appears only in the sixth place.

Figure 1. Preference-based Order of the Examined Factors - Present Clients



Source: own creation

Summarizing the results of all the questions (both the open and close ended ones) of the questionnaire, we proved our hypothesis; that is to say the present clients do not choose consultants according to their reputation.

4.3. Potential Clients

Besides the present clients, as it was mentioned earlier, we also studied the potential clients; at first with the help of questionnaires, in line with the examination of the present clients. In this case it is important to raise attention to the fact that our sample consisted of university students, as we believe that they are the ones who might use consultancy services in the future. We have received answers from 120 individuals, the average age of who were 22.5 years (50% females, 50% males). It was a great advantage that students of the Budapest University of Technology and Economics were easy to reach.

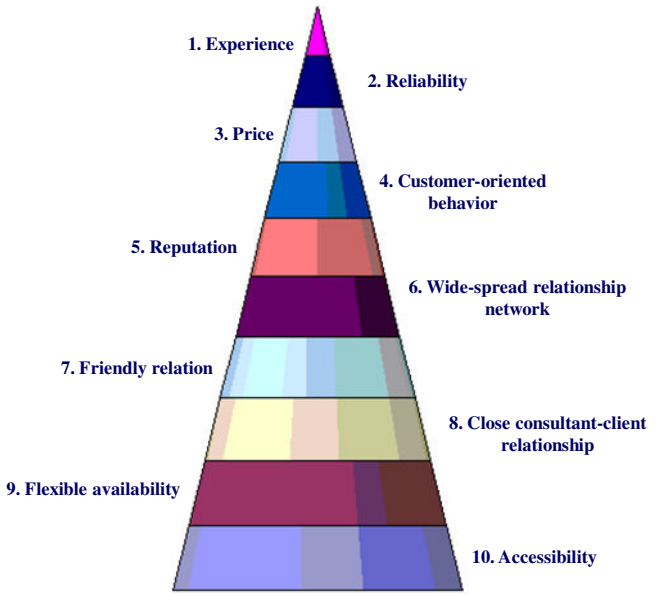
4.3.1. Questionnaire

120 people filled the final questionnaire at the end and the two most important questions are highlighted hereinafter.

In one of our open-ended questions we asked the participants to list those claimed, awaited features that should describe a consultant according to them. Among the 39 different answers reliability (76 times), experience (46 times), price (19) and reputation (16) were the most often mentioned factors.

In one of our close-ended questions the participants were asked to rank those ten factors that we have already examined among present clients. The results can be seen on Figure 2.

Figure 2. Preference-based Order of the Examined Factors - Potential Clients



Source: own creation

Experience turned out to be the most important factor and it was followed by reliability. (This was the other way around in the case of the above-mentioned open-ended question.) Besides these, price and customer-oriented behavior belong to the most determining factors. Reputation, which is important regarding our hypothesis, was only ranked as fifth regarding importance. (Figure 2)

There was a slight difference in the results of the open and closed-ended questions therefore we decided on the further examination of the potential clients. We used the method of conjoint analysis in order to screen out possible conscious and/or unconscious torsions in the participants' opinions.

4.3.2. Conjoint analysis

Conjoint analysis is an advanced, relatively new (developed since the 1970s) market research technique which gets under the skin of how people make decisions and what they really value in products and services. When using the conjoint analysis, results get really close to outcomes of real-life decision-making situations. The great advantage of conjoint analysis is that it does not directly ask one's opinion on certain features but the customers (clients) are asked to make a decision regarding product variants, in our case service variants. The principle behind is the following: products or services have to be broken down into their constituent parts then the possible combinations of these parts should be tested to look at what customers prefer. In practice, participants had to make their own preference-based order of the given service variants, where each variant consisted of a certain combination of the previously defined constituent parts. (Each variant was presented on a separate card.) This way the analysis examines several features together, which decreases the possibility of torsions in the participants' opinions – especially compared to the results when questionnaires are used.

At first, the features or constituent parts, which are to be examined, should be collected to let the analysis be carried out. As an outcome, the analysis will determine these factors' relative importance in decision making. Besides collecting the features, different levels of each of them have to be defined. Based on our previous results, we examined five features: consultant-client relationship; price; reputation; flexible availability; relationship and communication with the client. All the five factors and their levels can be seen in Chart 1.

Table 1. Factors and Levels – Conjoint Analysis

Factors & Levels	**Consultant-client relationship**	**Price**	**Reputation**	**Flexible availability**	**Relationship & communication with the client**
Low	formal consultant-client relationship, keeping off	free of charge service	HINT Consulting Group	defined „opening hours"	basic, necessary communication
Medium	-	1.000-2.000 HUF per occasion	-	-	-
High	Equal actors, alternate trust	20.000-40.000 HUF per occasion	BROKERNET	unlimited availability	relationship maintanance & communication with sufficient frequency

Source: own creation

In the case of the feature 'reputation', for instance, we had defined two levels, each of which was presented by firms; namely, the HINT Consultant Group and Brokernet. (Chart 1) The HINT Consulting Group is a non-existing, fake company which we made up only in order to make it sure that it is an unknown one for all of the participants. Brokernet is a well-known, Hungarian financial consulting firm, which was conducting an intense self-advertising campaign shortly before our research, therefore it was assumable to be one of the most well-known companies.

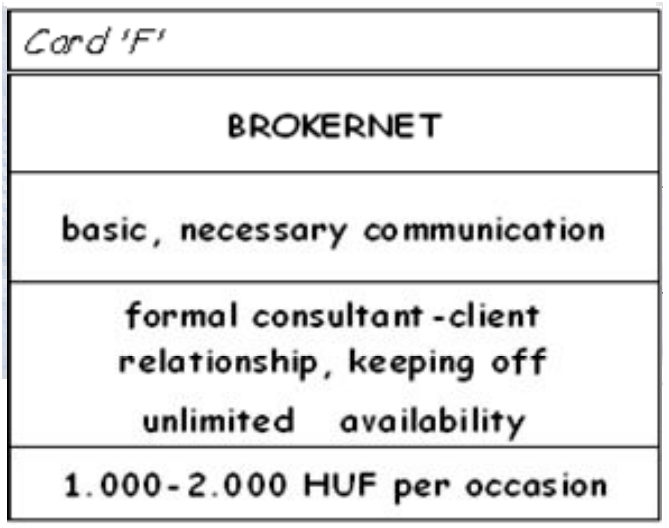
If one has a look at the previous chart (Chart 1) and multiplies the number of levels for each of the factors (2*3*2*2*2), it turns out to be 48, which means that there can be 48 different service variants in case we use all the possible combinations of the levels.

Obviously, this number had to be reduced, as it is almost impossible to make a decision on or rank 48 different variants – especially not within a relatively short time. Thus with the help of SPSS 16.0 we reduced the number of variants to 8, and we made 8 cards to present each of these service variants separately.

All the cards looked similar to the one on Figure 3, which shows Card F. This card, for instance, presented the service of Brokernet, which costs 1.000-2.000 HUF per occasion with unlimited availability. But the relationship between the consultant and the client is formal with only a basic amount of communication between them.

The task of the participants was to rank all the 8 cards according to their preferences – starting with the most preferred one.

Figure 3. An Example of the Conjoint Cards – Card F

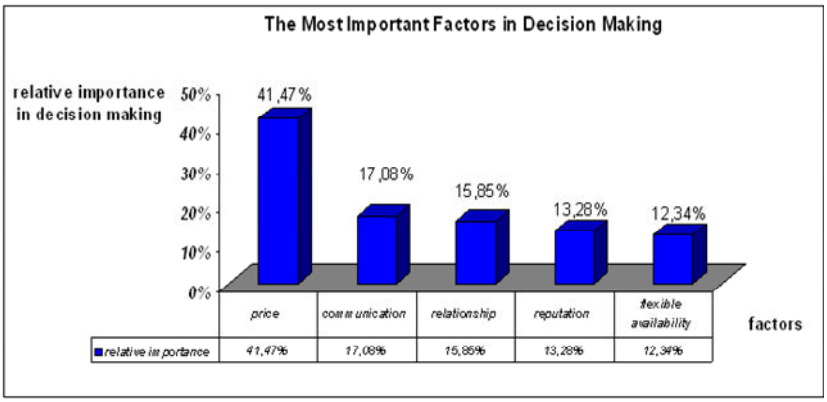


Source: own creation

We have collected answers in two waves (113 + 241), as the first steps of a possible longitudinal research. Altogether we could analyze the answers of 354 participants; the results can be seen on Figure 4. (A year later as a third wave of the sampling, 193 new potential clients' opinion was added to the sample. The results of the total (547 participants) sample affirmed the previous results.) Similarly to the case of the questionnaires, here we have also asked students of the Budapest University of Technology and Economics.

As the diagram (Figure 4) shows, price appeared to be the most important factor, which is not surprising knowing that many of the Hungarian consumers are price-sensitive. The factor reputation, which was important regarding our hypothesis, ended up to be the last but one in importance.

Figure 4. The Most Important Factors in Decision Making



Source: own creation

This way we proved our hypothesis, as it was confirmed by the results of the questionnaires and the conjoint analysis, thus we can state that reputation is not the most important factor in the clients' decision-making.

5. Comparison and overview of the results

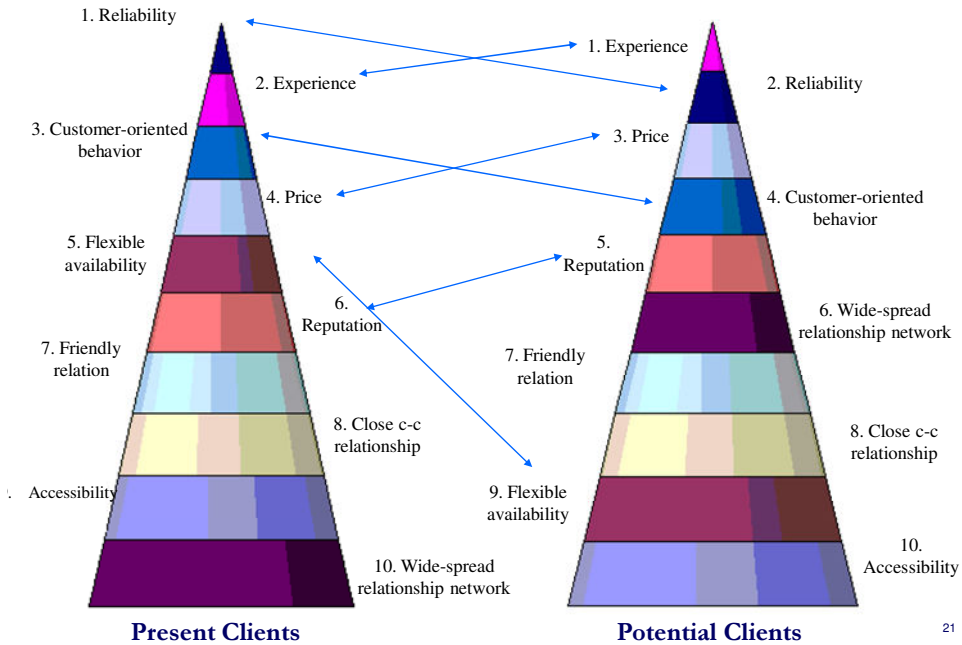
In order to reach a complete overview and better understanding of the results we have reached up until this point, we have found it necessary to compare the factors, which were said to be important by the consultants, to those that were said to be important by the clients.

Both consultants and clients were asked to list the awaited characteristics of a consultant. Expertise, customer-oriented behavior/"personalized service" and reputa-

tion were mentioned by both actors; however the clients also highlighted trust and price.

In the case of the clients we differentiated present and potential clients and examined their opinions separately. Their point of views have already been described separately (Figure 1, Figure 2) but in order to make a clear comparison it is useful to have a look at the results next to each other (Figure 5).

Figure 5. Preference-based Order of the Examined Factors Regarding Present and Potential Clients



Source: own creation

As it is visible on Figure 5, present clients ranked reliability first, while for the potential clients experience is the most important, right before reliability.

Thus our supposition that expertise is/should be basic and essential, appears to be proved and correct.

Having a look at Figure 5, one can also see that potential customers have found price more important than the customer-oriented behavior, which is exactly the other way around in the case of the present clients. The reason for this might be that the clients, who have already experienced, used the services of consultants, keeping their focus on the awaited successful outcome, find the customer-oriented behavior more important than the price. Furthermore, the fact that the price is very important for the potential customers is in line with the results of the conjoint analy-

sis, as there the price turned out to be the relatively most important factor – out of the five examined ones.

It also has to be mentioned that there is a significant difference in the relative importance of flexible availability if one wants to compare present and potential clients. It turned out to be noticeably more important for the present clients, but knowing that the sample of the potential clients consisted of university students who generally have a relatively flexible schedule, especially compared to the present clients who work fixed hours during the day, this result is not surprising and is in line with the expectable real-life behavior.

The last thing to be highlighted regarding Figure 5 is that reputation is slightly more important for the potential clients than it is for the present ones. If we take into account that the potential clients have never had the chance to experience the service itself (i.e. the customer-oriented behavior and its possible importance), it is understandable that reputation seems to be a factor which can be relatively objectively observed and therefore can be used as an important criterion in decision making. (Please note that reputation was given medium importance by both groups of the clients – which confirms our hypothesis.)

6. Conclusion

As a conclusion, we claim that some factors are more important for clients than the reputation of the company they finally choose. We believe that by emphasizing these factors, the consulting companies can fill a market niche. Consequently, by using the appropriate niche-marketing strategy they can become more successful – or in other words the gold collar workers can become gold medalists.

References

Aharoni, Y. (ed.) 1993: *Coalitions and Competition: The Globalization of Professional Service.* Routledge, London.

Biech, E. 2006: *A tanácsadás világa.* VHE Kft. Pécs.

Canback, S. 1999: *The Logic of Management Consulting*. Journal of Management Consulting, 10, 2, 3-11. p.

FEACO Survey of the European Management Consultancy Market 2005-2006, FEACO.

Grönroos, Ch. 2001: *Service Management and Marketing, A Customer Relationship Management Approach.* (2nd edition) John Wiley & Sons, Ltd, Chichester.

Gyöngyösy, Z. 2000: *A tanácsadási szolgáltatás marketingje*. Marketing és Menedzsment 2000, 4, 22. p.

Hoffmann-Kozák-Veres 2000: *Piackutatás*: Műszaki Könyvkiadó, Budapest.

Jakó, M. 2004: *Tanácsadók és szervezeteik Magyarországon* Easy Herald online folyóirat 2004. április www.kompetencia.hu, letöltés: 2007. október.

Kelley, E. R. 1985: *The Gold Collar Worker: Harnessing the Brainpower of the New Workforce.* Reading Mass, Addison-Wesley, Massachusetts.

Kéri J. 1971: *Az iparszervezetek és vezetés néhány elméleti és gyakorlati problémája Magyarországon a két világháború között*. (Kandidátusi értekezés.)

Kiss G. – Kornai G.: *A tanácsadóktól várják a megoldást a cégek a megszorításokra* Menedzsment fórum 2007. 03. 26, www.mfor.hu, letöltés: 2007. október.

Kotler, Ph. – Armstrong G. – Saunders J. – Wong V. 2004: *Principles of Marketing* (4th European Edition) Prentice Hall Europe, Essex.

Kubr, M. 1993: *Vezetési Tanácsadás*. Építésügyi Tájékoztatási Központ Kft., Budapest.

Kvale, S. 2005: *Az interjú – Bevezetés a kvalitatív kutatás interjútechnikáiba* Jószöveg Műhely Kiadó 2005, Budapest.

Ladó L. 1986: *Szervezéselmélet és –módszertan. A vezetés szervezési funkciója*. Közgazdasági és Jogi Könyvkiadó, Budapest.

Nádor, É. 2001: *Markáns változások a tanácsadói piacon – avagy marketingtanácsok a vezetési tanácsadóknak.* Marketing és Menedzsment 2001, 4, 68. p.

Nádor, É. 2005: *Professzionálisok-e az ügyfélkapcsolatok a professzionális tanácsadás piacán?* 15. Országos Marketing Konferencia Sopron.

Poór József (szerk.) [2000]: *Menedzsment tanácsadási kézikönyv* KJK-KERSZÖV Jogi és Üzleti kiadó KFT. Budapest.

Scipione, Paul A. 1994: *A piackutatás gyakorlata* Springer Hungarica Kiadó Kft. Budapest.

Szabó, K. 1993: *A tanácsadó cégek marketing-tevékenysége.* Ipar és Gazdaság 1993. május, 19-27 p..

Veres, Z. 1998: *Szolgáltatásmarketing*. KJK-KERSZÖV Jogi és Üzleti Kiadó Kft., Budapest.

Multi-Project Management

Nicole Bircks[1]

Multi-project management (MPM) is a highly contemporary issue for the conceptual design of an improved (multi-) project organization. It is based less on a uniform standard but rather on a variety of ideas for improved and integrated project work. Existing approaches vary in relation to their areas of application, consistency, and operability through which the creation of a uniform multi-project management design only partially exists (Grübler 2004).

Organizations are now facing the challenge to realize diverse and complex projects such as the introduction of a quality management, the implementation of multiple valuations and surveys of personnel, finance controlling, IT projects, and process optimizations.

Multi-project management can be the concept utilized to respond to these challenges. The purpose of this work is to demonstrate how this concept must be designed and used to support the project work and how it can result in an increase in efficiency and significant cost savings within organizations.

1. Introduction

Nowadays, one of the biggest challenges in project management is finding an answer to the question of how one can realize a variety of projects in an efficient and timely manner. The problem is that this challenge is hampered by many external and internal factors. Resources are not always available, we face increased market competition, both at a national and international level, and the pressure to perform as well as the cost factors increasing steadily.

Now, the question arises what can we do to counteract these problems? And finding an answer to that question is exactly what this work is about. The goal was to design a concept of multi-project management which is applicable, marketable and most notably efficient. In order to accomplish this goal, the project management itself was contemplated holistically, and the practical applicability of the concept was considered as stringent. Due to the fact that the complexity of individual projects has been tremendously increased during the last years, it is logical

[1] Nicole Bircks, Master of Business Adminstration,University of St. Thomas, USA, Diplom-Betriebswirtin, University of Applied Sciences, Trier, Germany

consequence that companies above a certain size must use a standardized methodology which will ensure a superior steering, coordination, and control of their projects. And such a holistic view of project work can be achieved by the implementation of multi-project management which should be clarified in the course of this work.

2. Problem Statement

This work is concerned, as the title of the work implies, with the problems of how one can integrate multi-project management, and of how one can make the concept itself workable. This involves answering the question of where should the multi-project management be situated in the organizational hierarchy (capacity for the integration) and on what components should the conceptual shape/structure of MPM be composed so that this concept "works" (functional capability).

The present study deals therefore with an analysis of appropriate tasks to be processed by the multi-project management. Both, the individual project approach and its associated issues as well as the basic sphere of activity of the multi-project management to the existing organizational units - in terms of the aspired integration - play an important role, because the sphere of activity of the MPM should help to eliminate, or at least to decimate the problems which may/definitely arise within the individual project management. Furthermore, these tasks must be set in a way which ensures that the MPM itself will be functional. As there are almost none identical-looking statements in the literature with regard to the task distribution of the MPM, this is one of the problems which has to be addressed in this work (Kunz 2007).

Another problem which has to be addressed arises from the distinction made in the literature about the definition of roles within the multi-project management. The description of roles is widely represented in the literature. However, due to the fact that different authors have different opinions about the role description, the problem arises that there is no consensus about what role has to be responsible for what task(s). This in turn leads to a subjective and oftentimes illegitimate allocation of tasks which does not guarantee an efficient workflow within the MPM. Similarly, the allocation of names of the respective roles is often different, so that there is again a contrary formation of the various players. But a common understanding about the allocation of roles between the project- and multi-project- initiative is very important, because it makes a substantial contribution to the functioning of the MPM. Only if there is a clearly defined role description, the interplay of all multi-project management parties can be guaranteed (Lomnitz 2004). Therefore, the most important roles and a description of their duties will be stated in this work.

3. Disambiguation and Definition of MPM

In the literature multi-project management is defined amongst others as “an temporal indefinite management tool that performs the implementation of all conceptual and informational activities as well as the coordination of duties which support the management and control processes of the entire project portfolio, thus creating favorable conditions for the successful implementation of individual projects." (Kwasniok 2007, 18.)

Accordingly, the individual project approach is explicitly included in this definition as it has to correspond with the overall multi-project management initiative in order to achieve the desired steering and control effects. Only this holistic view of the importance of individual projects in conjunction with the entirety of all projects enables an efficient and effective multi-project management.

Building on that, this work describes the term multi-project management as a holistic approach that improves the project work of a company. This approach is built on the basic tasks of the multi-project management, involves a clear division of roles and allows a (permanent) improvement or development of the multi-project management itself. This means that the MPM can already be conceptually designed by some basic conditions which again can be enriched through company specific innovations and adaptations. Thus, the multi-project management is understood as a concept that can adapt to the dynamic realities of our time and of the company, and is therefore able to make significant contributions to an improved, holistic project work.

Finally, multi-project management can be therefore defined as "a concept which allows through a detailed consideration of each individual project, a superior planning, steering and controlling of all project that are done within a company, which in return accelerates the efficiency and the effectiveness of the project management itself."

4. Problem Analysis and Method of Resolution

As already mentioned in the problem statement, it is important to identify individual problems that may arise within the individual project management in order to deduce the fields of activity of the multi-project management. It should be noted that the individual project management is in fact a subject matter of MPM. However, the individual tasks which arise in the actual project management do not fall naturally into the remit of MPM. This means that the tasks of multi-project management are generally stemmed from the (individual) project management problems. However,

there must be a draconian separation of the subject areas of the project- and the multi-project-initiative, because multi-project management should contribute to an increased efficiency and effectiveness of the project work. But these improvements can only be achieved when the multi-project management starts, where the individual project management reaches its limits or when problems arise which require a higher-level of control. Both, multi-project management and project management are complementary. However, the areas of activity but must be considered in isolation insofar as the multi-project management will never be engaged in the actual tasks of the project management. This statement will be justified in the course of this work as this mutual interplay between the project- and the multi-project management initiative will be clearly described and demonstrated.

4.1 Analysis of the Allocation of Tasks of the MPM

The following table shows in the first column individual problems which may arise within the project management. The second column states the necessary MPM measures which will help eliminating the presented problems of the project management and the last column includes the name of the respective MPM division responsible for the correct implementation of the individual tasks.

Table1. Problem areas, MPM duties, and MPM divisions

Project Management Problems	**Duties of the MPM**	**Responsible MPM division**
Project selection	• Implementation of an objective project selection process • Project evaluation on the basis of this approach	**MPM-Planning**
Overrun of budgets	• Implementation of an approach which calculates uniformly the project costs per project • Monitoring the compliance of this process • Calculating the overall project costs of the company • Determining the total project budget in consultation with the company management	
Disadvanteous allocation of project resources	• Determination of clearly defined decision-making authority with regard to resource allocation • Implementation and execution of a superior capacity planning	
Prioritization of projects	• Implementing a project portfolio management • Execution of the project portfolio management in terms of a superior steering and controlling of all projects • Evaluation and visualization of project interdependencies	**MPM-Prioritization**
Missing/No risk analysis (and its impact on the project landscape)	• Implementing a risk management • Safeguarding the adequate execution of the risk management	
No control of the project landscape	• Execution of regular controls of the entirety of all projects through the deposit and usage of relevant data	
No control of the progress of individual projects	• Implementing standardized reporting processes • Controlling the existence and completeness of the respective report	**MPM-Control**
No efficiency control	• Implementation and execution of an efficiency control for each individual project	
Missing/No decision making power	• Setting its own decision-making authority in consultation with the company management	**MPM-Steering**
Lack of know-how transfer	• Implementing an enterprise knowledge management	
No information supply	• Implementation of an appropriate information system (MPM tool) that enables the acquisition of data	
Lack of skill among the project participants	• Staff training in order to make it clear how to use the new documents and the MPM-tool	

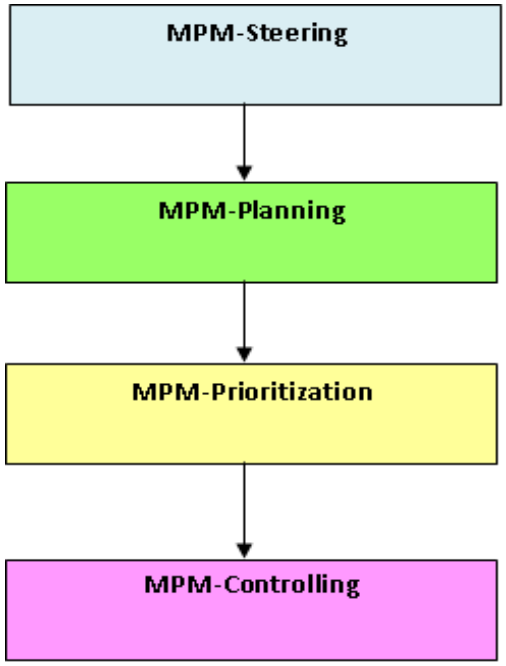
Source: Nicole Bircks, 2009, in dependence on *Kunz*, Strategisches Multi-Projektmangement, 2007

4.2 Interplay of the MPM-Divisions

The interplay of the different areas within the multi-project management can be chronologically presented as follows:

1. In order to implement and to build the MPM-department the MPM-steering comes to the fore of needed actions.
2. Afterwards the necessary prerequisites of the MPM-prioritization are created through the MPM-planning.
3. The MPM-prioritization ensures the efficiency and effectiveness of the project portfolio and is supported by the deposited reports.
4. The MPM-control then ensures that the completed projects will get a final check, so that the "circle" of an efficient and effective project work is closed.

Figure 1. Functional MPM-Divisions



Source: own creation

5. Analysis of the Role Allocation of the MPM

In order to ensure a common understanding regarding the respective actors which occur in conjunction with the multi-project management, a description of the respective roles must be made. The role description is an important aspect within the multi-project management theme as an unclear definition of roles can lead to conflicts and reconciliation measures, which in turn can lead to a waste of time and energy (Adler-Sedalczek 2005). Therefore, the question of what tasks and responsibilities the individual project team members should have, must be answered, because the interplay of all MPM players can only work if there is a clear delineation of all individual roles. Therefore, the individual tasks of the multi-project participants will be shortly described afterwards.

5.1 Duties of the Top Management within the MPM

Even with the introduction of a multi-project management, the top management is in their power to make decisions and is not restricted in giving instructions in any way. The company management ultimately decides which projects will be implemented and with what priority. Likewise, they can decide the postponement of larger projects (Lomnitz 2004). They also retain the final decision-making on continuation, suspension and removal of individual projects. As the multi-project management in direct contact with the top management, and as they can inform the top management of the current state of each project any time, the duty of the MPM incorporates giving advice and assistance to the company management.

5.2 Duties of the Multi-Project Executive Committee (MPEC)

This body has, unlike a "normal" steering committee, not the task of making a decision about a single project or program, but rather adjudicating upon priorities within the project portfolio and deciding on them. Their duties and competences should therefore include:

- Final decision on whether or not a project will generally be included in the project portfolio
- Prioritization of project proposals by dividing projects into project classes (priority A, B, C)
- Determining the composition of the (normal) steering committee for approved projects
- Determining project managers for respective projects
- Approval of budgets (the project portfolio budget/total budget will be approved by top management; the approval of certain amounts within the already approved budget is a duty of the MPEC)

- Superior approval of resources for project portfolios
- Decision-making in the case of resource shortages between projects, if problems cannot be solved by the MPM
- Conflict resolution among various projects, if they cannot be resolved by the MPM
- Supporting the implementation and development of the MPM

The relationship between the multi-project management and the MPEC is self-explanatory: The MPM provides the MPEC with relevant data/information so that this committee is able to make decisions.

5.3 Duties of the MPM-Manager

In the literature, tasks of MPM-managers are described very differently. Sometimes they are limited on only a few duties, such as the ongoing coordination and steering of the project portfolios, or they are presented in a way so far-reaching that the multi-project manager is responsible for almost all functions of the MPM-department.

At this point it is therefore proposed that the responsibilities of the multi-project managers are generally determined according to enterprise size, respectively to the number of projects to be coordinated. This will mean that in smaller companies, which want to control holistically a few projects, the scope of tasks of the multi-project manager has to be bigger, whereas in medium or large businesses his/her duties has to be restricted, especially because it has to be expected that in such companies the MPM department consists of a larger number of other employees who can take over some/many tasks for the multi-project manager. In those companies the multi-project manager therefore acts as a head of the department and has to ensure that its employees are properly managed. Among his duties as a supervisor he has to "ensure the achievement of the given and agreed objectives and has to foster the motivation of employees and ensure the preservation of groups led by them."(Olfert 2004, 37.) So he certainly has the power which legitimized him to place orders and instructions.

But finding an answer to the question of what tasks he has to take over within the multi-project management itself is obviously of greater importance. Therefore, it is essential to ensure that the range of tasks assigned to him, is designed in a realistic and achievable dimension.

Notwithstanding the literature, this work sees the multi-project manager as a person who especially has to ensure the smooth running of his department. Specifically, this means that all areas of responsibility which have been determined in consultation with senior management must be initiated by the multi-project manager. Thereby, it should be irrelevant whether the multi-project manager does

some of these tasks personally or delegates them to his staff. Focus should therefore be on a successful implementation of the required sphere of activities for which the multi-project manager is held responsible. The multi-project manager is therefore responsible for both, the success or non-success of the introduction of the multi-project management as well as for the performance of work after the implementation.

To get this working, first of all the rules about the decision-making authority should be defined in collaboration with the company management as well as with the project organization. This task must be done personally by the multi-project manager, because only if these basic conditions are set out, including the "limits" for each area, the multi-project manager can start his work.

Ideally, the multi-project management should consist of the previously mentioned organizational units: the MPM-planning, -steering, -prioritization, and -control. Through this differentiation of organizational units it is obvious that the multi-project manager should not be charged with a variety of specific tasks, but rather has to ensure that these tasks will be done properly (by his staff).

However, there are some duties that the multi-project manager has to carry out in person. First, he must personally gather important information and needed decisions from its own employees, from employees of the project organization, from senior management as well as from the multi-project executive committee. Second, he has to mediate in conflict situations, for example between members of the MPEC, or between project participants, which will be mainly achieved by the neutrality and objectivity of his position. Furthermore, he also has to identify existing problems between different projects, to analyze the relationships within the project portfolio, and to prepare relevant information, so that the management or the MPEC can use them as a useful decision memo (Adler-Sedalczek 2005).

It is apparent that the multi-project manager should be seen as a person, who accumulates all of the departmental information. He is the coordinator of the project work and enables the interplay between management, multi-project executive committee, project organization, and the MPM department itself. He is the one who creates with the help of his staff a well-functioning multi-project management, provides all necessary information and establishes all standards for the further development of the MPM. He is: the initiator, coordinator and decision maker within the multi-project management.

5.4 Duties of the MPM Personnel and Project Office Staff

The areas of responsibility of employees who work in multi-project management department are not described in the literature, except those who work in a so-called project office or the department of project management. The reason for this is probably because there is no concept of multi-project management as it is described

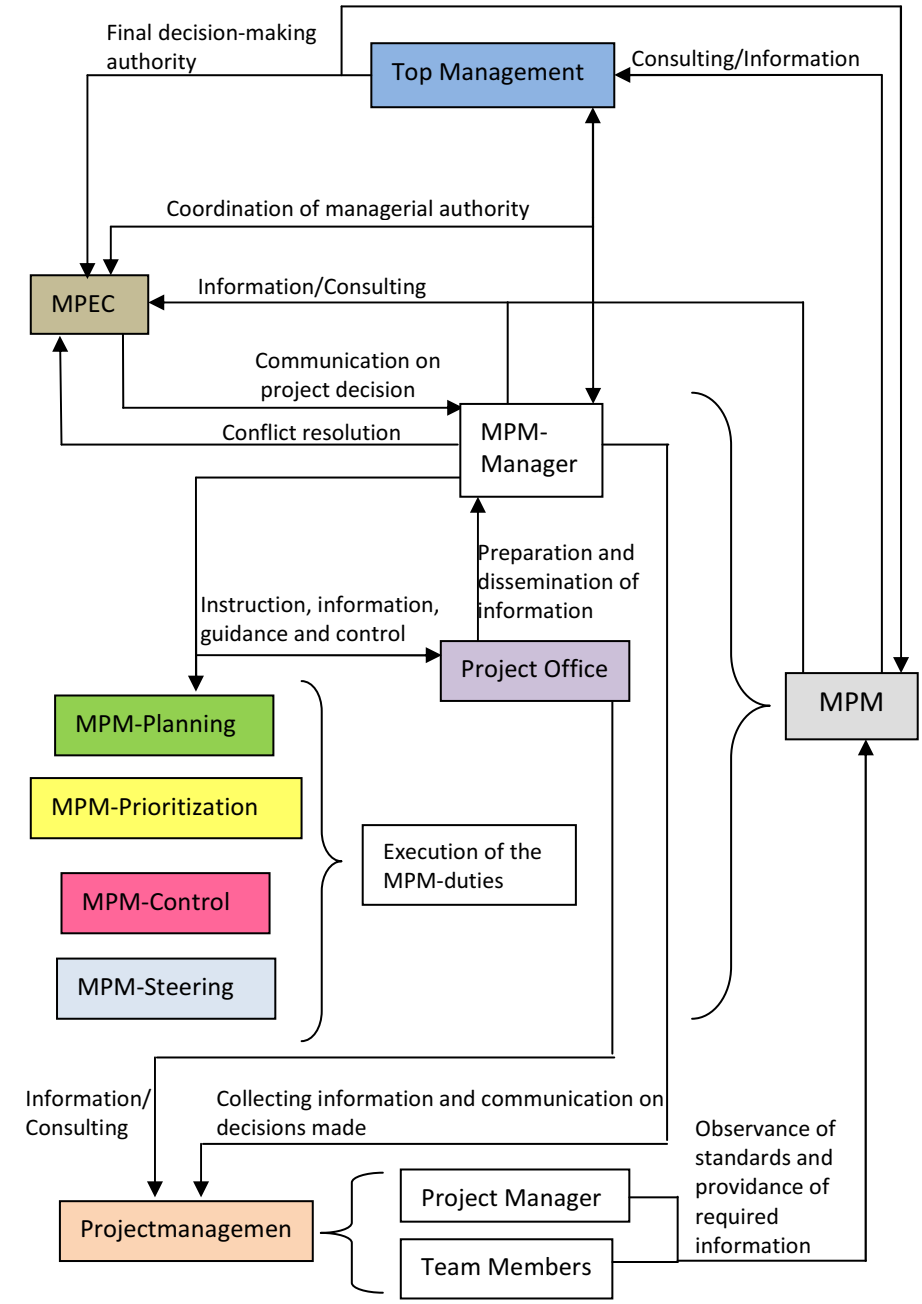
in this work. The duties of those employees who work in the MPM-department result from the previously described scope of tasks in the table. For example, those employees who work in the MPM-planning are obligated to take over the tasks of the division "MPM-planning", and those who are work in the MPM-prioritization have to take over the duties described therein, etc.

The project office, however, should also be seen as a permanent facility within the multi-project management department, which takes over the administrative activities for this area. This office should on the one hand ease the workload of the multi-project manager through the providence of necessary information, such as those which the multi-project manager intends to submit to the MPEC as a basis for their decision making. On the other hand, the project office should support the project manager as well as the project staff by providing assistance to project applicants in formulating their requests, or by answering questions regarding to the use and recruitment of data of the multi-project management tool. The project office can therefore be seen as a service unit that provides project information such as the current status of project applications and provides information about the current stand of a project. Furthermore, another administrative task of the project office is the provision and maintenance of a project management manual (Pfetzig-Rhode 2001).

5.5 Results of the Role Description

The main roles of the multi-project management and their respective duties have been described in the previous section. It is apparent that multi-project management builds on a close collaboration between the actors and that their effectiveness can only thrive if each person is familiar with the tasks assigned to him/her, edits them properly and knows the boundaries of each work area. The following figure will demonstrate the interplay between the different roles and their duties:

Figure 2. Interface of MPM Participants



Source: own creation

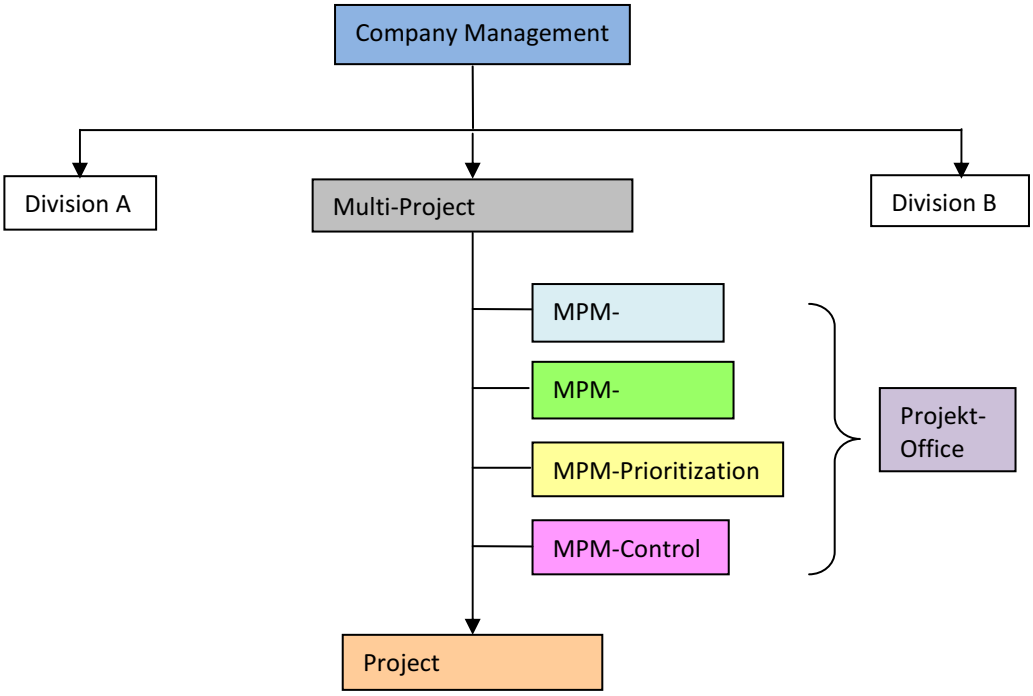
6. Implementation of Multi-Project Management

Since the main problems like defining the term of multi-project management, determining the duties of this department as well as the definition of roles has been described, it is essential to develop a solution for the introduction of a multi-project management.

6.1 Organizational Structure

To make it possible that the multi-project management can act as a separate division, it is important to embed this unit in the organizational hierarchy of a company. Since the multi-project management is a complex management, it must be positioned in the organizational structure in a way that a cross-project planning and steering can be ensured. The MPM department should therefore be located directly under the company's management (Lomnitz 2004). The company's organizational chart can be pictured as follows:

Figure 3. Organizational Structure



Source: own creation

6.2 Operational Structuring

Generally the operational organization is concerned with the design of work processes; in that case with those that occur within the multi-project management. To successfully implement the multi-project management, it is important to establish a chronology of the described tasks. It should be noted that in this work only the one-time tasks will be positioned in a chronological order as they the basic prerequisite for the implementation of a well-functioning MPM.

In the description of the duties it has already been mentioned that the first step for the implementation of a multi-project management is "the definition of the decision-making authority". This determination shall refer to the decision-making powers between the multi-project management, the project organization, and the line instances, as well as the related determination of the managerial authority over resource allocation.

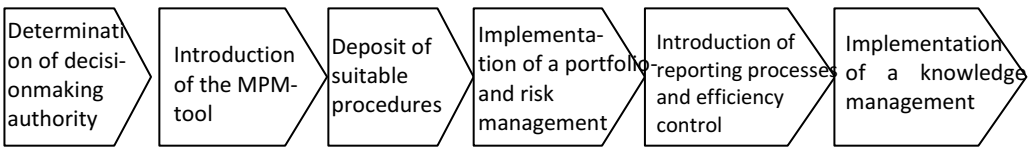
A next step would be the establishment of an appropriate multi-project-management tool, as the provision of such a tool enables the overall planning, steering and control of the MPM. In this tool the required, standardized forms for the project work will be deposited. Afterwards, appropriate procedures for the project selection, the capacity planning, and for the project cost evaluation must be deposited.

Thereafter, the project portfolio management and the risk management have to be introduced. In order to guarantee the multi-project management control, standardized reporting processes will be deposited and an efficiency control for each project will be introduced.

A final step would then be to establish a multi-project knowledge management, so that the project-relevant knowledge and experiences for future (multi-) project work will be available. Thus, the cycle to ensure efficient and effective project work will be closed.

Summarized the sequence of activities which are necessary in order to implement a multi-project management can be pictured as follows:

Figure 4. Operational Structuring



Source: own creation

7. Conceptual Implementation of the Sphere of Activities

This section deals with a detailed description of problem areas that need to be processed by the multi-project management. This description refers to the one-time tasks which need to be implemented in order to design a well-functioning MPM.

As the determination of the decision-making power as well as the introduction of a multi-project management tools, should depend on the individual notion of each company, these two fields of activity will not be described in this work.

The description of the scope of duties takes place after the designation of the respective, corresponding organizational unit within the MPM.

7.1 MPM-Planning

The division "MPM-planning" is responsible for all activities that are associated with an effective and efficient project selection. This means that this area introduces all measures and covers areas of responsibility that allow an objective selection of projects.

7.1.1 Deposit of Standardized Forms and Establishment of a Project Cost Evaluation System

The deposit of a uniform template management is the first performance of this division. The forms that need to be uniformly deposited are:

- The project proposal
- The authorization letter
- The rejection letter

Furthermore, it is the responsibility of this division to deposit a method for a uniform project cost evaluation. This involves the deposit of a procedure that can determine the project cost of a single project. Generally, one distinguishes two main categories: the global cost estimation procedures and the analytical costs estimation procedures (Patzak-Rattay 1998).

The difference between the two methods lies in the accuracy of their statements: the analytical methods are much more detailed and therefore more accurate, since they are based on individual work packages of each project. The project manager evaluates together with the respective responsible person the costs for each work package. The individual costs per package will then be aggregated to the total project costs. The cost estimation procedures, however, relate as the name suggests on estimates. In order to determine the cost of the project this procedure uses appropriate benchmarks or adequate metrics so that the total project costs can be estimated.

As at this point the multi-project manager just needs to get an overview of the possible incoming costs, it is suggested to use the cost estimation procedure in order

to save time and therefore money. It is just about getting an answer to the question of whether or not the project will be incorporated into the project portfolio which can be answered through this basic evaluation process.
In order to evaluate the project costs there are at least three factors which should be considered:

- "Transfer pricing for the internal project team
- Material costs, e.g. hardware, travelling, equipment, etc.
- Costs for external consultants" (Lomnitz 2004, 117.)

To avoid "uncertainty" by doing those calculations, it is still advisable to not only calculate one cost estimate, but to make an estimate by an optimistic, realistic and pessimistic value (Patzak-Rattay 1998). Through this distinction the project manager has on the one hand greater "elbowroom" because he does not need to refer his project proposal to a certain-size budget cost. On the other hand, the decision-makers can consider whether or not they are willing to take into account even the "worst scenario", the pessimistic value. It is apparent that the project costs can roughly be determined by this procedure and that the decision makers are able to make a decision based on the provided estimates. The duty of the multi-project management now is to ensure that this procedure for the project cost estimation is uniformly applied by the project managers and team members.

7.1.2 Implementation of a Standardized Project Selection Procedure

The selection process should review the project requests in terms of their "project appraisal", their "project classification", including the "project attractiveness", and their "project potential".

In the case of the project appraisal one determines whether or not the future project plans are generally projects in their true sense. Thereby, the individual project characteristics will be proved and separated from line activities. The project classification refers to a classification in terms of "must-do-projects", "shall-do-projects", and "can-do-projects".

Must-do projects are those projects that must be carried out, e.g. because of legal and technological constraints. They have therefore a supremacy over all other projects. Shall-do projects are determined after the determination of the must-do projects and are related to strategic and/or economic criteria, such as projects with the highest economic benefit or with the highest cost savings. Can-do projects are classified as those projects that are deferred from the current plan period. They will be conducted only if the necessary resources for their implementation are (additionally) available (Lomnitz 2004).

As the must-projects are defined by their necessity, only the remaining shall-do and can-do projects have to be proved in terms of their "project attractiveness". The project attractiveness should mainly appeal to the strategy compliance and the economic benefits of the respective projects. The attractiveness of the individual

projects will be visualized by a portfolio analysis. Thereby, the strategy conformity as well as the profitability of a particular project should be related to each other.
This classification of projects in the project portfolio may lead to additionally generated must-do projects. These kind of must-do projects do not result from a legal or technological necessity, but from a highly profitable contribution of the project itself. Therefore, these projects "should" also be in the category "must-do projects", and thus necessarily implemented.

The final step is identifying the "project potential". The identification of the project potential is already closely related with the project portfolio prioritization (MPM-prioritization), as the project potential is the company's ability to provide the required operational resources (resources and budget), and the general willingness to implement this project at all.

The general willingness to do a certain project is incumbent upon the multi-project executive committee and ultimately the company management. However, the multi-project management should ensure that the MPEC and the top management are able to make well-founded decisions. Therefore, the MPM has to provide sound documents and helpful information. Therefore, the MPM-planning must make two actions in the context of determining the project potential:

1. The determination of the overall project budget per period
2. The establishment of a comprehensive capacity planning

The determination of the overall project budget can be done through using three methods: "top-down", "buttom-up", and the "interative approach", whereas the iterative approach is the one recommended in here.

The budget planning based on the iterative process can be described as follows: The top management suggests a project budget according to their financial capabilities and their willingness to invest. The MPM-planning also suggests on her part a total project budget based on the aggregated project costs made by project managers in their project proposals. Subsequently, a vote takes place between the multi-project manager and senior management about the "actual" amount of the overall project budget. The multi-project management will underlie its proposals through detailed information, whereas the senior management will make the final decision about the amount of the project budget.

The last duty of the MPM-planning is now the deposit of an overall capacity planning. The capacity planning shall ensure that all necessary human resources for individual projects are available; respectively that in case of insufficient resources this scarcity is identified early. This aspect also plays a crucial role in the subsequent prioritization of projects by the MPM-prioritization.

The capacity planning should also be built on the iterative approach described above. This time, the MPM has to find an agreement between the line management and the respective project leaders. In terms of a top-down approach the line management must suggest how many employees can work in projects. The project applicants in turn must estimate in advance how many employees are needed for

their project. The task of the multi-project management is to add up the number of employees across all projects, so that the required "overall number of employees" for project work is known.

The coordination of human resources is designed a little more complex than the previously described project budgeting, since the potential project team members must be determined from all divisions. This means that each department needs to make a proposal on how many people could be working in projects.

The task of the multi-project management is then to collect the proposals made by line managers, to balance the available staff and the required number of staff, and to make, if necessary, appropriate adjustments.

7.2 MPM-Prioritization

The MPM-prioritization has basically the task to prioritize the (new) projects set by the MPM-planning. That means that this division has to make a ranking for project implementations, then set them into a project portfolio and visualizes the interdependencies within the portfolio, even to existing projects.

7.2.1 Implementation of Risk Management

Another task of this division is the introduction of a risk management. The risk management shall ensure that individual project risks will be estimated and evaluated before a project will be realized. The MPM-prioritization has to deposit a form that allows a standardized risk assessment for each project. The project managers are responsible for evaluating the risks of their projects on the basis of normative criteria. The filled-out risk checklist should then be forwarded to the MPM-prioritization, so that this division can make a clear risk presentation of all projects. The result shows how many projects have a high, medium and low risk (A, B, C), so that it can be decided across all projects which projects should/could be incorporated into the project portfolio.

Generally, there should be a balanced ratio of A, B, or C projects within the project portfolio. Because even though projects in risk category A (high risk) often promise the greatest opportunities, such as above-average sales increases or significant cost savings, they still contain the previously identified highest risk, which in turn can quickly transform the vaunted future "potential" in a future "disaster". If more than one A-project simultaneously does not keep its promise and many of the identified risks arise, the project landscape will be affected which in turn can lead to a dramatic situation within the whole company.

7.2.2. Implementation of Portfolio Management

First of all, it should be mentioned that the introduction of the project portfolio management is basically a static process that deals with the deposit of all required data. The execution of the project portfolio management, however, is a dynamic process and represents the main task of the MPM-prioritization.

The introduction is therefore static, because it creates the basic conditions to visualize the different project portfolios. These conditions may be regarded as the "planning" of the project landscape. The execution of the project portfolio management is a dynamic process, since it deals with the ongoing steering and controlling of the project portfolios. The project portfolio management shall ensure that the project landscape is permanently supervised in terms of interactions, inconsistencies, etc.

The overall objective of the project portfolio approach is the "clear presentation of all projects and their assessment based on comparable criteria." As the MPM-planning has determined the amount of the total project budget, the data of the general availability of capacity (human resources) as well as the expected cost per project, and has made a classification of the proposed projects, the MPM-prioritization should now design the project landscape. Therefore, it is important to bring together the various data of the MPM-planning as well as to take into account the data collected by the risk analysis in order to get an overview of all relevant data, which are important to the prioritization of projects within the project portfolio.

Before the MPM-prioritization pictures all data, the division must still set the likely cost of each project against the predetermined total project budget. Therefore, the MPM-prioritization puts the identified individual project budgets into a table while taking into account the individual project classification. It then distributes the overall budget to individual projects. So it can be determined for what projects the given overall project budget is sufficient, and what projects can therefore be included into the project portfolio. The continuous acquisition and processing of data has the dual advantage that the MPM-prioritization is always in a position to ensure the superior planning, steering and controlling of the project landscape. Secondly, the multi-project management is able to give advice on the totality of all projects to the company's management at any time.

Summarized one can say that through the introduction of the project portfolio management as well as through the introduction of the risk management, one could determine which (new) projects should be included into the project portfolio. Therefore, the static process is completed. Now the company knows what projects can and should be actually implemented and realized. These projects will now be placed in the project portfolio(s) and the planning process of the project portfolio management with respect to the totality of all projects is completed.

7.3 MPM-Control

The MPM-control deals first of all with a superior single project control, intending to ensure that the set objectives of the MPM-planning have been actually achieved by the completed project. In this regard, the division deposits and uses standardized forms and implements a generally applicable reporting process that must be used by the respective project leaders. Another task carried out by the MPM-control is executing a final efficiency control for each project. Here, the results of a project will be compared with what was originally intended, so as the actual improvement achieved by the project can be demonstrated.

Therefore one can say that the MPM-control is concerned with monitoring all currently running projects as well as with the final examination of all projects already undertaken. The multi-project management is therefore on the one hand always informed about possible drawbacks and is able to take appropriate countermeasures. On the other hand, the multi-project management is able to consider any interdependencies among projects, and thus is able to prevent that the course of individual projects will be jeopardized through others projects. Furthermore, the MPM-control is in a position to give advice and offer support to current or future project team members through the deposit of field reports and the usage of former project closing reports.

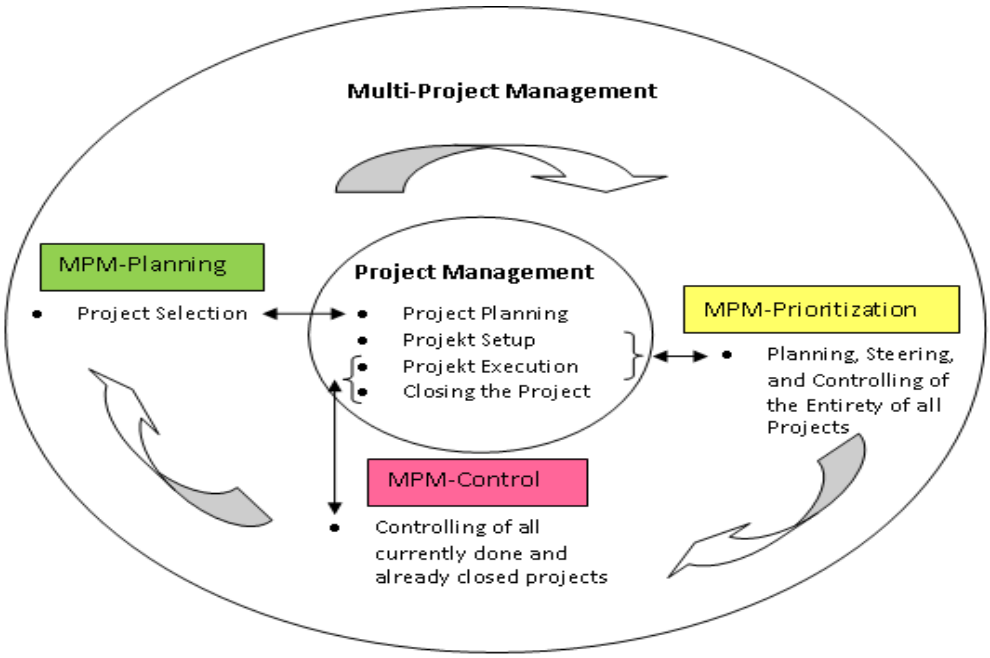
8. Conclusion

This work[2] has tried to demonstrate the various aspects of multi-project management and what activities must specifically be taken over. The aim has been to structure the concept of multi-project management in a way that the interaction of all stakeholders can be guaranteed in a logical and workable manner. The conceptual implementation began with the selection of projects which should be included in the project portfolio, and a standardized and appropriate selection process was deposited (MPM-planning). After the selection of potential projects had been completed, it was necessary to test their general feasibility and to set all the approved projects into the project portfolio. The permanent control and monitoring of the project landscape has been identified as the main task of this area, because of the associated complexity and variety of tasks (MPM-prioritization). Finally, it was pointed out

[2] One has to mention that this article is just a very small excerpt of the original work. The complexity and task identity can probably be better understood, if one reads/sees all corresponding processes, activities, tables, figures, etc. in detail.

how both, individual projects as well as the entirety of all projects have to be controlled. Through embedding the last division - the MPM-control - into the multi-project management framework, the cycle of "planning, steering, prioritizing and controlling" projects and the project landscape can be closed.

Figure 5. The MPM Cycle



Source: own creation

References

a) References mentioned in this paper:

Adler, Anna, Sedlaczek, Ralf (2005): Multi-Projektmanagement, Portfolioplanung und Portfoliocontrolling; in: Schott, Eric und Campana, Christophe (Hrsg.): Strategisches Projektmanagement, Springer-Verlag, Berlin Heidelberg, 2005

Grübler, Gerit (2004): Ganzheitliches Multiprojektmanagement – Mit einer Fallstudie in einem Konzern der Automobilzulieferindustrie, Dissertation,

Göttinger Wirtschaftsinformatik, 1. Auflage, Band 50, J. Biethahn, M. Schumann (Hrsg.), Göttingen, 2005

Kunz, Christian (2007): Strategisches Multiprojektmanagement – Konzept, Methoden und Strukturen, 2. Auflage, Deutscher Universitäts-Verlag, Wiesbaden, 2007

Kwasniok, Sascha (2007): Multiprojektmanagement – Organisation, Interdependenzen, Implementierung, VDM Verlang Dr. Müller e.K. und Lizenzgeber, Saarbrücken, 2007

Lomnitz, Gero (2004): Multiprojektmanagement – Projekte erfolgreich planen, vernetzen und steuern, Redline Wirtschaft, Frankfurt, 2004

Olfert, Klaus (2004): Personalwirtschaft – Kompakt Training, 4. Auflage, Friedrich Kiehl Verlag, Ludwigshafen (Rhein), 1999

Patzak, Gerold und Rattay, Günter (1998): Projekt Management – Leitfaden zum Management von Projekten, Projektportfolios und projektorientierten Unternehmen, 3. Auflage, Linde Verlag, Wien, 1998

Pfetzing, Karl und Rohde, Adolf (2001): Ganzheitliches Projektmanagement, 1. Auflage, Band 2, Verlag Dr. Götz Schmidt, Gießen, 2001

b) References mentioned in the original work:

Amelingmeyer, Jenny (2002): Wissensmanagement – Analyse und Gestaltung der Wissensbasis von Unternehmen, 2. Auflage, Deutscher Universitäts-Verlag GmbH, 2002

Dammer, Henning, Gemünden Hans-Georg, Lettl, Christopher (2006): Qualitätsdimension des Multiprojektmanagements, in: Zeitschrift für Führung und Organisation, 75. Jg., 2006, S. 148-154

Dobiéy, Dirk, Köplin, Thomas, Mach, Wolfram (2004): Programm-Management – Projekte übergreifend koordinieren und in die Unternehmensstrategie einbinden, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2004

Fiedler, Rudolf (2005): Controlling von Projekten, 3. Auflage, Vieweg Verlag, Wiesbaden, 2005

Glaschak, Stephan, A. (2006): Strategisches Multiprojektmanagement – Konzept, Unternehmensbefragung, Gestaltungsempfehlungen, Rainer Hampp Verlag, München und Mering, 2006

Hiller, Mark (2002): Multiprojektmanagement – Konzept zur Gestaltung, Regelung und Visualisierung einer Projektlandschaft, Dissertation, Lehrstuhl für Fertigungstechnik und Betriebsorganisation, Band 43, Prof. Dr.-Ing. G. Warnecke (Hrsg.), Kaiserslautern, 2002

Hirzel, Matthias, Kühn, Frank, Wollmann, Peter (2006): Projektportfolio-Managment, Strategisches und operatives Multi-Projektmanagement in der Praxis, 1. Auflage, Gabler Verlag, Wiesbaden, 2006

Herbolzheimer, Claus, W. (2005): Multiprojektmanagement, Personalentwicklung und Weiterbildung, Multiprojektmanagements als Promoter für die Entwicklung unternehmerischer Kompetenz von Mitarbeitern, Wissenschaftlicher Verlag Berlin, 2005

Kessler, Heinrich, Winkelhofer, Georg (2002): Projektmanagement – Leitfaden zur Steuerung und Führung von Projekten, 3. Auflage, Springer Verlag Berlin, Heidelberg, 2002

Litke, Hans-Dieter (2004): Projektmanagement – Methoden, Techniken, Verhaltensweisen, 4. Auflage, Carl Hanser Verlag, München, 2004

Lukesch, Christoph (2000): Umfassendes Projektportfoliomanagement in Dienstleistungskonzernen am Beispiel eines großen, international operierenden Versicherungsunternehmens, Universtitäts-Dissertation, ETH Zürich, 2000

Pradel, Michael und Südmeyer, Vera (1996): Multiprojektcontrolling, Planung des Projektportfolios bei Versicherungsunternehmen, in: Versicherungswirtschaft, 51. Jg., 1996, S. 1550-1555

Reschke, Hasso, Schelle, Heinz, Schnopp, Reinhardt (1989): Handbuch Projektmanagement, Band 2, Verlag TÜV Media GmbH, Köln, 1989

Rickert, Dirk (1995): Multiprojektmanagement in der industriellen Forschung und Entwicklung, Deutscher Universitäts-Verlag GmbH, Wiesbaden, 1995

Schindler, Martin (2001): Wissensmanagement in der Projektabwicklung, 2. Auflage, Josef Eul Verlag, Köln, 2001

Scheurer, Steffen (2000): Strategische Unternehmensentwicklung durch strategisches Multiprojektmanagement, in: Zeitschrift für Planung, 11. Jg., 4/2000, S. 379-409

Steinbuch, Pitter (2000): Projektorganisation und Projektmanagement – Moderne Organisation für Praxis und Studium, 2. Auflage, Friedrich Kiehl Verlag GmbH, Ludwigshafen (Rhein), 1998

Wöhe, Günter (2002): Einführung in die Allgemeine Betriebswirtschaftslehre, 21. Auflage, Verlag Franz Vahlen GmbH, München, 2002

An Analysis of Trends in Female Labor Force Participation in Turkey

Ceyda Özsoy [1]*- Sevilay Atlama*[2]

High female participation in the labor market implies both an advance in the relative economic and social position of women and also an increased utilization of human potential for economic development. The female labor force participation rates have increased substantially in developed countries. However, either many developing countries or Turkey the female labor force participation rates show a declining trend. In Turkey, from 1955 to 2009 labor force participation rate diminished from 72 percent to about 25 percent. This ratio is nearly one third of the rate for men in recent years. Additionally, Turkey has the lowest rate among OECD countries. Several factors can be taken to explain this trend. In this paper, we investigate the reasons for the declining trend in the female labor force in Turkey. Some of those: recent increases in enrollment rates at all levels of schooling delay entry into the labor market of younger, the changing composition of the labor force away from agriculture towards non-agricultural activities, cultural values against their participation in market work, their lack of education and marketable skills, unfavorable labor market conditions, huge wage gap and imparity promotion opportunities between female and male workers, exclusion from social security, and early retirement system.

Keywords: Female labor force, Female labor force participation rate, Turkey

1. Introduction

Before analyzing recent trends in the women's labor force participation rate, we need to define what we mean by female labor force participation rate. Female labor force participation rate indicates the ratio of the female labor force to the female population 15 year of age and over. Labor force refers to the share of the working age population who are either in a job or actively looking for one (Abhayaratna-Lattimore 2006).

[1] Ceyda Özsoy, PhD, Assistant Professor, Anadolu University, Faculty of Economics and Business Administration, Turkey
[2] Sevilay Atlama, Research Assistant, Anadolu University, Faculty of Economics and Business Administration, Turkey

The female labor force participation rates have increased considerably in the developed countries in recent years. In contrast, in many developing countries and in Turkey the female labor force participation rates show a declining trend (Tansel 2001). During the last decade, a number of studies have pointed to this decline and the strikingly low rates of labor force participation and have analyzed its consequences for both men and women (World Bank 1993, Özbay 1991 and 1994, Özar 1994, Ecevit 1998, Ecevit et al. 2000, Tansel 2001).

There are several reasons for the declining trends in the labor force participation rates in Turkey. Structural changing composition of the labor force away from agricultural towards nonagricultural activities and consequently rural-urban migration has rightly been considered as the main reasons for this decline. This is why the withdrawal of women from traditional agricultural activities in which they formerly engaged (Ecevit 2003). Second, recently, younger populations have been staying in school longer. This contributes to the declining trends in the participation rates of the young. Third, the early retirement scheme introduced in the early 1980s is another factor that contributed to the declining participation rates of the middle to older age groups (Tansel 2001). Besides, unemployment rates, educational attainments and skill levels, the "discouraged worker" effects, and prevailing cultural norms that exclude participation in market work may also play a role in the low participation rates of women in urban areas. In addition, the failure of official labor statistics to cover workers in the informal sectors needs to be taken into account when analyzing the patterns of the female labor force participation rate in urban areas where uneducated and unskilled women work in informal sectors (Palaz 2005). Women's personal and family characteristics include age, years of schooling and their marital status, number of children and children's age group also affect female participation rates (Dayıoğlu-Kasnakoğlu 1997).

In this paper, we focus on trends in the female labor force participation rate and examine the factors, which influence women's decision to enter the labor force. In addition to this analysis, we examine policies, which provide more and better work opportunities for women identified by Turkish State Planning Organization (SPO 2006) in 9th Development Plan.

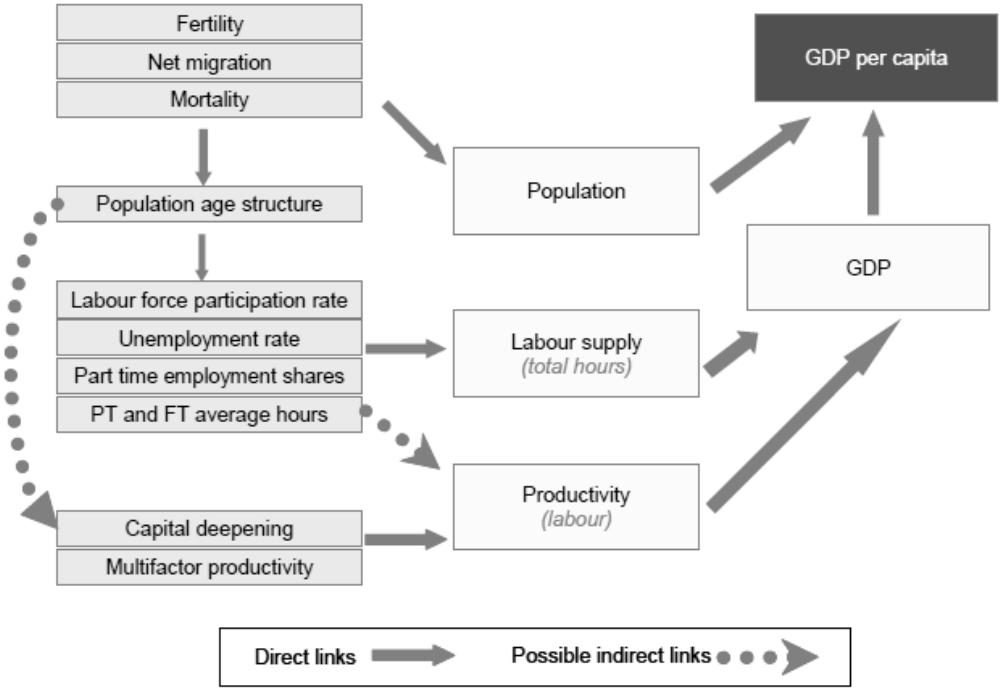
2. Why Does High Female Labor Force Participation Desirable?

A high female participation rate is desirable on several grounds. High female participation in the labor market implies both an advance in the relative economic and social position of women and also an increased utilization of human potential for economic growth and development (Kızılırmak 2005).

Higher female employment is instrumental in building capacity for economic growth and poverty reduction. Three key factors that impact on economic growth,

often referred to the three P's, are population, participation (or labor supply) and productivity (figure 1).

Figure 1. The 'Three Ps' of Economic Growth: Population, Participation and Productivity



Source: PC (Productivity Commission) 2005, Economic Implications of an Ageing Australia, Research Report, Canberra

The higher the number of people participating in the workforce, or the more hours worked, the higher the potential output produced and, assuming everything else unchanged, the higher the potential level of GDP per capita (Abhayaratna-Lattimore 2006). According to Argy (2005), joblessness (or non-participation) 'represents a big waste of national economic potential'.

But, economic growth is not the only reason why participation matters. Higher workforce participation can also reduce the fiscal pressures associated with providing welfare support and serve social inclusion and equity goals (OECD 2003a).

In addition to these outcomes, higher female employment increases labor supply, productivity and standard of living, reduce poverty among women and children. More and better jobs for women mean higher incomes and better life for them, their family and also society as a whole. Using women's talents and skills in

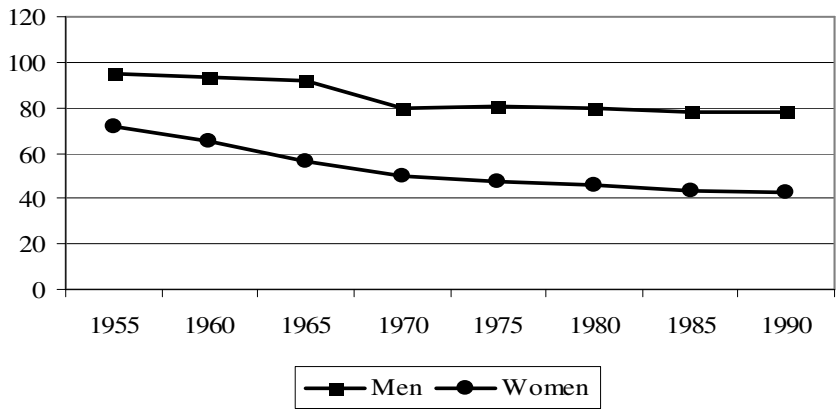
the labor market not only provides families with more economic independence, but also increases women's self confidence and social respectability. (SPO-World Bank 2009a).

At the meanwhile, higher levels of female employment allow government investments in education to be used more efficiently as women use their acquired talents productively in the economy. In countries with low levels of female employment, families often under-invest in girl's education. On the contrary, working women generally are more involved than non-working women in making decisions in relation to their children's education and health, which are externalities that positively affect the welfare of future generations (SPO-World Bank 2009b). As a last outcome of higher female employment is "gender equity".

3. The Trend of Female Labor Force Participation Rate in Turkey

According to the population census and household labor force survey results, overall labor force participation rates have been decreased gradually over the last 50 years in Turkey - from 95.4 percent in 1955 to 69.1 percent in 2009 among men, likewise from 72.0 percent in 1955 to 23.5 percent in 2009 among women (World Bank 1993). It can be easily seen on Figure 2 and Table 1, labor force participation rates are especially low level and diminishing nature both women and men.

Figure 2. Labor Force Participation Rates by Gender, Turkey (1955-1990)



Source: TURKSTAT, Statistical Indicators 1923-1990, Table 1-8, 38 (from Özbay, 1994)

Notes: The population census figures for the years 1955-1965 include population 15 year of age and over while for 1970-1990 they include population 12 years of age and over

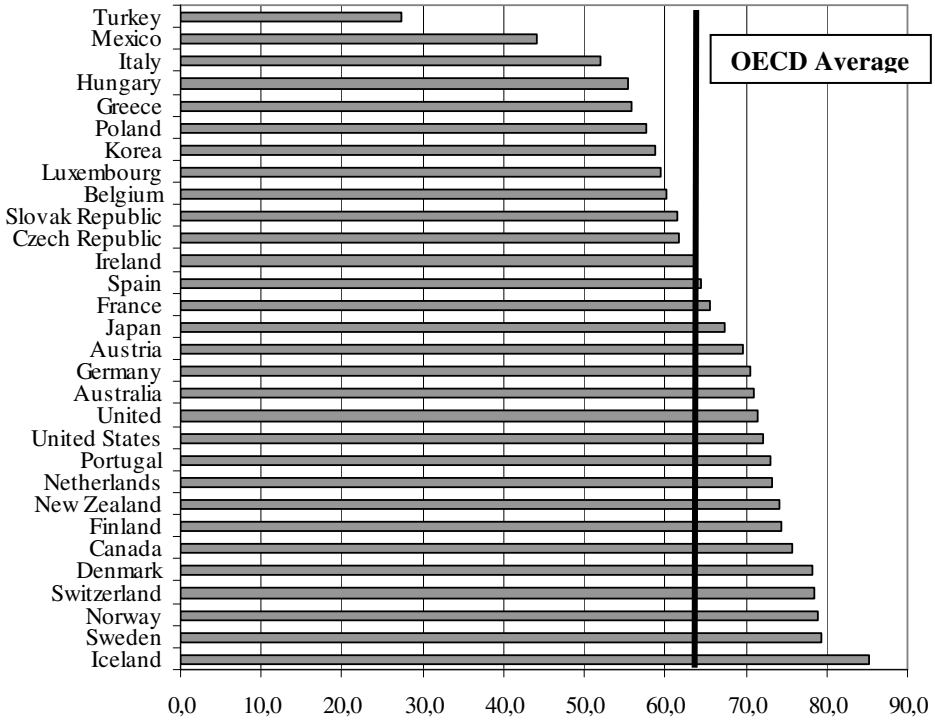
Table 1. Labor Force Participation Rates by Gender, Turkey (1991-2009)

Years	Total (%)	Men (%)	Women (%)	Years	Total (%)	Men (%)	Women (%)
1991	57.0	80.2	34.1	2002	49.6	71.6	27.9
1992	56.0	79.6	32.7	2003	48.3	70.4	26.6
1993	52.1	78.0	26.8	2004	48.7	72.3	25.4
1994	54.6	78.5	31.3	2005	48.3	72.2	24.8
1995	54.1	77.8	30.9	2006	48.0	71.5	24.9
1996	53.7	77.3	30.6	2007	47.8	71.3	24.8
1997	52.6	76,7	28.8	2008	46.9	70.1	24.5
1998	52.8	76.7	29.3	2009	45.9	69.1	23.5

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)
Note: Labor force participation rates include population 15 year of age and over

Figure 3 shows labor force participation rates among OECD countries. According to OECD Employment Report statistics, average labor force participation rate for OECD countries in 2008 was 63.2 percent for women. Iceland has the highest rate (85.4 percent), while Turkey has the lowest female labor force participation rate. In fact Turkey is only country stayed below 30 percent.

Figure 3. Female Force Participation Rates, OECD Countries (2008) (%)



Source: OECD Employment Report 2008

4. Factors Which Determine Low Participation to Labor Force of Women in Turkey

Table 2. Person Not in Labor Force by Reason and Gender, Turkey (15 years old and over), (Thousand)

Reasons for not being in labor force	1988				2008			
	Men	%	Women	%	Men	%	Women	%
Available for work but not seeking a job								
Discouraged	120	3,8	244	2,2	371	5,0	241	1,2
Other	77	2,5	178	1,6	460	6,2	778	4,0
Seasonal Worker	12	0,4	22	0,2	72	0,9	243	1,2
House wife	-	-	8.860	78,9	-	-	12.186	62,4
Student	998	31,9	564	0,5	2.086	28,0	1671	8,6
Retired	921	29,5	195	1,7	2.806	37,7	682	3,5
Disabled or ill	539	17,2	813	7,2	1.270	17,1	2.144	11,0
Family or Personal Reasons	-	-	-	-	107	1,4	1155	5,9
Other	458	13,7	353	3,1	268	3,6	427	2,2
Total	3.125	100	11.229	100	7.441	100	19.526	100

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

Table 2 indicates the reasons accounting for non-participation of women in labor force. According to the data, the most important reason not to enter the labor force for women is being a housewife (62.4 percent). Other reasons such as being a student, being retired or disabled and ill etc. have little affect on not entering the labor force. However for men, being retired or a student are the most important factors pertaining to why men do not join the labor force (Palaz 2005). While about 38 percent of males are out of labor force in their retired status, the proportion of women presently retired after having worked formally under a security scheme is extremely low (only 3,5 percent in 2008). Another striking imbalance is the fact that while 8,6 percent of women remain out of labor force due to student position, 28 percent for males.

What are the reasons for female labor force participation rates are low level and diminishing nature in Turkey? It can be explain to this question under three basic headings. Firstly; structural transformation from agriculture to industry and urban migration, Secondly; economic constrains and lastly; personal and family characteristics.

4.1. Structural Transformation From Agriculture to Industry and Urban Migration

Since the early 1970's, Turkey has experienced fast growth and a structural transformation from an agricultural economy to an industrial one. The changing structures of the Turkish economy and the nation's rapid urbanization have affected the composition of the labor force (Kasnakoğlu-Dayıoğlu 1997).

Despite the restructuring of the Turkish economy from agriculture to industry about 40 years, a considerable number of women are still employed in agriculture. Table 3 shows the distribution of employment by sector of economic activity for men and women. The agricultural employment constituted about 35.7 percent of the men's employment and about 76.9 percent of the women's employment in 1988. In 2008, 42.1 percent of women compared to 17.8 percent of men were working in this field. This high percentage share of agriculture within female employment also persists despite a large movement of rural population to cities. But generally migration from rural to urban negatively affected the female participation rate more than males. While women have traditionally had a source of employment in agriculture as unpaid family workers, they can not participate in urban labor force after the migration (Palaz 2005). This is because a majority of urban women are housewives; whereas, nearly all rural women work outside the home in agriculture. Hence, women become housewives or engaged in unregistered informal job when they moved to urban areas.

Table 3. The Distribution of Employment by Sector of Economic Activity and Gender, Turkey
(15 years old and over), (%)

Sector	1988		2008	
	Men	**Women**	**Men**	**Women**
Agriculture	35,7	76,9	17,8	42,1
Industry	25,2	8,7	22,4	14,9
Service	39,9	14,4	59,8	43,0

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

On the other hand, In Turkey, service sector employment has grown rather rapidly and become the largest employing sector for men and the second largest employing sector for women, after agriculture, in 2008. Actually, the decline of agriculture and the growth of the service sector are common for developing societies such as Turkey.

Between the periods 1988 and 2008, employment in industry and service increased; nonetheless, the representation of women in these sectors remained low when compared to men. Although women's participation to the service sector has increased in this period, their participation to industry sector is very low (at 43 percent and about 15 percent respectively in 2008). The reasons for women's low

participation in industry and increased participation in service are the rapid growth of the service sector and the creation of a greater number of jobs for women there than in the industrial sector. In that sense, at least in urban areas, the service sector is the most promising in terms of women's future employment. One reason for this increase might be the expansion of jobs that are considered suitable for women in various sub-branches of the sector (Ecevit 2003).

Table 4. Labor Force Participation Rates by Gender and Geographical Location in Selected Years (%)

Years	Men		Women	
	Urban	**Rural**	**Urban**	**Rural**
1988	78.1	84.7	17.7	50.7
1990	76.8	83.0	17.0	52.0
1995	74.1	82.6	16.8	49.3
2000	70.9	77.9	17.2	40.2
2001	70.6	76.3	17.4	41.7
2002	69.8	74.5	19.1	41.4
2003	68.9	72.9	18.5	39.0
2004	70.8	74.7	18.3	36.7
2005	71.5	73.5	19.3	33.7
2006	70.8	72.9	19.9	33.0
2007	70.6	72.6	20.2	32.7
2008	70.7	73.2	21.1	33.1

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

Reduction in the women's labor force participation rates from 50.7 percent in 1988 to 33.1 percent in 2008 is a clear sign of rural to urban migration. In Turkey, urbanization has been accelerated and it still carries on by increasing its speed. Nowadays, while 70 percent of the population has lived in cities, only 30 percent of them have lived in urban areas. Like in many developing countries, women in rural labor markets of Turkey mostly work as unpaid family workers in agriculture and in some non-market activities such as home production and voluntary jobs. However in 1988-2008 period urban labor force participation rates of women almost kept its level unchanged. This means that women who are largely unpaid family workers in rural labor markets drop out of the labor force when they migrate into urban areas. In other words, women who had migrated to cities became invisible (Biçerli-Gündoğan 2009).

Table 4 demonstrates the labor force participation rates by gender and geographical location. During the period from 1988 to 2008, there is a great disparity between the participation rate of rural and urban area. The labor force participation rate of women is especially low in urban areas. While in 1988, the labor force participation rate of urban women was 17.7 percent, the rate for their

rural counterparts was 50.7 percent. Although the gap between urban and rural participation rates of female labor force started to close, the gap between female and male labor force both urban and rural areas have been continued increasingly.

The lower participation of women in the urban labor market stems directly from the nature of employment in urban areas. Agriculture, where schooling is not a prerequisite to employment, plays an important role in the urban areas, whereas industry and service, which require a relatively more educated labor force, dominate the urban labor market (Dayıoğlu-Kasnakoğlu 1997).

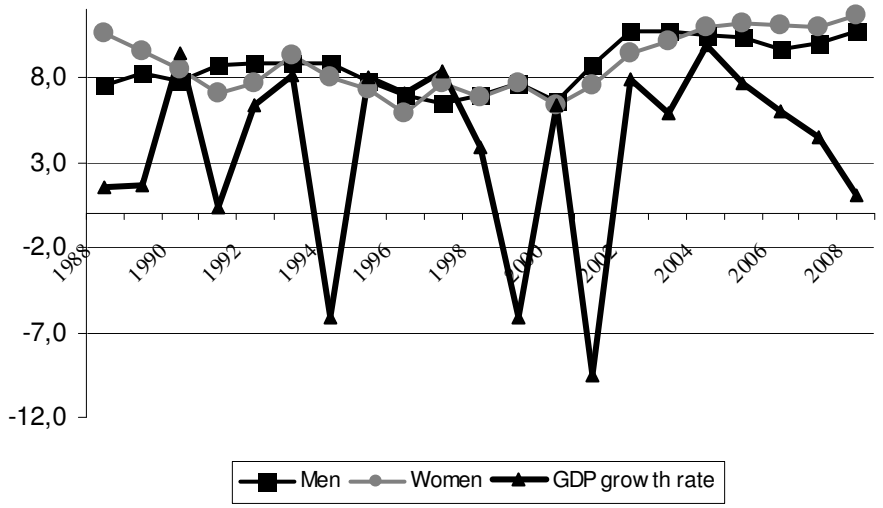
4.2. Economic Constrains

4.2.1. Economic Crises and Unemployment

Turkey has experienced many economic crises. The shrinkage occurred in real sector; because of the following one after another economic crises have been emerged higher unemployment rates. As can be seen on Figure 4, economic crises which occurred in 1994, 1999, 2001 and last crisis in 2008, GDP growth rate has extremely decreased. Actually, Turkish economy could not create enough job while high growth performance period. At present, the phenomenon of growth without employment creation is valid for Turkey too. However, it is not true to say, "firstly women are dismissed their job at the economic crises periods". In 2000-2002 women's labor force participation rate higher than that of men. Because of the influence of 2000 and 2001 crises, a significant rate of men was being unemployed.

At the economic hardship times, unemployment force, especially women who are not working previously, have to seek employment outside of the family. High employment rate causes a negative effect in expectation of finding jobs that seems especially in women who are available for work. They increasingly tend to give up seeking jobs and join the group of housewives and discouraged workers. Alternatively, women are trying to protect their families from poverty by working informal sector. Low educated women usually work at domestic cleaning, child minding or home-working (paid work at home) in order to overcome economic difficulties and thus become members of the informal sector (Ecevit 2003).

Figure 4. GDP Growth Rate and Unemployment Rate by Gender, (%)



Source: TURKSTAT (www.turkstat.gov.tr)

4.2.2. Unregistered Employment

Unregistered employment is at the core of understanding urban unemployment problems. In an environment where the formal sector has been unable to create enough jobs to absorb growth in the urban labor force resulting from both migrations to cities and births there, unregistered employment becomes a remedy for the unemployment. The formal sector offers only limited job opportunities for the unskilled and uneducated population. In addition to this, unequal income distribution and widespread poverty together with the costly process of searching for formal jobs appears as major barriers to entry into the formal sector. For economically deprived individuals, labor income is the main source of revenue (Gündoğan 2007).

Unregistered employment is also a wide spread problem in the Turkish economy. In 2006 almost half of the total employment (48.5 percent) was unregistered. Some of the common features of informal sector employment are lack of protection, limited provision or absence of social security benefits, and insufficient income (Ecevit 2003). It seems that women have a clear disadvantage in this regard. For the same year women's unregistered employment ratio (66 percent) was nearly 20 points higher than that of men's.

The increasing employment of women in the informal sector can be put down to various reasons. First of all, unregistered employment is especially a widespread event in the services sector and agriculture where production units are usually small and jobs require little qualification. Because small firms have limited operational capital, they are financially weak and therefore over-sensitive about costs. These are

also the sectors where the majority of working women are employed (Biçerli-Gündoğan 2009). Secondly, with the increase in unemployment, women are tending to lose ground in the formal sectors of the economy. Thirdly, stagnating and falling household incomes due to the poor performance of the economy also leads to the increased entry of women into the informal labor market (as explained formerly). Lastly, export-oriented industries, in the search for cheap labor, favour women's employment. These factors are leading to the increasing informalization of the female labor force. This informalization takes place broadly in one way: work is being pushed out of factories and formal work situations into small workshops (sweatshops), mall-scale ateliers in the garment and textile industry, homes and informal situations (Özdemir et al. 2004).

4.2.3. Unqualified Jobs and Unfavorable Working Conditions

Women often work 'feminized' sectors and professions and remain in lower job categories with less access to senior positions. Occupational and sectoral segregation has remained almost unchanged, indicating that the increase in female employment has taken place in sectors already dominated by women. A better gender balance across studies and occupations could contribute to meeting future skills and labor market needs (European Commision 2009).
In addition, broadening access for women to employment in an enlarged scope of industries and occupations will be important to enhancing opportunities for them in the labor market. In its latest report on global employment trends for women (2008), the International Labor Office (ILO 2008) states: "Society's ability to accept new economic roles for women and the economy's ability to create the jobs to accommodate them are the key prerequisites to improving labor market outcomes for women, as well as for economic development on the whole".

The poorer the region, the greater the likelihood that women are among the ranks of the contributing family workers or own-account workers. The two statuses together make up the newly defined "vulnerable employment". Vulnerable employment is a newly defined measure of persons who are employed under relatively precarious circumstances as indicated by the status in employment. Because contributing family workers and own-account workers are less likely to have formal work arrangements, access to benefits or social protection programmes and are more "at risk" to economic cycles, these are the statuses categorized as "vulnerable". The vulnerable employment rate, therefore, is calculated as the sum of own account workers and contributing family workers as a percentage of total employment. The indicator is highly gender sensitive since, historically, contributing family work is a status that is dominated by women. There is also a connection between vulnerable employment and poverty: if the proportion of vulnerable workers is sizeable, it may be an indication of widespread poverty. The connection arises because workers in the vulnerable statuses lack the social protection and

safety nets to guard against times of low economic demand and often are incapable of generating sufficient savings for themselves and their families to offset these times. The indicator is not without its limitations; some wage and salaried workers might also carry high economic risk and some own-account workers might be quite well-off and not vulnerable at all. But, despite the limitations, vulnerable employment shares are indicative of employment in the informal economy, especially for the less developed economies and regions, and the fact that a strong correlation has been established between high poverty rates for a region and high shares in vulnerable employment does substantiate the weight of the new indicator to measure progress towards the goal of decent employment for all (ILO 2008).

Table 5. The Distribution of Employment by Employment Status and Gender, Turkey (15 years old and over), (%)

Employment Status	**1988**		**2008**	
	Men	**Women**	**Men**	**Women**
Regular employee	39,4	18,1	56,8	47,6
Casual employee	8,6	4,6	7,0	5,6
Employer	4,9	0,2	7,5	1,4
Self employed	33,8	6,8	23,8	11,0
Unpaid family workers	13,5	70,2	4,9	34,4

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

As is seen in Table 5, the employment status of women and men also differs markedly. Most women work either in regular employment (47.6 percent) or in casual employment (5.6 percent); 34.4 percent work as unpaid family workers. There is a similar point regarding the employment status for men as well: the highest percentages of men (56.8 percent) work as regular employees (waged and salaried). However, while 31.3 percent of men are self-employed or employers, only 12.4 percent of women have this employment status. The most significant characteristic of the working status of women is the high ratio of unpaid family workers. Although the share of this group within the female labor force has decreased steadily from 70.2 percent in 1988 to about 35 percent in 2008, it is still high when compared to women in other employment categories. This originates mainly from the fact that almost all women working in agriculture are considered unpaid family workers whereas all rural heads of households (men) are considered self-employed. When compare to the number of men working as unpaid family workers is almost over one-sixth of the number of women working under a similar status.

4.2.4. Gender Differences in Earned Income

One of the consequences of gender segregation on the labor market is the persisting gender pay gap, (the difference between the wages earned by women and those earned by men) partly due to the fact that women are concentrated in less valued jobs and positions than men. Because women are more likely to work part-time and interrupt their career for family reasons, they are likely to face negative consequences in terms of pay, career advancement and accumulated pension rights. This also has an impact on the risk of falling into poverty (European Commision 2009).

The main cause of gender pay gap is gender difference in human capital. This simply means that women's levels of education, skill and experience, as their human capital, continue to remain lower than those of men. This is thus the widely accepted view of why women are paid less than men (Ecevit 2003).

Table 6. The Gender Gap in Earnings

	Estimated Earned Income (PPP US$) 2006		**Ratio of Estimated Female to Male Earned Income**
	Female	**Male**	**Female/Male**
Very High Human Development			
Norway	46.576	60.394	0,77
Japan	21.143	46.706	0,45
United States	34.996	56.536	0,62
United Kingdom	28.421	42.133	0,67
Greece	19.216	38.002	0.51
High Human Development			
Hungary	16.143	21.625	0,75
Cuba	4.132	8.442	0,49
Bulgaria	9.132	13.439	0,68
Romania	10.053	14.808	0,68
Turkey	5.352	20.441	0,26
Medium Human Development			
Iran	5.304	16.449	0,32
China	4.323	6.375	0,48
Indonesia	2.263	5.163	0,44
South Africa	7.328	12.273	0,60
India	1.304	4.130	0,32
Low Human Development			
Ethiopia	624	936	0,67
Mozambique	759	848	0,90
Sierra Leone	577	783	0,74
Afghanistan	442	1.845	0,24
Niger	318	929	0,34

Source: UNDP, Human Development Report 2009, Palgrave Macmillan, New York, 2009, p. 181-189

Table 6 demonstrates "gender pay gap" in selected countries. As it can be seen on table, the wages of working women are extremely lower than those of men in Turkey. In fact, it is known that even in countries with very high human development, women's earnings are generally lower than that of men's. According to the Human Development Report 2009, women in Turkey earn average 26 percent to men's. Although Turkey is in the high human development category to United

Nation Development Programme, the low ratio of estimated female to male earned income is a good indicator in terms of the gender gap.

4.3. Personal and Family Characteristics

4.3.1. Educational Attainment

People with lower educational qualifications are both less likely to be labor force participants and more likely to be without a job even if they actively seek one. Differences in unemployment rates for males and females generally decrease with educational attainment (OECD 2008). Similarly differences in labor force participation rates for males and females generally decrease with educational attainment. So, there is a positive relationship between an increased level of education and women's participation to labor force. While this correlation is less pronounced in rural areas, labor force participation of urban females increases clearly with their educational status. Education not only lays the ground for wage increases by raising productivity, but also gives social legitimacy to this participation by weakening patriarchal ways of thinking (Toksöz 2007).

Table 7. Labor Force Participation Rates by Educational Attainment and Gender, Turkey
(15 years old and over)

Educational Attainment	**1988**		**2008**	
	Men	**Women**	**Men**	**Women**
Illiterate	70,5	32,3	36,0	14,5
Literate-No Diploma	76,3	31,7	50,7	18,5
Primary School	88,9	34,3	75,1	21,1
Middle and Voc-Middle School	61,4	19,5	82,9	21,6
High School	75,5	45,7	66,2	29,1
Voc. High School	82,8	52,5	80,3	38,3
University	89,5	82,5	82,7	70,0

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

Table 7 shows labor force participation rates by educational attainment and gender. In 2008, the female labor force participation rate increase along with educational achievement (the lowest rate being recorded for illiterates and the highest rate for university graduates). For males, labor force participation does not show much change with respect to educational status. The relationship between schooling and labor force participation rate only becomes apparent at higher education levels for both genders.

4.3.2. Age structure

Table 8. Labor Force Participation Rates by Age Group and Gender, Turkey

Age Group	1988		2008	
	Men	Women	Men	Women
15-19	64,2	40,5	36,7	17,1
20-24	87,5	40,8	71,8	33,2
25-29	97,9	35,9	92,3	33,3
30-34	98,5	36,4	94,8	31,5
35-39	98,5	35,5	94,8	32,4
40-44	95,8	34,5	93,4	28,7
45-49	89,0	34,3	80,7	24,7
50-54	82,4	34,1	64,6	20,3
55-59	71,0	27,3	48,2	16,4
60-64	58,1	19,8	37,7	13,1
65+	33,3	10,1	19,4	5,8

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

Table 8 demonstrates the labor force participation rates by age groups and gender. Not surprisingly, prime working age (25 to 45 years) make-up the vast majority of the labor force. For both genders, participation rates increase by age, reach a peak around the ages of 25-45 for men and 20-40 for women and decline thereafter.

Tansel (2001) shows in her study that younger population (15-24 years old) have been staying in school longer recently. The 1997 law, which extended the duration of compulsory primary education from five to eight years, and the current increase in enrolment rates at all levels of schooling have helped to delay recent participation rates of the young population. Another reason for declining trends in the female force participation rates could be the early retirement system introduced in early 1980`s. According to these retirement arrangements women at age 50 or after 20 years of service and men after 25 years of service or at age 55 had the right to retire. This institutional arrangement contributed to the lower participation rates of the middle to older age groups. This retirement plan has since been changed, as longer years of service and higher age limits were introduced in 2001. The new law is expected to increase participation rates around 50 years of age (Tansel 2001).

4.3.3. Marriage and Child Care responsibilities of Women

The relationship between marital status and female participation reveals that single women and divorced women have much higher participation rates than married and widowed women (see Table 9). It is well known that financial motivation is the most important factor in women's participation for divorced women, because they are the

head of the household and need to support their family and children. Single women's labor force participation rates are likely to be higher than married women's because they do not have family obligations, but it is likely to be lower than the participation rate of divorced women because financial need is not as important to them (Palaz 2005).

Table 9. Labor Force Participation Rates by Marital Status and Gender, Turkey (15 years old and over)

Marital Status	**1988**		**2008**	
	Men	**Women**	**Men**	**Women**
Single	71,8	47,8	58,3	35,3
Married	86,4	32,0	76,2	22,4
Divorced	81,1	41,5	69,1	42,9
Widowed	30,1	16,0	19,3	8,6

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)

The effect of husband's employment situation on wife's participation is also a good indicator of women's status in the labor market, specifically of their "secondary worker" status. Husband's employment situation could affect wife's participation decisions in two ways. According to the "added worker effect" hypothesis, wives, inactive in the labor market, decide to temporarily participate in order to compensate for the loss of income due to husband's unemployment (Kızılırmak 2005). In contrast, the "discouraged worker" effect suggests that women drop out of the labor force during periods of recession because expected returns to search are low: wages are depressed and the probability to finding employment is small (Psacharopoulos-Tzannatos 1992).

On the other hand, being married and having children negatively affects the labor force participation of women due to their housekeeping chores and childcare responsibilities. Maternity is one of the main factors that influence female participation rates. Especially, a childcare responsibility is more crucial who has children below seven age and school age children between ages of seven and fourteen (Dayıoğlu-Kasnakoğlu 1997). The idea that young children may suffer if their mother works is relatively wide spread in all countries and has an impact on the labor force participation of women of a child bearing age (25 to 44 years) (OECD 2003). They rejoin only later in life, when their children are old enough to take care of themselves. Therefore, instead of the usual bell-shaped age participation profile observed for men, their interrupted careers give rise to an M- shaped pattern (Dayıoğlu-Kasnakoğlu 1997).

Table 10. Women Not in Labor Force by Marital Status and Age Group, Turkey (15 years old and over), (Thousand)

Age Group	1988				2008			
	Single	**Married**	**Divorced**	**Widowed**	**Single**	**Married**	**Divorced**	**Widowed**
15-19	1484	273	2	1	2183	252	2	0
20-24	351	1050	3	3	838	1108	11	2
25-29	78	1234	13	8	269	1829	24	7
30-34	24	1076	6	10	128	1749	27	13
35-39	18	981	10	19	88	1618	26	28
40-44	9	774	10	38	56	1488	36	48
45-49	4	652	5	58	43	1386	53	90
50-54	5	608	5	107	24	1197	48	142
55-59	9	529	13	149	25	923	37	192
60-64	6	384	8	195	12	691	22	251
65+	8	385	13	613	30	1136	37	1357

Source: TURKSTAT, Household Labor Force Survey Results (www.turkstat.gov.tr)
Note: It is accepted that the age of youths (15 to 24 years), child-bearing aged women (25 to 44 years) and older men and women (55 to 64 years)

Another reason for the low participation of women in the labor market in Turkey is that young boys are the second breadwinners, after the fathers, of the poor families in the urban areas. These boys, instead of mothers, work in the labor market (Bulutay-Tastı 2004). On the other hand, the male participation rate demonstrates different profiles; married men have the highest labor force participation rate (78.2 percent), divorced men have the second highest (68.2 percent), and widowed men have the lowest (22.4 percent) after single men (58 percent). All these figures support the idea that women are the secondary (supplementary) labor force in Turkey.

The participation rate of the widows is the lowest for both men and women. This may possibly be due to the fact that a high proportion of the widows are elderly and have reduced participation rates (Tansel 2001).

4.3.4. Cultural Norms Against Women

Cultural factors against women are also affect women's participation in market work. The most important cultural factor influencing women's position and participation to market work in Turkey is patriarchy (Gündüz-Smits 2006). Family pressure from husbands, parent and in-laws is an important constrains to employment faced by poorly educated women (SPO-World Bank 2009b). The head of the household and the main breadwinner is expected to be a man. On the other hand, within the framework of socio-cultural factors, that is gender-based division of

work and gender roles based on this division, women in Turkey, are considered homemakers who should take care of domestic duties and childcare and be responsible for the well being of other family members, including the elderly. According to Toksöz (2007), especially in case where the employment of unqualified women of low educational background is regarded as threat to the dominancy of male household heads as "breadwinners", women and young girls are not allowed to work out of their homes. Despite considerable changes in the last few decades, there are deep underlying cultural beliefs about gender roles. These beliefs are especially strong among people living in rural areas and migrants to the cities who continue to preserve many of their traditional attitudes (Ecevit 2003).

5. Policies Affecting Female Labor Force Participation

Under the 9th Development plan which covers the 2007-2013 period, it has been developed an "Action Plan for Gender Equality" by Turkish State Planning Organization (SPO 2006) for provide more and better job opportunities for women. In this framework, three policies have been indentified (SPO-World Bank 2009b).

- *Creating job opportunities for first time job seekers for young women: Short term work regulation and programs to enhance labor demand for women.*

Interventions should aim at promoting formal employment for women with low levels of education, and especially for those transitioning from school to work. More flexible labor market regulation may reduce the barriers of businesses hiring women. For instance, the Government of Turkey has introduced recently a program that subsidizes employers' social security contributions for new hired women for up to 5 years. Programs like this are likely to contribute to increase the rate of employment for women, especially in times of healthy economic performance.

- *Affordable childcare: Pre-school, Public and/or subsidized childcare.*

Many women could be encouraged to work by having access to affordable care for their children. This could be achieved by promoting early childhood development programs (ECD), such as preschool education and public/subsidized childcare programs (SPO-World Bank 2009b). The empirical evidence shows that childcare subsidies do increase female labor supply. Childcare subsidies reduce the relative price of formal childcare and, therefore, increase the relative return of market work (OECD 2003).

- *Sustaining investments on education: Increase attainment rates and invest in Vocational Education and Training (VET).*

Higher education attainment is associated with higher levels of female participation. Investment in VET are likely to prepare women, and especially young women, with the skills needed to quality for good jobs in the labor market. Turkish Government has scaled up of investment in VET for women though Turkish Employment Organization. International evidence indicates that investments on VET help women

to get formal jobs, and promotes gender equality in earnings and labor market opportunities.

By introducing positive action, creating job opportunities for first time job seekers for young women, affordable childcare and sustaining investments on education especially VET, the state may make a major contribution to the strengthening of the labor market position of Turkish women.

Turkish SPO stated that, in 8th development plan period (2001-2005) progress was not made in labor force participation and employment rates, which were lower than European Union averages. Such low rates are caused by insufficient participation of women to the labor force and employment. The female labor force participation and employment rates are around one third of the rate for men. On the other hand, labor force participation rate is expected to increase by 2.1 percentage points during the 9th Plan period (2007-2013) through raising the level of education, increasing employability by active labor market policies, and facilitating and encouraging entrance into the labor market. The main determinant of this increase is expected to be the increase in female force participation rate (SPO 2006).

6. Conclusion

This paper has examined the factors underlying female participation in Turkey. With in the context of globalization women's labor force participation and employment rates have increased all over the world whereas Turkey differs from most countries with decrease. Turkish labor market is characterized by high rates of population and labor force growth, declining rates of participation and exceptionally low levels of female participation in urban areas. Towards sustainable growth and development targets of Turkey, optimal resource usage is a necessity. In Turkey, almost one-half of the population is women, effective usage of women resource as labor force is important in respect to development targets. In addition to this, it is known that the role of women on the child care is also important for socio-economic development. So, women have an active role on Turkish economic development.

Turkey has experienced important structural and social changes that would be expected to facilitate women to seek jobs. In this sense several satisfactory progresses are:

- The social attitude toward working women have changed in recent years,
- Women are becoming more educated
- Women are getting married at a later age,
- Fertility rates are declining
- Micro finance applications are growing up.

After all educational attainment is on the rise, while fertility is on the decline, and other positive developments might impact female labor force participation positively in next years. It should not be forget to the importance of a sustainable and high level of economic activity in achieving higher female participation and employment.

References

Abhayaratna, J.- Lattimore, R. 2006: Workforce Participation Rates- How Does Australia Compare?**,** *Australian Government Productivity Commission Staff Working Paper, December,* Canberra.

Argy, F. 2005: An Analysis of Joblessness in Australia, *Economic Papers (Economic Society of Australia)*, March, Vol. 24. Issue 1, 75-96. p.

Biçerli, M. K.- Gündoğan, N. 2009: Female Labor Force Participation in Urbanization Process: The Case of Turkey, *The 10th Asian Urbanization Conference,* Hong Kong.

Bulutay, T.-Tastı, E. 2004: Informal Sector in the Turkish Labor Market, *Turkish Economic Association Discussion Paper,* No. 22.

Dayıoğlu, M.- Kasnakoğlu, Z. 1997: Kentsel Kesimde Kadýn ve Erkeklerin İşgücüne Katılımları ve Kazanç Farklılıkları*, METU Studies in Development*, 24/3 329-61. p.

Ecevit, Y. 1998 "Küreselleşme, Yapısal Uyum ve Kadın Emeğinin Kullanımında Değişmeler", In: F.Özbay (ed.): *Küresel Pazar Açısından Kadın Emeği ve İstihdamındaki Değişmeler,* İstanbul, İnsan Kaynağını Geliştirme Vakfı and T.C. Başbakanlık Kadının Statüsü ve Sorunları Genel Müdürlüğü.

Ecevit, Y- Tan, M.- Üşür, S. 2000: *Kadın-Erkek Eşitliğine Doğru Yürüyüş: Eğitim, Çalışma Yaşamı ve Siyaset,* Türk Sanayicileri ve İşadamları Derneği, Yayın No:TÜSIAD-T/2000- 12/290.

Ecevit, Y. 2003: *Women's Labor and Social Security*, In: Bridging the Gender Gap in Turkey: A Milestone towards Faster Socio-economic Development and Poverty Reduction, Poverty Reduction and Economic Management Unit Europe and Central Asia Region, September 16, 73-101. p.

European Commision. 2009: *Report on Equality Between Women and Men 2009,* Belgium.

Gündoğan, N. 2007: Urbanization and Labor Market Problems in Developing Countries: The Case of Turkey, *The 9th Asian Urbanization Conference,* Chun-Cheon City.

Gündüz, H. A.-Smits, J. 2006: Variation in Labor Market Participation of Married Women in Turkey, *NICE Working Paper,* Version 2, Autumn, Radboud University.

ILO 2008: Global Employment Trends for Women, March, International Labour office, Geneva.

Kasnakoğlu, Z.- Dayıoğlu, M. 1997: *Female Labor Force Participation and Earnings Differentials Between Genders in Turkey*, In: Rives, J.M.-Yousefi, M. (eds.): Westpost Economic Dimensions of Gender Inequality: A Global Perspective, CT and London, 95-117. p.

Kızılırmak, A. B. 2005: Labor Market Participation Decisions of Married Women: Evidence From Turkey, Mimeo., Ankara University. (http://www.policyinnovations.org/ideas/policy_library/data/01222/_res/id=sa_File1/).

OECD 2003a: Labor Force Participation of Women, *OECD Economic Studies,* No. 37.

OECD 2003b: Towards More and Better Jobs, *OECD Employment Outlook 2003,* Paris.

OECD 2008a: *Education at a Glance 2008.*

OECD 2008b: *OECD Employment Report 2008.*

Özar, S. 1994: Some Observations on the Position of Women in the Development Process of Turkey, *Bogaziçi Journal: Review of Social, Economic and Administrative Studies,* 8/1-2.

Özbay, F. 1991: Türkiye'de Kadin ve Çocuk Emegi, *Toplum ve Bilim*, No. 53.

Özbay, F. 1994: Women's Labor in Rural and Urban Settings, *Boğaziçi Journal,* Vol. 8. Issue 1-2. 5-19. p.

Özdemir, A.M.- Erel, D. - Yücesan-Özdemir, G. 2004: Rethinking the Informal Labor Market in Turkey: a Possible Politics for the Trade Unions, *South-East Europe Review,* 3/2004.

Palaz, S. 2005: Women's Labour Force Participation in Turkey", *Sosyal Siyaset Konferansları* (Prof. Nevzat Yalçıntaş'a Armağan Sayısı), 591-606. p.

PC (Productivity Commission) 2005: Economic Implications of an Ageing Australia, *Research Report,* Canberra.

Psacharopoulos, G- Tzannatos, Z. 1992: Latin American Women's Earnings and Participation, in the Labor Force, *World Bank Working Paper,* WPS 856.

SPO-World Bank 2009a: *Turkey's Greatest Untapped Potential: Women*, (www.worldbank.org).

SPO-World Bank 2009b: *Social and Economic Benefits of More and Better Job Opportunities for Women in Turkey*. (http://siteresources.worldbank.org/ECAEXT/Resources/Turkey_FemaleReport_3_pager_final_final.pdf).

SPO 2006: *Ninth Development Plan (2007-2013)*, Ankara.

Tansel, A. 2001: Economic Development and Female Labor Force Participation in Turkey: Time Series Evidence and Cross-Province Estimates, *Economic Resource Center Working Papers in Economics,* 01/05, Ankara.

Toksöz, G. 2007: *Women's Employment Situation in Turkey,* International Labor Office (ILO), Ankara.
TURKSTAT: *Household Labor Force Survey Results* (www.turkstat.gov.tr).
UNDP. 2009: *Human Development Report 2009,* Palgrave Macmillan, New York.
World Bank. 1993: Turkey Women in Development, *A World Bank Country Study,* Washington D.C.

Using statistical methods for analysing regional differences of labour market

László Lőkös, yr[1] – Klára Lőkös Tóth[2] – Emese Prihoda[3]

The objective of present descriptive, exploring analysis, was to demonstrate and illustrate regional (NUTS II) and county (NUTS III) level differences based on three dominant labour market indicators.

Accordingly, we pursued a statistical analysis on the 2004 and 2008 labour market indicators (basic data) of the 7 statistical regions, and the 19 counties of Hungary. These indicators are: unemployment rate, employment rate, rate of economically active population.

The comparison is made based on two years descriptive statistical data using various types of methods with taking into consideration the changes of the three main indicators of the labour market. By means of special scatter plots, we characterised and illustrated the relative changes of the regions by correlating them with each other, namely the more and more strong separation of the main groups, so as the increase of their homogeneity within groups.

Keywords: regional unemployment, rate of unemployment, rate of economically active population, employment rate, analysis, region, county

1. Introduction

According to many studies in the field of social sciences, the problem of high rate unemployment in European countries is one of the most serious phenomenon since many years now (Baranyai 2007; Cseh Papp 2008, Pál 2005, Lőkös 2005, Vajsz-Pummer 2006). The demand for reducing unemployment gets more and more attention in European general objectives, together with the propagation of a 'longer active career model', originated in demographic reasons, the so called 'ageing society' causing many stress in the labour market which enforced the reforms of the traditional pension scheme, so much as the remarkably growing appearance of unem-

[1] László Lőkös, yr PhD, associate professor, Károly Róbert Collage, Faculty of Economics, Department of Public Services Management, Gyöngyös

[2] Klára Lőkös Tóth, PhD, associate professor, Szent István University, Faculty of Economics and Social Sciences, Department of Economy Analysis, Gödöllő

[3] Emese Prihoda, PhD student, Szent István University, Faculty of Economics and Social Sciences, Department of Economy Analysis, Gödöllő

ployed young school-leaver population. In our opinion, the effective way to treat the above mentioned problem requires intervention in sub-national (regional, county, micro-regional) level, however, the evolvement of these programmes requires statistics-based regional labour market research, which may serve for the preparation of economic (labour market) decisions.

Our aspect shows and uses data of two time periods, year 2004, when Hungary joined the European Union, and year 2008, the most up-to-date data available. Our ambition is to demonstrate the changes in the field of regional labour market between the two examined periods.

2. Materials and methods

Data were downloaded from the website of Public Employment Service (PES). We drawn in territorial (NUTS[4] II: region, NUTS III: county) data from 2004 and 2008.

The three indicators we used in order to characterise the labour market situation are:

1. unemployment rate (UNR%)
2. employment rate (EMR%)
3. activity rate (Active%)

We applied different statistical methods to demonstrate the regional differences in Hungary. We started our investigation with basic descriptive statistical methods, later on, we implied the tools of hypothesis testing in order to be able to compare the data of the regions with each other. We measured the coherence between variables with correlation coefficient. In the course of different types of analysis, we used graphical illustrations, with which we supported the revealing of relationships.

3. Results

It is a widely spread and accepted statement, that the developing unemployment rate can be used to measure competitiveness. Therefore, we claim the need of analysing data which represents the status of the Hungarian labour market the most, like the employment rate, the unemployment rate and the activity rate.

Principally, we analysed the above mentioned labour market indicators in the 7 statistical regions (NUTS II) of Hungary.

[4] NUTS: Nomenclature of Territorial Units for Statistics

3.1. Regional analysis based on data form 2004 and 2008

Table 1. The changes of major labour market indicators within the examined time period

	2004	**2008**	d*	**2004**	**2008**	d*	**2004**	**2008**	d*
Region	UNR%	UNR%	UNR%	EMR%	EMR%	EMR%	Active%	Active%	Active%
CH	4.38	4.62	0.24	56.60	56.20	-0.40	59.20	58.90	-0.30
WTd	4.65	5.68	1.03	54.70	54.70	0.00	57.40	58.00	0.60
CTd	5.72	6.29	0.57	53.70	53.60	-0.10	56.90	57.20	0.30
STd	7.15	9.27	2.12	46.60	45.50	-1.10	50.00	50.20	0.20
SGP	6.81	9.20	2.39	47.00	47.60	0.60	50.40	52.40	2.00
NGP	7.46	12.15	4.69	44.80	44.30	-0.50	48.40	50.50	2.10
NH	9.16	12.26	3.10	44.50	43.90	-0.60	49.00	50.10	1.10
Total	6.08	7.81	**1.73**	50.50	50.30	**-0.20**	53.80	54.60	**0.80**

**difference 2008-2004*
Source: http://kisterseg.afsz.hu/index.php

NH: Northern Hungary, NGP: Northern Great Plain, SGP: Southern Great Plain, CH: Central Hungary, CTd: Central Transdanubia, WTd: Western Transdanubia, STd: Southern Transdanubia

Table 2. Descriptive statistical analysis of the major labour market indicators

	2004			**2008**		
	UNR%	EMR%	Active%	UNR%	EMR%	Active%
Mean*	6.48	49.70	53.04	8.50	49.40	53.90
Standard deviation	1.56	4.73	4.24	2.84	4.88	3.68
CV%	24.1	9.5	8.0	33.4	9.9	6.8

* simple average of regional means
Source: own creation

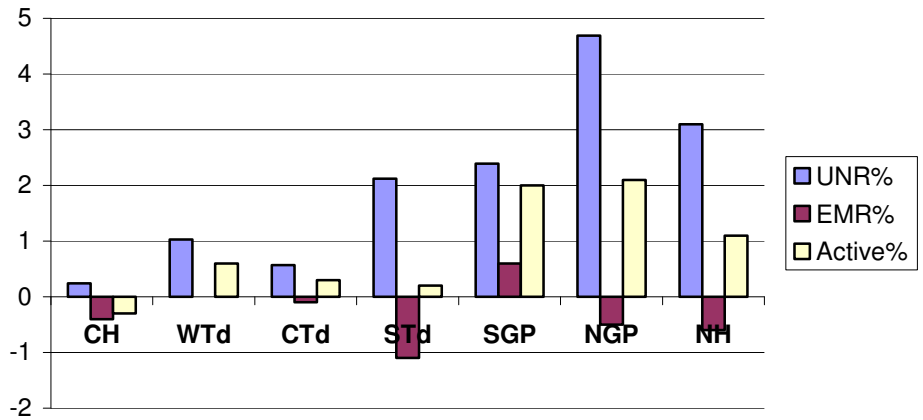
Between the two examined time periods, the unemployment rate shows the largest difference. As *table 2* shows, the unemployment rate (UNR%) has risen in all the regions, but its rate has large differences. The growth was only 0.24 percent score in Central Hungary, although in Northern Hungary it was 3.1, in Northern Great Plain it was 4.69.

The differences in the two other indicators are less remarkable. The employment rate (EMR%) has only risen in Southern Great Plain region (by 0.6 percent

score) while the rest of the regions represent a slight decrease. Southern Transdanubia is the only region where the decrease has been over 1 percent score.
Excluding Central Hungary, the activity rate (Active%) has slightly risen in all the regions, especially in Northern Great Plain (by 2.1 percent score) and in Southern Great Plain (by 2 percent score). The activity rates of regions do not show large differences, they vary between 50.1% (Northern Hungary) and 58.9% (Central Hungary).

In total (national economy level), the employment rate decreased by 0.2 percent score, the unemployment rate decreased by 1.72 percent score, while the activity rate showed a 0.8 percent score of increase. Although, we can observe larger changes in certain regions. The most dramatic change took place in Northern Great Plain, where the unemployment rate raised from 7.46% to 12.15% between 2004 and 2008. The changes of the values of each regions can be seen on *Figure 1*.

Figure 1. Illustration of the level of changes in each region (in percent score)

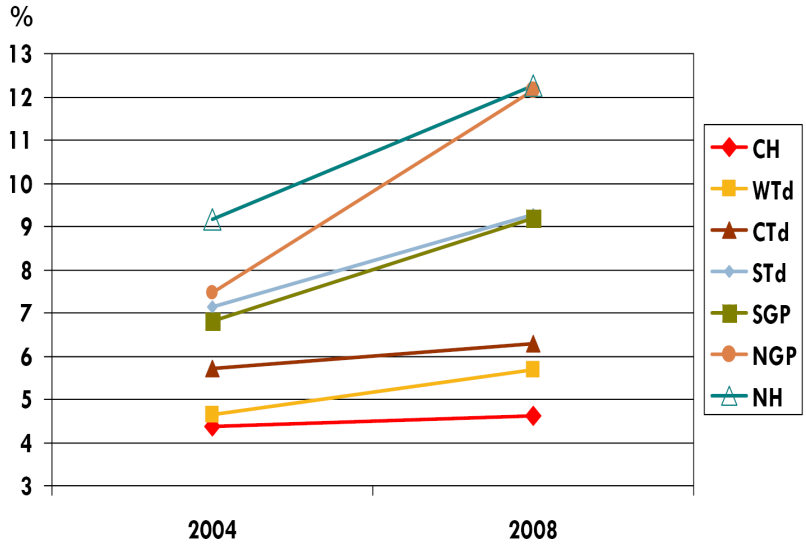


Source: own creation

In point of the three examined indicator, we can easily observe on *Figure 1*, that Central Hungary, Western Transdanubia and Central Transdanubia regions (CH, WTd, CTd) show slight change, Southern Transdanubia, Southern Great Plain and Northern Hungary (STd, SGP, NH) represents medium, while Northern Great Plain (NGP) shows large change.

Our analysis proved, that the values of unemployment rate have the largest change in time, that is why we would like to illustrate the changes on a line chart. (Figure 2)

Figure 2. Comparison of the rate of unemployment in the two examined year



Change in %score	
CH	0.24
WTd	1.03
CTd	0.57
STd	2.12
SGP	2.39
NGP	4.69
NH	3.1

Source: own creation

The line chart above clearly shows, that while unemployment rate in Central Hungary (CH) and Central Transdanubia (CTd) regions hardly changed, the rate in Northern Great Plain (NGP) shows the largest change. In the same time, the differences can also be observed in the vertical position of the lines.

Based on the unemployment rate (UNR%), we can make 3 groups:

1. Group with stable unemployment rate, with relatively low value: Central Hungary (CH), Central Transdanubia (CTd), Western Transdanubia (WTd) regions
2. Group of moderate increase: Northern Hungary (NH), Southern Transdanubia (STd) and Southern Great Plain (SGP) regions
3. Group of large increase: Northern Great Plain (NGP) region.

Table 3. Rank-score values of regions based on the major labour market indicators

	2004			2008		
Region	UNR%	EMR%	**Active%**	UNR%	EMR%	**Active%**
CH	7	1	1	7	1	1
WTd	6	2	2	6	2	2
CTd	5	3	3	5	3	3
STd	3	5	**5**	3	5	**6**
SGP	4	4	4	4	4	4
NGP	2	6	**7**	2	6	**5**
NH	1	7	**6**	1	7	**7**

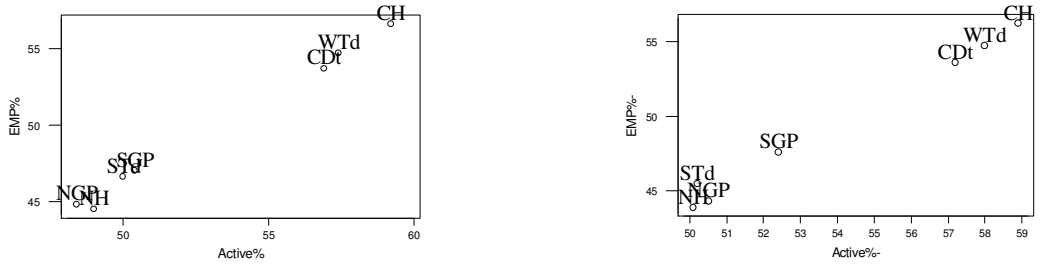
Grade 1 belongs to the region that has the highest value.
Source: own creation

Despite the fact that the value of the changes differ form each other, the rank-score values of the regions only differ in terms of the activity rate. Northern Great Plain improved its value from 7 to 5, while Southern Transdanubia and Northern Hungary regions have fallen back by 1-1 grade.

Henceforward, we analyse the position of the regions in a co-ordinate system based on indicators pairs.

Figure 4. Employment rate (EMR%) depending on active population (Active%)

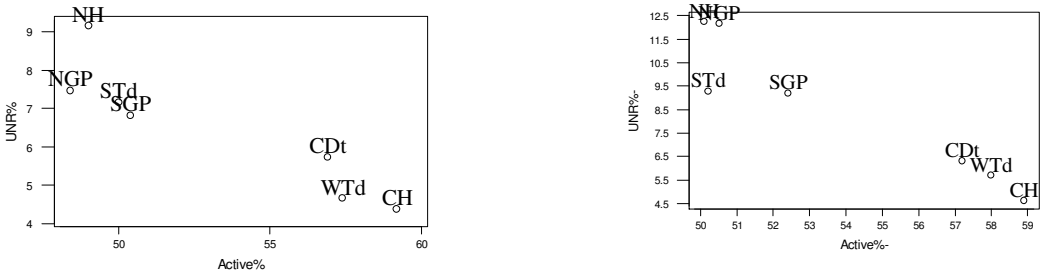
2004 **2008**



Source: own creation

Figure 5. Unemployment rate (UNR%) depending on active population (Active%)

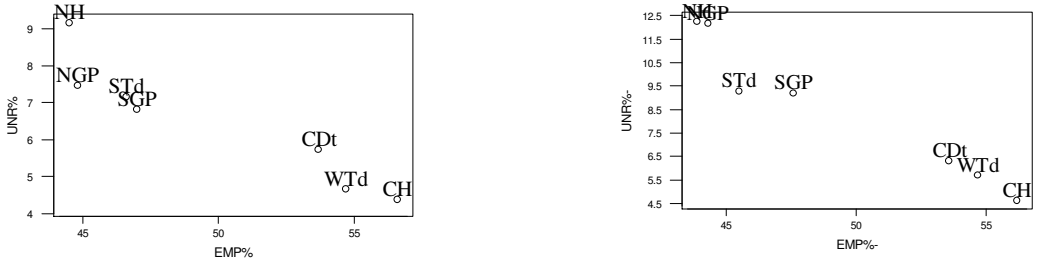
2004 **2008**



Source: own creation

Figure 6. Unemployment rate (UNR%) depending on employment rate (EMR%)

2004 **2008**

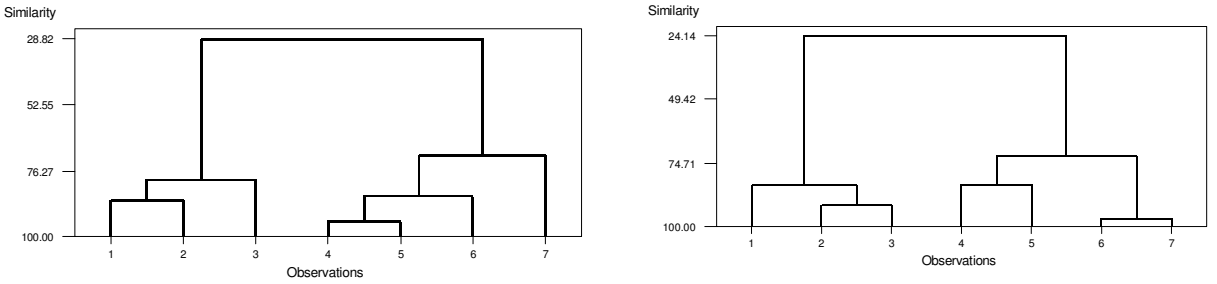


Source: own creation

The correlograms of indicator pairs consequently generate the same group of regions. *Group 1* includes Central Hungary, Western Transdanubia and Central Transdanubia (CH, WTd, CTd) regions. *Group 2* comprise Southern Great Plain, Southern Transdanubia, Northern Great Plain and Northern Hungary (SGP, STd, NGP, NH) regions. The separation of the two above mentioned groups is also confirmed by the dendogram of the cluster analysis.

Figure 7. Grouping regions with cluster analysis

2004 **2008**



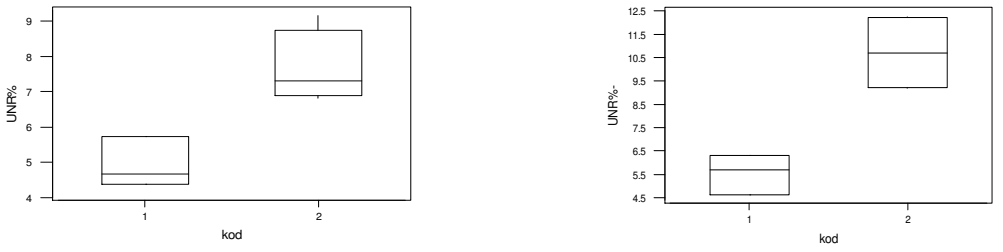
Legend: 1-CH 2-WTd 3- CTd 4- STd 5- SGP 6- NGP 7- NH

Source: own creation

We can observe on the dendograms of 2004 and 2008, that the two groups previously created remained the same, the homogeneity within the groups slightly increased. The examined attributes of regions became closer to each other within the groups. The same conclusion can be drawn if we have a look at the boxplots (*Figure 8*), which is limited to illustrate the unemployment rate.

Figure 8. Boxplots of the two groups of regions based on unemployment rate

2004 **2008**



Group 1*: CH, WTd, CTd,* ***Group 2:*** *SGP, STd, NGP, NH*

Source: own creation

County level analysis based on data form 2004 and 2008

Central Hungary has been excluded form the analysis because of the disconformities of its values.

Table 4. County level values of the examined 6 regions based on the three major labour market indicators

		2004			2008		
Region	County	UNR%	EMR%	Active%	UNR%	EMR%	Active%
CTd	Fe	6.63	53.20	57.00	6.60	52.50	56.30
	Ko	5.08	54.90	57.90	5.02	56.40	59.40
	Ve	5.21	53.10	56.10	7.09	52.30	56.30
WTd	Gyo	4.03	53.20	55.50	4.25	55.60	58.10
	Vas	5.14	56.70	59.80	5.65	54.40	57.70
	Za	5.09	55.00	58.00	7.77	53.50	58.10
STd	Ba	7.37	43.30	46.80	9.28	44.40	49.00
	So	7.01	46.30	49.70	9.51	45.30	50.10
	To	7.00	51.70	55.60	8.92	47.70	52.30
NH	BAZ	10.10	43.50	48.40	13.61	42.40	49.10
	He	7.84	46.50	50.50	10.27	47.00	52.40
	No	8.07	45.10	49.10	11.01	44.40	49.90
NGP	Ha	6.72	46.80	50.20	10.94	44.30	49.80
	Sza	9.03	41.90	46.00	15.59	41.30	48.90
	JN	6.41	46.00	49.20	9.23	48.70	53.60
SGP	BK	7.14	49.50	53.30	8.93	48.40	53.10
	Be	7.21	43.30	46.70	10.86	44.70	50.10
	Cs	6.01	47.00	50.10	8.15	49.10	53.50

Source: http://afsz.hu/

NH: Northern Hungary, NGP: Northern Great Plain, SGP: Southern Great Plain, CH: Central Hungary, CTd: Central Transdanubia, WTd: Western Transdanubia, STd: Southern Transdanubia

Fe: Fejér, Ko: Komárom-Esztergom, Ve: Veszprém, Gyo: Győr-Moson-Sopron, Vas: Vas, Za:Zala, Ba: Baranya, So: Somogy, To: Tolna, BAZ: Borsod-

Abaúj-Zemplén, He: Heves, No: Nógrád, Ha: Hajdú-Bihar, Sza: Szabolcs-Szatmár-Bereg, JN: Jász-Nagykun-Szolnok, BK: Bács-Kiskun, Be: Békés, Cs: Csongrád

Table 5. Descriptive statistical indicators

	2004			**2008**		
Indicators	UNR%	EMR%	Active%	UNR%	EMR%	Active%
Mean*	6.73	48.72	52.22	9.04	48.47	53.21
Standard deviation	1.47	4.60	4.31	2.78	4.54	3.51
s%	**21.80**	9.40	8.30	**30.8**	9.40	6.60
Min	4.03	41.90	46.00	4.25	41.30	48.90
Max	10.10	56.70	59.80	15.59	56.40	59.40
Range	**6.07**	14.80	13.80	**11.34**	15.10	10.50
Median	6.86	**46.90**	**50.35**	9.08	**48.05**	**52.75**

* simple average of regional means

The most significant change again belongs to the unemployment rate. As far as the unemployment rate concern, the differences between regions have been increasing, which can be proved by the growth of s%, it almost doubled, it changed from 21.8 to 30.8 percent. The activity rate shows positive development, its rate increased, while its standard deviation decreased.

By using county level statistical data, the separation of the two groups of regions can also be verified by analysis of variance. In the next session, we analyse weather there is a provable statistical difference between regions by analysing the data of their counties. (we consider the data of counties as repetition data of regions)

The outcomes of the labour market analysis of the 6 regions are summarised on *table 6* below.

Table 6. Comparison of regions mean by labour market indicators

6/1. Comparison of employment rate means

NGP	NH	SGP	STd	CTd	WTd
44.900	**45.033**	**46.600**	**47.100**	**53.733**	**54.967**

6/2. Comparison of activity rate means

NGP	NH	SGP	STd	CTd	WTd
48.467	**49.333**	**50.033**	**50.700**	**57.000**	**57.767**

6/3. Comparison of unemployment rate means

WTd	CTd	SGP	STd	NGP	NH
4.753	**5.640**	**6.787**	**7.127**	**7.387**	**8.670**

Source: own creation

The employment rate and activity rate does not show any difference between the separation of region. We can separate the same groups as before. (Group 1: WTd, CTd, Group 2: STd, SGP, NH, NGP)

The unemployment rate is more differentiated. We can not clearly observe the separation of groups, though we can state significant difference between regional means many times. (Those regions, which are underlined with a common line, do not show statistically admitted difference, while regions which do not have a common line, show statistically admitted difference.)

Since the analysis of variance proved the separation of regions in many cases, hereafter, we try to determine effect of territorial affiliation. We analyse how the county level values contribute to the standard deviation of the regions. In table 7, the variance relation reflects this relationship.

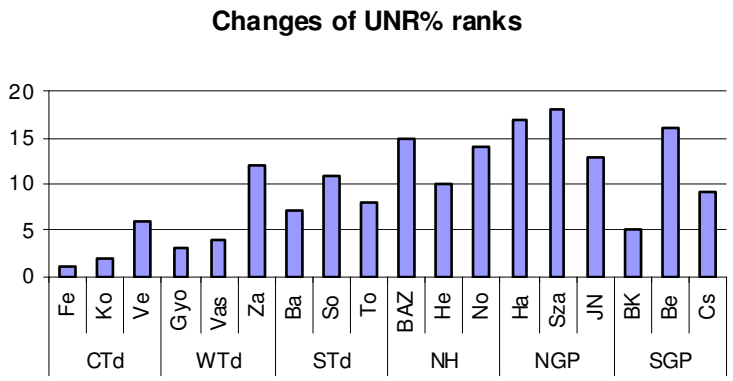
Table 7. The relationship between labour market indicators and regions

Indicator	Variance relation	Strength of relationship
UNR%	H^2= 0.731	H= 0.855
EMR%	H^2= 0.784	H= 0.885
Active%	H^2= 0.746	H= 0.864

Source: own creation

The variability of unemployment rates were determined by the region by 73%. The largest ration (H^2) belongs to the employment rate with 78.4%. In all the three cases, the relationship between the region and the labour market indicators are significant and strong. Consequently, the territorial affiliation strongly determines the labour market indicators, though other factors may also contribute to the outcomes. The rate (effect) of other factors is around 25%.

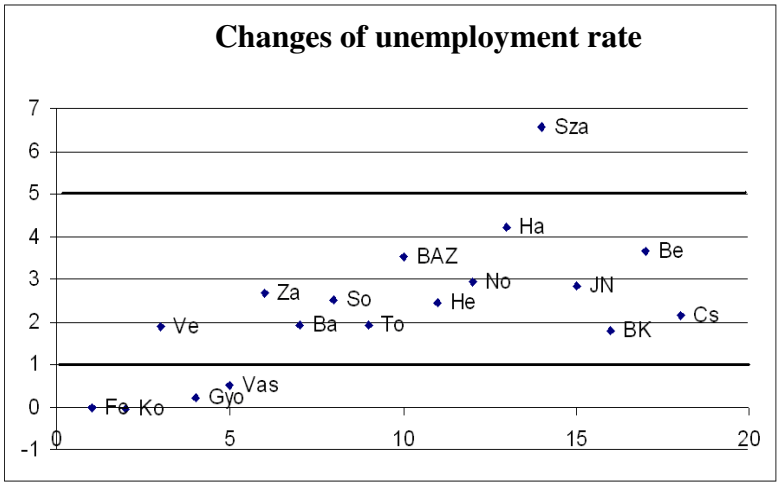
Figure 9. Ranking the changes of county level rank-score values



Source: own creation

Based on *Figure 9* we can state, that largest change occurred in Northern Great Plain and Northern Hungary regions, while values of Central Transdanubia region showed the slightest change.

Figure 10. Changes (by percentages) of the unemployment rate by county



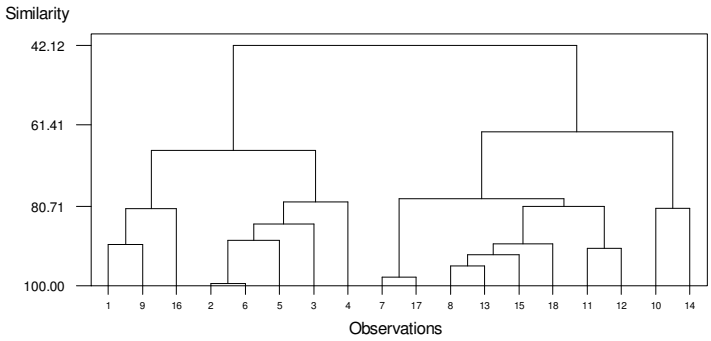
Source: own creation

The three groups of development can be clearly seen on the scatterplot above. The first group comprises those counties, where the unemployment rate changed by

less that 1%. Second group consist of counties, was between 1% and 5%, while the counties of the third group developed by more that 5%.

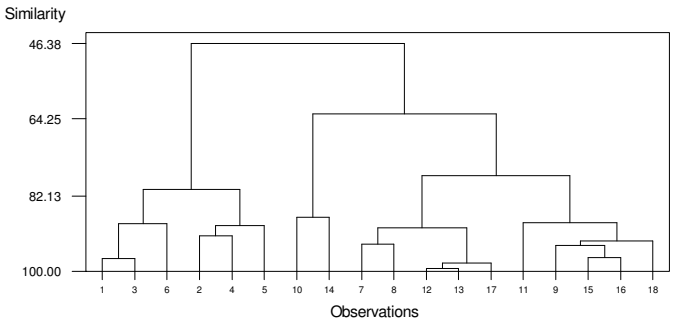
In the next part, we present the clusters of counties using the major labour market indicators. *(Figure 11)*

Figure 11. Clusters of counties using the three major labour market indicator (2004)



Source: own creation

Figure 12. Clusters of counties using the three major labour market indicators (2008)



Source: own creation

The two dendograms – based on cluster analysis - show the clusters of counties as well as the occurred changes. Based on the dendogram, we can form 4 groups, which is illustrated with boxplot on *Figure 13*.

The groups of the two time periods (2004 and 2008) comprise different counties (*Table 8*).

Table 8. Grouping of counties by the three major labour market indicators in 2004 and 2008

	Group members in 2004		Group members in 2008
1	**Fe, To, BK**	1	**Fe**, Ko, Ve, Gyo, Vas, Za,
2	Ko, Ve, Gyo, Vas, Za,	2	Ba, So, No, Ha, Be
3	Ba, So, No, Ha, Be, He, Jn, Cs	3	**To, BK**, He, Jn, Cs
4	BAZ, Sza	4	BAZ, Sza

Source: own creation

Table 9. Comparison of formed groups in 2004

Indicator	Group 1	Group 2	Group 3	Group 4
UNR%	Medium	Low	Medium	High
EMR%	High	High	Low	Low
Active%	Rather high	High	Rather low	Low

Source: own creation

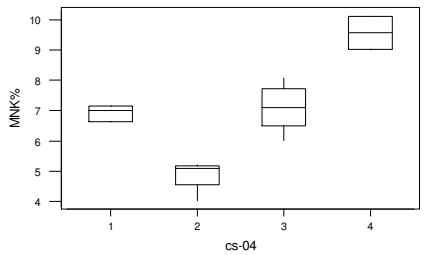
Table 10 .Comparison of formed groups in 2008

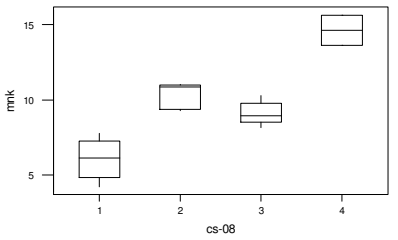
Indicator	Group 1	Group 2	Group 3	Group 4
UNR%	Low	Medium	Medium	High
EMR%	High	Worse than medium	Medium	Low
Active%	High	Low	Medium	Low

Source: own creation

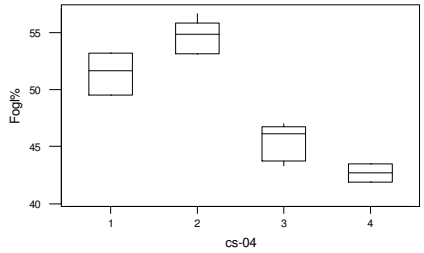
Figure 13. Boxplot illustration of the groups

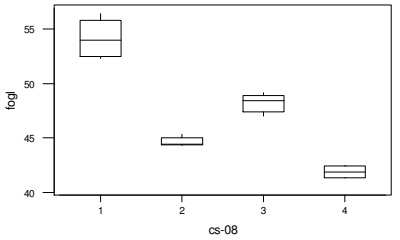
2004 **2008**



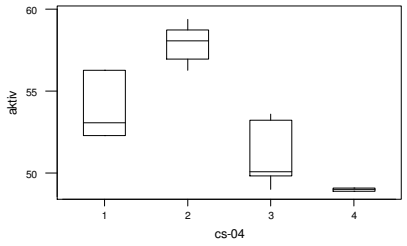


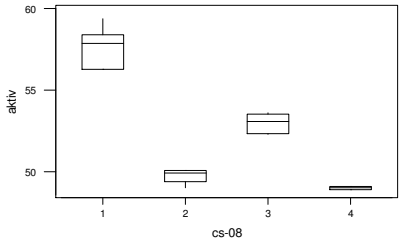
Unemployment rate





Employment rate





Activity rate

Source: own creation

4. Conclusion

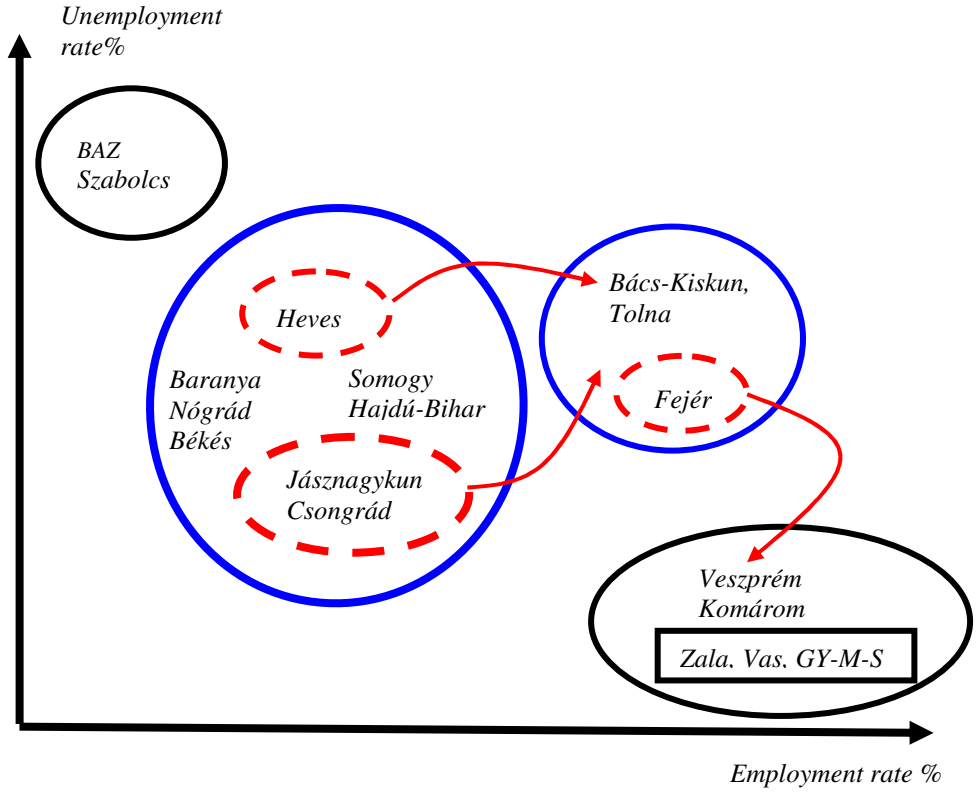
One of the main findings of our analysis is that specific clustering of the six examined EU regions can be observed in both years. The first cluster contains Central Transdanubia and Western Transdanubia and the second cluster with significant differences includes South Transdanubia, Southern Great Plain, Northern Great Plain and Northern Hungary.

The analysis of the data of NUTS 3 regions shows much more layered result. The NUTS 3 clusters based on labour market indicators do not match the clusters of NUTS 2 regions. The previously most developed cluster of NUTS 3 regions, i.e. the counties of Western and Central Transdanubia, continue to hold their high place by the two latterly analysed aspects, the labour market indicators and the administrative sectioning as well. On the other hand all the other clusters of counties includes NUTS 3 regions regardless the previous clustering of NUTS 2 regions. (*Figure 14*)

The round-shaped clusters of counties on the Figure 14 show the enclavement of year 2004. The rugged lined circles show the changes by year 2008. It can be clearly observed, that the cluster of less developed counties, i.e. Borsod-Abaúj-Zemplén and Szabolcs-Szatmár-Bereg, has not been changed.

The most advanced NUTS3 regions of Western Transdanubia, based on their labour market indicators, form a common cluster with the two counties of Central Transdanubia.

Figure 14. Groups of counties, and the changes of groups within the examined time period



Source: own creation

By the year 2008 the missing Fejér county also closes up to the cluster. Therefore, the cluster of most developed NUTS 3 region includes two entire administrative NUTS 2 regions.

In 2004 three counties of different NUTS 2 regions, Bács-Kiskun county (South Great Plain), Tolna county (Southern Transdanubia) and Fejér county (Central Transdanubia) form a cluster. From this cluster Fejér county closed up to the cluster of more developed Central Transdanubian counties, i.e. it moved to the cluster of its real administrative NUTS 2 region.

The main reasons of the advance of Fejér county is that, while its unemployment rate is mildly decreasing, the rate of the other counties of Western and Central Transdanubia (expect Komárom-Esztergom county) are distinctly increasing.

Since their rate of employment has raised, Heves, Jász-Nagykun and Csongrád counties have been moved to a better cluster, to the same group where Bács-Kiskun and Tolna counties are.

The differentiation showed in other studies of regional competitiveness and development indexes of Hungary is well observable in our present research on labour market indicators as well. The differences between more and less developed regions are increasing in Hungary.

References

Baranyai, J. 2007: Aktuális munkaerőpiaci ismeretek. pp 48. (http://www.z-szem.hu/letoltheto_elemek/baranyi_jozsef-aktualis_munkaeropiaci_ismeretek.pdf)

Cseh Papp, I. 2008: Spatial labour market processes. In: SZIE Bulletin special issue Part I. pp.45-53.

Pál, Gáspár szerk. 2005: A hazai régiók versenyképességi lehetőségei: a versenyképességi koncepció regionális vetületei. ICEG Munkafüzetek 12. pp.49.

Lőkös, L. yr – Lőkös, K. 2005: A regionális munkanélküliség kistérségi differenciálódásának vizsgálata a 2003. évi T-Star adatai alapján. Környezetvédelem, regionális versenyképesség, fenntartható fejlődés c. konferencia előadásai Pécs, I. kötet. pp.229-236.

Vajsz, T. – Pummer, L. 2006: A magyarországi régiók közötti különbségek vizsgálata. Gazdálkodás (year 50.) Vol. 4., pp.18-30.

Judgement of employment by the methods of systematic data collection

Dr. Zsuzsanna Dabasi Halász[1] *– Katalin Lipták*[2] *– Eszter Siposné Nándori*[3]

Our research aims at examining the opinion of the Hungarian population about employment. We would like to define what factors are mainly associated with employment and what the exact meaning of these factors is. To carry out our research we have used the techniques of systematic data collection.

This method, developed in the 1980s by anthropologist Susan C. Weller, and mathematician A. Kimball Romney in the United States, help researchers in the social sciences collect better interview or questionnaire data. This slowly spreading qualitative research method of systematic interviewing – where each informant is asked the same set of questions – diminished the sample size required in social science research in a revolutionary way.

In our paper we present a research overview about how to examine employment with the method of systematic data collection.

Key words: systematic data collection, interview techniques, employment, sampling

1. Introduction

In spite of many favourable macro-economic processes the man in the street feels that the nadir of global recession is still ahead. The Eurobarometer survey conducted in the EU 27 member states indicates that citizens regard unemployment as the main menace. At the same time, their anxieties concerning the economic situation have become somewhat less.

That is why communicating the objectives and tools of employment policy in the most comprehensible way possible receive heightened significance. It is important that the man in the street understands and makes use of the opportunities and

[1] Dr. Zsuzsanna Dabasi Halász Ph.D., assistant professor, University of Miskolc, Faculty of Economic, World- and Regional Economic Institute
[2] Katalin Lipták, Ph.D. student, University of Miskolc, Faculty of Economic, World- and Regional Economic Institute
[3] Eszter Siposné Nándori, junior assistant professor, University of Miskolc, Faculty of Economic, World- and Regional Economic Institute

The research of Katalin Lipták has been supported by the "Közösen a Jövő Munkahelyeiért" Foundation.

thus his/her sense of uncertainty decreases; (s)he helps economy get over the crisis by way of optimistic anticipations. We argue that the Hungarian employment policy, its institutional system, the official databases and the researchers use numerous definitions, while the society interprets the particular phases differently. It is the task of the legal system and the profession to use the concepts developed in a given era.

In this study we introduce a method for labour market research that enables faster reaction at lower investment in terms of time and money. It means a fast and flexible feedback between the hypothesis and research findings and formulating a more comprehensible economic policy.

It leads to stories, we can generate processes, which are effective and can correct the hypotheses, extremely far from reality, of theory and politics; or at least warn those people who try to seek connections between the professional theorems to the concrete processes that they are wasting their time.

The reason for choosing the method we use – briefly introduced in chapter 3 – is to attempt to introduce a fast, flexible and cost-effective process for labour market analyses. The method basically is a sampling survey, its information are not deeper, however it provides opportunity for quick report-like surveys.

2. Employment situation

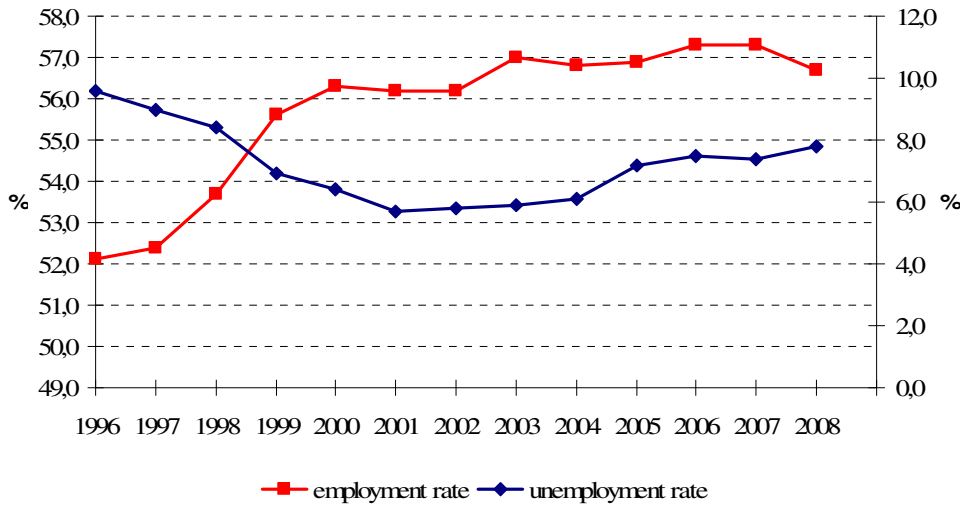
One of the mid-term priorities of the Hungarian employment policy, aiming at enhancing the long-term stagnating employment, is that the expansion of employment should be facilitated by way of letting inactive population return to the labour market and preventing unemployment to become permanent.

The token of success in the case of the European and thus the Hungarian employment policy is that the actors of the economy can understand and operate it. The labour market problem of the core member states of the European Union and those problems experienced in new joiners are similar. European labour markets – as opposed to the American one – are usually regarded as inflexible and sclerotic because the employment is lower, the regional and industrial crises are more permanent, it takes longer for the unemployed to find job, and the wages and salaries adapt to changes in the demand over a longer period of time. Hungary occupies an average position in the region according to the labour market indicator used in the comprehensive World Bank study (2002) and the Cazes-Nesporova (2003) book. Similarly to other countries in the region, the employment of those completed primary school is extraordinarily low, the regional differences are high and permanent, and the taxes on labour are relatively high.

The Hungarian economy got into a slump in the fourth quarter of 2008, as a result of the recession of the world economy that began to improve only in the second half of 2009. The nadir in 2009 brought about a decrease in living standard, it also increased income discrepancies. The setback of the internal demand, which was

even greater than that of the external demand, entailed considerable real sacrifices. The rate of unemployment was also at its lowest point, the number of long-term unemployed significantly grew.

Figure 1. Labour market situation in Hungary



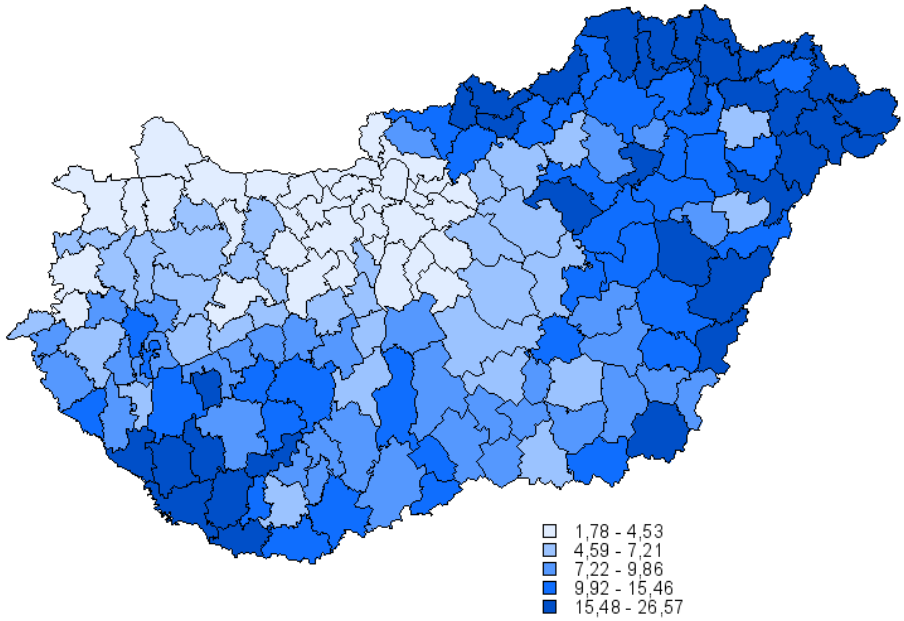
Source: own compilation on the basis of Eurostat data

The value of the two main labour market indicator, in the chart displayed in Figure 1, has showed a similar trend to the other ex-socialist countries. Putting the rates of unemployment and employment in the same system of co-ordinates demonstrates that the trend changed to opposite between 1998 and 1999.

Hungary is different from other countries of the region in two respects: the job-seeking intensity of unemployed men (with special regard to those with low level of education) is extraordinarily low; some sources mention that the costs of launching new enterprises are high. Apart from the general European objectives (supporting the disabled, fostering life-long learning and fighting tax-avoidance), employment policy has to concentrate on the above areas in Hungary including making efforts to find out, using more profound data-collection and analysis, that whether the picture, developing from the elementary statistics and the available research findings, is precise enough.

The rate of unemployment well demonstrates the regional differences in Hungary (Figure 2).

Figure 2. Rate of unemployment in Hungarian subregions (2007)



Source: Resource map, Hungarian Academy of Science

First of all, significant changes are necessary in the institutional system of labour-market policies, in the inter-institutional co-ordination, and in the integrated operation of the organisation that puts the employment policy into practice. It is common that the institutional system of the labour policies is, similarly to other countries, scattered in Hungary and there is much to improve in the area of inter-institutional co-ordination as well. The various organisations use different concepts, and the impacts of the various measures, which are well known and properly interpreted by experts, fail because of the social incomprehension and misunderstandings. The various concepts are not only mixed up in the surveys conducted in the specific member states but also in the international comparative researches.

3. Methodology

Items expressing the general belief about the terms of employment can be found using qualitative research methods. First, a list of items related to employment is elicited from informants, and then they are asked to rank these items to find the ones most related to the field of interest.

The first step in any study is to get a clear understanding and boundaries of what is being studied. To do this, informants have to be asked to list the items be-

longing to the domain of interest. Free listing is a good way to ensure that the domain and the items are culturally relevant. The most important use of free listing is to ensure that the researcher is dealing with relevant items and to find the boundaries of the domain (Weller - Romney 1988). In this study, free listing is going to be used to get a list of items for further research, and to obtain understanding on what people think of the terms of employment. The sample size necessary for free listing was determined by taking into account the amount of agreement in the responses of the informants. Stability in order usually can be found at around 20 or 30 informants, so a sample size of 20 or 30 is in general sufficient. Informants are going to be chosen using multistage cluster sampling. The list of unemployment-related items is going to be elicited by asking the following questions: 'Who do you consider to be unemployed in general?', 'Do you know unemployed people?', 'Why do you think they are poor?'. Similarly, items related to social aids can be collected by asking: 'What kind of social aids do you know?', 'Do you know of anybody who is entitled to social aids', 'What kind of aid does he/she get?', 'Why do you think he/she is entitled to it?', 'Do you think that social aid can help him/her in solve his/her problem?'. Based on their answers, further answers can be asked if necessary.

This study has two aims: finding out (1) what the factors most related to the term unemployment and social aid are; and (2) what the exact meaning of long and short term unemployment are. The first question can be answered by ranking the items from free listing. As ranking a large number of items could be difficult and could take a lot of time, special methods are used to simplify the task. The method of the analysis is thus systematic data collection, developed in the 1980s by the anthropologist, Susan C. Weller, and the mathematician, A. Kimball Romney in the United States. This method helps researchers in the social sciences collect better interview or questionnaire data. The goal of social sciences, that is the better understanding of experimental and observational data, requires careful analysis of data. Increased understanding requires systematic observation, classification, analysis and evaluation. Structured interviewing formats help to reach this goal (Weller - Romney 1988).

The qualitative research method of systematic interviewing – where each informant is asked the same set of questions – diminishes the sample size required in social science research in a revolutionary way while the reliability of the results is still as high as in case of traditional techniques. This is possible by taking into consideration cultural competence of the population in defining the sample size.

Consensus theory is used when the researcher does not know what the answers are or what they should be and, instead, tries to discover the "culturally correct" responses to the questions. This theory helps to determine the number of informants necessary to get reliable answers in cases when the answers to the questions are not known ahead of time. It gives the possibility to measure the cultural competence of informants (the probability that the informant knows the answer to a given question) and it allows to reconstruct the "culturally relevant" answers to a

specific question (Romney et al. 1986). It is assumed that the correspondence between the answers of any two informants is a function of the extent to which each is correlated with the truth, for example with the culturally correct answers (Nunally 1978). Cultural competence is the result of the socialization process, which refers to the incorporation of social effect and beliefs to the individual's behaviour and personality (Vajda 1999). Roberts (1964) states that in any culture information is stored in the minds of its members and in artefacts. This study focuses on the part of the culture that is stored in the minds of its members.

According to consensus theory, the number of informants needed depends on the average level of competence, the confidence level and a minimum rate of questions we would like to classify correctly (Weller - Romney 1988). The average level of competence can be defined on the basis of free listing and then consensus theory can help to define the sample size.

If at least 90% of the questions should be classified correctly at the 0.95 confidence level, the reference to Table 1 shows the sample size.

Table 1. Minimal Number of Informants Needed to Classify a Desired Proportion with a Specified Confidence Level for for Different Levels of Competence

Proportion of Questions	*Average Level of Competence*				
	0.5	0.6	0.7	0.8	0.9
0.95 Confidence Level					
0.80	9	7	4	4	4
0.85	1	7	4	4	4
0.90	13	9	6	4	4
0.95	17	11	6	6	4
0.99	29	19	10	8	4
0.99 Confidence Level					
0.80	15	10	5	4	4
0.85	15	10	7	5	4
0.90	1	2	7	5	4
0.95	3	14	9	7	4
0.99	*	0	3	8	6

* Well over 30 informants needed

Source: Weller, S. C - Romney, A. K.: Systematic Data Collection. p 77.

Informants then can be selected using the method of multistage cluster sampling with stratification (Maxfield - Babbie 2009).

The general belief about unemployment and social aid can be investigated with the methods of triadic comparisons, quicksort or balanced-incomplete block designs. The choice of the method to be used is determined by the number of items elicited in free listing.

After establishing the list of items, triad data may be collected by first enumerating all sets of size three for the items. The next step is to randomize the triads, both by position within the triad and by triad order. With a triad form ready, you can collect data either orally or with a questionnaire. Informants are asked to order the items within triads from "most" to "least" on some attribute. As for unemployment, informants would be asked to choose the item most related to unemployment. Then, from the remaining two items they are asked to choose again the one that is most related to unemployment. The total number of triads, however, goes up very rapidly with an increase in the number of items, so this data collection format is only practical with 10 or fewer items. The total number of triads can be reduces by using a balanced incomplete block design (Weller - Romney 1988).

Balanced-incomplete block (BIB) designs systematically compare subsets of items. The designs work by controlling the number of times that each pair is compared. By reducing the number of times each pair is compared to other items, the total number of subsets is reduced, while still maintaining comparisons among all items. BIB designs are identified with three parameters: n, the number of items; lambda, the number of times each pair occurs; k, the number of items in each set or block; and b, the number of sets or blocks. Even using this data collection format, the method becomes impractical with more than 25 items (Weller - Romney 1988).

Comparison of a large number of items can be done with quicksort. It minimizes the number of paired comparisons by assuming transitivity. It means that if an informant has judged A to be greater than B and B greater than C, then you assume that A is greater than C and you do not ask any questions about that pair. Names of items are written on cards which are first randomized and a card is selected as a 'standard'. All cards are compared to the standard and are divided into two piles: the cards 'greater than' and those 'less than' the standard. This process is repeated for each pile, until all items are ordered (Weller - Romney 1988). In case of unemployment related items, informants would be asked to divide them into two piles: items that are related more and related less to unemployment.

To answer the second question, namely the exact meaning of short and long term unemployment, rating scales can be used. They are the most widely used methods to collect data in written format as they work best with literate informants. Scales are usually expressed as four- to eleven-point scales. The more points a scale has, the more reliable it is said to be (Weller - Romney 1988). To find out the general belief about long term unemployment, informants may be asked the following question "What is the minimum time period for which somebody has to be unemployed in order to be considered long term unemployed?". A similar question can be asked for short term unemployment ("For how long can you be unemployed at longest to be considered short term unemployed?") to control for the potential inconsistency in the answers.

Eszter Siposné Nándori tested the model referring to poverty. In fact, there is no single exclusive interpretation of poverty and that is the reason why both the size

and proportion of poverty calculated on the basis of the various theoretical concepts are different. The essence of the systematic data collection methods is that every interview respondent receives the same questions. The objective of her research with these types of questions was to identify those factors the respondents would consider to have the closest (also less close or no) relationship with poverty (Siposné 2008). She conducted a traditional data collection as well in order to test the reliability of the estimation and hypothesis testing, in the course of which she used a sample size generally accepted in social sciences. She examined, by way of contrasting the findings of the two data collections, whether the systematic data collection using small sample can reach the same level of reliability as the traditional methods.

Research findings: the sample she used was extremely small; she tested the reliability of the method by a similar research conducted on a larger sample. Having contrasted the two research findings it has become clear that they have lead to the same result in the case of all variables examined. Therefore, the assumption behind the systematic data collection, that is one can draw reliable conclusions from smaller samples than usual if the average competences of the respondents are taken into consideration, is reasonable. Average competences of respondents mean here the extent to which they think in a similar way of the concepts used in the survey.

A great advantage of systematic data collection regarding the clarification of employment-related issues, is that it decreases the necessary sample size, and thus, it reduces the time and cost to be spent on the research. Whether or not the method can really be applied in surveying employment and unemployment could be ascertained once we have the result of the, so called, free-listed enumeration. The average competence of data providers, that has to reach a critical value so that later phases of the method are applicable, can be established only on the basis of this.

4. Research plan aiming at the clarification of the concepts of employment policy

4.1. Concptualization, operationalization and Itelszon's hypothetic-deductive method

By our research plan we use the conceptualization and operationalization. Conceptualization is a process by way of which we define what mean by particular expressions. In the course of conceptualization we specify what mental pictures are related to our concepts and we choose the observations and measurements are appropriate for our research purposes.

The following concepts need precise definition in our research:

- Unemployment
- Employment
- Economic meaning of employment policy tools

- Effect of employment policy tools on national and regional development

While conceptualization means refining and specifying abstract concepts, operationalizing means developing those concrete research procedures that take us to real empirical observations using these concepts, that is, we define concrete empirical procedures leading to measuring the variables.

Having set the main objectives, we have to identify the most appropriate method to reach them. In this case we chose the previously introduced method of systematic data collection. The research is new also in that this method has mainly been used for sociological researches, and not for labour market and economic analyses.

We used Itelszon's hypothetic-deductive method in our research, this method includes six phases. (Itelszon 1967)

1. We assume that the economic performance of the national economy relates to the efficiency of regional employment policies. The central tools do not get to the heavily problematic segments in the case of specific countries.
2. We are going to examine the unsuccessfulness of employment policy.
3. According to our hypotheses the weight of regional employment policies has to be significant and the have to respond to individual problems. Communication faults play role in the efficiency of the employment policy tools. The labour market definitions do not cover the population's concepts. The concept of employment, activity and unemployment are undefined in the society.
4. We can critically argue that regional competitiveness is improved by strategies focusing on individual problems. Its success requires strategy-awareness.
5. In a crucial experiment we are going to make a comprehensive comparison using regional employment policy tools.
6. As a theorem, we wish to assert that the various regions in Europe are in different labour-market situations; therefore, not only EU and national employment policy are necessary but also more explicitly defined regional employment policy.

In ordinary people's thinking employment means that one has a job or not. On the contrary, unemployment means the lack of job – in over-simplified terms. That is why we hold it important to ask, by way of the method introduced above, the man in the street about the most important questions and definitions relating employment and unemployment.

The purpose of our research is to match the meaning of the concepts emerged in employment policy with the colloquial language of the man in the street. Our reason for establishing this purpose was that, according to our hypothesis, the professional definition and the population's definition concepts mismatch. It also

leads to the insufficient efficiency of the economic policy measures because the population cannot make use of these measures due to misunderstanding. The axiom taken from business management suggests that every strategy is worth the extent to which it is implemented.

The indirect result we wish to reach is to explore information that can help economic policy actors communicate their measures more precisely to the population. The direct result will be the population's interpretation of the labour market definitions.

4.2. Definitions relating unemployment

The literature knows and uses many and basically similar definitions for both employment and unemployment. Below give an incomplete overview of the most well known ones.

"Unemployment, as an overall social phenomenon has existed in the world since capitalism became a dominant form of production" – suggests Bánfalvy (1997) in the introduction of his book.

Unemployment has gradually reached its modern meaning which demonstrates the evolution of the concept: "That person can be regarded as unemployed who would occupy a job, but currently (s)he has not got a paid job". (Galasi 2003)

Those bereft of hope are also worth mentioning who, although unemployed, do not seek job because of the lack of employment opportunities. Hidden unemployed often get out of the traditional unemployed category and are regarded as inactive then.

Theoretically an individual is unemployed if (s)he, among given labour-market conditions, is willing but unable to undertake a paid job. The situation is somewhat more complicated from an empirical perspective. A person is unemployed, according to ILO, if the following conditions apply:

- (s)he did not have a job during the week before the survey,
- although (s)he actively sought job and
- would have been available for work. If (s)he spent at least 1 hour work for wage or salary the week before the survey then (s)he is regarded as employed. If (s)he had a paid job for shorter than 1 hour or did not seek job, then (s)he is inactive.
- A person is registered unemployed if (s)he has been registered with a job-agency. The one who is eligible for unemployment allowance is called insured unemployed. Their proportion might show a distorted picture because this category may include those who do not want to undertake job or may exclude those who seek job but do not register.

4.3. Range of tasks of employment policy

Employment is, on the one hand, shaped by the market and, on the other hand by the governmental policy influencing it. Making employment policy work is the joint

competence of the governmental, local-governmental and business sphere, in which the government relies on the contribution of the actors of the economy in order to accomplish the employment goal. The tasks of the employment policy can be illustrated by the enumeration below. Employment policy is an external regulatory system of the labour market aiming to influence the labour market in a way that:

- its operation becomes more flexible and harmonic,
- it fosters the establishment of labour market equilibrium,
- the implementation of social functions also receives attention,
- the size and professional/educational structure of the labour supply entering the labour market meets the requirements of the demand,
- the labour supply decreases, e.g. the increase of the number of participants in education or the length of education,
- it continuously adjusts labour supply to demand as perfectly as possible,
- it helps develop the size and structure of labour market demand that is capable of employ workforce appearing as labour supply to the greatest extent possible,
- it supports the living, on the basis of the principle of social care and solidarity, of those who exited the labour market through no fault of their own.

4.4. Questions asked in the interview

The method of systematic data collection can start with the questions below. Since the objective of our study is to introduce the methodology, we have not conducted the actual survey yet; we are going to elaborate on the findings of our future research in a subsequent paper.

- Who do you think is unemployed?
- What is the minimum period of time spent unemployed by those regarded as permanently unemployed?
- Up to what age can someone regarded as of working age?
- What does “black work” mean for you?
- Up to what age do you consider someone to be an entrant in the labour market?
- Who do you think are bereft of hope in terms of employment?
- What do you mean by the application of active employment policy tools?
- What are the characteristics of the traditional forms of employment?
- What is the minimum working time of full-time employees in terms of hours?
- Below how many hour of working time per week can be somebody regarded as part-time employee?

5. Conclusions

We have introduced the theoretical background of the systematic data collection in our paper, we approached the issue only from a methodological perspective in our study. This method, developed in the US, is used to define the concepts of social sciences, including economic sciences, that are difficult to be defined.

The advantage of this method is that it decreases the necessary sample and, at the same time, the level of reliability of the data is the same as that of those research techniques using large samples.

We have developed the research plan for employment policy and unemployment, and we intend to use the systematic method of data collection in our future research.

We regard the method as acceptable, as Eszter Siposné Nándori has tested systematic data collection for poverty surveys and she received similar findings as using the traditional data collection. We suggest that the method of systematic data collection may make researchers' work easier, although the focus of the research is of key importance.

References

Bánfalvy, Cs. 1997: *A munkanélküliségről.* Magvető könyvkiadó, Budapest.

Cazes, S. – Nesporova, A. 2003: *Labour markets in transition: Balancing flexibility and security in Central and Eastern Europe.* International Labour Office, Geneva.

Galasi, P. 2003: *Munkanélküliségi indikátorok és az állásnélküliek munkaerő-piaci kötődés.* Budapesti Munkagazdaságtani Füzetek.

Global 2002: *Global Competitiveness Report*, The World Bank, Washington DC.

Itelszon, L.B. 1967: *Matematikai és kibernetikai módszerek a pedagógiában*. Tankönyvkiadó Vállalat, Budapest.

Maxfield, M. – Babbie, E. 2009: *Basics of research methods for criminal justice and criminology* (2nd ed.). Belmont, CA: Wadsworth.

Nunally, J. C. 1978: *Psychometric Theory*. McGraw-Hill, New York.

Roberts, J. M. 1964: The Self-Management of Cultures In Goodenough, W.H. (ed.): *Explorations in Cultural Anthropology.* McGraw-Hill, New York. 433-454. p.

Romney, A. K. -Weller, S. C. – Batchelder, W. H. 1986: *Culture as Consensus: A Theory of Culture and Informant Accuracy*. American Anthropologist, New Series, 88/2. 313-338. p.

Siposné Nándori, E. 2008: Judgement of Poverty in Borsod-Abaúj-Zemplén County According to the Method of Systematic Data Collection. *MicroCad 2008*, 187-193. p.

Vajda, Zs. 1999: Somlai Péter: Szocializáció. *Szociológiai Szemle*, Vol 2., 166. p.

Weller, S. C. - Romney, A. K. 1988: *Systematic Data Collection.* Qualitative Research Methods, Volume 10, Sage Publications.

Comparing native versus immigrants' occupational choices of the Italian labour force: a generalized linear mixed model approach

Luca Bartoli[1] -MariaChiara Pagliarella[2] -Laura Palombo[3] -Renato Salvatore[4]

In this paper a short-term analysis of the occupational behaviour of the Italian labour force is proposed by analysing official Istat micro-data of 2006 through a logit model with subject-specific intercept. The aim is to compare the labour choices of autochthon workers to those of immigrants also considering the native country. Many national and international authors have been dealing with this interesting topic mainly in an economic key, treating the problem of defining the phenomenon of labour migration as a cost or as a benefit for the destination Country. Moving from the assumption that the complexity of the phenomenon and the lack of qualitatively acceptable data create such big limitations and delays in the formulation and the elaboration process of meaningful solutions, our proposal is an "a posteriori" analysis of labour choices taking into consideration the intrinsic demographic characteristics of native and non native individuals. The idea is to clarify if native and non-native workers share the same labour choices and expectations in terms of qualification grade as a proxy of their satisfaction. To this extent, in this paper firstly the hypothesis that native and non native workers show the same expectation will be tested, secondly the occupational behaviour of immigrants and autochthons will be compared on the basis of derivation country.

Keywords: Immigrants, labour, mixed model

1. Introduction

This contribution analyses the sample of the Italian Labour Force with specific emphasis on the distribution of workers per country of origin. The goal is to highlight the presence of systematic relations between the socio-demographic characteristics of workers and their occupational choices; each research question of the study is investigated, underlying the differences in the behaviours of immigrant and autochthons.

[1] Luca Bartoli, University of Cassino
[2] Maria Chiara Pagliarella, University of Cassino
[3] Laura Palombo, University of Cassino
[4] Renato Salvatore, University of Cassino

Migration is one of the most urgent phenomenons as demonstrated by the most recent data issued by Istat on the international presence in the Italian labour population, showing once again in 2008 and in the first three months of 2009, the decrease in autochthon workers and the increase in immigrants. Although, the study of this social phenomenon recalls definitional and motivational questions. The first ones regard the difficulties in the individuation of the definition of "migrants" (Willekens 1984) mainly due to the following factors: quality of data; system of data collection; differences among countries; missing and unknown information (Kuijsten 1995). Moreover, different kinds of problems occur when dynamic (Natale - Strozza 1997) or stock data are analysed; in fact both approaches face meaningful limits whose nature changes on the basis of the specific object of analysis.

Many researchers have been dealing with the topic of international migration and their implications on local occupational levels. Those studies are mainly directed towards the analysis of income with the goal of analysing the effect of immigrants on local level of employment and the question point is if their presence can be considered as complementary or competitive. Even the estimation of the salary differentials and the determinants of the relative causes has been treated by several international (Borjas 1999) and Italian authors (Acocella - Sonnino 2003). Other studies concern the impact of the presence of international workers on the occupational opportunities of local population in terms of employment rates and living conditions and also the role of illegal immigration (Tronti 1995, Venturini 1996). The object of this contribution differs from those here reported; it can be contextualized in the demographic and economic research area of "migration theories" (Piore 1979, Massey 1990, Arango et al 1993). Those theories were formulated with the aim of specifying a socio-economic motivation with a scientific basis, towards the occurring of migration flows.

In this paper the definition of migrant is based on the concept of permanent or semi-permanent change of residence[5] (Lee 1965). The focus of the study relies on the analysis of the composition of the Italian Labour Force moving from the following assumption of the "Dual Labour Market Theory": foreign workers cover the shortages of the labour market of the receiving country in the lowest positions of the job hierarchy. For this reason, the main variable object of this analysis is the typology of profession and it was aggregated in two modalities: capital intensive and labour intensive professions. In order to explicit the socio-demographic factors influencing the decision to accept a capital intensive or a labour intensive profession, a specific preliminary descriptive analysis has been undertaken on the variables already tested in previous researches (Acocella - Sonnino 2003, Lutz 2008). The following variables were introduced in the analysis: gender, education level and age class. After the descriptive analysis of the socio-demographic variables included in

[5] Coherently with the Lee's definition, in this study, all foreigners (for birth or for citizenship) were considered immigrants for professional reasons.

the model as covariates, three models were estimated: the first is a classic logistic model aimed at investigating which socio-demographic variable, among those considered, has an affect on the probability to be involved in a labour intensive profession (which is complementary to the probability of a capital intensive job); the second and the third models add to the previous information the explicative function of the random intercept and the random slope respectively. Those models are aimed at observing the intra-clusters variability. The goal of the generalized mixed model estimation is to observe if the effect of the covariates on the dependent variable changes when the origin of workers is used as variables of aggregation.

2. "Dual labour" migration theory

The "Dual labour" theory proposes a different approach for the explanation of international migration in alternative to neo-classical theories. Its basic assumption is that demographic and social changes currently affecting modern societies of developed countries - such as the decline in t birth rates, the increase in divorce and legal separation, the ageing phenomenon, educational expansion, the economic level of development and the well-being of people - generate job vacancies in the lower positions of the job hierarchy (Piore 1979).

The dual labour theory describes the occurring of the migration phenomenon as the effect of such demographic and social dynamics. More precisely the decrease in fertility levels in industrialized countries might have led to smaller inflows of teenagers who are more willing to accept less satisfying job positions; the increase in the emancipation of women and the rise in divorce rates could have led to the increase in female employment rates causing the lack of workers able to substitute the female role in the household services. Finally, a higher instruction level created an increase in the supply of more qualified personnel profiles and, as a consequence of a higher perception of the life social status, the decrease in the availability to accept economically and socially less satisfying jobs.

This theory divides the labour market into a primary segment, characterized by a capital-intensive method of productions and a secondary segment characterized by labour intensive production process. By consequence, it states that skilled workers, holding social status, higher income and employment conditions in higher consideration, are in the first segment; on the contrary, unskilled ones are in the secondary segment. The "dual labour" theory hypothesizes that foreign workers migrate for the dual labour market structure of the receiving country where they are pulled to cover the lower positions of the job hierarchy. By using the Labour Force Istat Dataset this paper aims to test this basic assumption of the "Dual Market Theory" and, at the same time, to clarify the effects of some demographic characteristics on the final job decision of workers. This methodological approach

moves from the consideration that this specific database gives us *a-posteriori* information on employees' choices. In this sense our final considerations on the one hand can be or cannot be a confirmation of the hypothesis of the cited theory, on the other hand could add some other interesting points of research.

3. Informative instruments of the analysis

The study was conducted on the data of the Istat database of Italian Labour Force survey of the third trimester of 2006. The statistic unit of this survey is the family of the interviewed worker, for this reason the database also contains data of the other members of the worker's family and the observations refer to each family member in professional conditions and currently working. The final database[6] contains 57866 cases. The variables and the relative modalities included in the model are reported here.

The dependent variable:

1. profession (dichotomised) having the following modalities:
   - labour intensive;
   - capital intensive.

The following variables are included in the model as covariates and are reported here with the relative modalities:

2. age classes:
   - 15 – 24 years old;
   - 25 -34 years old;
   - 35 - 44 years old;
   - 45 – 54 years old;
   - 55 – 64 years old;
   - 65 – 74 years old;
   - 75 and over;
3. gender:
   - male;
   - female;
   - level of education:
   - none or primary school;
   - middle school or High school graduation;
   - academy for Fine Arts and Bachelor's degree;
   - degree studies;

[6] The database obtained after the elimination of all the units with at least one missing value in the analysed variables.

- master post graduate or doctorate.

Finally the last variable is the classification variable used for the aggregation of the observations in clusters. It aggregates the observations on the basis of Country of birth and Country of citizenship and counts the following eight modalities:

4. origin;
    - Italians;
    - born abroad and Italian citizenship;
    - Italian birth and foreign citizenship
    - born abroad and European citizenship (except Italy);
    - born abroad and Asian citizenship;
    - born abroad and African citizenship;
    - born abroad and North and Central American citizenship;
    - born abroad and South American citizenship.

4. Description of the sample

In this section, a brief description of the analysed sample on the basis of the distribution of the socio-demographic variables is proposed, in order to trace a first representation of the relations tested in the models. The demographic composition of the analysed sample provenance and citizenship is illustrated in the following table 1. It shows that the presence of immigrants in the sample is around 6.3%, 2.5 of which could be defined as “long term immigrants” being immigrants with Italian citizenship. The remaining percentage summing up to 3.9 %, is represented by “ more recently immigrated people”; this definition derives from the fact that they are those immigrated that still did not acquire the Italian citizenship. This last typology of immigrants was disaggregated by country of birth in order to observe if their behaviour differs throughout the clusters.

Table 1. Number and percentages of workers per country of provenance and citizenship

Origin/citizenship	**N. of presence**	**% of presence**
1 Italian	54195	93,7
2 Foreign born and Italian citizenship	1427	2,5
3 Born in Italy and Foreign citizenship	6	0,0
4 Born abroad and European Citizenship (except Italy)	1238	2,1
5 Born abroad and Asian citizenship	313	0,5
6 Born abroad and African citizenship	495	0,9
7 Born abroad and North and Central American citizenship	32	0,1
8 Born abroad and South American citizenship	160	0,3
Total	**57866**	**100,0**

Source: own creation

Table 2 illustrates a cross classification of the units on the basis of the following two variables: age class and the Country of origin and of citizenship. It shows that the largest proportion of workers concentrates in the central classes. The survey collects data on the national labour forces: by consequence, most of the units refer to workers not younger than 15 and not older than 64 , which is the conventional occupational age interval.

Table 2. Percentages of workers by classification of origin and age class

Origin	**15 - 24**	**25 - 34**	**35- 44**	**45- 54**	**55- 64**	**65- 74**	**75 and over**	**Total**
1 Italian	6	21	31	28	12	2	0	100
2 Foreign born and Italian citizenship	4	21	44	23	7	1	0	100
3 Born in Italy and Foreign citizenship	33	33	17	17	0	0	0	100
4 Born abroad and European Citizenship (except Italy)	10	33	32	21	4	1	0	100
5 Born abroad and Asian citizenship	7	24	40	23	5	0	0	100
6 Born abroad and African citizenship	7	28	46	18	2	0	0	100
7 Born abroad and North and Central American citizenship	9	28	28	28	6	0	0	100
8 Born abroad and South American citizenship	4	31	33	22	10	1	0	100
Total	6	21	31	28	12	2	0	100

Source: own creation

Table 3 shows the male presence ratios that refers to the percentage of males out of the females. It is very interesting to notice that immigrants from South and North America are mostly female, on the contrary, Italians, Europeans, Africans and Asians immigrated workers are more often males .

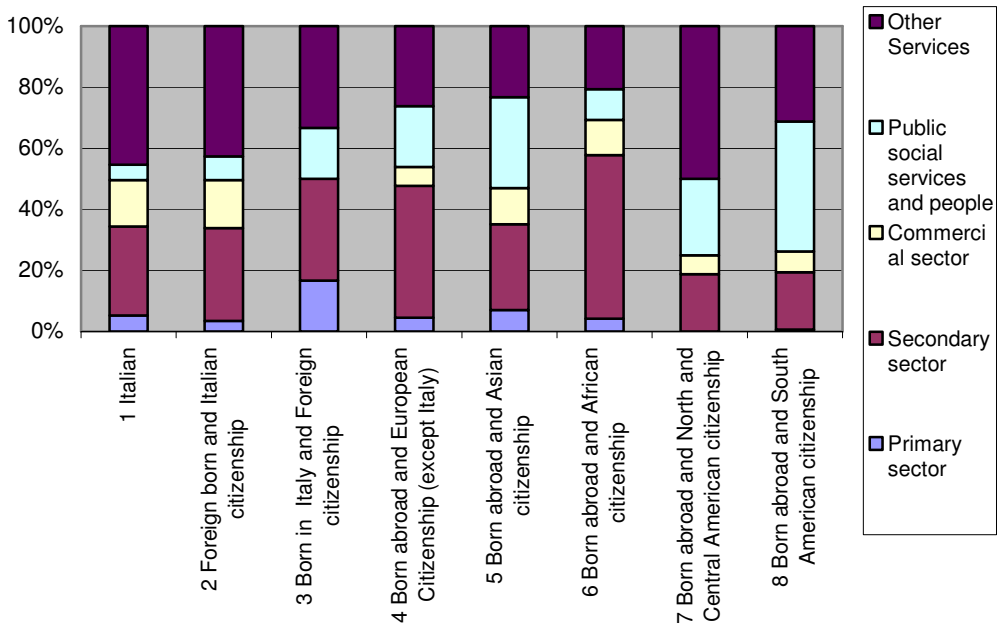
Table 3. Percentages of workers per Country of origin and citizenship

Origin	**Male gender ratio**
1 Italian	150,4
2 Foreign born and Italian citizenship	113,3
3 Born in Italy and Foreign citizenship	500,0
4 Born abroad and European Citizenship (except Italy)	113,4
5 Born abroad and Asian citizenship	192,5
6 Born abroad and African citizenship	358,3
7 Born abroad and North and Central American citizenship	52,4
8 Born abroad and South American citizenship	61,6
Total	149,1

Source: own creation

Figure 1 represents the distribution of the Italian labour forces per sector of activity; it shows that the larger proportion of immigrants are in the sectors of services to the persons and in the secondary sector. Both the sectors regard intensive labour jobs and the higher frequency of the variable has probably to be read considering the professions of cleaning ladies, old and baby sitters, and, in the secondary sector, mainly those in the construction services.

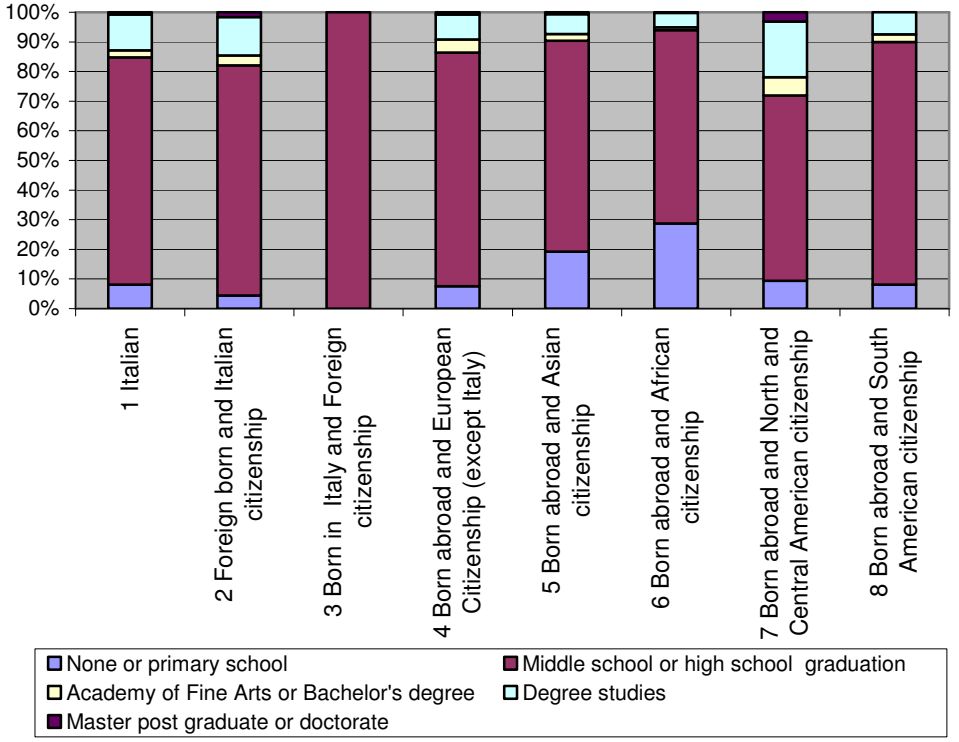
Figure 1.Percentages of workers per Country of origin and citizenship and economic sector of activity



Source: own creation

The following graph shows that the proportion of workers with lower education level is bigger in the clusters of immigrants, among which especially African and Asians show the most considerable percentages associated to the modality "none or primary school" instruction level. On the contrary, immigrants particularly from North and Central America are characterized by the biggest percentage of graduated and post-graduated workers.

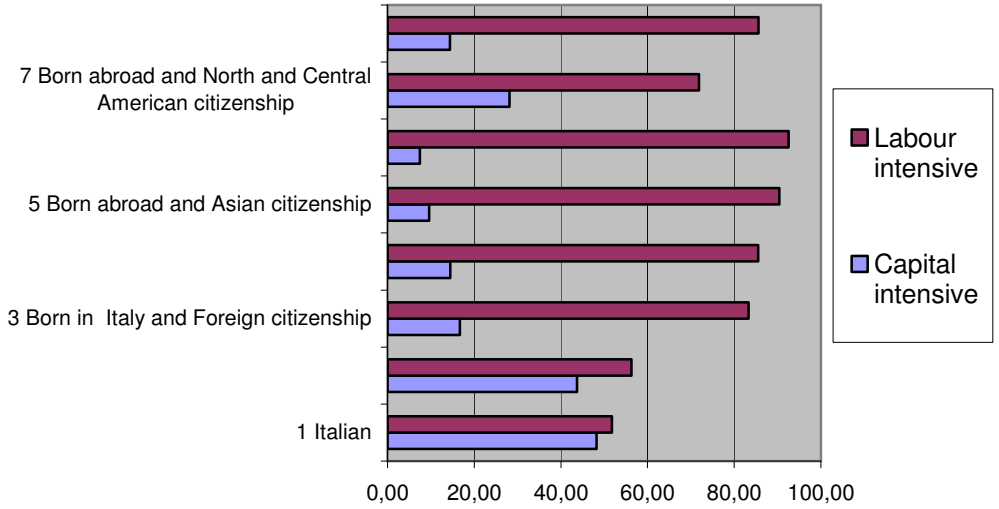
Figure 2. Percentages of workers per Country of origin and citizenship and education level



Source: own creation

The last figure shows that Italian workers strongly differ from immigrant workers considering the intensity of labour or capital: Italian native and citizens mainly work in capital intensive jobs.

Figure 3.Percentages of workers per Country of origin and typology of profession



Source: own creation

The descriptive analysis pointed out that more than one third of the immigrated workers has Italian citizenship. Moreover, the composition of the sample shows that immigrants without citizenship behave in a different way with respect to the country of origin; on the contrary immigrants with Italian citizenship have more similar characteristics to Italian workers. Such differentiations are investigated in the second part of this paper with particular attention to the role of the level of instruction.

5. The professional choices of Italian Labour Forces

The challenge of determining the relations between migration and labour market is ambitious and hard for several reasons, among which, limitations due to informative and methodological nature prevail. In this contribution the effort to avoid the problem of the absence of data was reached by using stock official data; the methodological efforts to use a model able to give a robust and wide interpretation of the social problems deriving from the occupational structure of the Italian labour market, was attempted by the use of multilevel analysis. It is a specific methodology for the analysis of data with complex patterns of variability, with a focus on nested sources of variability.

5.1. Research hypothesis

The logistic model tests the hypothesis that age and education level have a negative effect and male gender a positive effect on the probability to perform a labour intensive job

The research question of the multilevel models relies on the hypothesis formulated on the basis of the "dual labour market theory". This principle asserts that the probability of being involved in labour intensive jobs is bigger for foreign workers than it is for autochthon workers. In addition, the study has the goal to investigate to what degree, having Italian birth and citizenship, or being immigrated (and from where), changes the effect of socio-demographic characteristics on the probability to perform a capital intensive or a labour intensive job.

5.2. Methodological description of the models

The methodological part of this research consisted on the estimation of three models:

- a simple logistic model;
- a logistic model with a random intercept effect;
- a logistic model with a random intercept and a random slope effect.

It is supposed that there are k independent observations $(y_1; \dots; y_k)$, and that the $i-th$ observation can be treated as a realization of a random variable Y_i. It is assumed that Y_i has a binomial distribution

$$Y_i \sim B(n_i; \pi_i) \qquad (1.1)$$

with binomial denominator n_i and probability π_i. This defines the stochastic structure of the model. The distribution of Y_i is a *Bernoulli* distribution with parameter π_i, and can be written in a compact form as $\Pr\{Y_i = y_i\} = \pi^{y_i}(1-\pi)^{y_i}$.

The response variable y_i is binary, assuming only two values that for convenience we code as one or zero. Note that if $y_i = 1$ the probability is π_i, otherwise if $y_i = 0$ the probability is $1 - \pi_i$.

The expected value and variance of Y_i are respectively: $E(Y_i) = \mu_i = \pi_i$ and $var(Y_i) = \sigma_i^2 = \pi_i(1-\pi_i)$. Note that the mean and variance depend on the underlying probability π_i. Any factor that affects the probability will alter not just the mean but also the variance of the observations. This suggests that a linear model

that allows the predictors to affect the mean but assumes that the variance is constant will not be adequate for the analysis of binary data.

Suppose further that the *logit* of the underlying probability π_i is a linear function of the predictors $x_i'\beta$ as follows:

$$logit(\pi_i) = log\left[\frac{\mu_i}{1-\mu_i}\right] = x_i'\beta \qquad (1.2)$$

where x_i is a vector of covariates and β is a vector of regression coefficients, where β_j represents the change in the *logit* of the probability associated with a unit change in the $j-th$ predictor holding all other predictors constant.

The model defined in equations (1.1) and (1.2) is a generalized linear model with binomial response and link function *logit* with only a fixed effect.

The model used in this work is a mixed model for categorical response data including a usual fixed effects for the covariates plus the random effects on the intercept.

It is assumed that the sample get from N different clusters. Let $i = 1, \dots, N$ and $j = 1, \dots, n_i$ are respectively the indicators of the $i-th$ cluster and the $j-th$ individual observation. Let n_{ij} denote the total number of observations within the $i-th$ group. The response variable y_{ij} is binary, it can take only two values one or zero according to the presence or absence of the phenomenon of study and x_{ij} are the covariate (explanatory) variables.

Assuming that the incidence of a case of study differs from groups, an appropriate model that takes account of a random effect on intercept is a model with varied intercept from clusters, as follows:

$$\Pr(y_{ij} = 1|b_i) = \mu(b_i + \beta' x_{ij}). \qquad (1.3)$$

This is the *logit model*, a conditional probability model where:

- μ represents the probability function of y, and it is $\mu = \mu(s) = e^s/(1+e^s)$, function defined for all $s \in (-\infty, \infty)$;
- $b_i = (a_i - \alpha) \sim N(0, \sigma^2)$, where a_i is the iid random variable of the intercept, is normally distributed $a_i \sim N(\alpha, \sigma^2)$ with both parameters of population average intercept α and intercept variance σ^2 unknown;

- β is the $1 \times m$ vector of parameters to be estimated using the data (y_{ij}, x_{ij}) and the first component of this vector is the average population intercept;
- x_{ij} is the $1 \times m$ vector of the explanatory variables and the first component is 1 due to the presence of the random intercept.

This model also implies that observations y_{ij} and y_{ik} are independent within a cluster.

The random effect represent the influence of subject (o cluster) i on its repeated observations that is no captured by the observed covariates.

Including the random effects, the expected value of the response variable y, which is related to the linear predictor via the link function, is given by: $\mu_{ij} = E\left(Y_{ij} \middle| b_i, x_{ij}\right)$. This is the expectation of the conditional distribution of the study variable given the random effects, that here equals $\Pr\left(y_{ij} = 1 \middle| b_i, x_{ij}\right)$ the conditional probability of a response given the random effects and covariates.

5.3. Results: Italian labour market. Comparisons between the behaviour of autochthons and foreigners

The following tables show the results of the statistical part of the analysis, where the research hypothesis has been tested. The logistic equation obtained in all the three models confirmed the expected relations of the effect of education, gender and age on the probability of employees to be involved in a labour intensive job. More precisely, higher instruction and age levels are negatively correlated to labour intensive jobs; on the contrary, being a male worker, increases the probability expressed by the dependent variable.

Table 4. Logistic model without random effects

Dependent variable: Labur Intensive Profession (y=1)					
Intercept and Covariates:	**Beta**	**S. E.**	**T-Value**	**Alpha**	**Pr >|t|**
Intercept	5.02	0.06	80.06	0.05	<.0001
Instruction level	-0.84	0.00	-95.20	0.05	<.0001
Gender (male)	0.47	0.02	23.26	0.05	<.0001
Age	-0.34	0.00	-37.46	0.05	<.0001

Source: own creation

Table 5. Logistic model with random intercept

Dependent variable: Labur Intensive Profession (y=1)

Intercept and Covariates:	**Beta**	**S. E.**	**T-Value**	**Alpha**	**Pr >|t|**
Intercept	6,80	0.41	16.29	0.05	<.0001
Instruction level	-0.86	0.00	-95.67	0.05	<.0001
Gender (male)	0.49	0.02	23.92	0.05	<.0001
Age	-0.33	0.00	-36.13	0.05	<.0001
Variance	1,17	0.61	0.09	0.05	0.09

Source: own creation

Table 6. Logistic model with random slope (level of education)

Dependent variable: Labur Intensive Profession (y=1)

Intercept and Covariates:	**Beta**	**S. E.**	**T-Value**	**Alpha**	**Pr >|t|**
Intercept	4.59	0.20	22.49	0.05	<.0001
Instruction level	-0.43	0.09	-4.76	0.05	0.0031
Gender (male)	0.49	0.02	23.68	0.05	<.0001
Age	-0.35	0.00	-37.44	0.05	<.0001
Error term on intercept random effect	0.15	0.10	1.45	0.05	0.19
Covariance	-0.06	0.04	-1.36	0.05	0.22
Between clusters variance	0.05	0.02	1.87	0.05	0.11

Source: own creation

As shown in the tables, the three equations show the same effects of the covariates on the dependent variable. Although, the second model adds some more information to the study: it includes a first element of distinction among the eight groups through which the units of the sample have been classified per country of origin and country of citizenship. With the introduction of the intercept random effect, the different positions of the intercepts of each of the eight clusters can be observed and compared with the average of the overall intercept value. Finally, with the introduction of the random slope effect, for each group, the intra-clusters difference in the effect of the variable "education level" (that is the so called "between effect") on the probability of belonging to the capital intensive or to the

labour intensive segment of the labour market, is expressed by the increase or decrease in the slope of the function associated to each of the eight clusters. The predicted values obtained in the estimated models are those reported in the table 7.

Table 7. Predicted intercept and slope values for each group

CLUSTERS	PREDICTED VALUE		
	Intercept (R.I.M.)	**Intercept (R.S.M.)**	**Slope Ed.L. (R.S.M.)**
1 Italians	-1,83	0,70	-0,49
2 Foreign born and Italian citizenship	-1,37	-0,29	-0,15
3 Born in Italy and Foreign citizenship	-0,42	0,02	-0,03
4 Born abroad and European Citizenship (except Italy)	0,85	0,14	0,07
5 Born abroad and Asian citizenship	0,98	-0,19	0,19
6 Born abroad and African citizenship	0,73	-0,23	0,21
7 Born abroad and North and Central American citizenship	0,35	-0,20	0,09
8 Born abroad and South American citizenship	0,52	0,03	0,09

Source: own creation

The estimations of the predicted values gave a useful comprehension key to the analysis, adding information in terms of differences of occupational choices among workers coming from different Countries of emigration. The first reflection comes from the evident difference between the first three groups and the others. The intercept values as well as the slopes decrease if compared to the average value (shown in table 5 and in table 6) : it means that, in general, the probability of being involved in labour intensive professions is highly effected by education levels in the clusters of Italians, immigrants with Italian citizenship and Italian natives with foreign citizenship compared to the effect on the groups of European, Asian, African and American immigrants. The sense of this result is that those workers, belonging to higher levels of education, are not attracted by less satisfying jobs like labour intensive ones. On the contrary, for all the groups of immigrants without citizenship or not Italian natives, the effect of the instruction level changes: even if, as expressed by the first model, it remains negatively correlated to the labour intensive professions, in those populations it seems to be less strong. Moreover, the introduction of the random slope on the education level, diminishes the variability caused by the random intercept. The last notation refers to the single values of the

predicted values of the random slopes: the negative effect of the education levels on the dependent variable is lowest for African and Asian workers (the slope passes from – 0.43 to – 0.22 for Africans and to -0.24 for Asians); it is still lower for North and Central and South Americans (the slope passes from – 0.43 to -0.34) and finally it is lower for European (mainly coming from the East, (the slope passes from – 0.43 to -0.36).These results seem to suggest a sort of hierarchic trend of the relations between instruction level and typology of profession passing from richer to poorer Countries. Moreover these numbers report important information on the different occupational decisions of immigrants offering useful instruments for the formulation of public policies and business strategies. To conclude, the hypothesis of the dual labour market theory was confirmed by the results of this study, suggesting that the division of the labour market into autochthon and immigrants opens an interesting and complex object of analysis.

6. Concluding remarks

Migration phenomenon has many common points of study with economic subjects like the socio-demographic composition of the Labour Force. This contribution tried to investigate one of those aspects consisting of the differences between natives and immigrant occupational behaviour. The methodological base from which this study moved is the "Dual labour Market Theory" which states that people from poorer countries are pulled to cover the bottom positions of the job hierarchy of richer countries. Moreover this theory motivates this assumption with some demographic trends characterizing richer countries in the last decades. In this paper some demographic variables are introduced in order to formulate the dual theory and to test it. The models have been estimated in order to test the relations between some demographic variables and the probability for a worker to be involved in a capital intensive or in a labour intensive profession and if such effects change between native and immigrated employees.

The results confirmed the initial research hypothesis consisting in the formulation of the assumption of the "Dual Labour theory" .Although, other innovative and interesting points of this contribution have arisen the first is the useful informative value of the mixed models in the study of the relations between migrations and labour market. In addition, the use of an economic database for a socio-demographic problem like the relations between migration and employment made it possible to deal with the migration phenomenon from a different perspective avoiding information limitations and ensuring the contents of the data. Another interesting result is the possibility to disaggregate labour forces in autochthons and immigrants: it showed the meaningful additive function of this demographic perspective especially when the origin of immigrants is also considered. To conclude, another interesting point to further explore is the difference in behaviour between

immigrated workers with Italian citizenship and other immigrants, especially for the interesting value of the study of their social integration and adaptation, after some years, to the occupational behaviour of autochthon workers.

References

Acocella - Sonnino 2003: Movimenti di persone e movimenti di capitali in Europa, il Mulino, Bologna.

Arango, J. – Hugo, G. – Kouaoci, A. – Pellegrino, A. – Taylor, J.E. 1993: Theories of international migration: a review and appraisal, Population and Development Review, Vol. 19 Issue 3.

Borjas, J. G. 1999: The Economic Analysis of Immigration. In Anshenfelter, O. - Card, D. (a cura di): Handbook of Labour Economics, vol. 3 A, North –Holland, pp. 1697-1760.

Kuijsten, A. C. 1995 : International Migrations: harmonization of data collection. In Collects et comparabilitè des donneés démografique et sociales en Europe, Chaire Quetelet 1991 : a cura di Institut de demografique dell'Universite Catholque de Luvain, Academia e L'Harmattan, Louvain-la-Neuve, pp. 593-624.

Lee 1965: Theory of migrations, University of Pennsylvania, USA.

Lutz, H. 2008: Gender in the migratory process, Conference on Theories of Migrations and Social Change, July.

Massey, D. 1990: The social and economic origins of immigration, Annals of the American Academy of Political and Social Science, Vol. 510, No. 1, 60-72 p.

Natale, M. – Strozza, S. 1997: Gli immigrati stranieri in Italia. Quanti sono, chi sono, come vivono?Cacucci Editore, Bari.

Piore, M.J. 1979: Birds of passage: migrant labour in industrial societies, Cambridge University Press.

Tronti, L. 1995: The effects of Immigrations on the Economic System. A Survey, mimeo.

Venturini, A., 1996: Extent of competition between and complementary among national and Third-World migrant workers in the labour market: an exploration of the Italian case, in AA. VV. , The jobs and the effects of migrant workers in Italy. Three essays, International Migration Papers, 11, Geneva, ILO, pp. 23-42.

Willekens, F. 1984: Comparability of Migration. Utopia or reality? Working paper, Nidi, 47, Voorburg.

Use of statistical methods in sociology

Katarina Čobanović[1] - Valentina Sokolovska[2]

Data analysis in human sciences requires the use of statistical methods. Statistical techniques could provide explanations of how and why some statistical methods are used. The aim of the paper is oriented to the possibilities of using statistical methods both in education and in research in the field of sociology. Sociology and other human sciences, such as psychology, demography, economics and others, do require the use of statistical techniques.

This paper is focused on the analysis of data and the use of adequate statistical methods. Some statistical methods are more suitable for quantitative, some others for qualitative data. The paper is oriented to data analysis and the choice of an appropriate method. In sociology, like in other fields of science, it is necessary to deal with some simple and some more complex statistical procedures such as: descriptive statistics, inferential statistics, based on the sample survey, dealing with hypothesis testing, significance tests and analysis of variance, linear and multiple correlation and regression techniques, analysis of variance, linear and multiple correlation and regression techniques, analysis of frequency data using chi-square, nonparametric statistical procedures, multivariate methods etc.

The emphasis in the paper is placed on experience in using statistical methods in education and teaching of statistics and research work in sociology at the University of Novi Sad. The presented examples serve as an illustration of wide application of statistical methods.

Keywords: statistical methods in sociology, education, research work, types of data, analysis of data

[1] Faculty of Agriculture, Department of Agricultural Economics and Rural Sociology, University of Novi Sad, the Republic of Serbia, katcob@polj.ns.ac.yu
[2] Faculty of Philosophy, University of Novi Sad, the Republic of Serbia, valentinas@neobee.net

ACKNOWLEDGEMENT: The paper has been supported by the Ministry of Science of Republic of Serbia (Projects No. 149013D, No. 149007).

1. Introduction

"Statistics is a branch of scientific methodology. It deals with the collection, classification, description and interpretation of data obtained by conducting surveys and experiments. Its essential purpose is to describe and draw inferences about numerical properties of populations" (Ferguson 1966).

The experience in statistics application indicates that a single statistical method can be used in different research areas for dealing with different types of problems. In this respect, we can talk about the universality of a statistical method (Bethea et al.1985; Čobanović et al. 1997). For instance, an experiment conducted in agriculture or in laboratory conditions is an equivalent to a sociological survey (Blalock 1960). Various possibilities which application of a statistical method provides in certain research areas should be considered as an adequate way of dealing with a problem of the research. However, it should be noted that statistics is not a method by which it is possible to solve all the problems in a research. Likewise, statistics is not a mere gathering of data and their presenting (Blalock 1960). Statistical texts cannot be treated as literary texts, since statistical instruments are presented in a very condensed form (Blalock 1960).

When applying a statistical method, it is common to differentiate between quantitative and qualitative features and variables. Nominal and ordinal variables and data are usually considered as qualitative (attributive), while interval variables and ratio variables are considered as quantitative (Ferguson 1966, Krneta 1987). Also, it is common to apply nonparametric statistical methods on nominal and ordinal data, while parametric methods are used for the interval and ratio data (Ferguson 1966).

Statistics and statistical methods have highly significant application in sociology. Functions of statistics are numerous: the methods of descriptive statistics have an important application for describing natural phenomena; inferential statistics is used for inductive reasoning about unknown properties of a larger group using the known indicators of the causes; hypothesis testing most frequently refers to the results of one, two or more causes, on the basis of which it is possible to draw conclusions on the problem of the research, by accepting or refuting an initial hypothesis; regression and correlation analysis, in the most simple case, examines the influence and dependence between two or more variables. If the relationship of a greater number of variables is examined, it is multiple regression and correlation. For a detailed analysis of a multiple regression model it is necessary to introduce partial correlations and explain the relations between variables using partial correlation coefficient; application of chi-square distribution and chi-square test is important in cases with qualitative variables for which it is known or assumed that

are interrelated. Chi-square test is quite a common test based on determining the sum of the quotient of the square of the difference between the observed and expected frequencies and expected frequencies. Chi-square value has multiple applications, for instance, testing equality of distributions, independence test, contingency coefficient C, and so on (Hadživuković 1991, Hinton 2004); nonparametric tests for independent and dependent samples have also important application in research in sociology. These tests primarily refer to attributive variables (Hinton 2004). What should also be mentioned is the importance of nonparametric tests in application of the variance analysis method with a single factor, on the basis of ranked data. The most significant tests are Kruskal-Wallis test i Friedman test (Hadživuković 1991, Hinton 2004). Regarding nonparametric statistics, when examining the correlation and association between variables, it is important to mention Spearman rank correlation coefficient, contingency coefficient C, coefficient fi, which represents an indicator of association and is applied on discrete and dichotomous variables (Ferguson 1966, Hadživukovič 1991).

Application of quantitative methods in sociology by the World War II referred to descriptive statistics and simple methods (Raftery 2000). After the World War II, with the increase of the scope of data, more complex statistical methods started being used in sociology (Raftery 2000). For a great number of more recent statistical methods developed in 20th century it was sociology where they first found their application. **Thus, for example, correspondence analysis, a statistical technique useful in the analysis of categorical data, especially in sociological research, was emphasized in Greenacre's book (Greenacre 2007).**

The aim of this paper is to point to the importance of applying statistics in sociological research and statistics teaching. The paper presents the experience acquired while teaching statistics to sociology students at the University of Novi Sad. Moreover, the paper also presents some experience of applying statistical methods in sociological and demographic research**. Used statistical methods in the paper are results of the previous authors experiences. The authors wanted to describe the use of appropriate statistical methods in different research problems. There were used classical statistical tests in the explanation of Census data results. In the explanation of relationship between different variables, on the sample survey results base, there were used correspondence analysis and chi-square test. The authors wanted to outline the importance of use of statistical methods in social and similar sciences.**

2. Statistical methods in teaching

At the Department of Sociology of the Faculty of Philosophy in Novi Sad, the course in statistics is studied at undergraduate, graduate and doctoral studies. Within

the curriculum for the undergraduate academic studies of sociology, there are two compulsory one-semester courses in statistics on the first year of the studies.

The first course is entitled Descriptive Statistics, and its content is intended to introduce students to the problems of collecting and presenting quantitative data on the sociological issues; descriptive statistical measures; measures of central tendencies, and measures of variability; probability theory and sample method, as well as hypothesis testing. The objective of the course is mastering the basic statistical methodology used for planning quantitative research, for systematizing research results and reaching valid conclusions and decisions.

Within the second compulsory course entitled Statistical Methods, students are introduced to the basic methods of statistical analysis. Special emphasis is placed on the methods of nonparametric statistics: chi-square test, contingency coefficient, Wilcoxon rank test, Kruskal-Wallis, Friedman test, Spearman and Kendall rank correlation coefficient. Furthermore, the course encompasses simple linear regression and correlation, multiple regression and correlation analysis, time series analysis and variance analysis. The objective of the course is mastering statistical methods, including their adequate application and interpretation of obtained results.

During the academic undergraduate studies, students also participate in scientific-research projects of the Department of Sociology. Within the surveys conducted on the territory of the Autonomous Province of Vojvodina, students are engaged primarily to conduct surveys, where they gain fieldwork experience. Afterwards, they store the gathered data in software for statistical analysis and are present when preliminary research results are announced. Thereby, they can have access to a number of phases of research procedure, as well as gain insight into its complexity.

At graduate and doctoral academic studies in sociology students can opt for two optional courses in statistics. The course at graduate studies is entitled Applied Regression Analysis in Sociology, and it is intended to introduce students to the methods of applied regression analysis, as well as to point to the possibilities of its application in sociological analyses. The students also acquire knowledge on determining quantitative dependence of sociological categories, examining the validity of sociological theories hypothesis, as well as anticipating future movements of social phenomena on the basis of the assessed quantitative dependence.

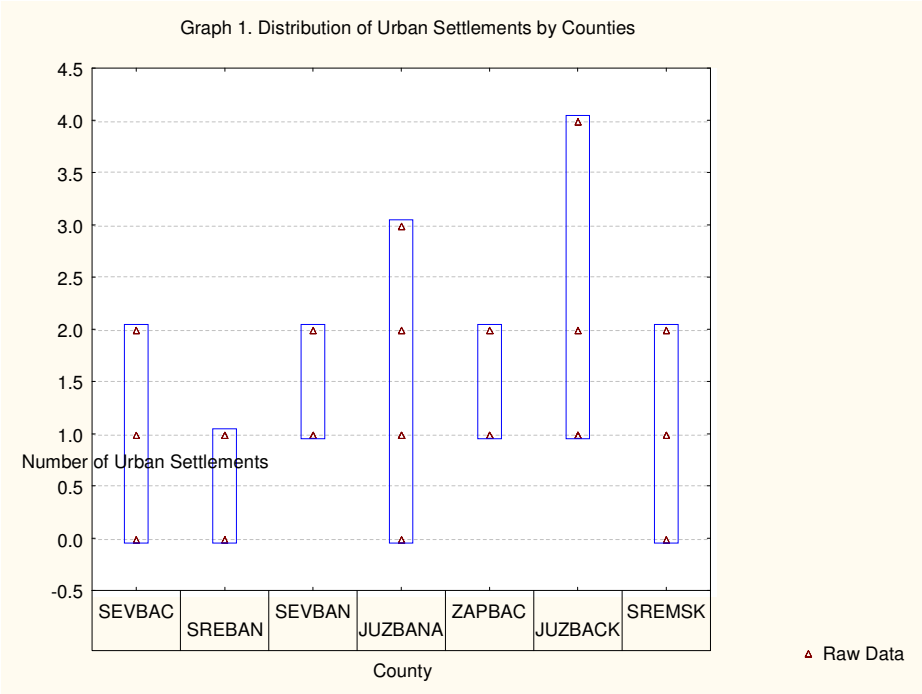
At doctoral studies of sociology, students can study multivariate techniques, also as optional subjects. The course encompasses canonical correlation analysis, multivariate analysis of variance, discrimination analysis, principal components analysis, factor analysis, and grouping and correspondence analysis. Studying and research work includes gathering of adequate empirical material for thesis and its statistical analysis; examining the literature and problem studies from multivariate techniques, using statistical software for data processing, discussions on applied techniques, their scientific scope and limitation in sociological problems analysis.

3. Statistical methods in research

The research on marital status of Vojvodinian population according to the census results from 2002 indicates that the changes of marital status were influenced by socio-political factors, economic factors and a number of cultural factors (Čobanović et al. 2008). Two types of families were analyzed: families with children and families without children at the level of the Autonomous Province, as well as at the level of Vojvodinian counties. Furthermore, the two examined types of families were analyzed also in relation to the types of their settlements: urban settlements and other (rural) settlements. In order to determine the existence of statistically significant differences between the two types of families and types of settlements, t-test was used for 2 independent samples. The number of families according to the type (with and without children) was expressed per capita. The graphical illustration was made using a "box-plot" diagram. In this example, a "box-plot" diagram, as a means of presenting data analysis, confirmed the results of the t-test, i.e. the existence of statistically significant differences between the two types of families and the types of settlements.

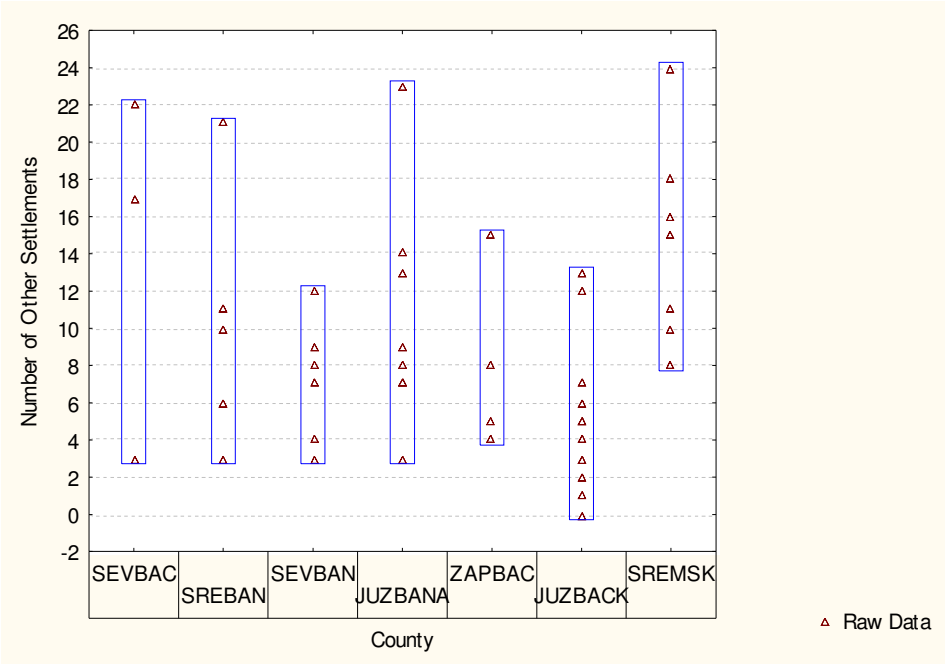
This research, based on the results of the population, households and dwellings census 2002, also included the analysis of the number of municipalities according to the type of settlements in the counties of Vojvodina. The research encompassed the analysis of municipality distribution in relation to the overall number of settlements, to the number of urban settlements and to the number of other settlements in 7 counties in Vojvodina. The municipality distribution (45 municipalities) in counties is graphically presented using a "variability plot" diagram, which in this case proved to be a very suitable way of presenting and analyzing of discontinued variables (the number of settlements) (Graph 1, Graph 2).

Graph 1. Distribution of Urban Settlements by Counties



Source: own creation

Graph 2. Distribution of Other Settlements by Counties



Source: own creation

Regression analysis (for 45 municipalities) in this research, based on the model of multiple linear regression, was used for examining the influence of certain variables on the number of families with children and the number of families without children at the level of urban and rural settlements. The results of regression analysis showed that the number of families with children from urban settlements is significantly determined by the overall number of families (with the positive correlation), the number of inhabitants of rural settlements (with the negative correlation) and the number of families without children of rural settlements (with the negative correlation). The regression model is as follows:

$$\hat{Y}_i = 45.31810 + 0.54817X_1 - 0.44021X_2 - 0.31989X_3$$

$$\left(R_A^2 = 0.999\right) \quad (t = 181.14) \quad (t = 12.011) \quad (t = -2.915)$$

The regression results of the number of families without children in rural settlements on the number of families with children in urban settlements (with the positive correlation), and the overall number of families (with the positive correlation), on the number of families without children in urban settlements (with the positive correlation) and the overall number of inhabitants in urban settlements (with the negative correlation). The regression model is as follows:

$$\hat{Y}_i = 52.79073 + 0.30519X_1 + 0.05356X_2 + 0.64293X_3 - 0.51155X_4$$
$$\left(R_A^2 = 0.984\right) \quad (t = 35.61) \quad (t = 0.26) \quad (t = -4.45) \quad (t = 3.74)$$

In this case, the following explanations can be provided: by the increase of the number of urban inhabitants (X_4) the number of families without children in rural settlements significantly decreases, which can be the result of movement of young population from rural to urban settlements, while in rural settlements there are only old people left (families without children). At the same time, the increase of the number of families without children in urban settlements (X_3) has positive effect on the increase of the families without children in other settlements. This relation can be explained by the assumption that the families without children, in both urban and rural settlements, are the result of the birth and death rate decrease among the population, as well as the process of senilization of Vojvodinian population, which according to a lot of indicators, belongs to the type of old population.

The results of the research, conducted in 2006 by the Department of Sociology of the Faculty of Philosophy in Novi Sad, on the sample of around 1,200 units of observation (respondents), and conducted on the territory of the Autonomous Province of Vojvodina, provided the material for the analysis of socio-economical and cultural aspects of multiculturality (Čobanović et al. 2007).

This research includes 21 selected variables. Some of them are defined as variables expressing grades of socio-economical situation, cultural views, ethnic aspects and the attitudes and opinions about many other features of multiculturalism in Vojvodina. Some variables refer to age, sex, national and confessional affiliation, education, occupation, activity sector, migration characteristics, county, household income, household size and the number of households with members without income. The variables are divided into 2 groups set by the author, keeping in view the assumptions on possible relations that can appear between the variables. In this way, a group of potential "independent" variables and the group of "dependent" variables were formed.

The first group of 6 dependant variables refers to the following phenomena: Evaluation of the family life quality (5 modalities); Evaluation of the most important family moments (7 modalities); Attitude towards joining EU and NATO (6 modalities), Evaluation of the culture preservation method (8 modalities); Attitude of the Serbian nationality members towards members of national minorities (7 modalities); Attitude of the national minority members towards national majority membership (7 modalities).

The second group consisting of 15 independent variables refers to the following phenomena: Sex (2 modalities); Age (6 modalities); National affiliation (11 modalities); Confessional affiliation (11 modalities); Education (6 modalities); Employment status (4 modalities); Occupation (28 modalities); Agency sector (8

modalities); Sojourn (2 modalities); Sojourn Character (2 modalities); County (7 modalities); Previous sojourn vacancy (2 modalities); Household monthly income (20 modalities); Household size (10 modalities); Number of household members without income (10 modalities).

In this research the majority of the analyzed variables are qualitative, given descriptively, with a larger number of modalities. In order to perceive the relations of certain combinations of variables the method of correspondence analysis was applied. Correspondence analysis describes the character of the relations between the occurrences and the structure of attributive (categorical) variables (STATISTICA 7.0, Introductory Overview). The correspondence analysis in this research was based on two-dimensional contingency tables and on determining the overall chi-square value. On the basis of chi-square value the statistical significance of the examined relations, i.e. the combinations of variables, was determined. The results of the correspondence analysis were generated by the statistical programme STATISTICA 7.0. This paper states some interesting results:

- The relation between the level of education and national identification was statistically significant with the chi-square value 132.64^{**}, $(df = 50; p = 0.0000)$;
- The relationship between the national status and employment status was statistically significant $(\chi^2 = 60.473^{**}; df = 30; p = 0.00081)$. It is worth outlining that the categories of employment (employed, unemployed, economically dependants and retired) were very heterogeneous;
- The relationship between the national status and age category and the relationship between the national category and education category were not statistically significant;
- The relation between the aspect of Serbian nationality according to national minorities and employment status was statistically significant $(\chi^2 = 29.645^{**}; df = 8; p = 0.04112)$;
- The contact of national minorities to majority (Serbian) population and the category of sex was not statistically significant $(\chi^2 = 8.0816; df = 6; p = 0.23220)$;
- The relation between the attitude in favour of joining EU and NATO and the age category was not statistically significant $(\chi^2 = 31.20379; df = 25; p = 0.18240)$;
- The relation between the attitude in favour of joining EU and NATO and other groups and the category of education level was statistically significant $(\chi^2 = 50.427^{**}; df = 25; p = 0.00019)$;
- The relation between the decision to join EU and NATO and nationality was statistically significant $(\chi^2 = 117.30^{**}; df = 50; p = 0.00000)$;

– However, the relation between the attitude in favour of joining EU and NATO and type of settlement was not statistically significant;

– The relation between the attitude in favour of joining EU and NATO and employment status was statistically significant $\left(\chi^2 = 26.0610^{*}; df = 15; p = 0.03744\right)$;

– The relation between the category of keeping culture of the nation and the age category was statistically significant, while the relation between the category of keeping culture of the nation and national affiliation was not significant;

– The relation between the category of keeping culture of the nation and the category of confessional affiliation was statistically significant $\left(\chi^2 = 231.80^{**}; df = 70; p = 0.00000\right)$;

– The relation between the category of keeping culture of the nation and the category education level was statistically significant $\left(\chi^2 = 62.001^{**}; df = 35; p = 0.0033\right)$;

– The relations between the categories of perceiving current lifestyle and educational attainment, age groups, employment status and nationality were not statistically significant.

In the doctoral thesis "Acultural Processes of Ethnic groups in Vojvodina" – defended in May 2009 at the Department of Sociology of the Faculty of Philosophy in Novi Sad – correspondence analysis, among others, was used for analyzing the above mentioned survey in 2006. One of the analyzed attitudes is the way the respondents perceive Vojvodina. For the question "How do you personally perceive Vojvodina?" the respondents of Serbian nationality opt mostly for the reply that it is a province of Serbian people and national minorities; the Montenegrins perceive it equally as a province of Serbian people and national minorities, and as a province of its citizens living on its territory; the Croats, Hungarians and Romanians perceive it most as a province of its citizens living on its territory, while the Slovaks see it as a province of different nations and ethnic groups (Graph 3). According to the Tables 1 and 2, which present the results of the correspondence analysis, it can be observed that almost 90% of the responses to this question are assigned to the modality 1, which replaces the reply "as a province of Serbian people".

Table 1. The table of correspondence analysis for How do the respondents perceive Vojvodina regarding the national (ethnic) affiliation of the respondents (survey of 2006)

How do you personally perceive Vojvodina	national (ethnic) affiliation of the respondents						
	Serbian	Montenigrin	Croat	Hungarian	Slovak	Romanian	Active Margin
1	46	2	1	2	1	1	53
2	349	10	4	30	6	5	404
3	135	2	11	56	15	3	222
4	248	10	19	82	13	9	381
5	5	0	0	1	1	0	7
Active Margin	783	24	35	171	36	18	1067

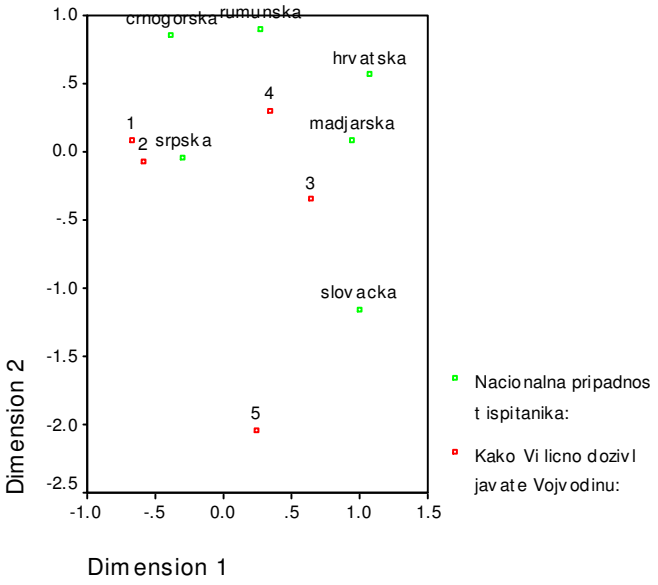
Source: own creation

Table 2. The overview of correspondence analysis for How do the respondents perceive Vojvodina regarding the national (ethnic) affiliation of the respondents (survey of 2006)

Dimen-sion	Singular Value	Inertia	Chi Square	Sig. (20 degrees of freedom)	Proportion of Intertia		Confidence Singular Value	
					Accoun-ted for:	Cumu-lative	Standard Deviation	Corre-lation 2
1	0.280	0.078			0.898	0.898	0.027	0.030
2	0.088	0.008			0.089	0.987	0.036	
3	0.032	0.001			0.012	0.999		
4	0.010	0.000			0.001	1.000		
Total		0.087	93.016	0.000	1.000	1.000		

Source: own creation

Graph 3. How do the respondents perceive Vojvodina regarding the national (ethnic) affiliation of the respondents (survey of 2006)



Source: own creation

The variable "How do you personally perceive Vojvodina?" has the following modalities: 1) As a province of Serbian people. 2) As a province of Serbian people and national minorities, 3) As a province of different peoples and ethnic groups, 4) As a province of its citizens living on its territory, and 5) I do not know.

4. Conclusion

The paper states the importance of statistical methods in sociology. Due to the significance of statistical data analysis it is possible to apply adequate statistical methods, on which the value of acquired results and conclusions is based. The paper presents the experience of applying statistical method in teaching sociology students. There are also some author's results, related to applying statistical methods in the field of sociological and demographic research. Thus, the marital status of Vojvodinian population was examined on the basis of the data from the population, households and dwellings Census 2002 in Serbia. The paper also presents the results of correspondence analysis, referring to the relations of the most significant socio-economic, cultural, ethnic and other aspects of living on the territory of Vojvodina. These results are based on the sample of around 1200 respondents on the territory of

Vojvodina in 2006. In sociology, like in many other disciplines, the use of statistical methods in practise was justified, because of its great importance in explaining the character of data, in explaining the behaviour and relations of investigated phenomena (social phenomena, demographic phenomena and so on) and in making conclusions. Both parametric and nonparametric statistical methods are suitable for use in sociology because of the nature of data. In sociology are common both quantitative and qualitative data.

References

Bethea, R. M.-Duran, B. S.-Boullion, T. L. 1985: *Statistical Methods for Engineers and Scientists*. Marcel Dekker, Inc., New York.

Blalock, H. M. 1960: *Social Statistics.* McGrow Hill Book Company, New York.

Ferguson, G. A.1966: Statistical Analysis in Psychology and Education, 2nd Edition, McGraw-Hill, London.

Greenacre, M. 2007: Correspondence Analysis in Practice, Second edition, Chapman & Hall, Boca **Raton**, FL, USA.

Hadživuković, S. 1991: Statistički metodi s primenom u poljoprivrednim i biološkim istraživanjima, Drugo prošireno izdanje, Poljoprivredni fakultet, Novi Sad.

Hinton, R. R. 2004: Statistics Explained, 2nd Edition, Routledge, New York.

Krneta, M. 1987: Statistika za sociologe, Ekonomika, Ekonomski institut

Raftery, A. E. 2000: Statistics in Sociology, 1950-2000. *Journal of the American Statistical Association*. 95, 450; ABI/INFORM Glober, pp 654-661

Sokolovska, V. 2009: *Acultural Processes of Ethnic groups in Vojvodina*., Doctorial Thesis, Faculty of Phylosophy , University of Novi Sad: (unpublishied material).

STATISTICA 7.0 2007: StatSoft, University Licence, University of Novi Sad

Čobanović K, -Nikolić-Đorić, E.-Mutavdžić, B. 1997: The Role and the Importance of the Application of Statistical Methods in Agricultural Investigations. *Agrieconomica*. Vol. 26. Pp. 457-470.

Čobanović, K.-Sokolovska, V.-Nikolić-Đorić, E. 2007: Statistical Analysis of Multiculturalism Research in Vojvodina, Applied Statistics 2007, Program and Abstracts, Ribno (Bled), Slovenia, p.40

Čobanović, K.-Sokolovska, V.-Nićin, S. 2008: Statistical Analysis of Marital Status of the Population of Vojvodina, Applied Statistics 2008, Program and Abstracts, Ribno (Bled), Slovenia, p. 48

Application of Statistical Data and Methodology in Practice and Legal Education

Ildikó Szondi[1] *– Adrienn Princz*[2] *– János Mészáros*[3]

The science of statistics in regards of its methods and tools has a special relationship to the social sciences, thus sometimes escapes attention of the legal students and the members of the legal profession. In order to understand the structure of the Hungarian legal system, the knowledge of the statistical sciences is essential. For students learning private law the knowledge of statistics is helpful, showing the importance of certain legal statues or the necessity of legal rules in governing social networks. In order to show the mechanism of application of law in everyday life we also need statistical tools. Our lecture will show the statistical distribution of cases filed to the Hungarian Court System. Out of the total number of filed court cases every second one will be completed within a year. In he second part of our lecture, we will discuss some aspects of corporate law, showing the incidence of establishing business partnerships and corporations. In Hungary in 2008 the established business partnerships, corporations mainly consisted of limited responsibility and financial asset based partnerships. The science of statistics will help to better understand social life-networks by showing average wages, employment, or workplace contracts, their incidence and frequency. For our students it would be difficult to bet, to find out the real ratio and nature of work-contracts, in reality only 8% of work-contracts will be established with a pre-defined time-limit. It is also interesting to know, that out of 8 employees only 3 are women, the rest of it (5) are men. In the field of family law, in respect of the marriage data it is surprising, that in Hungary since 2004, the ratio of people older that 15 years, living in marriage, decreased, and now they are a minority. In 2008 almost 74 000 marriage ended up, by death of the partner, or divorce, and only 40 000 new marriages were registered. The followings also will show the Hungarian social and legal situation: the changing number of criminal cases, the special structure aspect of felonies, and felons, and the abovementioned and also the evaluation of the law-enforcement, court system, the work of judges, In the modern age of globalization, and connected world, it is of utmost importance that the statistical system, the criminal statistics should work with great precision and with the highest professionalism. As the Internet took over every aspect of the modern societies, it is

[1] Ildikó Szondi, Graduate of Law, Associate Professor, Department of Statistics and Demography, Faculty of Law, University of Szeged
[2] Adrienn Princz, Graduate of Law, PhD Student, Department of Statistics and Demography, Faculty of Law, University of Szeged
[3] János Mészáros, Fifth-year law student, Teachers Assistant, Department of Statistics and Demography, Faculty of Law, University of Szeged

providing easily and readily accessible communication channels, for professionals and lay people alike. Our lecture will attempt to yield an insight of the recent state and importance of criminal-statistics, taken into account the special circumstances of Hungary. We also would like to emphasize once more, how important is the statistical knowledge for the legal profession and legal student nowadays.

Keywords: Private Law, Criminal Statistics, Legal Education

1. Introduction

Our every-day world is characterized by economic and social difficulties. Hence, it is especially important that social sciences be able to provide a precise analysis of our circumstances. Statistics might prove as an efficient tool in this struggle for professors and students in every respect of legal education. Considering the methodology, analysis and the means system of statistics it has special relationship with social sciences, however, it often ignores the interest of university students who study law and are more prone to humanities rather than mathematics (Steiger 2008).

For decades, proficient knowledge of the fields of literature and history were the most important among admission criteria for law students. Those who could prove sufficient knowledge in these areas became eligible for admission as law students; those, who proved their proficiency in mathematics and history, chose economy as a career path. This policy resulted in loss of contact between economy and law in terms of educational methodology, and as a result of this trend law students had no useful mathematical knowledge. The system of structured studies has changed since 1989. Present law students are more aware that statistical science is helpful for their better understanding the practice, the implementation of law and the development of law.

Our Department of Statistics and Demography at the Faculty of Law of the University of Szeged considers the dissemination of knowledge in the fields of general and applied statistical methods one of its most important goals, in order to help students to acquire complex and insightful knowledge in these fields. This policy resulted in the increasing number of students, who select law and statistics for their fifth year dissertation or thesis. An average of hundred students undertakes such an enterprise each year, twenty to forty of them dissert about the connection and interrelation between law and statistics.

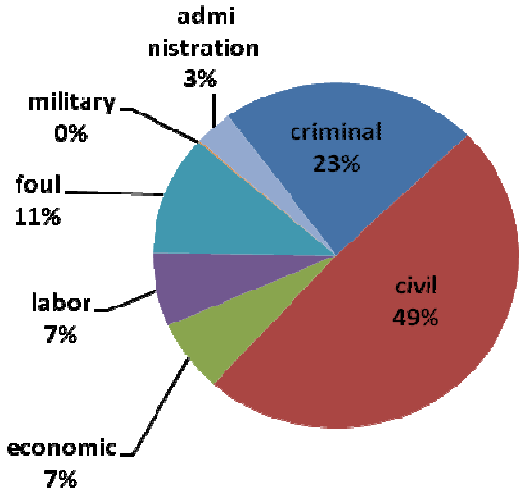
2. Statistics in Law. Application of Statistics in Private Law

Below, we have to provide some insight into the workings of the Hungarian justice system. The largest practical subsystem of the Hungarian legal system, besides administration, that should use statistics more often is the judicial system. Insight in

the justice system via the toolbox of statistics is necessary before explaining how statistical data can improve the level of legal education.

To understand the system and function of Hungarian jurisdiction, statistics is inevitably important. People feel that like other European courts Hungarian courts are also overloaded and procedures are slow. Law students should look behind these statements to what extent the Hungarian courts are busy with cases and how long the court reaches the final judgment (Juhászné 2009). In 2008, local, county and regional courts have received 398.430 cases with the following ratio.

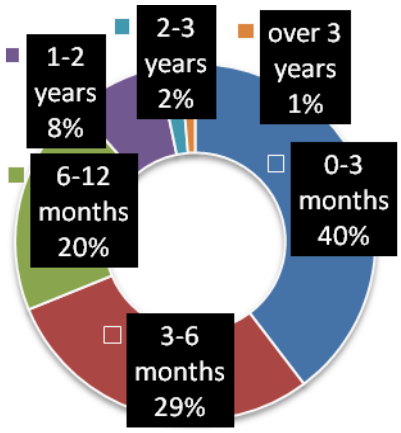
Figure 1. Cases arriving in 2008 to local, county and regional courts divided by cases



source: www.birosagok.hu

As it turns out, 89% of civil cases arriving to the local courts are finished in a year. The need that the cases be finished in a reasonable time is one of the basic human rights (right to a hearing within a reasonable time). To establish effective, fast, simple and cheap ways of deciding court cases has long been one of the most frequent aims of civil law codification.

Figure 2. Duration of prosecution at local courts
(finished in 2008 civil cases)

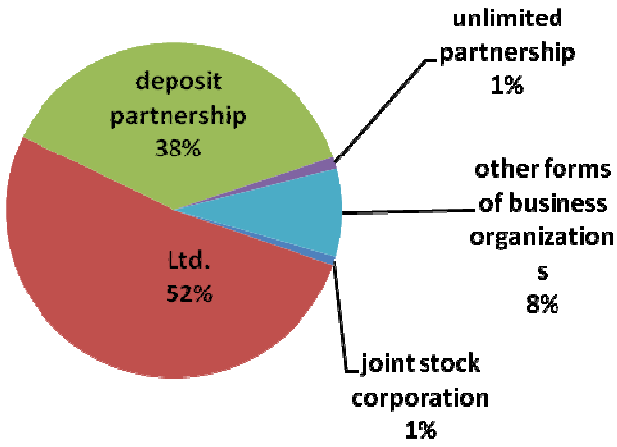


Source: www.birosag.hu

3. Corporate Law

The importance and unavoidability of norm of corporate law is clearly shown by the number of business organizations that were over 500.000 in 2008. The primary demographic data source of enterprises is based on the register of business organizations published and operated by the Central Statistics Institute. In 2008 the most popular form of enterprise was the limited liability company. Their number had exceeded 292.000. It is one and a half times larger than in 2000 (Farkas et al.). The low initial capital explains the popularity of limited liability company form. Among business organizations the second most popular form was the deposit partnership. In 2008 there were 211.000 deposit partnerships in Hungary exceeding year 2000 with 60 thousand.

Figure 3. Number of business organizations in 2008



Source: www.ksh.hu

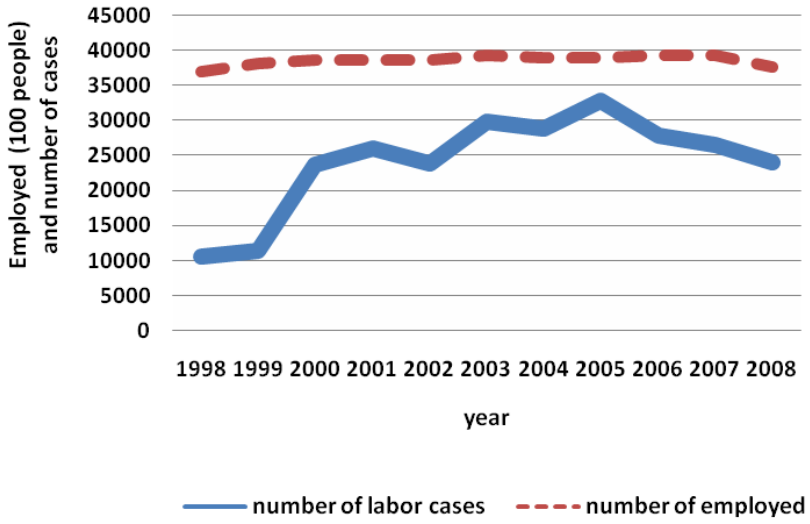
The number of joint stock corporations is relatively low beside the two ruling company forms, but it doesn't reduce the importance of it.

4. Labor Law

Law students will feel the importance of studying labor law on themselves when they have labor dispute with their employer. Students having their parents support do not really feel the usefulness of labor law. In case of facing the number of labor law cases they will study Labor Code with bigger interest (Berki 2008).

Based on the first quarter of 2009 labor force survey between the age of 15-74 4.167 million people appeared on the labor market. Among these 3.764 people were employed and 403.000 were unemployed. On national level the average net income was 121.400 HUF for blue-collar workers the average wage was 89.400 HUF and for intellectuals' the average wage was 154.200 HUF.

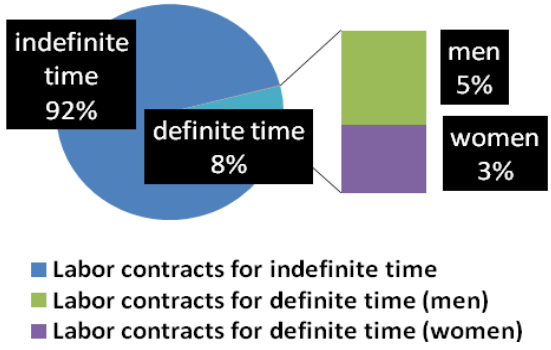
Figure 4. Labor cases and the number of employment (from 1998 through 2008)



Source: www.ksh.hu

One may ask why did the number of labor cases increase between 1999 and 2005 while the number of employment did not change significantly. Law students have no time to be absorbed in all those fields of life which regulations and norms they study. In this case the greatest help is statistics with its numbers and diagrams that shows the most important characteristics of the social relations regulated by a given branch of law (in case of labor law for example the average wage, the number of unemployed people's cases).

Figure 5. Labor contracts based on duration in 2008



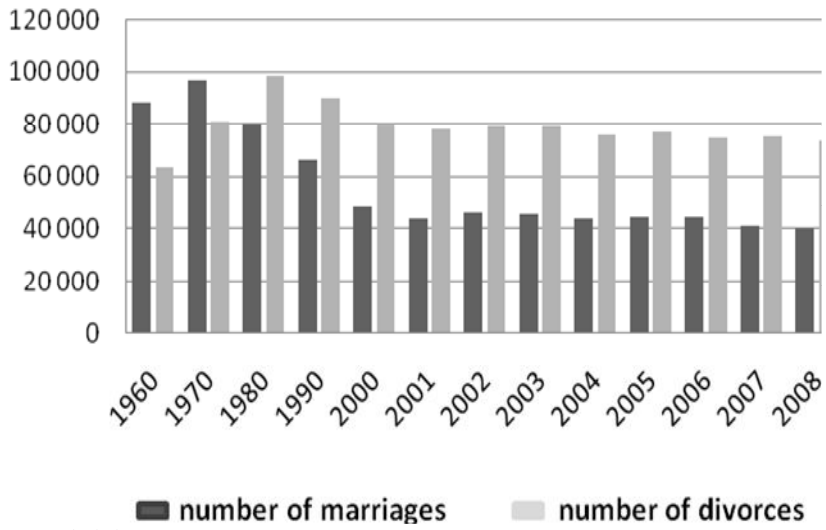
Source: Ministry of Labor

Out of 100 labor contracts 92 indeterminate contracts were made in 2008 in Hungary. Out of the low number 8 determinate contracts 5 were made for men and 3 for women.

5. Family Law

The responsibility of the family is to secure the continuity of human race, giving birth and raising children as long as they take their places in the society. The most important legal foundation and basis of the family is the marriage regulated by the family law, which is a voluntary alliance of a man and a woman (Hegedűs 2006). Its importance can be expressed with one number. In 2008 with 40.000 more people got married than the previous year. Unfortunately we can say the same about the marriage property law. In 2008 25.000 marriages ended (Gyémánt et al. 2005).

Figure 6. Marriages and divorces from 1960 through 2008



Source: www.ksh.hu

The diagram shows well the changes of the image of marriage, value judgment, and lifestyle of the society. The 44-46 thousand marriages made at the turn of the century do not reach the number of marriages made in the 70's that was 90-100 thousand. The number of marriages made in 2008 is below the previous years with 1.8 percent –apart from the I. World War- it was the lowest in the 130-year-old history of peoples' moments. The rare marriages and frequent divorces hard hit the scale of marriages (Kőrös 2007). Invariably more marriages end with becoming widower or

because of divorces than new marriages are made. In 2008 73.800 people widowed and 40.100 got divorced. Since this process is stable, the consistency population based on family status changes considerably.

Since 1999 the number of single people has increased from 20.3 % to 31.7 % and the number of divorced people from 7,4% to 10.5%. Parallel with this the number of marriages has decreased from 61.2% to 46.5%. During the past few years it can be witnessed that among the population of the age of 15 and above married people are in minority. Since 2004 less than half of the adult population is married (KSH 2007).

Such an introduction was indispensable in order to be able to comprehend the importance of statistics, and the help it can provide in practice. Now, we would like to point out some implications relevant to the use of statistics in legal education, as a basis for practical applications, through the lens of criminal statistics.

6. Criminal Statistics in Legal Education. The Example of Criminal Law.

"The only statistics you can trust are those you falsified yourself"
(Sir Winston S. Churchill)

We have to admit: Churchill's slightly cynical motto is of perpetual truth. There might be some who nod understandingly hearing it; others might raise their voices anxiously finding its content offensive. However, it is an undisputable fact that the result of statistical analysis can always be approached from two angles: Truth of Falsehood.

It is of utmost importance that every research domain be examined, analyzed objectively, with the most accurate state-of-the-art methods possible and that the adequate conclusions be drawn from the research. This is true especially these days when -- in an extent never seen before -- chaos and pessimism characterize public opinion in economic and social terms.

Why exactly is it important to abide by this supreme moral-ethical law? If one does not abide by this principle then Churchill's words might come true as abusing data leads to chaos and pessimism detrimental not only to the individual but to the collective as well. On the other hand, however, this moral compass should be used as a mirror in order to provide answers for certain questions.

Modern day statistics are much appreciated and used in every aspect of research. Its methods shall not only provide accurate information for domestic users, but also for international users and statistical data has to comply with requirements set forth by the Eurostat and other international organizations. Given these requirements, data is collected, treated and processed from the public and private sector professionally and adjacent analyses, evaluations provide useful help to

legislators, courts and professors when preparing laws, verdicts and or teaching strategies, or in every other aspect of day-to-day life.

The integration efforts of Hungary towards the EU were manifold both before and after 2004. Legal integration is one of many and has several branches, one of which is criminal law. Seeing through the integration in criminal law could be to a great extent assisted by examining and evaluating criminal statistics, but this field of statistics is not only important in this respect. The role of education is also to be stressed here: statistics (i) help students gain a more comprehensive outlook on the different fields, areas of law, (ii) help them achieve an integral practical knowledge-base as part of the University studies. That is why statistics must be taught in connection with main subjects in the curriculum, along with several other branches of social sciences, this way helping the students develop a complex legal way of thinking.

In the following, I wish to comprehensively present the place of criminal statistics in legal education, primarily with respect to substantive law. I will discuss the possibilities of its practical application -- while presenting the current status of criminal law in Hungary --, with particular focus on stressing the areas where statistics should be applied with increased willingness and frequency.

7. Short Outlook on the Current Situation

The Statistics Act, No. XLIV. of 1993 is safeguarding the principles referred to in the introduction and guaranteeing the accuracy and professionalism of the data collected. It contains provisions on the methods and goals of statistical data-collecting and on the organizations that are authorized to collect data for statistical purposes. As for criminal statistics, the 59/2007 (XII.23.) Order of Ministry of Justice and Law-Enforcement serves as a unified code for both law enforcement and prosecutorial statistics. Law enforcement agencies and the Prosecutor's Office are not just entitled but also legally obliged to continuously collect such data and compile criminal statistics periodically.

The appearance of the Internet and the introduction of different information systems and databases had a beneficial effect on all different areas of criminal statistics. Collecting and systematizing data became significantly quicker, and the results of the statistical analysis are only 'one click away' for those interested. The bodies mentioned above conducting statistical surveys shall forward their findings to the Statistical and Analytical Division of the Ministry of Justice and Law Enforcement's Criminal Policy Secretariat (Igazságügyi és Rendészeti Minisztérium, Büntetőpolitikai Szakállamtitkárság, Statisztikai és Elemző Osztály) who is responsible for publication of these data at crimestat.b-m.hu, also known as the Hungarian Criminal Information System, a website, which contains up-to-date domestic law enforcement information and is accessible to anyone.

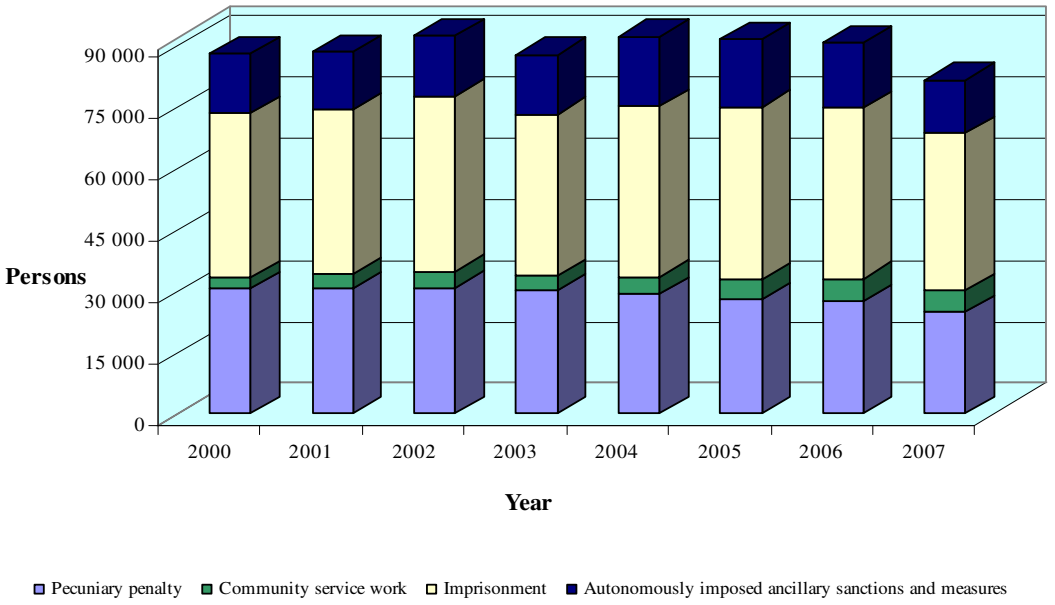
(On an adjacent note, we also have to mention those institutes that conduct surveys of public opinion that provide feedback to law enforcement agencies as to their activities by asking collecting relevant data from the public.)

8. The Appearance of Criminal Statistics in Substantive Criminal Law

As for the importance of criminal statistics in legal education, in the following I would like to point out the role of statistics in legal education and its current status at my University. I will use comparison of teaching materials at different domestic universities as an explanatory tool for my reasoning, and will also refer to domestic criminal law when pointing out areas that lack statistical basis in education.

Students at the University of Szeged, Faculty of Law familiarize themselves with the basics of statistics right in the first semester and they can broaden this knowledge by applying the basics in optional courses of the curriculum throughout the remaining four years of their studies: they can choose to learn -- among others -- justice and criminal statistics as well. On the other hand, introduction to substantive criminal law is part of the third semester's curriculum, where students first learn about the „general provisions" of the domestic Criminal Code (criminal law). It is noteworthy at this point that a significantly low percentage of the currently used law books on „general provisions" apply statistical data as means to support statutory texts in spite of the fact that there are certain chapters that allow; moreover, call for the application of criminal statistics (Földvári 2006, Bárd et al. 2002, Balogh-Kőhalmi 2007, Belovics et al. 2006, Nagy 2008, Görgényi et al. 2007). One of such domains is e.g. „the Doctrine of Criminal Law Consequences", which -- among others -- gives a distinct overlook on the trends in the numbers of principal penalties domestically. (See: Diagram 1)

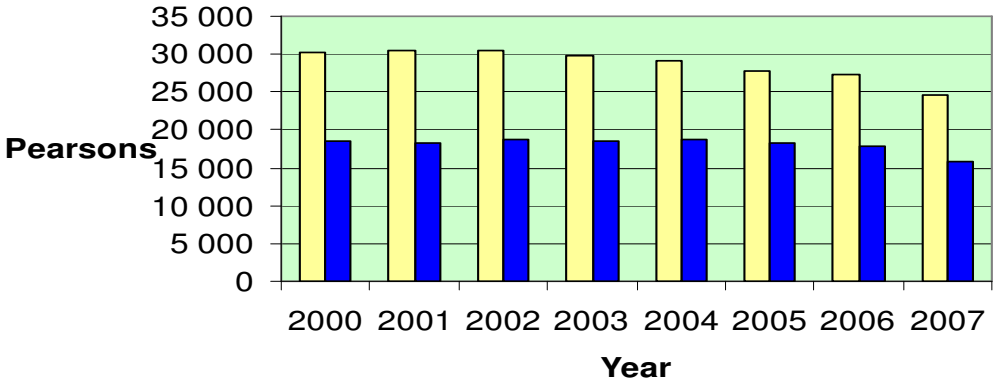
Figure 7. Principal penalties 2000-2007



Source: Hungarian Statistical Yearbook, 2007

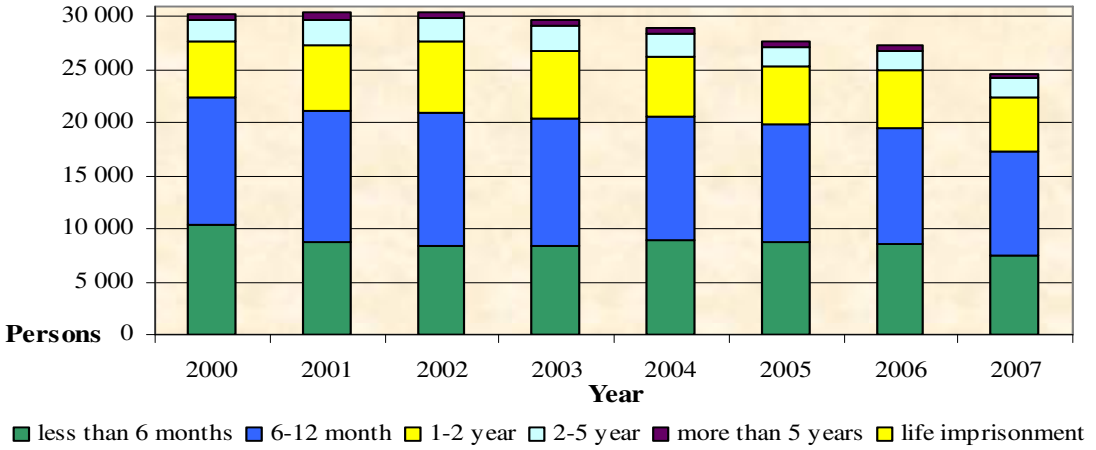
The rate of principal penalties and the proportion of the types of penalties have to be taken into account from several points of view. We can find herein useful information on how and why they change in Hungary, and this is also helpful in international comparison. However, based on international treaties, it is an obviously declared goal that the use of imprisonment shall only be exceptional; while examining the rates of incarceration the much-anticipated decline cannot be experienced (Nagy 2005). In light of this liberalization: (i) the number of those, sentenced to imprisonment (the execution of which was suspended), and (ii) the time of punishment, in case of punishments involving incarceration show the following tendencies:

Figure 8. Crimes punished by incarceration and the number of sentences suspended



Source: Hungarian Statistical Yearbook, 2007

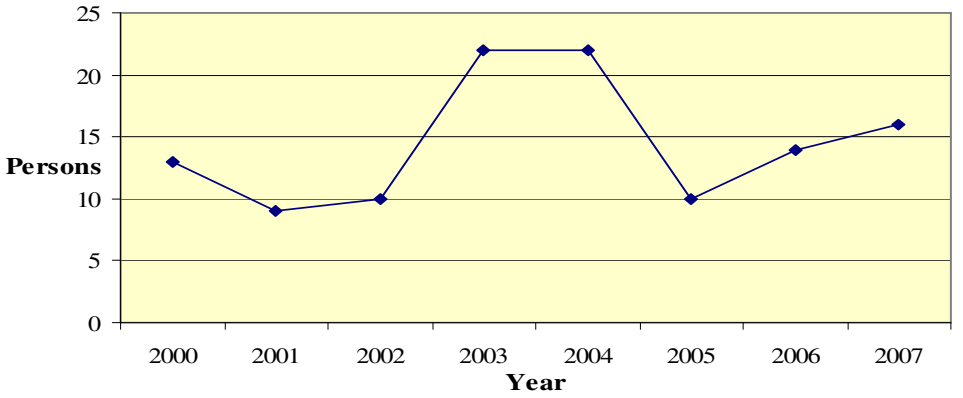
Figure 9. The number of persons sentence to incarceration



Source: Hungarian Statistical Yearbook, 2007

The number of perpetrators, who have been sentenced to life imprisonment, is significantly low. The following diagram shows the changes in the numbers from the year 2000.

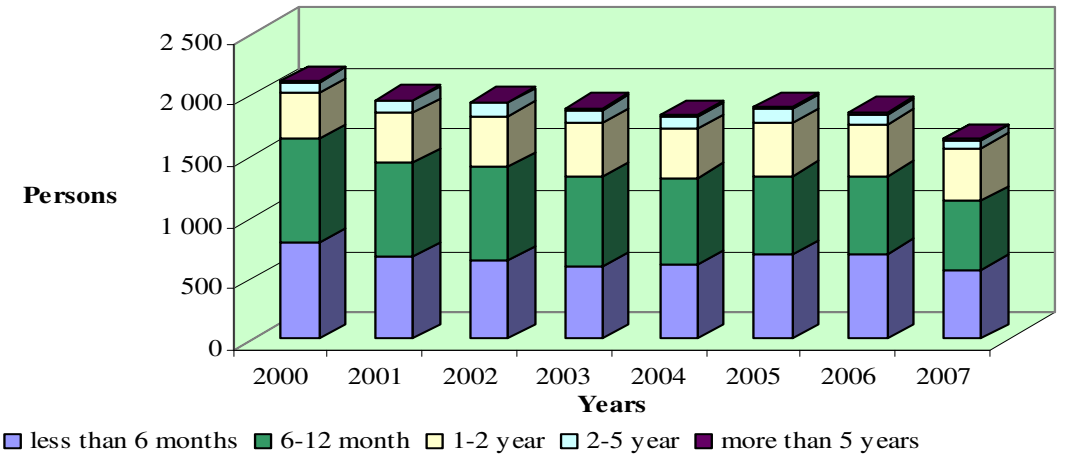
Figure 10. Change in the number of life of crimes punished by life imprisonment



Source: Hungarian Statistical Yearbook, 2007

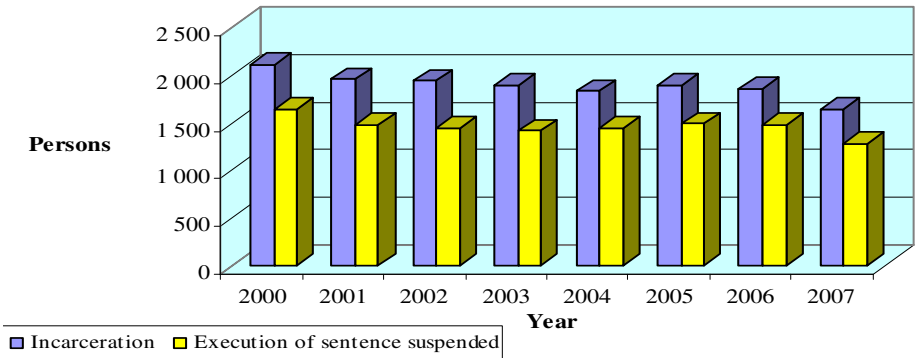
Statistical methods presented hereinbefore seem useful in representing the actual status of youth offenders as well.

Figure 11. Number of imprisonments against youth offenders



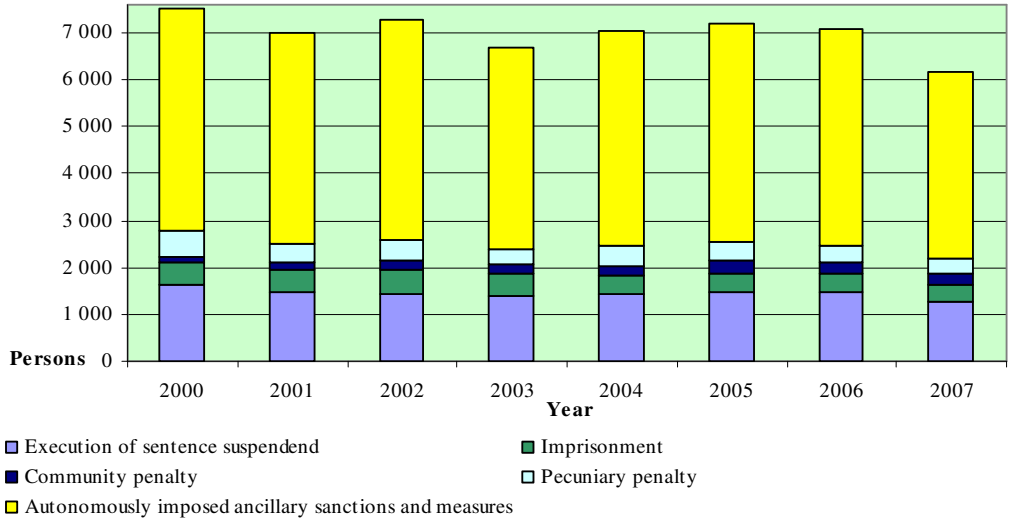
Source: Hungarian Statistical Yearbook, 2007; CD Appendix in carceration

Figure 12. The ratio of suspended sentences of youth offenders to the aggregate number of imprisoners



Source: Hungarian Statistical Yearbook, 2007; CD Appendix

Figure 13. Numbers of sanctions against youth offenders



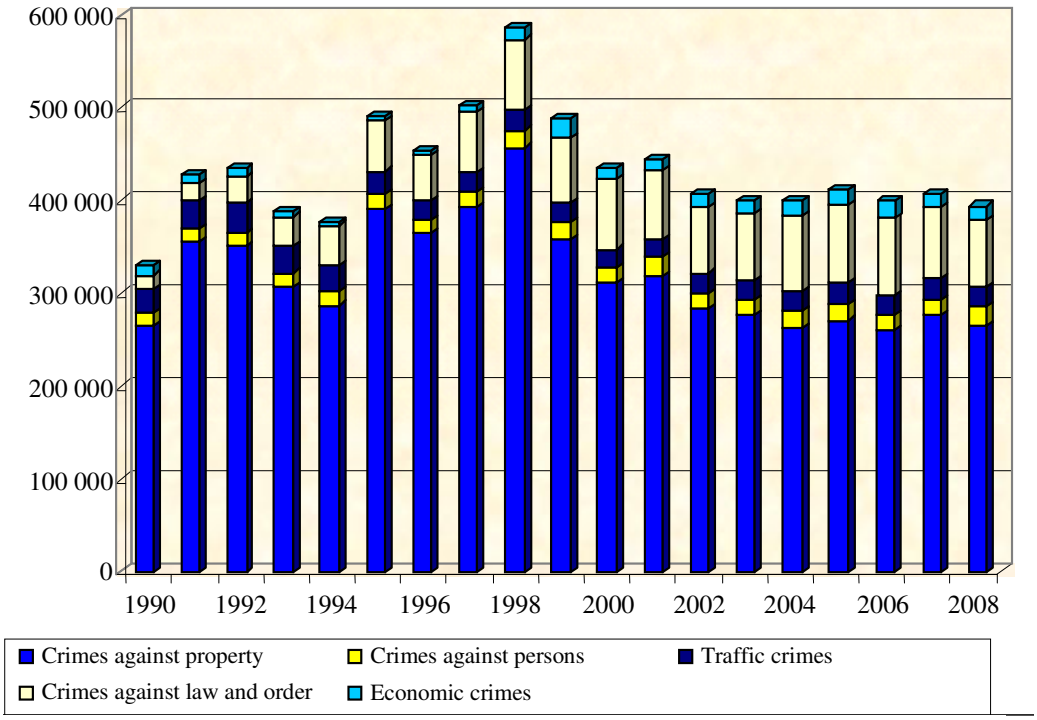
Source: Hungarian Statistical Yearbook, 2007; CD Appendix

In terms of the „specific provisions" of the Criminal Code (part of the fifth semester curriculum) there are many more opportunities to apply criminal statistics. Law Enforcement agencies, courts and executive organs might conduct parallel research and surveys on the cases in their respective jurisdictions. Since we have no knowledge on the exact extent of crime, these analyses are of great value serving as primary guidelines in getting to know domestic law enforcement efforts. Why could this be important? It is, mostly, because the number of law books on „specific provisions" applying statistics in explanatory materials is even smaller than that of „general provisions" law books in the same respect (Nagy 2009, Balogh 2008, Belovics et al. 2007, Blaskó et al. 2008, Blaskó et al. 2006, Erdősy et al. 2007, Fehér et al. 2001).

While studying special provisions, students get to know different facta of crime and become able to analyze these in detail to the fullest extent. However, this means nothing more to them than a mere aggregate of statutory articles, paragraphs and subsections and only a few of them will be able to see the 'big picture': the system of these provisions and the different elements of those, for the first time. Statistics can help them and the others to tie the dots and provide practical aids for their understanding of the dry statutory language with data based on real life examples. How exactly would this be possible?

In Hungary, the material covering "special provisions" starts with the introduction to crimes against persons, which is the part of the material that shall be the most emphatic in the course of criminal law studies. However, if we step out of criminal law onto the field of criminal statistics and compare the number of crimes against persons to all crimes committed; we find that their share, compared to the whole, is virtually insignificant: it is closer in its rates to the rate of traffic crimes, which is a smaller and less emphatic part of criminal legal education. I would like to point out the detrimental effects of mass-media in this respect that leads to exaggerating occurrences of murder, bodily injury, kidnapping.

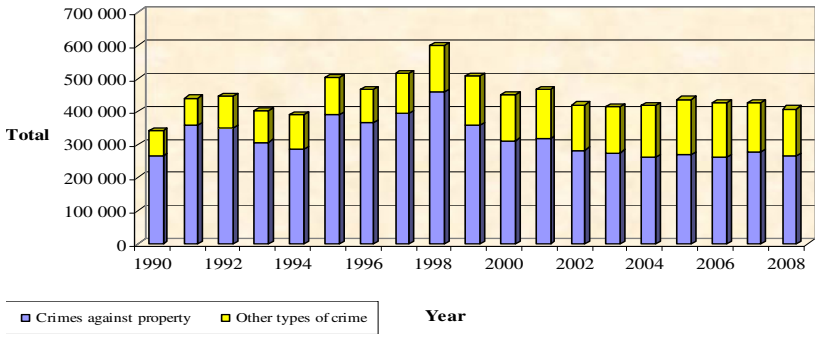
Figure 14. Patent (discovered) crimes (publicly prosecuted) 1990-2008



Source: http://portal.ksh.hu/pls/ksh/docs/hun/xstadat/xstadat_eves/tabl2_08_02i.html (in Hungarian)

If we analyze the previous table, from a different aspect, we find that the number of crimes against property significantly exceeds the number of all other types of crime:

Figure 15. Crimes against property within publicly prosecutes patent crimes compared to the total of other crimes 1990-2008



Source: http://portal.ksh.hu/pls/ksh/docs/hun/xstadat/xstadat_eves/tabl2_08_02i.html (in Hungarian)

Compared to the number of publicly prosecuted patent crimes, the rate of crimes committed against property was the highest (78.39%) in 1992. However, regarding the rates, with respect to the total of crimes committed (600.621) and also within this, the crimes against property (457.188) those were the highest in 1998. After 1998, a significant decline can be experienced. Analysis of these numbers might raise several questions in the students, regarding the emphases of the curriculum, and the importance of certain crimes. If these data are examined and used correctly by an expert in criminal statistics, it leads to objective insight, into the proportions of patent crimes.

Further analyzing the diagram above, we can only indirectly conclude, that in terms of crimes against property, the significant change in the statutory limits of value (increasing from 5.000 HUF first to 10.000 HUF, then to 20.000 HUF) had some effect on the numbers. (NB the peak points overlap with the dates of the changes in the statutory limits of value.)

This means, that crimes against property, not exceeding these limits of value, will not be governed by criminal law anymore, but by the law of misdemeanors. This of course leads to their incorporation in misdemeanor statistics, instead of criminal statistics. Pursuant to this, there appear to be fewer crimes included in criminal statistics, which the government comments, as a success in crime prevention. However, following from my arguments, this was clearly not the case, in light of the right interpretation of the numbers. It would be worth examining, how the analysis of criminal statistics disregarding misdemeanors, appears in criminal politics. [E.g. if we consider, that criminal statistics does not confirm the emphasis put on homicide (including murder and manslaughter) in the course of criminal law education.] (NB Following from the above, it is quintessential to deal with data

pertinent to misdemeanor statistics if we talk about criminal statistics, but we will not touch upon further questions of this problem in this current article.)

Summing up, it shall be pointed out, if we properly apply the methods of criminal statistics and then our efforts amount to a much clearer picture on the status of crime; also -- more importantly -- we can enable students to integrate their knowledge by having an objective system at their disposal to refer to. This system should be based on accurate measurements, compiled in a way meeting strict requirements, in order to make possible the understanding of criminal legal matters e.g. why certain crimes will be classified as life threatening assault and battery and why not as attempted murder.

Criminal law without criminal statistics is blind and criminal statistics without criminal law is pointless. - said Prof. Hans-Heinrich Jeschek. We could agree with his assertion on criminology, applying it to criminal statistics. The two fields cannot exist without each other, and their close relation is important for the different methods of criminal statistics help the development of the complex legal thinking needed for those interested in criminal law to become the best professionals possible. Criminal statistics has several methods that cannot be fully applied without sufficient knowledge on the basics of criminal law, on the other hand, however, we have to consider those areas as well that are perfectly compatible with the dogmatic system of domestic criminal law.

These areas and methods shall be inserted in the current materials of legal education with intent of an extended application, and they shall also be applied in the course of seminars and lectures as well. The major fora shall also be better promoted where those interested can find statistical data and broaden their knowledge on certain aspects of e.g. criminal law. As a last point of interest, it is to be emphasized as a pivotal advantage that this branch of statistics could put those interested in possession of not only domestic but also European and international information. In light of the all-reaching integration these days, this could only serve to the benefit of the students and might lead to several different advantages for them in the course of their professional life.

References

Balogh, Á. – Kőhalmi, L. 2007: *Büntetőjog I. Általános rész,* Dialóg Campus Kiadó, Budapest.

Balogh, Á. 2008: *Büntetőjog II. Különös Rész*, Dialóg Campus Kiadó, Budapest.

Bárd, K. – Gellér, B. – Ligeti, K. – Margitán, É. – Wiener, A. I. 2002: *Büntetőjog Általános Rész*, KJK Kerszöv, Budapest.

Belovics, E. – Békés, I. – Busch, B. – Molnár, G. – Sinku, P. – Tóth, M. 2006: *Büntetőjog Általános Rész*, HVG-Orac, Budapest.

Belovics, E. – Molnár, G. – Sinku, P. 2007: *Büntetőjog Különös Rész*, HVG-Orac, Budapest.
Berki, E. 2008: *Munkaügyi kapcsolatok közszolgálatban*, Pólay Elemér Alapítvány, Szeged.
Blaskó, B. – Miklós, I. – Schubauer, L. 2006: *Büntetőjog Különös Rész II.,* Rejtjel Kiadó, Budapest.
Blaskó, B. – Miklós, I. – Pallagi, A. – Schubauer, L. – Zentai, Á. 2008: *Büntetőjog Különös Rész I.*, Rejtjel Kiadó, Budapest.
Complex DVD Jogtár Plusz.
Erdősy, E. – Földvári, J. – Tóth, M. 2007: *Magyar Büntetőjog Különös Rész,* Osiris kiadó, Budapest.
Farkas, Cs. – Papp, T. – Strihó, K.: *Társasági jog III./1*, A gazdasági társaságok közös szabályai, SZTE Állam- és Jogtudományi Kar, 2007. pp. 7-46.
Fehér, L. – Horváth, T. – Lévay, M. 2001: *Magyar Büntetőjog Különös Rész,* Bíbor Kiadó, Miskolc.
Földvári, J. 2006: Magyar Büntetőjog Általános Rész, Osiris kiadó, Budapest.
Görgényi, I. – Gula, J. – Horváth, T. – Jacsó, J. – Lévay, M. – Sántha, F. – Váradi, E. 2007: *Magyar Büntetőjog Általános Rész,* CompLex kiadó, Budapest.
Gyémánt, R. – Katona, T. – Szondi, I. 2005: *Demográfia*, Pólay Elemér Alapítvány, Szeged.
Hegedűs, A. 2006: *Családi jog*, Kis-ker Bt., Szeged.
http://portal.ksh.hu/pls/ksh/docs/hun/xstadat/xstadat_eves/tabl2_08_02i.html
http://www.birosag.hu
http://www.birosag.hu/engine.aspx?page=Birosag_Statisztikak
http://www.hvg.hu
http://www.ksh.hu
http://www.magyarorszag.hu
http://www.otthonbiztonsag.hu/ob/index.php/mbiztonsagrol/52-vagyelbuncselek
http://www.police.hu/elemzesek/bunuldozes/bunterkep.html?pagenum=3
http://www.police.hu/statisztika
http://www.securitron.hu/bunozes_2004_ig.htm
http://www.szmm.gov.hu/
http://www.ugyeszseg.hu
Hungarian Central Statistical Office: *Hungarian Statistical Yearbook 2007*, and CD Appendix.
Juhászné Zvolenszki, A. 2009: *Polgári eljárásjogi ismeretek*, Pólay Elemér Alapítvány, Szeged.
Kőrös, A. 2007: *A családjog kézikönyve*, HVG-Orac, Budapest.
KSH, Szociális és Munkaügyi Minisztérium: *Nők és Férfiak Magyarországon 2007*.
Nagy, F. 2005: *Az európai uniós országok fogva tartotti rátájának alakulásáról*, Börtönügyi szemle, 2005/3.
Nagy, F. 2008: *Büntető Jog Általános Rész*, HVG – Orac, Budapest.

Nagy, F. 2009: *Büntető Jog Különös Rész*, HVG – Orac, Budapest.

Steiger, J. 2008: A jogászok, a közgazdászok és a számok; ELTE Állam- és Jogtudományi Kar – Központi Statisztikai Hivatal: A statisztika és a közigazgatás elkötelezettje, ünnepi kötet a 60 éves Katona Tamás tiszteletére, Budapest.

The preliminaries and introduction of the course "Technical Mathematics", and the examination of the students' workload

Krisztina Lászlóné Kenyeres[1]

The course Technical Mathematics was introduced during the academic year of 2006/2007 at the College of Dunaújváros. I compare the course with those courses taught at a time period preceding the introduction of the new course. I examine what kind of changes happened in the average grades by contracting several courses and forcing the students to accomplish a similar course curriculum within a shorter time period.

With statistics, I illustrate the success of the course. I examine the syllabus of the course in terms of its achievability, and investigate and analyze the students' workload. With the help of the analysis of the results of the students – the 'learning outcomes' –, I search for solutions and compensations to provide better achievability of the course. At the time of this lecture I can present only a summary of the results of the national survey carried out during the Fall term, I supplement it later with the results themselves, which are connected to the topic of my future Ph.D. thesis.

Using the viewpoint of ECTS, I examine the structure of the syllabus, and propose innovations in methodology.

Keywords: Technical Mathematics, Learning Statistics, examination of students' workload, national survey

1. Introduction

I examine the kind of changes occurred due to the merge of several courses into one, and due to the fact that the students had to complete this course during a shorter time period.

I examine the syllabus of the course in terms of its achievability, and investigate and analyze the students' workload.

[1] College of Dunaújváros.

2. The introduction of the course Technical Mathematics

The course Technical Mathematics was introduced in the academic year 2006/2007. This course is one semester long and it contains those subjects which are thought separately: Probability Theory, Mathematical Methods, Mathematical Statistics. The course is composed of lectures of two hours or two-times-two hours per week (15x2 hour lectures on probability theory and statistics, 4x2 hour lectures on linear algebra), furthermore two-hour practice sessions per week (8x2 hour practice sessions on probability theory, 7x2 hour practice sessions on mathematical statistics, and 12x2 hour computer lab sessions (DF1 2006, DF2 2006, DF3 2006,).

2.1 The aim of the module "Technical Mathematics"

The main aim of the module is that the students are cognizing and acquiring the technical and practical bases of the applied mathematics subjects (linear algebra, probability theory, mathematical statistics) on such level which is indispensable for the investigation of the special subjects.

Through the using of MATLAB conceived as an aim to cognize an up-to-date mathematical software package. By the help and use of it many - related to the fields of linear algebra, analysis, and numerical methods- mathematical exercises should solve.

I won't detail the weekly exchanging and scheduling of the subject cause in 2008 I wrote it down in my gossip which issued in the publication of the Weeks of Science Conference at the College of Dunaújváros (DF4 2008, DF5 2009).

3. Experience

The fulfillment of the course depends on several factors, the course requirement is rather complex. In the next part, I introduce the average grades in the semesters before and after the introduction of the new course structure. I consider the academic year right before the introduction of the new course as the base for the comparisons for the old courses. In the following tables I summarize the distribution of the achieved grades for both semesters from that year.

Table 1. Grade credits of students in Semester 1 of 2005/2006

Semester 1 of year 2005/2006	failed (1)	satisfactory (2)	mean (3)	good (4)	excellent (5)	total number of students
Mathematical Methods	159	138	80	39	24	440
Mathematical Statistics	6	12	12	3	0	33
Probability Theory	126	153	126	95	48	548

Source: own creation

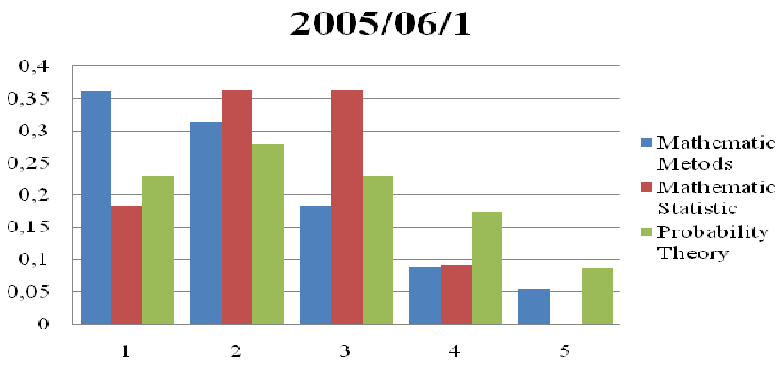
Table 2. Grade credits of students Semester 2 of 2005/2006

Semester 2 of year 2005/2006	failed (1)	satisfactory (2)	mean (3)	good (4)	excellent (5)	total number of students
Mathematical Methods	217	165	74	45	20	521
Mathematical Statistics	13	39	18	6	6	82
Probability Theory	115	172	138	98	60	583

Source: own creation

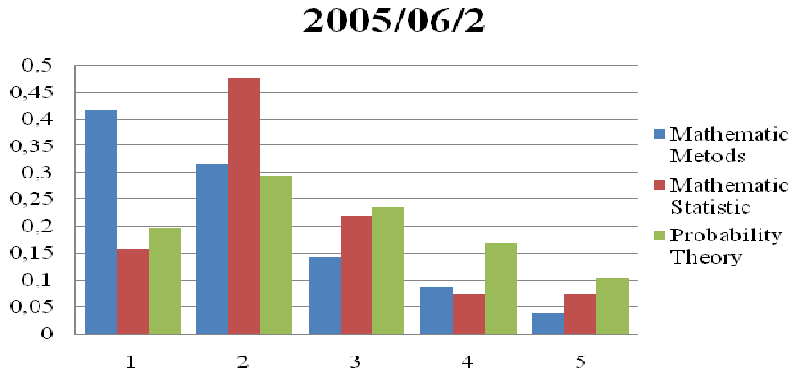
One can clearly see that the courses were attended by different numbers of students. Thus the courses could be compared only with some difficulties. Therefore, we show the next graphs with the same data, but now in terms of percentage (with respect to the total number of students) for each course.

Figure 1. The graphs for Semester 1 of year 2005/2006



Source: own creation

Figure 2. The graphs for Semester 2 of 2005/2006



Source: own creation

3.1 The results of Technical Mathematics

The course requirement is fairly complex, and hence the fulfillment of the course depends on various factors. In the next section I compare the grades of the students obtained since the introduction of the course.

For the course "Technical Mathematics" the five completed semesters give the base for comparison with the predecessor courses:

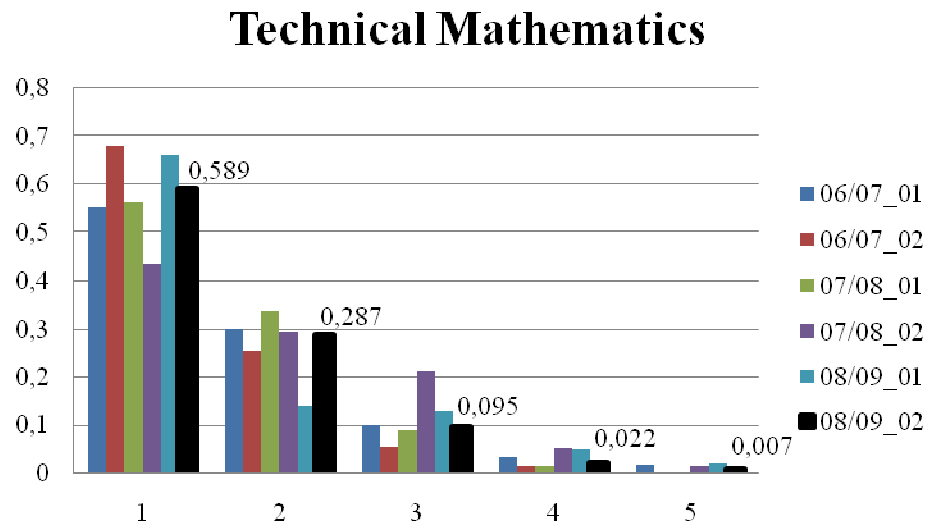
Table 5. Summary of the five completed semesters

	failed (1)	satisfac-tory (2)	mean (3)	good (4)	excellent (5)	total number of students
Semester 1 2006/07	0,55	0,3	0,1	0,03333	0,016667	60
Semester 2 2006/07	0,68	0,2533	0,0533	0,0133	0	75
Semester 1 2007/08	0,561	0,338	0,0878	0,0135	0	148
Semester 2 2007/08	0,433	0,293	0,21	0,051	0,013	157
Semester 1 2008/09	0,66	0,139	0,129	0,047	0,019	108
Semester 2 2008/09	0,587	0,287	0,095	0,022	0,007	139

Source: own creation

I graphed the data in the following figure to make it easier to compare them to one another.

Figure 3. The graphs for the last five semesters



Source: own creation

The 'total number of students' column shows that in the first academic year the course was taken only by a small number of students. The reason for this is that in that academic year only B.Sc. students were allowed to take the course. In year

2007/2008 not only B.Sc. students could take it, but all others who were not able to fulfill the course requirements for at least one of the predecessor courses (Probability Theory, Mathematical Methods and Mathematical Statistics). The following rule was introduced: Those who could not pass in the predecessor courses may pass by passing in the new course. In this way, the number of students taking the course increased dramatically.

Despite the large increase in the number of students, the improvement in the quality of the grades is clearly visible. The number of students with grade 'failed' decreased, the number of the other grades increased. This increased number thanked to two things – on the one part the semester repeaters are increase the number – the another part is that those students who are in the conventional training should take up this subject instead of the older subjects which they don't completed.

Taking into account the fluctuation of the number of students, we can observe a slight decline in the number of students who passed, but after reaching a balance state, we can immediately see an improvement. According to the structure of the subject they could easier completed the curricular. In the last two semesters again there has been a decline. From what I can see the cause of it is that the students are "accommodate" to the new requirements and as they see to complete the subject easier than expend less time to learn it. Analyzing the dispersion of the marks we can see that the highest was 0,26 in the first semester, the lowest dispersion we got from the II. Semester of 2007/2008 which was 0,15. In the last semester which I investigated (I. semester in 2008/2009) I counted 0,22. This value gave exactly the average of dispersions. The data for the last semester are highlighted in the graph. We can notice that it is close to the mean values of the number of grades in the previous semesters. Fortunately, even though minimally, but the ratio of students getting 'failed' has decreased. The number of students getting 'satisfactory' has more than doubled, but, unfortunately, there has been a significant decrease in the number of better grades.

Among the causes of this decline, a major problem is that the students bring less and less knowledge from high schools. There are more and more students in higher education, and thus their level of knowledge varies on a large scale. We can do nothing else but get adapted to this new situation and shape and re-form the course syllabus so as to be able to transfer the required knowledge to the students. We have to take into consideration that some parts are omissible from the curricular or not and/or maybe we teach it according to the base of the actual theme. The students know less but this knowledge is certain.

4. The attempts made to improve performance

In the first semester of the academic year 2009/2010 we made a try with a new syllabus. We did not reduce the course material, but re-distributed the topics among the

practice lessons and the laboratory exercises. We started the practice lessons with probability theory as before, but from the 9th week we continued with linear algebra, and mathematical statistics was transferred to the labs. Due to the change, we did not need computers on the practice lessons.

We decreased also the number of quizzes. There were three quizzes in practice lessons (on weeks 4, 9, and 14), and two quizzes in labs (on weeks 7 and 14); each for 20 marks. Hopefully students can achieve better results after this step (DF5 2009).

3.1. The future

Unfortunately, the lack of knowledge brought from high schools should be taken into consideration. It is therefore necessary to assess the knowledge of first-year students. This is necessary not only because of the course "Technical Mathematics", but also for the prerequisite courses Calculus I and Calculus II.

3.1.1. National survey

Due to the initiative of the Budapest University of Technology, this year was the first one in which a standardized assessment test has been written on a national level.

The objectives of the 2009 test:

- To examine the knowledge of the students entering higher education, and determine if it satisfies the requirements of their chosen degree programs.
- To investigate whether the recruitment score provides sufficient information on the student's knowledge?
- To verify our former conclusions, and support them with further empirical data.[2]

Execution of the examination:

The students wrote a sixty minute test during registration week; that is, at a time when the higher education institute had not yet "interfered" with their training. We at the College of Dunaújváros could have the test been written on September, 3rd with 500 first year students. The thing that the people preparing the test had in mind during their work was to examine the knowledge of the students important for higher education and necessary for the students' successful advancement. The test mapped especially the knowledge learnt in and brought from high school.

The questions were chosen centrally, and a solution with detailed instructions for markers was also enclosed to make marking as uniform as possible. Each institution itself organized the test and its marking, following the central instructions. Our

[2] http://members.iif.hu/rad8012/index_elemei/kriterium.htm

fellow instructors recorded the results in a centrally prepared Excel sheet, and sent it back for process. Peer comments on the results of a centrally prepared and recorded in an Excel table, they were sent back for processing.

The test consisted of 12 multiple-choice questions, two exercises about formulas and three computational problems.

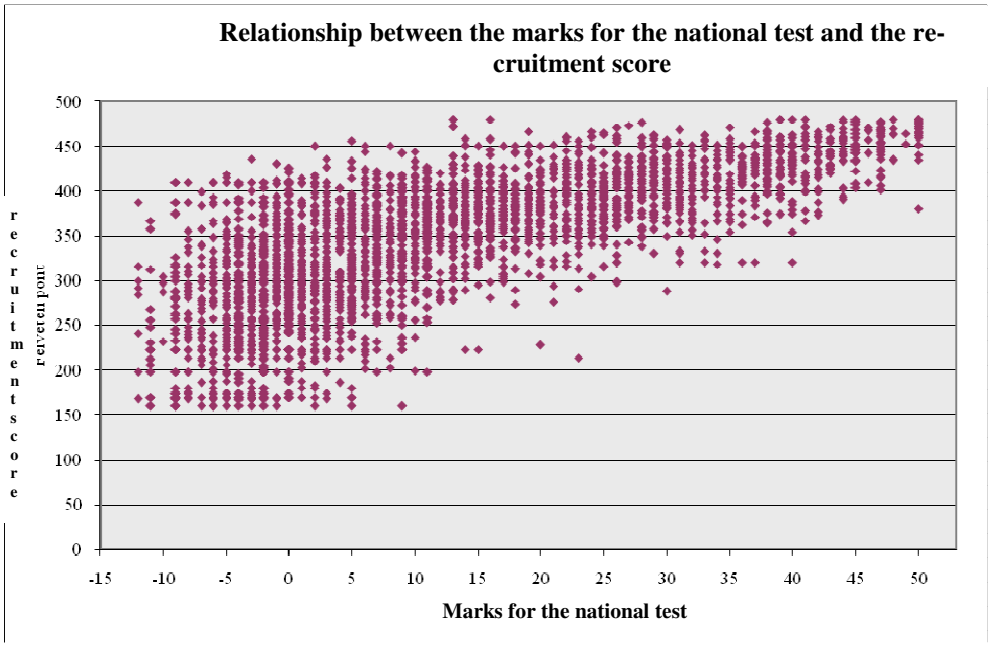
A very high percentage of students dealt only with the multiple-choice questions. They did not touch the two exercises requiring the knowledge of formulas, or used only basic formulas and did not care that they used the formulas with data different from those given in the exercises.

The third part contained the computational problems. Only ten percent of the students dealt with the last part. Even among those who made an attempt, only a few got maximum marks for them, many listed only the data gathered from the text, and made no real work.

In general, we could find numerous types of errors in the solutions.

In Figure 4, a diagram containing more than 3000 data is shown. Note how much the marks of this test and the students' recruitment scores are unrelated. I made a correlation calculation from the data where it was 0,723. One can see from the correlation coefficient that there is a correlation between the results of the qualifying examination and test but it isn't a strong connection.From the fact that someone had a high recruitment score does not follow that they got good marks on this test. Unfortunately, it seems that, among those with recruitment score between 300 and 400, we can find all the kinds of marks from the lowest to the highest. A question is raising that the process of the marticulation or perhaps the measuring isn't good. The Cronbach alpha of the measuring was calculated by the authors (Katalin Radnóti) which was 0,88 and this is a real good value. This confirms that hypotesis that the process of matriculation have to revise.

Figure 4. The summarized data of the national test



Source: Radnóti (2010)

Unfortunately, now we know what we have only suspected: that the knowledge of students coming from high school is less and less. My hope is that, due to this test, we become more aware of the knowledge of the students, and by using our experience we can make the syllabus and our teaching methods better.

In the Engineering Mathematics course, the practice lessons and labs allow us to avoid the frontal – instructor centered – teaching methods, and to bring forward practice oriented – student centered, knowledge-based – methods supporting individual work.

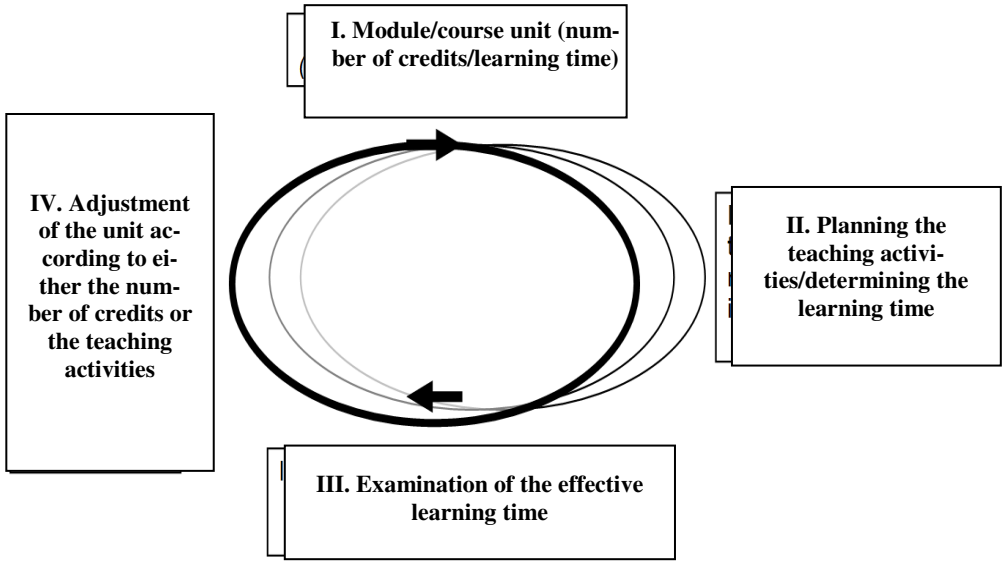
In the future, we intend to strengthen the changes that we have already started, with the involvement of students. Within the framework of quality assurance, we already surveyed the students at the end of each semester. We asked for their opinion of the course and the instructors. We would like to make a questionnaire in which we can measure also the success of the course. Among the types of questions, there would be ones aimed at the required goals of the training, at gaining skills, and at the time necessary for acquiring these skills.

With this questionnaire, we would get better feedback on how effectively we teach. We could make it more effective with reorganizing the quizzes. Quizzes and tests provide information on the knowledge of the students, but there are skills that

must be acquired till the end of the semester. With the questionnaire, the students could put it into words more accurately with which part of the course material they have more problems, and to what would we need to devote more time.

We cannot neglect to determine the students' workload. In the future, this should be taken into account and should be reviewed constantly. A circular investigation should be started, which I represent in the following figure.

Figure 5. Determining the load and task of the students



Source: Kadocsa (2004)

With this method, we receive a constant feedback from the students, by means of which we can redistribute the course material so as to make it more achievable for the students.

The results of the national test surveying the students should be assessed not only by us, instructors. The experts and politicians, who develop the recruitment system and determine the recruitment scores, should also draw the conclusions. Unfortunately, in many cases we should draw the conclusion that the scores achieved by the students do not reflect their real knowledge.

5. Summary

We should continue the trend that started in the country and at our College. We should take into account the students' knowledge and learning skills, and form the

syllabus accordingly. We should consider what and how we teach in order to make it easier for the students to acquire the necessary knowledge. We should transfer 'usable' knowledge: one that makes it easier to learn the material of the special courses of their own profession.

References

DF1, Valószínűségszámítás Tantárgyi program (MA-004) 2006: Dunaújvárosi Főiskola Matematika Intézet

DF2, Matematikai statisztika Tantárgyi program (TE-005) 2006: Dunaújvárosi Főiskola Matematika Intézet

DF3, Matematikai módszerek Tantárgyi program (TE-007) 2006: Dunaújvárosi Főiskola Matematika Intézet;

DF4, Műszaki matematika Tantárgyi program (MA-009) 2008: Dunaújvárosi Főiskola Matematika Intézet.

DF5, Műszaki matematika Tantárgyi program (MA-009) 2009: Dunaújvárosi Főiskola Matematika Intézet.

Kadocsa, L. 2004: A kreditalapú moduláris rendszere, mint a curriculumfejlesztés új minősége III. Felsőoktatási Kreditfórum, Dunaújváros 2004 jún. 3-4.

Radnóti, K. 2010: magánközlemények

How do test methods affect reading comprehension test performance?

Agnes Loch[1]

The paper describes how statistics were employed in language testing research to explore the effect of two test method variables of foreign language reading comprehension tests on test takers' reading comprehension performance. Statistical procedures were applied at three stages of the research: in the validation of the main research instruments, i.e. the reading tests (1), in grouping the participants into four comparable groups (2), and in analysing the participants' test performances on two reading comprehension tests (3).

Statistics and qualitative data analyses show that task type and native language use as test method variables, rarely have a statistically significant affect separately, but may rather exert a joint effect on performance.

Keywords: language testing, test method variables, test method effect

1. Introduction

The research explores the effect of two test method variables of foreign language reading comprehension tests - task type and native language use - on test takers' reading comprehension performance. The aim of the investigation is either to support or to reject the hypothesis that neither task type nor native (L1) or target language (L2) use influences reading comprehension performance significantly if the tasks target the same construct.

As communicative competences cannot be measured in any other way than by observing an individual's language performance, which is supposed to reflect the underlying competences, it is essential to consider all the possible factors that, besides actual reading comprehension ability, may influence performance and test results. Bachman (1991) sets up three categories to classify these contaminating factors: personal attributes (e.g. age, gender, occupation), test method facets (e.g. task type, dictionary use), and random factors (e.g. weather conditions, the test taker's physical or emotional state). Personal attributes and random factors are beyond the control of the examiner whereas method facets can be manipulated to make the assessment procedure and the results valid and reliable.

[1] Loch, Agnes, PhD, senior lecturer, Budapest Business School, Budapest, Hungary

Test method facets are a set of factors that specify the actual method of the assessment procedure. They cover the following categories: testing environment, test rubric, input, expected response, and the relationship between input and response (Bachman 1991). Test method facets can be carefully designed and controlled in order to minimize their distorting effects on an individual's test results. As the present study aims to investigate the effects of test method facets on reading comprehension performance, the crucial facets are the facets of the reading text and the input. The input includes the task and the use of L1 versus L2 in the input (and the expected response).

Several comparative studies have been conducted to investigate the effects of task types and native language use on reading comprehension. In a recent study Liu (2009) compares three task types and claims that gap-filling tasks have a significantly negative effect on test-takers' reading comprehension performance. Other researchers are more cautious in their conclusions. Shohamy's (1984), Wolf's (1993), and Gordon and Hanauer's (1995) studies are particularly remarkable because besides comparing short-answer questions and multiple choice tests, they also examined the effects of native language use. Based on their findings it is clear that the items or questions in the task provide additional information for the reader that may help comprehension. The amount and the quality of this information may substantially differ in the case of different task types. Native language use in the items and the expected response resulted in higher performance in each of these studies. However, due to weaknesses in research methodology the authors could not draw any general conclusions. It is still not explained whether improvement in performance in the studies was due to better understanding the questions, or to using the additional information in the questions to better understand the text itself.

It is worth noting that, although many theoretical works discuss the criteria for validation and reliability in detail (e.g. Bachman 1991, Bárdos 2002, McNamara 2000, Popham 1990), researchers rarely provide any information about the validation methods of the research instruments in their empirical studies.

2. Research questions

The broad research area of the present study is the investigation of how two testing variables – task type and the language of task and task completion – affect reading comprehension performance. The formulation of the exact research questions was based on the taxonomies in the literature (Alderson 2000, Urquhart-Weir 1998) as well as the findings of a teacher's questionnaire (Loch 2007, 2009a) and the statistical results of a Pilot Study including 185 participants. Two task types, short-answer questions (SAQ) and multiple choice items (MC), were selected for comparison in the Main Study. Thus, the main research questions focussing on the mutual relationship of task type and language use, were formulated as follows:

- How does the use of the native language in task rubrics, test items, and task completion influence reading comprehension performance in a short-answer questions test and in a multiple choice test?
- How do short-answer questions and multiple choice items as task types influence reading comprehension test scores when the task and the expected response are in English as the target language and in Hungarian as the native language?

3. Research method

The study compared the participants' performance on two reading comprehension tests including either short-answer questions (SAQ) or multiple choice items (MC): one in English as the target language, and one in Hungarian as the native language. Two sets of texts were selected, and four different tests were developed for each text: a short-answer questions test with rubrics and questions in English and in Hungarian, and a multiple choice test with rubrics and four options in English and in Hungarian. Thus, there were two sets of two texts and eight reading tests altogether.

Four groups of minimum fifty students each were involved in the research. Each group completed two tests. The participants in the same group worked with the same task type in the two tests, with language as the changing variable. Table 1 shows the groups and which versions of the tests they completed.

Table 1. The research matrix

	Group A	Group B	Group C	Group D
Test 1	SAQ test in English	MC test in English	SAQ test in Hungarian	MC test in Hungarian
Test 2	SAQ test in Hungarian	MC test in Hungarian	SAQ test in English	MC test in English

Source: own creation

Statistical analyses were employed at three different stages of the research:

1. in the validation procedure of the main instruments,
2. in forming comparable groups of participants,
3. in analysing data from test results and from questionnaires.

4. Statistical procedures

4.1. Validation of the main instrument

The validity of the tests was ensured in several ways. Besides qualitative methods, the statistical analysis of test results in two pilot studies (involving 185 and 202 students altogether) and correlating test scores with TOEFL[2] scores as a validated third measure (concurrent validity) helped ensure the validity and the reliability of the research instruments.

The data from the tests were processed using SPSS software (Version 11.0). Classical item analysis was carried out to calculate means, standard deviation, item test correlations, and reliability coefficients. The purpose of the analysis was to gain information about the tests as a whole, and to identify items for deletion or modification. Statistical results were expected to help validate the tests and decide which texts and items could be included in the final test booklets for the Main Study.

Poorly performing items were identified and modified after Pilot Study I. Besides modifying the wording of the questions, new items and new distracters were devised when necessary. After administering the tests in the second pilot stage, descriptive statistics and reliability analyses were carried out (Table 2). The results showed that the reliability of the tests increased considerably. Reliability for the SAQ test increased from $\alpha = .7399$ to $\alpha = .8398$, and for the MC test from $\alpha = .4327$ to $\alpha = .6631$ in the case of *Test 2*. In the case of *Test 1*, for the SAQ test it was $\alpha = .8149$, and for the MC test $\alpha = .7012$. The lower reliability coefficients of the MC tests were assumed to be related to the fewer number of items: the first version of both SAQ tests contained 30 items, whereas the MC tests contained 16 items only.

Table 2. The statistical analysis of the tests in Pilot Study II

Test	*M*	Facility value (%)	*SD*	Reliability (Alpha)	Adjusted reliability
Test 1 SAQ -E	21.98	73.2	4.8739	.8149	-
Test 1 MC -E	8.25	51.5	3.1057	.7012	.8148
Test 2 SAQ -E	18.87	62.8	5.6149	.8398	-
Test 2 MC -E	8.45	60.3	2.8559	.6631	.8082

Source: own creation

As reliability increases as items are added (Henning 1987, Csapó 1993), it was assumed that the reliability of the MC tests would increase if the number of items in the tests were increased to a specified length. The Spearman-Brown Proph-

[2] Testing English as a Foreign Language - the most widely accepted English language test developed by ETS (Educational Testing Service) US.

ecy Formula states the relationship between reliability and test length mathematically based on the assumption that the added items are of similar quality to other items in the test. Using the Spearman-Brown Prophecy Formula it was possible to calculate what the reliability of the MC tests would become if they contained the same number of items as the respective SAQ tests. The formula says

$$r_{ttn} = \frac{nr_{tt}}{1+(n-1)r_{tt}}$$

where, r_{ttn} = the reliability of the test when adjusted to n times its original length
r_{tt} = the observed reliability of the test at its present length
n = the number of times the length of the test is to be augmented.

By using the Prophecy Formula, in the case of *Test 1,* the estimated reliability of the MC test version was α = .8148, which corresponded to the respective SAQ test reliability (α = .8149). In the case of *Test 2* the calculated reliability for the MC test was α = .8082, which is also above the .8000 level. Although it was not possible to lengthen the MC tests to that extent, using the Spearman-Brown Formula was still relevant, and its results were reassuring. In an indirect way these results provided information about the items and confirmed their appropriateness for testing purposes.

Besides considering the reliability of the tests, the means and the facility values (calculated from the means) were also considered (Table 2). The analysis of the statistics helped to identify items which were particularly difficult or easy for the pilot population. By deleting problematic items it was possible to set the difficulty (facility value) level of the tests. After deletions, the item number of the SAQ tests was set at 24.

4.2 Forming comparable groups of participants

In order to compare performances on different test versions and draw conclusions on method effects, it was of crucial importance to set up four groups of participants, and to ensure that the groups were equivalent regarding their language proficiency.

Two-hundred and sixty-seven first-year students participated in the Main Study from Budapest Business School. On the basis of their TOEFL tests results (Phillips 1990), the participants were arranged into four groups of comparable language proficiency. As raw scores might not be regarded as interval data, the scores were converted by using the TOEFL Conversion Table. Then, the means and the standard deviations of the four groups were computed (M_A= .439.8, SD =63.6; M_B = 440.7, SD = 61.4; M_C =440.2, SD = 68.9, M_D = 441.8; SD = 63.9), and the means were compared using analysis of variance (ANOVA), which confirmed that there was no significant difference between the group means ($F_{3,234}$ = .168, p = .918), and thus, the groups were comparable. In addition, the participants' ability logits were

computed in a Rasch analysis, and were also compared in an analysis of variance ($F_{3,234}$ = .422, p = .737). The result showed that the groups were highly comparable (Loch 2009b).

4.3 Analysing data from test results

Following the traditional line of Classical Test Theory (CTT), the scores were regarded as interval data and were processed accordingly. For the statistical analyses the Statistical Package for Social Sciences software[3] was used. As the procedures of Item Response Theory (IRT) are recommended for much larger sample sizes, their application was limited and complementary in the present study (Bachman 2004, Baker 1997, Horváth 1997).

The test takers' performances on the eight test versions were compared by using both parametric and non-parametric statistical computations because distribution on one of the eight tests was slightly skewed. The procedures applied are shown in Table 3.

Table 3. Statistical procedures employed in the data analysis

Type of analysis	Non-parametric tests	Parametric tests
Checking for distribution	Chi-square	Chi-square
Comparing means (two data sets)	Wilcoxon test Mann-Whitney U test	Paired-samples *t* test Independent *t* test
Comparing means (more than two data sets)	Kruskall-Wallis test	ANOVA
Relationship between variables	Spearman rank order correlation	Pearson product moment correlation
Relationship among variables		Regression analysis Univariate analysis of variance

Source: own creation

Inferential statistics were run at three levels. First, *Test 1* and *Test 2* versions were compared to see if they were the same difficulty level. Secondly, the English (L2) and the Hungarian (L1) versions of the same tests were compared to check them for language effect. Finally, the short answer question version and the multiple choice version of the same tests were analysed to investigate task type effect. As the four groups were highly comparable concerning language ability, group differences were excluded from the possible reasons for potential differences.

[3] SPSS Inc. (1989-2003). *Statistical Package for Social Sciences* (Versions 11.0, 12.0)

When comparing the difficulty level of the tests, it was found that in the case of the SAQ tests the difference between test means was significant at the $p < .001$ level, with *Test 2* being more difficult for the participants than *Test 1*. However, in the case of the MC tests: in one group it was *Test 1*, whereas in the other group *Test 2* that proved to be significantly more difficult for the students. Questionnaire data and group interviews seemed to suggest that the test takers' insufficient language knowledge did not allow them to choose the correct answer from the only slightly different options provided in the MC items, which resulted in inconsistent test-taking behaviour. In spite of these results, positive correlation was found between the students' scores on the two tests.

Next, the English and Hungarian versions of the same tests were compared to investigate native language use effect on test performance. Both the means and the facility values showed that the Hungarian versions were easier and elicited higher performance (Table 4) although in the case of the SAQ tests the difference did not reach statistical significance at the .05 level.

Table 4. Comparative data about the SAQ and the MC tests

	Test 1 SAQ		Test 2 SAQ	
	in English	in Hungarian	in English	in Hungarian
Group	Group A	Group C	Group C	Group A
N	64	60	62	60
M	13.47	13.67	9.87	10.70
Range	22	18	20	20
SD	4.8500	4.7929	4.8332	4.8198
Variability	23.523	22.972	23.360	23.231
Facility value (p)	.5625	.5688	.4113	.4458
	Test 1 MC		Test 2 MC	
	in English	in Hungarian	in English	in Hungarian
Group	Group B	Group D	Group D	Group B
N	56	66	68	58
M	6.89	8.36	6.57	7.31
Range	12	10	10	9
SD	2.4913	2.4970	2.5934	2.5902
Variance	6.206	6.235	6.726	6.709
Facility value (*p*)	.4308	.5625	.4373	.4908

Source: own creation

In the case of the MC tests, however, the test takers performed significantly better on the Hungarian version. The mean difference between the English and the Hungarian versions of *Test 1* was significant *(t_{120}=-3,245, p = .002)*, and there was a medium effect size (d = .59) (Dancey-Reidy 2004). The adjusted R squared (R^2=

.080) showed that eight percent of the variation in test scores could be explained by the different languages. In the case of *Test 2* the non-parametric Mann-Whiney U test also indicated significant difference (z = -2,054, Asymp. Sig.= 0,40) (Table 5). This indicates that, at least in some cases, the language of the task had a decisive influence on the response.

Table 5. Test of significance in relation to the Mann-Whitney U test statistics of test scores on the L1 and L2 versions of the MC tests

	Test 1	Test 2
Mann-Whitney U	1223.500	.1451.500
Wilcoxon W	2819.500	3662.500
Z	-3.230	-2.054
Asymp. Sig. (2-tailed)	.001	.040

p<.05

Source: own creation

The third comparison focused on investigating the effect of task type. As the English multiple choice version of *Test 2* was slightly positively skewed, in this case the Mann-Whitney U test was applied. Significant difference was found in one case (Table 6): between the means of test scores on the SAQ and MC versions of *Test 1* in English ($t_{113} = 3.800$, $p < .001$), with an effect size $d = .72$. Performance on the SAQ test highly exceeded performance on the MC test. Task type explained 10 percent of the variation in performance ($R^2 = .105$).

Table 6. Results of Independent Samples t test and Mann-Whitney U test to compare means from different test formats

Test	Version	Groups	t	df	Sig. (2-tailed)	Z	Asymp. Sig. (2-tailed)
1	English SAQ + MC	A + B	3.800	113	.000*		
	Hungarian SAQ + MC	C + D	1.288	121	.200		
2	Hungarian SAQ + MC	A + B	1.201	113	.232		
	English SAQ + MC	C + D				-.682	.496

* *p*<.001

Source: own creation

It was also important whether task type and language use exercised any joint effect on the participants' performance. In the case of *Test 1*, the two variables jointly did not show a significant relationship with the test scores although had a significant effect on means separately (Table 7). This fact as well as the findings that the MC format affected performance on *Test 1* negatively and on *Test 2* positively indicate that task type and other factors may interfere.

Table 7. Tests of significance in relation to the Analysis of Variance for the joint effect of task type and language use

Variable	df	F	Significance	Partial Eta squared
Test 1				
Corrected Model	3	6.925	.000	.082
Task type	1	13.204	.000	.053
Language	1	5.183	.024	.022
Task type*language	1	3.382	.067	.014
Total	238			
Test 2				
Corrected Model	3	1.384	.248	.017
Task type	1	1.595	.208	.007
Language	1	2.504	.115	.011
Task type*language	1	.243	.623	.001
Total	238			

Source: own creation

Statistics confirmed the expectations that the participants' reading scores would significantly correlate at the .01 level with their scores on the TOEFL papers on receptive skills (r = .584). However, the overlap between scores was not particularly large, which indicates the distinctiveness of reading skills

5. Summary of findings

In the light of the results, it is obvious that the research hypotheses gained partial verification only. The statistics showed that in most cases no significant difference was found. However, there were exceptions, both in the task type and the language use comparisons, when the mean differences reached the statistically significant level. Due to these mixed results, no general conclusions can be drawn.

Although results were not consistent, some of the findings strongly suggest that test method variables may exert a joint effect with other factors such as text difficulty or test takers' characteristics. As the mixed results

gained about test format effects in this study do not provide a comprehensive conclusion, further research is needed in the area, especially in two directions: how task types influence performance at different levels of proficiency, and how task type effect is related to the conceptual and linguistic difficulty of a reading text.

References

Alderson, J. C. 2000: *Assessing reading.* Cambridge University Press, Cambridge.

Bachman, L. F. 1991: *Fundamental considerations in language testing*. Oxford University Press, Oxford.

Bachman, L. F. 2004: *Statistical analyses for language assessment.* Cambridge University Press, Cambridge.

Baker, R. 1997: *Classical test theory and item response theory. Theory in test analysis* [Special report 2, Language testing up-date]. Department of Linguistics and Modern English Language, Lancaster University, Lancaster.

Bárdos, J. 2002: *Az idegen nyelvi mérés és értékelés elmélete és gyakorlata.* Nemzeti Tankönyvkiadó, Budapest.

Csapó, B. 1993: Tudásszintmérő tesztek. In Falus, I. (ed.): *Bevezetés a pedagógiai kutatás módszereibe.* Keraban Kiadó, Budapest, pp. 277-317.

Dancey, C. P. - Reidy, J. 2004: *Statistics without maths for psychology*. Pearson/Prentice Hall, Harlow, UK.

Gordon, C. M. - Hanauer, D. 1995: The interaction between task and meaning construction in EFL reading comprehension tests. *TESOL Quarterly,* 29, pp. 299-322.

Henning, G. 1987: *A guide to language testing*: *Development, evaluation and research*. Newbury House, Cambridge, MA.

Horváth, Gy. 1997: *A modern tesztmodellek alkalmazása.* Akadémiai Kiadó, Budapest.

Liu, F. 2009: The effect of three test methods on reading comprehension: An experiment. *Asian Social Science, 5,* No. 6. pp. 147-153. Retrieved on 01/05/2010 from http://ccsenet.org/journal/index.php/ass/article/view/2491.

Loch, A. 2007: Task types in assessing reading comprehension. *BGF Tudományos Évkönyv.* BGF, Budapest, pp. 358-362.

Loch, A. 2009a: Native language vs. target language use in language testing. *Acta Linguistica* 8, Vol. 2, University of Matej Bela, Banska Bystrica, Slovakia, pp. 64-70.

Loch, A. 2009b: Minták összehasonlíthatóságának bizonyítása statisztikai eljárásokkal. Paper presented at *New Methods in Applied Linguistics Research Conference*. 26-27. October 2009, Budapest.

McNamara, T. 2000: *Language testing.* Oxford University Press, Oxford.

Phillips, D. 1990: *Longman practice tests for the TOEFL*. Longman, New York.
Popham, W. J. 1990: *Modern educational measurement*. Allyn and Bacon, Boston.
Shohamy, E. 1984: Does the testing method make a difference? The case of reading comprehension. *Language Testing,* 1, pp. 146-169.
Urquhart, A.-Weir, C. J. 1998: *Reading in a second language: Process, product and practice.* Longman. London.
Wolf, D. F. 1993: A Comparison of assessment tasks used to measure FL reading comprehension. *Modern Language Journal*, 77, pp. 473-489.

How can statistical analysis prove the existence of an independent foreign language skill?

Ágnes Dévény[1]

The aim of my research was to prove that foreign language mediation is an independent language skill that can be measured by specific criteria and as part of a foreign language examination it contributes to a more complex assessment of the test-takers' language knowledge.

Besides other methods that are beyond the scope of this paper, the most important part of the research was the descriptive and mathematical statistical analysis of the results of the 18 examination periods covering more than 27000 test results from Budapest Business School Examination Centre. I wanted to present in this paper how I used statistical analyses to prove the existence of an independent foreign language skill.

My research verified my hypothesis that written mediation examination task measures a segment of language knowledge independently; skills (subskills) appearing in it are not present in other examination tasks. Foreign language mediation is an independent language skill; therefore bilingual language examinations can measure skills that are not measureable by monolingual examination systems.

Keywords: assessing foreign language competence, language testing, pedagogy

1. Introduction

Is foreign language mediation an independent language skill? Is it part of our conception of language proficiency and a segment of the language learners'/users' language competence? There have been heated debates in professional circles on the role and function of mediation and the debates have not been settled yet (Szabari 2001; Bárdos 1997, 2005; Heltai 2001). Experts interpret and explain the concept of mediation in many different ways and their opinions vary, reflecting the insecurity of specialists in the field of bilingualism/multilingualism. Debates become even more heated when experts have to decide whether a mediation task (e.g. written mediation test) can be considered as legitimate language examination subtest (task) (Klaudy 1984, 1986a, 1986b, 1990; Fekete 2001, 2002). The scientific investigation of bilingualism/multilingualism in education and testing, as well as the interpretation

[1] Ágnes Dévény, PhD, associate professor, Budapest Business School, Institute of Foreign Languages and Communication, Faculty of Commerce, Catering and Tourism, Hungary (Budapest)

of the notion of foreign language mediation, are especially delicate topics as they conceal the diversity and distinctness of opinions and theories of language skills and abilities. Opponents of bilingual examinations do not regard any type of mediation as an independent language skill and an element of language knowledge, and do not accept the concept that by testing and measuring this skill we can get a more complex idea of the candidate's foreign language proficiency. On the contrary, many of them fear of the negative washback effect of the mother tongue on foreign language acquisition (Nikolov et al. 1999). Contradictory views are reflected in the fact that whereas mediation task was completely left out from school-leaving (maturity) examination tasks, several bilingual and monolingual examination systems have been accredited in Hungary since the beginning of the accreditation procedures in 1999 (Einhorn 1998; Nikolov et al. 1999; Alderson 2001). The acceptance of bilingual examination systems was justified only by some needs analyses and Hungarian language examination traditions, but there was not any scientific research verifying that foreign language mediation is an independent language skill, and it is part of the candidate's language competence.

The principal argument of the opponents of the mediation task is its negative washback on teaching practice, and the difficulties of its evaluation arising from the complex nature of the skill, which definitely includes several subskills.

The negative impact cannot be considered as a generally acceptable counter-argument in all cases. According to Bachman & Palmer (2000) a language test can be useful and reasonable if it corresponds in demonstrable ways to language use in a specific target language use domain and is based on certain procedures of needs analysis.

A national survey (Teemant et al. 1993) and some surveys of smaller scope (Major 2000; F. Silye 2004; Dévény – Szőke 2007) were conducted mainly in the field of professional language usage during the last decade. The results of these surveys show that mediation is verified as a real-life domain of language usage, so it is justifiable as a test task, but they do not produce any evidence of acknowledging it as an independent language skill.

2. Posing (raising) the problem

The aim of my research was to prove that foreign language mediation is an independent language skill. It is part of the candidate's foreign language competence, and the written mediation task in the language examination system, examined in my survey, is suitable for evaluating this skill just like other subtests each using its measurement criteria respectively. This language skill and its measurement behaves fundamentally in the same way in the test battery as other subtests, and its application in the examination system contributes to obtaining a

more complete and more objective picture of the candidates' foreign language proficiency.

I strove to answer two sets of questions in my study: *1. from the point of view of the examination; 2. from the point of view of language teaching*. In this paper I will only deal with the first point of view and present my research questions concerning this point:

– What is the construct validity of the written mediation task in the language examination?

– Is mediation a reliable examination task in the language examination system?

– Are there any subskills in written mediation tasks that appear in other subtests as well?

– What does the inter-subtest correlation of written mediation task with other examination tasks show?

– Based on these research questions the following *hypothesis* was created:

– The special language skills that appear and can be measured in written mediation tasks are not present in other subtests, therefore bilingual language examinations can measure skills that are not measureable by monolingual examination systems.

In my research I used statistical analyses of the examination scores to examine the relationship between the results of the subtests. Although different types of statistical analyses e.g. confirmatory factor analysis (CFA) has been used in psychology and other social sciences to support measuring abilities, personality traits etc., the concept of using statistical analyses to prove the existence of an independent foreign language skill is not a usual approach in language pedagogy as until now in classical and modern test theories statistical analyses were used for solving different testing and test construction problems but did not serve as a method of proving the existence of a language skill.

3. The background of the research

For the research I chose the examination system of Budapest Business School, a bilingual, criterion-referenced, LSP[2] examination system, where along with the traditional language skills (speaking, writing, reading comprehension, listening comprehension) mediation skill is also measured at different levels of the examination. Examination tasks of complex intermediate (B2) exam at the time of the research were the following:

[2] LSP = language for specific purposes

1. Written examination:
   - grammar test
   - reading comprehension test
   - writing test (generally business letter)
   - written mediation test from Hungarian to foreign language
2. Listening comprehension test
3. Oral examination:
   - introductory conversation
   - oral mediation task (mediation of a newspaper article into Hungarian)
   - speaking test (conversation on vocational topics
   - situation

The Written mediation test under research as an intermediate examination task meant the following: the examinee had to mediate in written form a Hungarian text of approximately 150 words, with an output of approximately 100 words. The requirement was not a word for word translation of the text but it involved a summarizing element. Candidates were not allowed to use a dictionary.

4. Methods of the research

4.1. The characteristics of the sample

For the statistical analysis of the examination results I used the *examination scores* of the candidates who took the intermediate LSP exam from 2000 to 2007 in BBS[3] Examination Centre. The members of the sample were 17-35 year-old men and women, typically college and university students (97-98%), 2-3% of them were employees from different areas of the economy. The size of the sample was slightly different in each examination period, but it ranged typically from 1000 to 2000, summing up to 27,832 candidates in 18 examination periods.

4.2. Methods of data analysis

As a method of data analysis I used descriptive and inferential statistics calculating by the help of SPSS (inter-subtest correlations, subtest – final test score correlations, reliability analysis (index) of the subtests, factor analysis, multiple regression analysis, cluster analysis).

[3] Budapest Business School

5. The results of the research

Based on the statistical analyses I wanted to support my hypothesis claiming that written mediation examination task can measure a skill or a complexity of subskills that cannot be measured with other examination tasks, thus besides with the traditional language skills an *independent foreign language mediation skill* exists. Statistical analyses in this paper will be illustrated on the result of one examination period (January 2005) and the results of all 18 exam periods will be summarized separately. Figure 1 shows the main characteristics of the chosen examination period.

Figure 1. Descriptive statistics of the examination test scores (Sample period: January 2005) (n=1669)

	Grammar test	Reading Comp.	Writing test	Written mediation	Listening Comp.	Introct. conv.	Speaking	Situation	Oral mediation
Mean	10.78	12.12	12.31	5.81	9.52	15.53	13.21	14.72	7.26
SE of Mean	.088	.090	.093	.049	.096	.076	.101	.100	.050
Median	11	12	12	6	9	16	14	15	8
Mode	10	13	12	6	9	18	14	20	8
SD	3.610	3.680	3.814	1.994	3.917	3.093	4.116	4.073	2.038
Variance	13.030	13.544	14.549	3.975	15.341	9.564	16.943	16.588	4.153
Skewness	-.170	-.148	-.103	-.207	.385	-.643	-.455	-.757	-.585
SE of Skewness	.060	.060	.060	.060	.060	.060	.060	.060	.060
Kurtosis	-.561	-.415	-.414	-.280	-.412	,085	-.284	.268	-.110
SE of Kurtosis	.120	.120	.120	.120	.120	.120	.120	.120	.120
Range	19	19	20	10	19	16	20	20	10
Minimum	1	1	0	0	1	4	0	0	0
Maximum	20	20	20	10	20	20	20	20	10

Source: own calculations using SPSS

5.1. Internal correlations as a way of assessing the construct validity of tests

Construct validity is a form of test validation which essentially involves assessing to what extent the test is successfully based upon its underlying theory. (Alderson et al. 1995) In order to reveal the construct validity of the mediation task I calculated the inter-subtest and the subtest – final test scores correlation coefficients of the examination tasks.

5.1.1. Inter-subtest correlations

The idea behind inter-subtest correlation of test components is that if two test-components would correlate very highly with each other, we might assume that the two test components are *not* testing different traits or skills. It means that one of them might be superfluous.

At first I used Pearson Correlation to see the overlapping between the subtests. The optimal value in the case of inter-subtest correlation is between .3 – .5, that means a 9 to 25% overlap of variance between two subtests (Alderson et al. 1995). The results (Figure 2.) show that except for *Grammar test* all other subtests' inter-subtest indices are in the ideal interval. But even in the case of *Grammar test* the overlap is only about 30% of variance of the two subtests.

The problematic point of the measurement is that test results do not always show a normal distribution, as examinees who sit for a criterion-referenced proficiency exam are supposed to be well-prepared – so the distribution curve can be negatively skewed (see Figure 1). That is why the non-parametric correlation was counted as well. The correlation matrix showed nearly the same results as the parametric analysis.

Figure 2. Inter-subtest correlation matrix of written mediation (L1 → L2)[4] task. Comparison of the results of parametric and non-parametric analysis. (January 2005) (n=1669)

			Parametric analysis				Non-parametric analysis			
Subtests	M	SD	r Pearson Corr.	p	R^2	Overlap of variance %	r Spearman's rho	p	R^2	Overlap of variance %
Grammar test	10.78	3.610	**.561****	.000	.315	31	**.550****	.000	.302	30
Reading Comp.	12.12	3.680	**.424****	.000	.179	18	**.416****	.000	.173	17
Writing test	12.31	3.814	**.459****	.000	.210	21	**.445****	.000	.198	20
Listening Comp.	9.52	3.917	**.461****	.000	.212	21	**.450****	.000	.202	20
Introductory Conv.	15.53	3.093	**.367****	.000	.134	13	**.361****	.000	.130	13
Speaking test	13.21	4.116	**.341****	.000	.116	12	**.345****	.000	.119	12
Situation	14.72	4.073	**.386****	.000	.148	15	**.394****	.000	.155	16
Oral mediation	7.26	2.038	**.338****	.000	.114	11	**.340****	.000	.116	12

** Correlation is significant at the 0.01 level (2-tailed).

Source: own calculations using SPSS

The means of inter-subtest correlations of 18 examination periods show a desirable overlapping between the different subtests – an approximately 18 – 25% overlap of variance (Figure 3). The only exception is the *Grammar test* which shows an average of 36% overlap of variance. *Grammar test*, as it does not measure an independent language skill, behaves in the same way in all examination periods.

[4] L1=Language1 (generally mother tongue); L2=Language2 (foreign language)

Figure 3. Overlapping of skills measured by written mediation (L1 → L2) task with skills measured by other subtests
[Mean of inter-subtest correlations of 18 examination periods (%)]

Subtests	r	p	R^2	Overlap of variance (%)
Grammar test	.598	.000	.357	**36**
Reading Comp.	.493	.000	.243	**24**
Writing test	.507	.000	.257	**26**
Listening Comp.	.461	.000	.212	**21**
Introductory Conv.	.417	.000	.173	**17**
Speaking test	.426	.000	.181	**18**
Situation	.437	.000	.190	**19**
Oral mediation (L2-L1)	.420	.000	.176	**18**

Source: own calculations using SPSS

5.1.2. Subtest – total test scores correlation

According to classical test theory of language testing the correlations between each subtest and the whole test might be expected to be higher – possibly around +.7 or more – since the overall score is taken to be a more general measure of language ability then each individual component score (Alderson et al 1995). Subtest – total test scores correlation (Figure 4) is calculated because it shows the impact of the subtest on the total (final, overall) test scores of the examination.

Figure 4. Subtest – total test scores (minus itself) correlation matrix
(January 2005) (n=1669)

	Max. test score	r Total test score minus itself	p	M of subtests	SD of subtests	M Total test score minus itself	SD Total test score minus itself
Grammar test	20	**.671****	.000	11.81	3.427	88.75	22.276
Reading Comp.	20	**.691****	.000	12.34	4.150	88.21	21.654
Writing test	20	**.619****	.000	12.33	3.469	88.23	22.406
Written mediation (L1-L2)	10	**.726****	.000	5.55	1.996	95.00	23.218
Listening Comp.	20	**.637****	.000	10.24	3.656	90.31	22.214
Introductory Conv.	20	**.778****	.000	14.90	3.617	85.66	21.788
Speaking test	20	**.780****	.000	12.38	4.484	88.17	21.047
Situation	20	**.773****	.000	14.24	4.364	86.31	21.177
Oral mediation (L2-L1)	10	**.763****	.000	6.76	2.349	93.80	22.866

** Correlation is significant at the 0.01 level (2-tailed).
Source: own calculations using SPSS

The correlation coefficients in case of Grammar test (r = .671), Reading comprehension test (r = .691), Writing test (r = .616) and Listening comprehension test (r = .637) are on the low side. In the case of all other tests including Written mediation test (r = .726) are above the optimal .7 value. These results show a rather strong correlation between the subtests and the total scores and show how important these subtests are in the test battery. Non-parametric analysis in each case shows a

lower correlation coefficient of Written mediation test but it is near to the desired value.

Figure 5. shows the Subtest – total score (minus self) correlation of Written mediation task (L1 →L2) in 18 examination periods. The correlation coefficients are generally between .6 and .7, a bit lower than the optimal, but convincingly higher than inter-subtest correlation coefficients.

Figure 5. Subtest – total score (minus self) correlation of Written mediation task (L1 →L2) in 18 examination periods

		r	p	M	SD	M Total test score minus Mediation score	SD Total test score minus Mediation score	N
1	Sept. 2000	**.676**	.000	4.88	2.237	89.95	26.069	229
2	Jan. 2001	**.668**	.000	5.88	2.241	91.92	27.253	360
3	May 2001	**.642**	.000	6.27	2.134	98.46	21.735	2072
4	May 2002	**.646**	.000	5.75	2.140	97.63	21.599	2754
5	Jan 2003	**.605**	.000	5.86	2.182	96.21	20.480	1630
6	May 2003	**.595**	.000	5.70	2.034	97.86	20.306	2577
7	Sept. 2003	**.608**	.000	5.56	2,065	94.02	19.519	1180
8	Jan. 2004	**.601**	.000	5.20	2.156	96.91	20.378	1532
9	May 2004	**.584**	.000	5.39	1.928	94.90	20.367	2540
10	Sept. 2004	**.606**	.000	6.07	1.927	93.23	19.681	1280
11	Jan. 2005	**.726**	.000	5.55	1.996	95.00	23.218	1669
12	March 2005	**.581**	.000	5.81	1.994	95.45	20.551	319
13	May 2005	**.596**	.000	5.75	2.047	97.95	21.670	2286
14	Sept. 2005	**.656**	.000	5.91	1.9.50	89.99	19.449	1399
15	Jan. 2006	**.632**	.000	5.92	2.193	97.81	22.802	1596
16	May 2006	**.706**	.000	6.16	2.124	98.41	22.931	1922
17	Sept. 2006	**.608**	.000	5.24	2.191	90.28	24.404	1103
18	Jan 2007	**.672**	.000	5.74	2.226	98.62	23.655	1384

Source: own calculations using SPSS

As a result of the correlation analyses it is observable that the Written mediation test does not behave in a different way when compared with the other examination tasks, and does not show such a high common proportion of variance with them that would indicate that the mediation task measures the same skills as the other examination tasks (subtests). It was obvious that neither the written mediation task nor the other examination tasks can be omitted from the test battery.

5.2. Reliability analysis of the subtests

Reliability of tests is the extent to which test scores are consistent. According to classical item analysis in test construction it is useful to calculate the reliability index (coefficient Alpha) of test items in order to know which test item increases or decreases the reliability of the whole test. Those items that decrease the reliability of the test should be omitted.

In this case I considered each subtest as an item of the whole test (examination) battery and calculated the reliability indices of subtests (Figure 6).

Figure 6. Reliability analysis of subtests (January 2005) (n=1669)

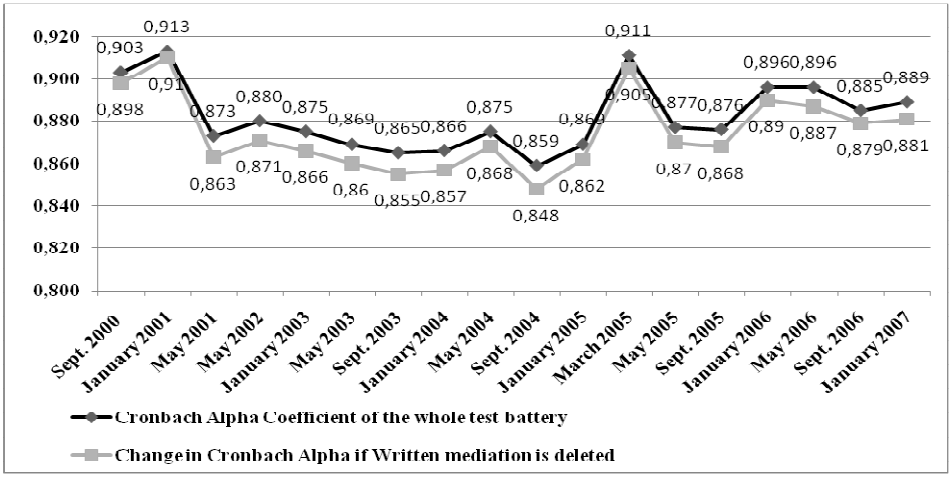
Item-Total Statistics	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Grammar test	90.47	369.586	.656	**.850**
Reading Comp.	89.13	377.324	.579	**.858**
Writing test	88.94	381.048	.524	**.864**
Written mediation (L1-L2)	95.44	422.064	.581	**.862**
Listening Comp.	91.73	376.122	.541	**.862**
Introductory Conv.	85.72	382.069	.678	**.849**
Speaking test	88.04	351.327	.683	**.848**
Situation	86.53	350.619	.698	**.846**
Oral mediation (L2-L1)	93.99	413.115	.681	**.856**

Full Test Reliability Statistics			
	Cronbach's Alpha	N of Cases	N of Items
	.869	1669	9

Source: own calculations using SPSS

As the results show (Figure 7) if *Written mediation test* were deleted from the test battery the reliability index (Cronbach alpha) of the test batteries would decrease in all examination periods. It means that the total scores of the whole examination and the examination itself would be less reliable.

Figure 7. Omitting written mediation (L1 → L2) task from the test battery reduces reliability of the language examinations (18 examination periods)



Source: own calculations

5.3. Factor analysis

The procedure of factor analysis is used to reduce the number of variables accounting for test performance by identifying the common underlying factor (or factors) shared by a series of tests in the test battery.

The results of the KMO (KMO≥0.8) and Bartlett's Test showed that the data were suitable for factor analysis. I used PCA method as it „gives us a way of discovering factors that underlie language performance and ways of testing the relationship among them." (Hatch – Lazaroton 1991) PCA initially attempts to fit as much of the data from the correlation matrix of all variables entered, into a single principle component, in other words, it attempts to explain through the first factor as much of the variability in the data as possible. Once it has done this it trawls through data again, looking for the second component which will explain as much of the remaining variance as possible (Green – Weir 2001).

In the case of the presented examination period all subtests loaded positively on the first factor with .642 or above (Figure 9), which can be considered to indicate a substantial link between them. They all load on the same factor as the first factor represents *general linguistic ability* (Green – Weir 2001).

Figure 8. Factor analysis (January 2005) (n=1669)

Component Matrix(a)

	Component				
	1	2	3	4	5
Grammar test	**.763**	.274	-.152	-.359	-.040
Reading Comp.	**.687**	.267	.441	-.379	.198
Writing test	**.642**	.326	-.523	.166	.398
Written mediation (L1-L2)	**.690**	.371	-.151	.067	-.563
Listening Comp.	**.656**	.319	.422	.488	.076
Introductory Conv.	**.751**	-.453	-.012	.158	-.009
Speaking test	**.759**	-.481	-.049	-.088	.030
Situation	**.776**	-.466	.028	.012	-.050
% of Variance	51.435	14.320	8.683	7.187	6.564
% Total Variance Explained by 5 factors 88.188					

Extraction Method: Principal Component Analysis. KMO = .863

a. 5 components extracted.

Source: own calculations using SPSS

More complex procedures can be followed such as rotation of the factors to see if any clearer solutions present themselves in order to reveal the underlying factors or components. In analysing test results Varimax rotation is the most commonly used procedure (Green – Weir 2001). The result of the Varimax rotation of test scores (Figure 9) showed that the different subtests representing different language skills fell on different factors with rather high factor loadings, except for *Grammar test* that does not represent a separate language skill.

Figure 9. Factor analysis – Varimax rotation (January 2005) (n=1669)

Rotated Component Matrix(a)

	Component				
	1	2	3	4	5
Grammar test	.287	**.586**	**.485**	**.388**	-.019
Reading Comp.	.218	**.875**	.116	.084	.294
Writing test	.191	.140	.187	**.925**	.169
Written mediation (L1-L2)	.199	.159	**.900**	.175	.230
Listening Comp.	.200	.239	.206	.167	**.887**
Introductory Conv.	**.842**	.064	.131	.144	.208
Speaking test	**.858**	.223	.106	.147	.013
Situation	**.862**	.179	.161	.078	.127
% of Variance	30.423	16.206	14.875	14.012	12.673
% Total Variance Explained by 5 factors					
88.188					

Extraction Method: Principal Component Analysis. KMO = .863
Rotation Method: Varimax with Kaiser Normalization
a. Rotation converged in 6 iterations
Source: own calculations using SPSS

Written mediation appears in the 3rd factor in the observed examination period (January 2005) and accounts for 15% of variance of the test battery.

In the 18 examination periods foreign language mediation tests appear in the 2nd to 5th factors. Most of the times it falls on the 2nd and 3rd factor. Factor loadings are between .730 and .911. On average, in the 18 examination periods, foreign language mediation tests account for 17% of variance. In all cases the factors correlate well with total test scores.

Summarising the results of the factor analysis show that foreign language mediation skill represented by Written mediation tasks is one of the components of general language ability. It can be clearly separated from other language skills with its high loading of variance and on average it explains 17% of the variance of the language skills in the test battery so it represents an important part of language knowledge.

5.4. Multiple linear regression analysis (Method FORWARD)

Multiple regression analysis reveals the common subskills between mediation tests and other subtests.

The problematic point was to keep reliability of variables in the test battery on nearly the same level. (The test battery contains subtests that are subjectively scored, so it was important to elaborate and use a system of evaluation criteria that increases the objectivity of evaluation of these tests. Surveying the elaboration of this system is beyond the scope of this paper.)

In this paper I cannot present the whole process and all the steps of multiple linear regression analysis of the test scores of the examination period under research,

so I will show only the Model Summary (Figure 10) representing the most important result of the analysis.

Figure 10. Model summary of regression analysis (January 2005) (n=1669)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.561(a)	.315	.315	1.650
2	.613(b)	.376	.375	1.576
3	.630(c)	.397	.396	1.549
4	.637(d)	.405	.404	1.539
5	.638(e)	.407	**.405**	1.537

a Predictors: (Constant), Grammar test
b Predictors: (Constant), Grammar test, Listening Comp.
c Predictors: (Constant), Grammar test, Listening Comp., Writing test
d Predictors: (Constant), Grammar test, Listening Comp., Writing test, Situation
e Predictors: (Constant), Grammar test, Listening Comp., Writing test, Situation, Reading Comp.
f **Dependent Variable: Written mediation (L1-L2)**

Source: own calculations using SPSS

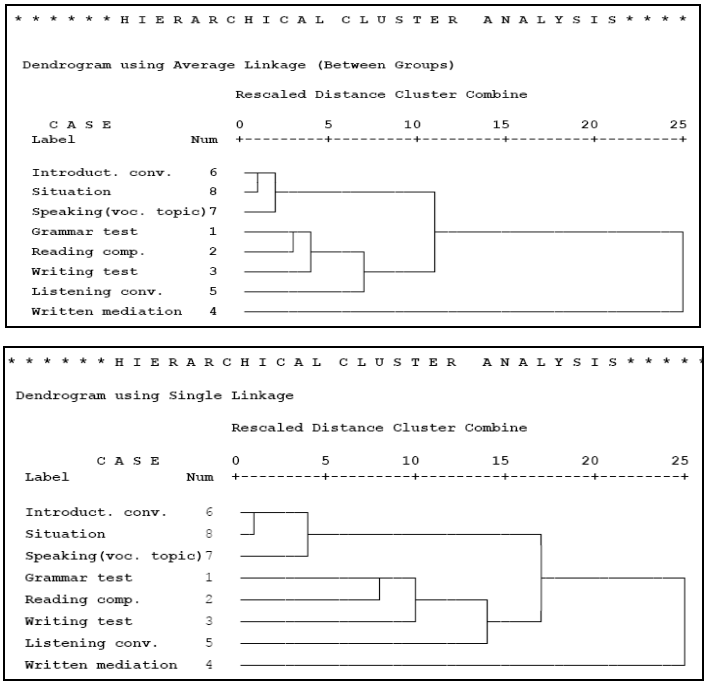
In this case (January 2005) Introductory conversation, and Speaking test on vocational topic were dropped out from the model. The Model Summary reveals that approximately 41% of the variance of Written mediation test can be explained by the help of other variables. It indicates the presence of common subskills. The remaining 59% indicates the existence of subskills that are present only in mediation skill.

Multiple regression analyses revealed the common subskills (the proportion of variation in the dependent variable explained by the regression model) between mediation tasks and other tasks. On average 54% of the variance of the mediation task is not explained by other examination tasks so it represents subskills that can be attributed only to foreign language mediation skills.

5.5. Cluster analysis

Clustering is the assignment of a set of observations into subsets (called *clusters*) so that observations in the same cluster are similar in some sense. Cluster analysis sorts through the raw data and groups them into clusters. A cluster is a group of relatively homogeneous cases or observations. Objects in a cluster are similar to each other. They are also dissimilar to objects outside the cluster, particularly objects in other clusters. So I decided to use this method to see how the different subtests relate to each other (Figure 11).

Figure 11. Cluster analysis – Dendrograms using Average and Single Linkage (January 2005) (n=1662)



Source: own calculations using SPSS

When observing the 18 examination periods in each case *Written mediation test* formed an independent cluster.

Summarising the results of cluster analyses, they also confirmed that foreign language mediation is an independent language skill as mediation tasks are sharply separated from other examination tasks forming an independent cluster.

6. Conclusions

My research verified my hypothesis that written mediation examination task measures a segment of language knowledge independently; skills (subskills) appearing in it are not present in other examination tasks. Foreign language mediation is an independent language skill; it is as legitimate as traditional basic language skills and its usage as an examination task contributes to a more complex notion of the test-takers' language knowledge therefore bilingual language examinations can measure skills that are not measureable by monolingual examination systems.

References

Alderson, Ch. J. 2001: The lift is being fixed. You will be unbearable today. Script of Plenary Session. 10th Hungarian Macmillan Conference, Budapest.

Alderson, Ch. J., Clapham, C. és Wall, D. 1995: *Language Test Construction and Evaluation*. Cambridge University Press, Cambridge.

Bachman, L. F. – Palmer, A. S. 2000: *Language Testing in Practice*. Oxford University Press, Oxford.

Bárdos, J. 1997: *A nyelvtanítás története és a módszerfogalom tartalma*. Veszprémi Egyetemi Kiadó, Veszprém.

Bárdos, J. 2005: *Élő nyelvtanítás-történet.* Nemzeti Tankönyvkiadó, Budapest.

Dévény, Á. – Szőke, A. 2007: Milyen nyelvtudást vár el a munkaadó a munkavállalótól. *Stratégiák 2007 és 2013 között. Tudományos Évkönyv 2006.* BGF, Budapest, 319-330. p.

Einhorn, Á. (szerk.). 1998: *Vizsgatárgyak, vizsgamodellek I*. Országos Közoktatási Intézet, Budapest.

F. Silye, M. 2004: A szaknyelvoktatás és a szaknyelvi tudást felhasználók igényeinek elemzése és megjelenítése egy angol szaknyelvoktatási program-modellben. PhD. Doktori Értekezés. ELTE, PPK. Neveléstudományi Doktori Iskola, Budapest.

Fekete, H. 2001: Az írásbeli nyelvvizsga a számok tükrében. Az ITK ORIGÓ angol középfok. In Fekete H. (szerk.):*Nyelvi Mérce,* Idegennyelvi Továbbképző Központ, Budapest, 2001/1-2, 54-58. p.

Fekete, H. 2002: Az írásbeli nyelvvizsga a számok tükrében. Az ITK ORIGÓ német középfok. In Fekete H. (szerk.):*Nyelvi Mérce,* Idegennyelvi Továbbképző Központ, Budapest, 2002/1-2, 45-56. p.

Green, R. – Weir, C. 2001: Using statistics for the analysis of test an evaluation data. (Manuscript).

Hatch, E.M. - Lazaraton, A. 1991: *The research manual: Design and statistics for applied linguistics*. Newbury House, New York.

Heltai, P. 2001: Iskolai tantárgy-e az idegen nyelv? *Nyelvi Mérce, ITK, Budapest* Vol. 1 Issue 1-2. 13-19. p.

Klaudy, K. 1984: Hogyan alkalmazható az aktuális tagolás elmélete a fordítás oktatásában? *Magyar Nyelvőr*, Vol. 108 Issue 3, 325-333. p.

Klaudy, K. 1986a: A fordítás helye és szerepe a nyelvoktatásban: a „valódi" fordítás a középiskolai nyelvoktatásban. In Lengyel, Zs. (szerk.): *Az idegen nyelvi nevelés-oktatás néhány iránya és lehetősége*. Oktatáskutató Intézet, Budapest, 353-371. p.

Klaudy, K. 1986b: A fordítás oktatásának elméleti alapjai. *Felsőoktatási Szemle* XXXV. 5, 277-283. p.

Klaudy, K. 1990: Átváltási műveletek a fordításban. *Idegen Nyelvek tanítása* 2., 33-45 p.

Major, É. 2000: Milyenfajta angol nyelvtudásra van szükség a nyelvigényes munkakörökben? Vizsgálat a munkaadók elvárásainak felmérésére. *Modern Nyelvoktatás*. 2000, Vol. 6 Issue 1., 33-49 p.

Nikolov, M. - Fekete, H. - Major, É. (eds.). 1999: *English language education in Hungary. A Baseline Study.* The British Council Hungary, Budapest.

Szabari, K. 2001: Anyanyelvtudás, idegennyelvtudás, nyelvi közvetítés. *Nyelvi Mérce.* Nyelvvizsgáztatók és Nyelvtanárok Lapja. ITK, Budapest, Vol. 1 Issue 1-2, 7-13. p.

Teemant, A. – Varga, Zs. – Heltai, P. 1993: *Hungary's Nationwide Needs Analysis of Vocationally-Oriented Foreign Language Learning: Student, Teacher, and Business Community Perspectives.* Művelődésügyi és Közoktatási Minisztérium, a Munkaügyi Minisztérium, az Egyesült Államok Tájékoztatási Hivatala és az Európa Tanács Modern Nyelvek Szekciója, Budapest.

Statistical indicators of poverty in Serbia

Tinde Dobrodolac Seregelj [1]

Half of the global population is regarded as poor to the extent that the very lives of a large number of people are endangered. Poverty is a multi-dimensional worldwide problem of modern society. In addition to similarities with other countries in transition, Serbia shows some specific features. The Government of Serbia adopted the Poverty Reduction Strategy Paper in 2003.

Keywords: poverty, sustainable development, poverty reduction strategy, transition

1. Introduction

Launched as a concept as early as 1983, sustainable development acquired its full significance in Agenda 21, adopted in 1992 at the United Nations Conference on Environment and Development (UNCED) in Rio de Janeiro. At the Summit on Sustainable Development held in Johannesburg, South Africa in 2002, elimination of poverty was adopted as the first of the three main goals.

Poverty refers to the inability to adequately satisfy some of the essential needs, such as food, clothing and accommodation, but it also includes other needs such as employment, adequate healthcare, welfare, education, cultural needs, and other components featuring in the contemporary lifestyle.

Half of the global population is regarded as poor to the extent that the very lives of a large number of people are endangered. Poverty is present in developed as well as underdeveloped countries.

At the dawn of the new millennium, poverty is present even in the most developed countries. According to an OECD report, the proportion of the poor in the total population ranges from about 6% in the Netherlands, 7% in Sweden, 10% in Germany, 11% in Canada, to 14% in the USA and nearly 20% in the UK (OECD 2002). The poverty threshold in Western Europe amounts to about 400 EUR a month.

As regards countries in transition, the World Bank's 2002 survey states that, with the exception of Slovenia and the Czech Republic (without data for our

[1] **PhD student** Faculty of Economics, Subotica, Serbia

country), more than half of the population of these countries were poor, i.e. lived off about 2 US dollars per day per person.[2]

The extent of poverty in the underdeveloped countries of Africa, Asia and Latin America is illustrated by the fact that the annual per capita GDP in 192 countries of the world ranges from 110 to 45,000 USD, the lowest being in some of these countries.

2. Poverty in Serbia

Poverty did not feature significantly in Serbia before the early 1990s, when the per capita GDP amounted to over USD 3,000. GDP was halved in the 1990s, combined with high unemployment rates, plummeting salaries, pensions and other personal revenues, and unprecedented hyperinflation. Compounded by economic sanctions, a large number of refugees and the 1999 air raids, this clearly lead to sudden impoverishment of the majority of population, but also to a considerable wealth amassed by a small number of people.

Faced with this problem, the international community adopted documents aimed at reducing the poverty in the world. The Fifth Principle of Agenda 21 states that 'cooperation of all countries on detecting poverty is a prerequisite for sustainable development'.[3]

The implementation of the project entitled *The Sustainable Development of the Republic of Serbia* is in progress in Serbia. A special significance is placed on the 2000 Declaration of the General Assembly of the UN on the eight millennium development goals adopted by the leaders of 189 UN member countries, with a primary focus on eliminating extreme poverty and hunger, i.e. halving the proportion of population with income under $1 a day, as well as the percentage of people affected by starvation.

In October 2003 the Government of Serbia adopted the Poverty Reduction Strategy (PRS), setting a goal to halve the poverty in Serbia by 2010, and **every tenth citizen of Serbia (10.5%) is known to have been poor in 2003. The poverty line for this year amounted to 4,970 dinars per consumer unit** (RSO 2003)**.**

The PRS defines poverty as a multidimensional phenomenon which, in addition to inadequate income to satisfy the basic needs, also includes lack of

[2] Transition, The First Ten Years, Analysis and Lessons for Eastern Europe and the Former Soviet Union, 2002., World Bank

[3] Narodi ujedinjenih nacija za bolju životnu sredinu 21. veka (1992), Rezultati svetskog samita u Rio de Žaneru, Brazil [We the peoples of the United Nations united for a better environment in the 21 century (1992), Outcome of the World Summit in Rio de Janeiro, Brazil], Belgrade: Federal Ministry of Environment

employment opportunities, inadequate accommodation and access to welfare, healthcare, educational and cultural services.

Poverty in Serbia is measured:

– based on the Survey of Living Standards (SLS) carried out in 2002 and 2003. In order to secure data comparability and gain clear insights into trends related to poverty, the Republic Statistical Office (RSO) conducted a new survey of living standards on the territory of Serbia in May and June 2007;

– in addition to the SLS, poverty in Serbia is also measured with the Household Consumption Surveys (HCS) regularly carried out by the Republic Statistical Office. Due to problems in conducting the HCSs in 2004 and 2005, this research cannot be used for calculating poverty trends for the period 2004-06.

It is generally accepted that consumption in transition countries is a better indicator of the living standards.

According to the latest survey, the HCS for 2007, **all households with monthly consumption under 8,883 dinars per consumer unit were regarded as poor. The consumption of 6.6% of the population of Serbia lies below this level, meaning that about 490,000 inhabitants of Serbia are poor.**

The gap between the richest and the poorest in Serbia is somewhat wider than the average of some Eastern European countries in transitions. One of the gap indicators used for measuring poverty levels is the share of consumption in 25% poorest households in the population's total consumption. Their consumption accounted for 11.1% in 2006.

In comparison with other countries in the region, poverty in Serbia in 2003 was at a level similar to Romania, lower than in Albania, and higher than Bulgaria and Poland.[4]

Poverty in Serbia increased dramatically in the 1990s. The middle class disappeared, the number of poor population grew 2.5 times, and an increasing number of people live only slightly above the poverty line. The increase in poverty is mostly due to a sharp decline in the GDP over the past decade. In late 2000, the total GDP amounted to 45%, and the per capita GDP was 40% lower than 1989, which is the greatest decline of GDP in Central and Eastern Europe. Such a situation had negative consequences on all the aspects of life in Serbia.

In 2006, poverty levels in the Republic of Serbia were calculated based on household consumption. Table 1 shows poverty indicators across the regions of Serbia in 2006.

[4] ECAPOV, II World Bank, 2005.

Table 1. Poverty indicators in the Republic of Serbia in 2006 (%)

	Percentage of poor population	**Total population structure**	**Poverty structure**
Belgrade	**4.3**	**21.4**	**10.5**
urban area	3.8	17.2	7.5
other	6.3	4.2	3.0
AP Vojvodina	**8.7**	**26.7**	**26.5**
urban area	4.9	14.9	8.3
other	13.6	11.8	18.2
Central Serbia	**10.7**	**51.9**	**63.0**
urban area	6.6	23.8	17.9
other	14.4	28.1	45.1
TOTAL	**8.8**	**100.0**	**100.0**

Source: www.prsp.sr.gov.yu (Sept. 3, 2008.)

As we can see, central Serbia is the region with the highest proportion of poor population.

3. Living standard indicator in Serbia in the transition period

Among others, macroeconomic trends in Serbia in the transition period from 2001 to 2007 are marked by a growth in economic activity and personal revenue, as shown in Table 2.

Table 2. The basic indicators of macroeconomic trends in Serbia

	2001	**2002**	**2003**	**2004**	**2005**	**2006**	**2007***
per capita GDP (EUR)	**1748.4**	**2241.5**	**2401.4**	**2629.7**	**2814.0**	**3354.0**	**4058**
real net wage growth (%)	**16.5**	**29.9**	**13.6**	**10.1**	**6.4**	**11.4**	**17.4**
GDP real growth (%)	**4.8**	**4.2**	**2.5**	**8.4**	**6.2**	**5.7**	**7.0**

** estimated value*
Source: Ministry of Finance of the Republic of Serbia, 2007

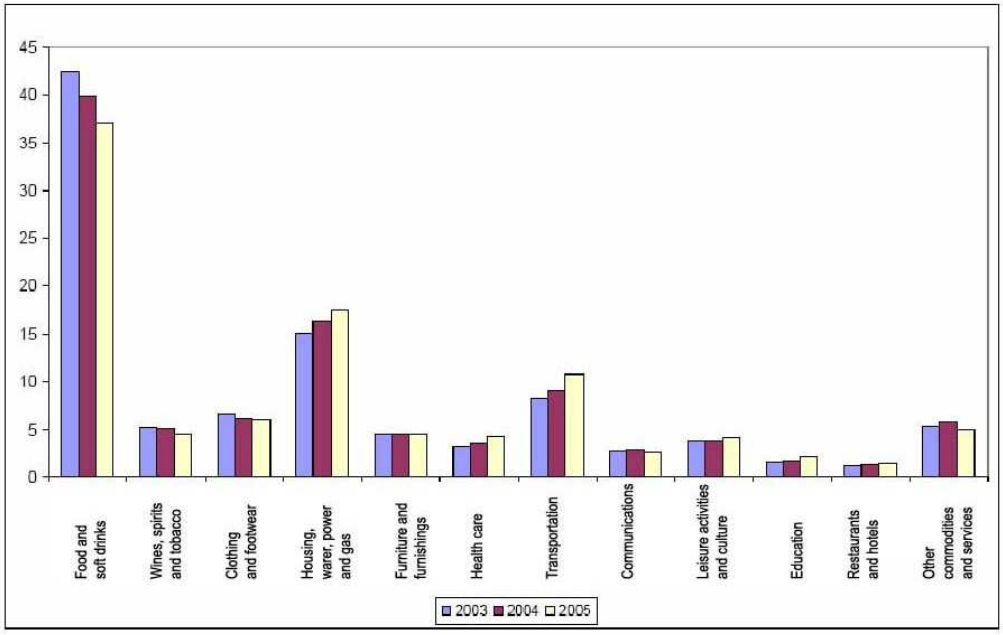
The economy of Serbia achieved a **growth of 5.7% in 2006**. This growth in GDP is somewhat lower than in 2004 and 2005, when it amounted to 8.4% and 6.2%

respectively. **Average net wages in the first seven years of transition, especially in the first three-year period, showed real growth**, generating a rise in the population's living standards.

The share of expenditure on food in the total household consumption can be used as an indicator of living standard. The structure of personal consumption in Serbia for the period 2003-05 shows that **expenditure on food and soft drinks fell from 42.5% to 37.1%,** as opposed to costs of accommodation, water, electric power, gas and other fuels (from 15.1% to17.5%), transportation (from 8.2% to 10.7%), shown in the chart below.

Such a trend in the structure of personal consumption in households is a characteristic of transition countries. In most countries of the EU, the highest share of personal consumption is accounted by housing or accommodation, water, electric power, gas and other fuels and transport costs, while expenditure on food and soft drinks is on the third place (RSO 2007).

Table 3. The structure of personal expenditure in the Republic of Serbia, 2003-2005



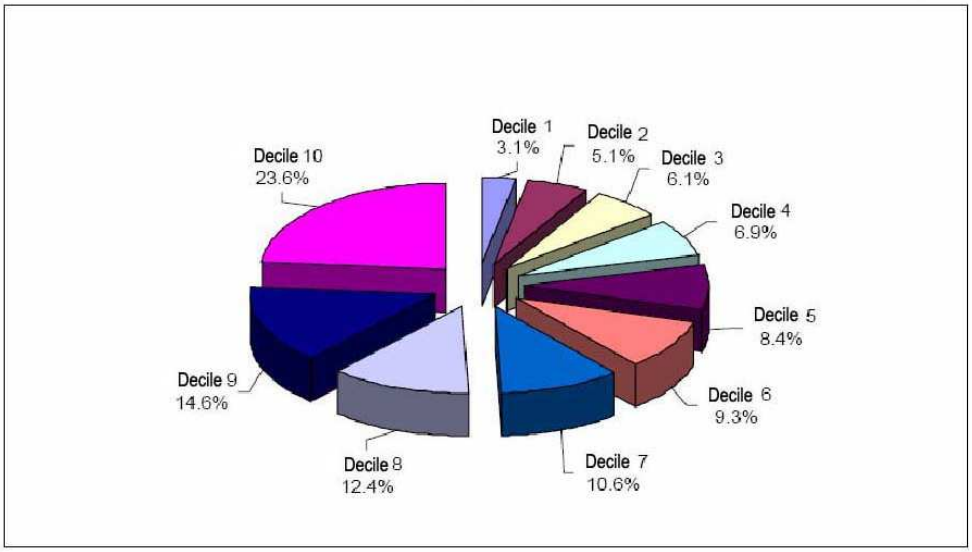
Source: RSO Bulletin 470, Belgrade 2007.

Based on a decile analysis from 2005 (RSO Bulletin 470, Belgrade, 2007, p. 46), it can be concluded that in the households of the first decile (**the poorest households**), the highest share of total personal consumption is accounted for by **expenditure on food and soft drinks (59%)**. As we move to **higher deciles, the share of**

expenditure on food and soft drinks is reduced (amounting to **28.2%** in Decile 10).

The household consumption of Deciles 9 and 10 **(the richest households) accounts for approximately 40% of the total expenditure of all households in the Republic of Serbia**, as illustrated by the chart below:

Table 4. The share of total expenditure in deciles in total expenditure in the Republic of Serbia, 2005



Source: RSO Bulletin 470, Belgrade 2007.

4. Conclusion

Although macroeconomic trends in the Republic of Serbia are favourable and the population's living standards show a growing tendency, our country cannot be content with the fact that 9% of Serbian population are living in poverty, let alone the fact that poverty levels are 2.5 times higher in Central Serbia than in Belgrade.

According to the Sustainable Development Strategy defined in the Second Report on the Implementation of the PRS in August 2007, the strategic guidelines aimed at halving the number of the poor in Serbia by 2010[5] are:

– dynamic economic growth and development with a focus on generating new jobs in the private sector;

[5] www.prsp.sr.gov.yu

– preventing the emergence of new poverty resulting from restructured economy and rationalised public administration;

– efficient application of the existing and defining new programmes and activities targeted at the poorest and vulnerable social groups.

Sustainable development cannot be viewed isolated from global inequalities. The concept of sustainable development must take into account the specific problems of poor countries, because an excessively wide range of changes is hardly feasible in these countries.

References

Minić V., Ajduković G. 2006: „Siromaštvo kao ključni problem Srbije u tranzicionim procesima“ [Poverty as a key problem of Serbia in transition processes], Collection of papers: *Tehnologija, kultura i razvoj* [Techology, culture and development] issue 13, Belgrade, Technology and Society Association, Tehnologija i društvo, 247-256 p.

The Government of Serbia 2006: *National Millennium Development Goals in the Republic of Serbia*, Belgrade.

The government of Serbia, *Poverty Reduction Strategy for Serbia,* Belgrade, www.prsp.sr.gov.yu (accessed Sept 3, 2008).

The Government of Serbia, *The Second Report on the Implementation of the PRS in Serbia*, Belgrade, www.prsp.sr.gov.yu (accessed Sept 3, 2008).

The Republic Statistical Office 2005: *Household Consumption Survey*, Beograde.

The Republic Statistical Office (2997), *Empirical consumer goods basket and living standard indicators, 2001-2006,* Bulletin 470, Belgrade, www.rzs.sr.gov.yu (accessed Sept. 3, 2008).

Housing allowance: Subsidy to landlords?*

Matti Viren[1]

This paper deals with incidence of housing subsidies. This property is analyzed using Finnish panel data from biggest Finnish cities. The main data set includes 50 000 households that have received housing allowances for the period 2000-2008. Preliminary results suggest that a part of subsidies will indeed go rents. A conservative estimate of the size of shift is 15 per cent but it is well possible that the number is even 25 per cent. On the other hand, the subsidy seems to have increased housing demand more than the subsidy-induced income effect would
have implied which is in accordance with the goals of the subsidy program. Our results seem to be consistent with other research results that have also ended up with relative high rent effects. If this is indeed the case, it is well founded to reconsider the need for reforming the system of housing subsidies at least with respect to the share of costs that remains on households' own account.

Keywords: Housing market, income transfers, incidence, housing allowance

1. Introduction

This paper deals with the incidence problem of government income transfers. The problem of incidence is very common in taxation but obviously it is also relevant for government transfers and subsidies. It becomes apparent when one tries to answer to the question of who in the very end is going to pay the tax or receive the transfer. The idea is that the burden of taxes (similarly, the benefit from transfers and subsidies) may shift from the original agent to the market counterpart even though the original agent has the legal obligation to pay the tax (or receive the subsidy) and take care of the proceedings of the payment or the receipt. The issue becomes quite complicated when we consider it in a general equilibrium setting which is illustrated in

[1] Department of Economics and the Public Choice Research Centre of University of Helsinki, and the Monetary Policy and Research Department of Bank of Finland

Acknowledgement: I am grateful to the Finnish National Pension Fund for the data and the Savings Banks' Research Foundation for financial support. Samuli Hiekka, Kari Kekola and Jose Hämäläinen have provided valuable help in organizing the data

Figure 1. Housing allowance does not only affect income of the poor but also income of the landlords, rents, tax rates, labour supply, housing production and so on. Because the pros and cons of housing allowance are so complicated, housing allowance represents also a typical *public choice* problem. Here, we ignore these elements and focus on the incidence question only.

In the case of taxation, we know the basic results of incidence pretty well. Very shortly, all depends on supply and demand elasticities of the respective aggregate supply and demand curves when we face perfect competition. In the case of monopoly, things are more complicated so that even over-shifting may take place (Musgrave 1959)[2].

As for income transfers or subsidies, we know a lot less. The reason is that national systems are very different and institutions play a more important role (cf. e.g. the MISSOC Data base of European Union which gives an overview of national social security systems and Ditch et al. (2001) for a comparison of housing allowance systems).

Housing assistance in Finland is quite extensive. Currently it includes 160 000 households (out of 2450 000 households) and the expenditures come close to 500 Mil. Euros (which is roughly 0.3 per cent of GDP).[3] Housing assistance is also related to social assistance which is distributed using the Housing assistance rules. Within social assistance all housing expenditures (which qualify these rules) are paid to poor households that have no or very little income.

The effects of housing assistance are analyzed in relatively numerous studies due to the fact that in a form or another it is used in most countries. A useful survey is provided by Rosen (1985). Evidence from more recent policy experiments is provided by Gibbons and Manning (2003) using the UK data. Gibbons and Manning arrive at the result that even 50 % of assistance shifts to rents. An even higher estimate is obtained by Fack (2006) who found the representative number for the French housing assistance reform was as high as 78 %. In the United States, Susin (2002) estimated the rent effects of housing vouchers and found the effect to be about 16 per cent. In Finland, there as two studies by Kangasharju (2003 and 2008) which have produced quite different results. The first one arrived at an estimate of 15 per

[2] In Finland, relative few studies of (tax) incidence have been carried out. See, however, Viren (2009) for VAT taxes and Korkeamäki and Uusitalo (2006) for employers' social security expenses.

[3] The Finnish housing allowance follows the rule: allowance = 0,8*(min(max_expenses, rent) – base deduction), where base deduction depends on income and household characteristics. For instance, in the case of single person household base deduction is 0, when 0 < income < 541, base deduction = 298 when income is 1245 and when income exceeds 1245, no allowance is granted. The maximum allowance (max_expenses) depends on the regulated norm space and rent levels. The latter depends on the location of the municipality and the age of the house.

cent while the latter got a much higher value (60-70 per cent). Both studies used single changes in the assistance parameters in the *differences in differences* framework where receivers of the housing assistance were controlled.[4] Finally a study and Hiekka and Viren (2008) ought to be mentioned. Basically, it used similar data as the current study but the sample for limited to 1000 households in the Turku city area. The study arrived at the tentative results that one fourth of housing allowance shifts to rents.

This study mainly uses panel data for Finnish households who have received housing assistance (sometimes) during the period 2000-2008. Altogether 50 000 households are included in the data. Basically, 9 major changes in the housing assistance are considered during the period but because different households have different possibilities of getting housing assistance (due to income, location and so on) we have also quite a lot of cross-section heterogeneity in statutory housing assistance levels which altogether provide much more variability in both the rents levels and (exogenous) housing assistance.

2. Background for empirical analyses

To find out the magnitude of incidence problem we consider the conventional supply-demand framework. In the first place, we assume that the demand for rented houses takes the following from:

$$Q = D(Y/P, A/P, N, P_H/P) \tag{1}$$

Where Q denotes the demand for housing services (space, quality, location and so on), Y income (excluding housing allowance), A housing allowance, N the size of the household, P_H the rent level and P the general price level. The supply of rental housing is, in turn, determined by the following function:

$$Q = S(P_H/MC), \tag{2}$$

where MC denotes the relevant marginal cost of the rental housing.

To measure these costs we use house prices (in fact, the regional indexes of house prices). Basically, we should use some sort of user of housing which after all is relatively easily constructed. This time, we however, use this simpler proxy.

By setting (1) and (2) equal, we may solve the system in terms of P_H which takes the following form:

[4] The properties of the differences in differences (DD) approach are surveyed in e.g. Bertrand, Duflo and Mullainathan (2004).

$$P_H = P_H(Y/P, A/P, N, MC/P) \quad (3)$$

In practice, this is estimated in a (log) linearized from. The coefficients obviously reflect the demand and supply elasticities which determine the incidence values in a way explained by e.g. Susin (2002). Needless to say, if supply is completely inelastic and demand perfectly elastic housing allowance shifts completely to rents (and thus goes to landlords). By contrast, if the supply side functions well in a sense being very price elastic (and competitive, of course) housing allowance would just increase the tenants' income and show up increase demand for housing. Intuitively, one might expect some sort of 50 – 50 outcome although the only way to find out the truth is to estimate (3) scrutinize the parameter values of Max

In what follows, equation (3) is estimated in the form:

$$(p_H/p)_{it} = a_{0i} + a_1(MC/p)_t + a_2 N_t + a_3 Space_t + a_4 Age_t + a_5(Y/p)_t + a_6 Max_t + u_{it}, \quad (4)$$

Where p_H denotes the rent level, p the (Consumer Price Index) CPI, MC the house price index, N the size of household (number of people), $Space$ the size of flat, Age the age of the flat, Y income (prior to housing allowance) and Max to the maximum achievable housing allowance. u_{it} is the random term where i denotes the i:th household and t the period (year).

The model is estimated using the fixed effects model as the basic specification although we scrutinize all versions fixed effects models, that is: no fixed effects, cross-section fixed effects, and both cross-section and time fixed effects).

The essential feature in the analysis is that the key variable Max does not directly depend on the actual rent level. It only measures the potential maximum achievable allowance that can be received for the flat in question (given the age of the flat and the location of flat). Obviously, the choice of the flat (in addition to the tenure choice) is endogenous which makes the Max variable to some extent endogenous as well. But it is hard to take the consequences of this choice problem into account.[5]Obviously there is some simultaneity link between rents, allowances and house prices as well due to capitalization of rents. For the time being, we ignore this problem.

Before we turn to estimation results some comments on data merit note.

The main data source is the Finnish National Pension Fund (FNPF) that distributes the housing allowance. The data sample that the FNPF has kindly provided to us is based on register data which are based on housing allowance applications which turn are based on actual lease agreement contracts. Thus, the data are pretty accurate compared e.g. to various survey measures.

[5] The results of Lyytikäinen (2006) suggest that these simultaneity problems are not particularly severe in the Finnish data).

The FNPF data cover 9 years (2000-2008) and include 50 000 households from 345 municipalities. Most of them come from biggest cities, thus about one half come from the Helsinki metropolitan area. Potentially, the number of observations is 450 000 but in practice it is much less because the there are no data for the periods in which housing allowance is not paid. The effective number of observations (after taking account all missing observations) is hence "only" 140 000. Although the data are good there is one problem with data: the data include only households that have received housing allowance. Thus, in the data we have no proper control group. In many cases the same households have received housing allowance during the whole sample period so that it is also a bit difficult to distinguish "new" rent level and "existing rent levels". Fortunately the data are large enough to facilitate comparison of new and existing (old) contracts.

The Finnish Income Distribution Data (FIDD) is much smaller (the number of observations is only about 26 500 consisting of 10 500 households) although the number of variables is much larger (760) because the data include all possible variables for income, taxation, income transfers and related items. The data are partly in a panel form but here we have just used the newest cross-section to see whether it makes any difference if we focus on households that receive housing allowance and households which do not receive housing allowance. Thus, basically we try test the "law of one price" in the Finnish rental markets. In addition, we use these data revisit the demand equation for rents housing.

Some idea of the Finnish system can be obtained from the enclosed graphs. Figures 2 and 3 illustrate the mean values of rents and the Max variable, Figures 4 illustrates the relationship between rental price and the demand for rental housing and finally, Figure 5 gives some idea of the presentence of rents (showing the annual frequency of changes in rents)

3. Interpretation of results

The estimation results are presented in Tables 1-3. Table 1 deals with the rent level equation (4) which is estimated from the Finnish panel data. From the same data, we have estimated also a demand equation for rental housing. Finally, we estimated a rent and rental housing demand equations from the 2007 cross-section data of the Finnish income distribution survey.

The demand equation is quite simple (double log) model of the following form:

$$Space_{it} = b_{0i} + b_1(p_H/p)_{it} + b_2 N_t + b_3(A/P)_t + b_4 Age_t + b_5(Y/p)_t + b_6 Space_{it-1} + u_{it}, \quad (5)$$

where *Space* denotes the quantity of the housing that is in practice measured by the space of the apartment while *A* denotes the housing assistance.

Before considering the estimates it is worthwhile to consider the dynamics of rents that is illustrated in Figure 5. Quite clearly, a considerable proportion of rents is not changed every year. The rents which have been decreased are also in most cases "constant rents": reductions are usually very small (something like rounding errors).

As for the result in table, we see that they are relatively robust in terms of panel data estimation procedure and variable transformations (level vs. log). The estimates of the shift parameter are all highly significant suggesting that the value of elasticity parameter is somewhat between 0.2 and 0.4 so that a very conservative estimate is just 0.2. Rents increase along with the price of houses and household income. By contrast, the age of house has a negative impact. Real rents also decrease if the tenant remains in the same flat for longer time. To some surprise, the role of the household size remains somewhat ambiguous. This is probably due to the fact that the MAX variable already includes the impact of the household size (a bigger family gets automatically a bigger household allowance).

The demand curve (5) estimates in table 2 are also easy to interpret: higher income increase demand and higher price lowers it. In the same way, larger household size has a positive effect. The role of housing allowance seems clear: it has a positive effect on the size of the flat and the effect seems to be larger than the effect of wage (other) income. This could, of course, reflect simultaneity between the size of the apartment and the housing allowance but even though we use just an allowance dummy or lagged allowance we get (ceteris paribus) a positive effect.

We also scrutinize the demand behaviour using the Income Distribution Database cross-section data for 2007 (Table 3). The data allows testing the importance of housing allowance in the case where only a part of households receives housing allowance. Quite clearly, rents are related to housing allowance, the coefficient of this variable is much higher the coefficient of wage income or other income transfers. The result is in fact well in accordance with the basic aim of the housing allowance, that is, an improvement of the housing conditions of the poor.

The cross-section data quite clearly suggests that the price of one law holds in the sense that rents (per squares meters) for housing allowance receivers and non-receivers are the same (see e.g. column 2 in Table 3). Actual (total) rents do indeed differ but this probably due to the fact that those who receive housing allowance live, ceteris paribus, in larger apartments.

4. Concluding remarks

It is all clear that a part of housing allowance shifts to prices. The questions only of the magnitude of the tax shift. In this study, a conservative estimate is 0.2. Thus one fifth of an increase in housing allowance shifts to market rents. That is, the rents of

all households in rental housing do increase. It is also well possible that the true number is larger than 0.2.

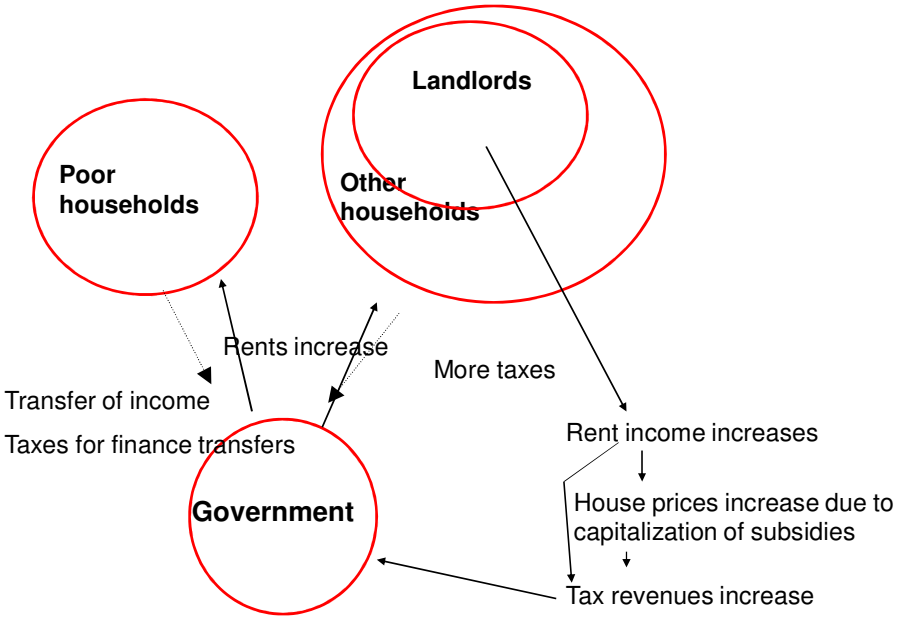
One has to keep in mind that the whole story does not end here. Housing allowances (in Finland, at least) are in practice indexed to market rents so that the allowances are adjusted annually according to developments in rental markets during the preceding 12 months. This creates a multiplier effect that can even double the short-term impact[6].

It has to be kept in mind that the increase in rents hurt those who are not eligible to housing allowance, basically the middle income households. Their real income will decrease and their housing demand will decrease even more. The situation is deteriorated still more because they have to pay additional taxes to finance the housing allowance. Thus, in all, housing allowance leads to larges changes in income distribution and so cannot even be sure that at the limit the Pigou & Dalton principle holds. Moreover, strongly income related housing allowance leads to poverty traps because the effective tax rates at least some households goes to 100 per cent. This, in turn, leads to adverse effects on labour supply which are not, of course, good thing thinking about the functioning of the labour market and the rate of inflation.

Therefore, we should very carefully scrutinize the general equilibrium effects of housing allowance and in particular consider possibilities of reducing the eventual shift of allowance to markets rents and (de facto) indexation of housing allowances.

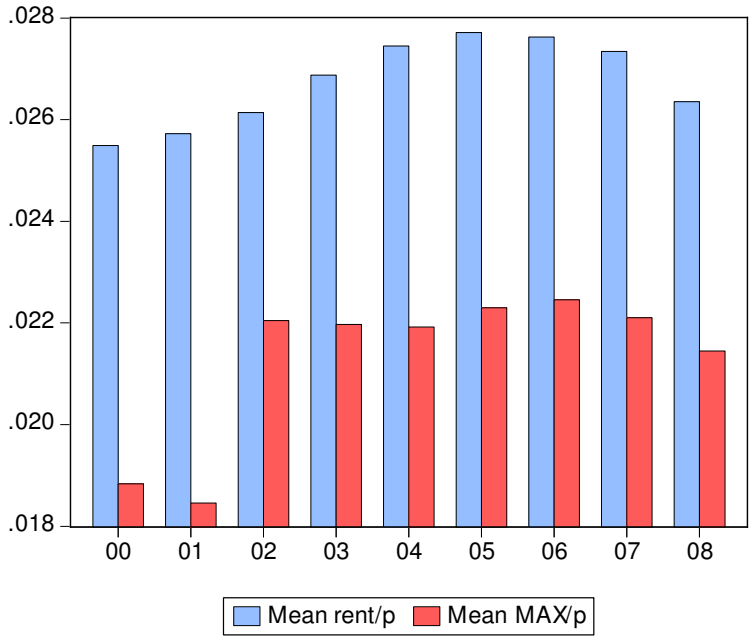
[6] Assume e.g. that we have a system where the rent = α (= exogenous factors) + 0.2*allowance and allowance is indexed as: allowance = rent(-1) – β (=some constant) we will arrive at the long run solution of the form: rent = (α-.2*β)/.8. So the long-run shift factor is 0.25 instead of 0.20. If the estimate of *Max* were 0.5 the long-run shift factor would in fact be 100 %.

Figure 1. Some key linkages of housing allowance



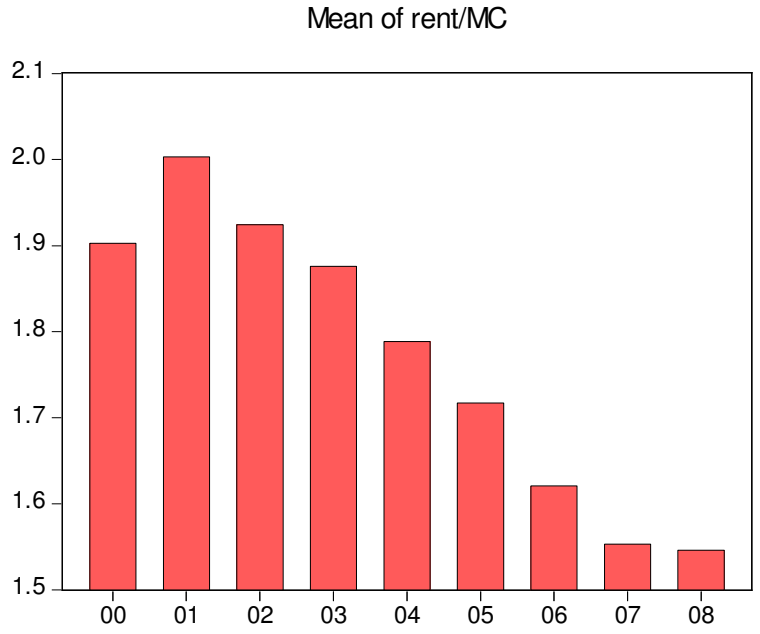
Source: own creation

Figure 2. Mean values of the rent and the Max variable.



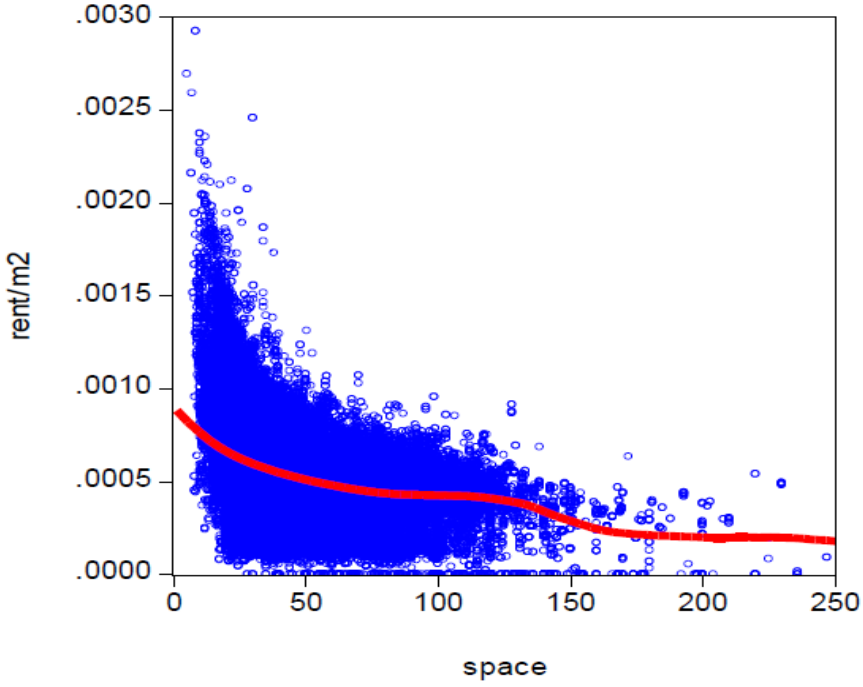
Source: own creation

Figure 3. Maximum assistance (*Max*) in relation to house prices



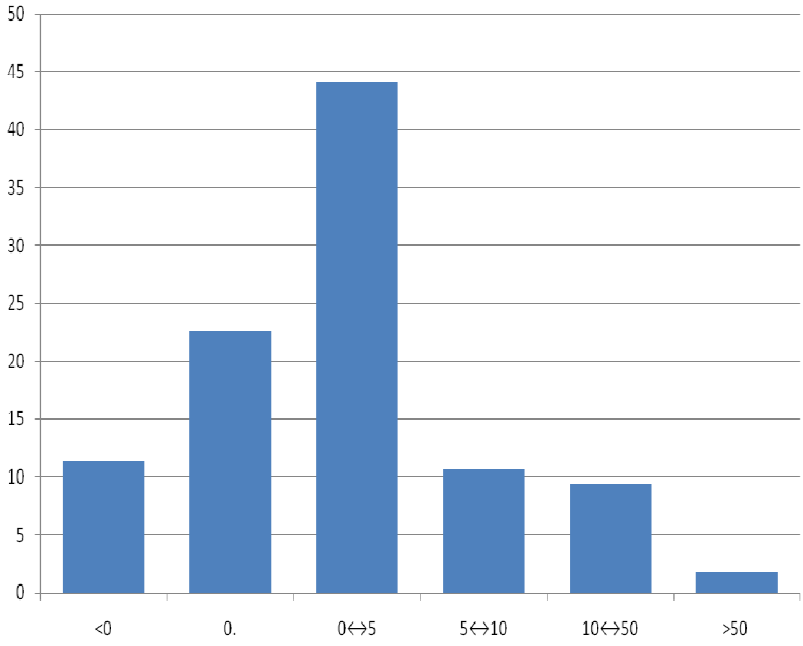
Source: own creation

Figure 4. Scatter plot between the apartment size and the rent level



Source: own creation

Figure 5. Frequency of changes in rent levels



Source: own creation

Table 1. Estimates of the rent level equation form the panel data

	1 *level*	*2 level*	*3 level*	*4 log*	*5 log*	*6 level rent/m²*
Max/P	.435 (42.02)	.429 (26.65)	.269 (14.67)	.345 (36.36)	.204 (20.35)	.583 (41.69)
MC/P	.352 (51.00)	.389 (47.66)	.342 (43.92)	.235 (49.78)	.219 (48.03)	4.267 (25.17)
Space	.185 (37.06)	.194 (24.69)	.226 (23.95)	.361 (45.58)	.434 (53.72)	-.002 (27.80)
N	-.185 (9.53)	-.310 (8.25)	-.006 (1.67)	-.022 (4.96)	.018 (4.02)	-.006 (7.71)
Y/P	.038 (30.74)	.042 (8.56)	.033 (5.42)	.009 (8.96)	.005 (5.25)	.040 (5.15)
Age of the house	-.013 (15.46)	-.020 (8.88)	-.027 (12.24)	-.005 (6.06)	-.008 (9.43)	.035 (6.36)
No change		-.057 (16.61)	-.045 (14.02)	-.020 (16.49)	-.017 (14.23)	-.008 (10.11)
R^2	0.935	0.952	0.956	0.952	0.956	0.895
DW	1.28	1.48	1.49	1.45	1.46	1.47
Fixed E	CS	CS	CS +local	CS	CS + local	CS + local

Source: own creation
The dependent variable is the real rent level. Numbers inside parentheses are corrected t-values

Table 2. Estimates of the housing demand equation from the panel data

	1 level	*2 log*	*3 log*	*4 log*	*5 log*	*6 log*	*7 log*
Assistance/p	.503 (23.53)	.186 (70.09)	.215 (84.39)	.068 (49.69)	.078 (52.72)		.006* (10.13)
Assistance Dummy						1.323 (50.20)	
Y/P	.101 (18.01)	.139 (48.26)	.172 (50.75)	.046 (34.78)	.057 (36.89)	.044 (32.29)	.013 (14.57)
N	.008 (61.16)	.328 (123.7)	.295 (99.89)	.088 (56.24)	.080 (50.52)	.468 (379.5)	.124 (74.67)
$Rent/m^2/p$	-4.024 (122.1)	-.421 (121.5)	-.441 (125.9)	-.155 (61.58)	-.172 (63.40)	-.333 (96.71)	-.131 (55.27)
No change	-.002 (21.57	-.019 (12.73)	-.020 (13.62)	-.058 (31.04)	-.060 (32.47)	-.018 (11.10)	-.062 (32.36)
$Space_{-1}$				.721 (213.4)	.711 (205.9)		.739 (223.1)
Panel	No FE	No FE	Period FE	No FE	Local	No FE	Local
R^2	0.704	0.716	0.723	0.905	0.906	0.684	0.901
DW	0.37	0.32	0.33	1.56	1.54	0.27	1.58

Source: own creation

*) Lagged value of assistance/ p is used instead of the current value. The dependent variable is the size of the apartment. Otherwise, notation is the same as in Table 1

Table 3. Estimates from the 2007 cross-section data

Number *Dep.var*	*1* *Rent/m²*	*2* *Rent/m²*	*3* *Rent*	*4* *Space*	*5* *Space*	*6* *Log(Space)*
Assistance	.169 (2.51)			1.515 (4.06)	1.776 (4.62)	
Assistance Dummy		.127 (0.84)	28.656 (3.47)			.088 (6.16)
Y (W/Tr)	.024 (6.07)	.022 (5.57)	.025 (6.42	.253 (8.57)	.296 / .542 (9.05/6.01)	.167 (13.81)
N	.401 (5.59)	.444 (6.40)	45.896 (7.91)	10.030 (18.95)	9.270 (15.72)	.366 (27.09)
Space	-.079 (15.85)	-.079 (15.88)	1.799 (5.65)			
Rent/m²				-2.799 (15.74)	-2.773 (15.60)	-.355 (16.77)
R^2	0.242	0.240	0.398	0.566	0.577	0.616
SEE	2.914	2.918	155.0	17.30	17.19	0.273

Source: own creation

Y denotes here household gross income, W wage income and Tr income transfers. Otherwise, notation is the same as in Table 1. Space is measured by the number of quadrate meters

References

Bertrand, M. - Duflo, E. - Mullainathan, S. 2004: How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119 /1, 249–275. p.

Ditch, J. - Lewis, A. - Wilcox, S. 2001: Social housing, tenure and housing allowance: an international review. In-house report 83, Department Work and Pensions.

Fack, G. 2006: Are housing benefits an effective way to redistribute income? Evidence from a natural experiment in France. Labour Economics13/6, 747-771. p.

Gibbons, S. - Manning, A. 2003: The incidence of UK housing benefit: evidence from the 1990s reforms. CEP Discussion Paper 597.

Hiekka, S. - Viren, M. 2008: Does housing allowance increase rents? Finnish Labour Institute's Working paper 241.

Kangasharju, A. 2003: Does the receivers of housing allowance pay a higher rent (in Finnish)? Government Institute for Economic Research (VATT) Discussion Paper 101.

Kangasharju, A. 2008: Housing allowance and the rent of low income households. VATT –Discussion Paper 458.

Korkeamäki, O. - Uusitalo, R. 2006: Employment effects of a payroll-tax cut – evidence from regional tax exception experiment. IFAU Working Paper 2006:10.

Lyytikäinen, T. 2006: Rent Control and Tenants' Welfare: The Effects of Deregulating Rental Markets in Finland. Government Institute for Economic Research (VATT) Discussion Paper 385.

Musgrave, R. 1959: *The Theory of Public Finance.* McGraw-Hill, New York.

Rosen, H. 1985: Housing subsidies: Effects on housing decision, efficiency, and equity. In Acerbic, A. - Feldstein, M. (eds): *Handbook of Public Economics I.*, Amsterdam, North-Holland, 374-420. p.

Susin, S. 2002: Rent vouchers and the price of low-income housing. *Journal of Public Economics* 83, 109– 152. p.

Viren, M. 2009: Does the value-added tax shift to consumption prices. AUCO Czech Economic Review 3/2, 123-142. p.

Fertility of immigrants and non-immigrants in the United States

Marcin Duszyński - Aleksander Suseł - Tomasz Wołowiec

The paper deals with fertility determinants in the populations of immigrants and non-immigrants in the United States. We consider determinants as follows: age, marital status, education, religion, and race and Hispanic origin. The analyses show that a relation between fertility and place of birth was statistically significant in both populations, women and men. We prove that migration is a major source of variation in fertility between immigrants and non-immigrants, which may lead to the postponement of maternity or marriage. Finally, we find that a duration of residence is positively correlated with fertility among US immigrants.

Keywords: migration, USA, reproductive, MCA, ordinal regression

1. Introduction

An analysis of fertility among US immigrants is one of the areas of study by social scientists. Rumbaut - Weeks (1986) pointed out that fertility among refugees from the Indochina region was negatively correlated with a duration of first marriage and efficiency of English language. Moreover, they found out that fertility was positively correlated with a duration of residence in the United States. A vast gap in fertility was shown by the total fertility rate which for the Indochinese population was 5,61 whereas for the US-born women was 1,80 birth per woman. Ford (1990) showed that a length of residence was positively correlated with fertility. The analyses indicated that a maximum of fertility was observed in 5 to 10 years after immigrants settled down in the US. According to Ford (1990), a postponing of marriage and cumulation of births due to immigration determined the distribution of births among immigrants. Kahn (1994) considered live births and planning children in the fertility analysis. With respect to the first measure, the analyses showed that number of live births among immigrants was significantly higher compared to non-immmigrants. Moreover, education and income were negatively correlated to number of live births. Kahn proved that a second generation of immmigrants was characterized by lower number of planning children compared with the third and higher generations. Finally, together with an increase of length of residence in the US, a gap in the level of fertility between immigrants and US-born women has diminished. Hwang - Saenz (1997) analyzed fertility among the immigrant Chinese women. They proved that

women who were born in China had a significantly lower fertility than foreign-born Chinese women. If children born only in the United States were considered, a gap in the level of fertility was small. Besides, the Authors indicated that an average of live births among US citizens was lower in relation to non-US citizens. On the other hand, an average number of births for US citizens was higher compared with non-US citizens. The Authors explained that the Chinese immigrants were still affected by birth control in their country of origin.

The paper deals with fertility determinants in the populations of immigrant women and men and non-immigrant women and men in the United States. In this study, we pose the following hypotheses: 1) higher fertility among immigrants occurs because most of them come from the countries with higher fertility than recorded in the US, 2) higher fertility among immigrant women is caused by postponing of motherhood owing to immigration, 3) immigration is a crucial determinant of fertility among immigrant women and men 4) duration of residence is positively correlated with fertility.

2. Methods and data

In the study, we employed the ordinary regression model and multiple classification analysis (MCA). The regression models are classified within the GLZ class models which were introduced to the literature by Nadler - Wedderburn (1972). The models are based on the family of exponential distributions such as the Bernoulli, Poisson or gamma distribution. The ordinary regression models do not assume linearity between dependent and independent variables. Mostly used link functions are identity, logit, probit and logarithmic link function which depend on a distribution, restriction and type of datasets. While, the initial results showed that other link functions gave much a worse goodness of fit of model, we applied the logit link function. The ordinary regression models are estimated using the iterative methods of estimation. Mostly used method is the Newton-Raphson method (Albert - Anderson 1984, Jennrich - Sampson 1976) and the Fisher's method (McCullagh - Nelder 1989, McCullagh 1989). We used the PLUM (Polytomous Logit Universal Model) procedure (McCullagh 1980, 1989) in the SPSS statistical package. The technical details on the ordinary regression model present (see, Agresti 1984, Armstrong - Sloan 1989, Hosmer - Lemeshow 2000, Long - Freess 2006, McCullagh 1980), whereas a wide spectrum of practical applications (see Chen - Escarce 2007, McNamee 2007, Meagher 2008).

The MCA is an additive model which is much less restrictive in comparison with the multiple regression or discriminant analysis. Firstly, dependent and independent variables do not have to be from an interval scale. Secondly, a researcher can control an influence of independent variable on depedent variable, before and after other variables are inluded into a model. The MCA model is

estimated using the mean square error minimization technique. The coefficients can also be estimated by solving a set of normal equations. More details on the MCA show papers (see Adrews *et al.* 1973, Retherford - Choe 1993) whereas applications (see Goldschneider - Mosher 1991, Mosher *et al.* 1992, Thornton 1979, van Ginneken - Razzaque 2003).

The source of a data was the National Survey of Family Growth, Cycle 6, which was conducted in 2002. A total number of respondents was 11671, 7643 of women and 4028 men. The data on 1079 immigrant women and 6155 non-immigrant women, and 709 immigrant men and 3901 non-immigrant men were retrieved from that database. We retrieved the data on age, marital status, religion, race and Hispanic origin for each respondent. In addition, the information about duration of residence in the United States was collected. We defined the level of fertility by means of two measures; there were live births for women population and biological children for men population. More details on that survey, and sample design, weighting, imputation, and variance estimation are presented in (Lepkowski *et al.* 2006).

3. Recent immigration in the United States

Recent legal immigration data shows that a total of 8 061 486 immigrants were registered in the United States between 2000 and 2007. The top five countries from which come the highest number of legal immigrants in the United States are presented in table 1. As we notice, the largest number of immigrants, about 1 352 084 (16,77%) came from Mexico. The top five origins of illegal immigrants in the United States are shown in table 2. A total of 11 780 000 illegal immigrants entered the United States in 2007. Again, the largest number of illegal immigrants, 6 980 000 (59,25%) were the Mexican citizens.

Table 1. Legal immigration in the United States by country of origin between 2000 and 2007

Country of origin (women and men)	N	%
Mexico	1 352 084	16,77
India	476 376	5,91
China	455 405	5,65
Philippines	434 965	5,40
Russia	390 493	4,84
Total	8 061 486	100,00

Source: the US Department of Homeland Security

A legal and illegal immigration data indicates that immigrant fertility depends mostly on the level of fertility of immigrants of Hispanic origin. Moreover, the countries of origin of immigrants are characterized by a significantly higher fertility compared to the level of fertility recorded in the United States (see table 1 and table 2).

Table 2. Illegal immigration in the United States by country of origin in 2007

Country of origin (women and men)	N	%
Mexico	6 980 000	59,25%
Salvador	540 000	4,58%
Guatemala	500 000	4,24%
Philippines	290 000	2,46%
China	290 000	2,46%
Total	11 780 000	100,00

Source: the US Department of Homeland Security

Table 3 shows the total fertility rates for selected countries of origin of legal and illegal immigrants in the United States.

Table 3. Total fertility rates in the countries of immigrants in 2007

Country of origin	Total fertility rate
Guatemala	3,59
Philippines	3,32
Salvador	3,04
India	2,76
Mexico	2,37
China	1,77
Russia	1,40

Source: the US Department of Homeland Security

The data provided by the NCHS shows that the TFR was 2,10 births per woman in the United States in 2007. The level rate for the US population was below the levels registered in the countries of immigrants (except China and Russia) (table 3). Based on the above presented analyses, the first hypothesis can be verified, that is, higher fertility among immigrants occurs because largely part of them come from the countries with higher fertility than recorded in the US.

4. Socioeconomic background and characteristics of fertility

In general, the demographic processes, including fertility analyses are determined by a social and economic background of those populations. We describe the socioeconomic background for the populations of immigrant women and men and non-immigrant women and men to understand these processes much better. In addition, our aim was to compare those populations in the light of socioeconomic factors. To enhance the analyses, race and Hispanic origin of respondents were considered. The fertility characteristics were used to estimate the fertility level before and after immigration to the United States. We employed three measures of fertility such as number of live births, number of planning children and total number of children (see Thornton 1979, Mosher *et al.* 1992). What is more, we collected and compared data on number of pregnancies before and after immigration.

We examined in our study a total of 1079 immigrant women and 6155 US-born women. The population of immigrants consists of 729(67,6%) women of Hispanic origin, 192(17,8%) of non-Hispanic white and 158(14,6%) of non-Hispanic black women. The group of Hispanic origin women was the oldest one at the time of immigration to the United States. While, the average age was 18,57 years, it could suggest that a common woman completed some secondary school before immigration. Meanwhile, the average length of education was 11,24 years which indicates that Hispanic women were educated below secondary school. What is more, this group was likewise the worse educated amongst other group. The highest education level was registered for non-Hispanic white women with the average years of education equaled 14,48. We see that religion was important for 81,5% of non-Hispanic black woman, and 67,9% of Hispanic origin women. The highest percentage (69,2%) of ever married women was observed among non-Hispanic white women, whereas the longest duration of first marriage (10,86 years) was recorded for US-born women. The lowest education level of Hispanic immigrants may be a key factor leading to the lowest income and the highest unemployment level. The wealth statistics show that the average total gross income was 24 230 USD, whereas the average IAE was 13 349 USD for Hispanic origin immigrants. In case of US-born women, the averages were 39 004 USD and 25 458 USD, respectively. The employment characteristics indicate that only 52,7% of Hispanic women had a full-time or part-time job and 43,1% of them was forced to ask for a public assistance. In the end, only 32,1% of Hispanic origin immigrants had own house or apartment. The figures for non-Hispanic white immigrant and non-Hispanic black immigrant were 47,9% and 38,0%, respectively. The highest ratio was noted for US-born women; 52,4% of them had house or apartment. More details present table 4.

Table 4. Socioeconomic characteristics of immigrant and non-immigrant women in the United States

Characteristics (averages or percentages)	Race and Hispanic origin of immigrants †			US-born women (N=6155) †
	Hispanic (N=729)	Non-Hispanic white (N=192)	Non-Hispanic black (N=158)	
Time and age				
# of years on immigration till 2002	12,04	15,21	12,79	-
Age at time of immigration to the US	18,57	15,84	17,44	-
Social background				
# of years of education	11,24	14,48	13,40	13,07
Duration of first marriage	10,56	9,08	9,27	10,86
% of ever married	64,30	69,20	49,30	52,28
% of declaring an importance of religion	67,90	45,50	81,50	57,85
Employment				
% of having full time of part time job	52,70	60,40	65,80	62,11
% of receiving public assistance	43,10	8,90	32,30	26,50
Wealth				
Total gross income (USD)	24 230	42 142	36 859	39 004
Income per adult equivalent (IAE)	13 349	27 133	24 585	25 458
% of having own house or apartment	32,10	47,90	38,00	52,40

Note: †194 immigrant women and 189 US-born women other races and Hispanic origins were excluded from the analyses

Source: own creation

We repeated the analyses in relation to the population of immigrant and non-immigrants. We examined a total of 711 immigrants, 501(70,5%) of Hispanic origin, 115(16,2%) of non-Hispanic white and 95(13,4%) of non-Hispanic black men. A number of US-born men was 3901 individuals. Considering the social background, we see that the highest education level was registered for non-Hispanic white men,

whereas Hispanic origin immigrants were the worse educated group of men. The averages years of education were 14,33 and 11,30, respectively. Moreover, a religion was important for 64,1% of non-Hispanic black and 59,2% of Hispanic origin immigrants. The lowest ratio of ever married men, 33,09%, was registered among US-born men, whereas the highest one, 55,1%, among immigrants of Hispanic origin. Similarly to the population of immigrant women, the group of Hispanic origin men was the oldest one at the time of immigration to the United States. The average age was 18,10 years. The employment and wealth characteristics point out that immigrants of Hispanic origin were the poorest immigrant groups. The mean of total gross income was 28 815 USD and 27,7% of them received a public assistance. For non-Hispanic white and non-Hispanic black the mean of total gross income was 44 533 USD and 38 548 USD, respectively. Finally, 26,1% of Hispanic immigrants had own house or apartment. In case of US-born men, 54,30% of them had any property. See details in table 5.

Table 5. Socioeconomic characteristics of immigrant and non-immigrant men in the United States

Characteristics (averages or percentages)	Race and Hispanic origin of immigrants † Hispanic (N=501)	Non-Hispanic white (N=115)	Non-Hispanic black (N=95)	US-born men (N=3901) †
Time and age				
# of years on immigration till 2002	12,26	14,55	12,77	-
Age at time of immigration to the US	18,10	15,23	17,92	-
Social background				
# of years of education	11,30	14,33	13,46	12,69
Duration of first marriage	3,71	2,75	3,62	3,27
% of ever married	55,1	35,7	46,2	33,09
% of declaring an importance of religion	59,20	35,0	64,1	44,72
Employment				
% of having full time of part time job	80,0	73,9	87,1	70,22
% of receiving public assistance	27,7	7,8	12,9	18,10
Wealth				
Total gross income (USD)	28 815	44 533	38 548	41 029
% of having own house or apartment	26,1	40,0	35,5	54,30

Note: †126 immigrant men and 147 the US-born men other races and Hispanic origins were excluded from the analyses

Source: own creation

The above analyses show that Hispanic origin immigrants, both women and men, were the most numerous group of immigrants in the United States. They were the worse educated group with the lowest salary. They needed a public assistance and only some of them had house or apartment, regardless of having a full time or part time job.

The second part of this chapter deals with the fertility characteristics. The characteristics were presented for a time before immigration and from immigration until year of 2002 (table 6). We present the fertility characteristics for US-born women for comparative purposes.

Table 6. Fertility characteristics

Characteristics (averages or percentages)	Place of birth		
	Outside the US (N=1079) †		In the US (N=6155) †
	Before immigration	After immigration	
% of fecund*	68,00		60,13
% of contraception users**	66,40		58,75
% of pill users***	27,37		18,34
% of condom users***	29,09		15,58
# of live births	0,45	1,13	1,18
# of pregnancies	0,56	1,52	1,73
# of planning children	0,96		0,98
# of total children****	2,54		2,16

Notes: †194 immigrant women and 189 the US-born women other races and Hispanic origins were excluded from the analyses. *Applicable if a respondent was married or cohabiting. Figures base on the populations of 675 immigrant women and 2927 the US-born women. **In the last 12 months before the survey. Applicable if a respondent had a sexual intercourse with a male in the last 12 months before the survey. Figures base on the populations of 875 immigrant women and 3616 the US-born women. *** The most used contraceptive methods. **** For immigrants, figure bases on the average number of live births equals 1,58

Source: own creation

Based on table 6, we can formulate two important conclusions. Firstly, the average number of live births and number of pregnancies before immigration was 0,45 and 0,56, whereas after immigration was 1,13 and 1,52. Secondly, the average number of live births for a whole population of immigrants was 1,58, whereas the average number of pregnancies was 2,08 (not shown in table 6). The averages for US-born women equaled 1,18 and 1,73, respectively. These figures suggest that the motherhood decisions were taken in the population of immigrants frequently. It could be explained by a fact that immigrants come from the countries where fertility is higher than in the US.

Table 7. Average number of live births before and after immigration by age of mother

Age groups	Live births	
	Before immigration	After immigration
15-19	0,01	0,14
20-24	0,18	0,66
25-29	0,29	1,22
30-34	0,48	1,28
35-39	0,70	1,44
40-44	0,83	1,46

Source: own creation

The above conclusions allow us to verify the second hypothesis. That is, immigrant women tend to postpone of motherhood when they plan to immigrate to the US. Furthermore, the average number of live births was higher before immigration in each of 5-year age group (table 7). This could indicate that there is a pattern of motherhood behavior among immigrants entering the United States.

5. Basic level analysis

We employed the ordinary regression to determine an impact of immigration on fertility. The analyses were performed among immigrants and US-born respondents regardless of race and Hispanic origin. The level of fertility was measured by number of live births and number of biological children in the populations of women and men, respectively.

We present the parameter estimates and standard error for number of live births and number of biological children in the regression models in tables 8 and 9. There is evidence that a relation between fertility and place of birth was statistically significant in both models. Respondents who were born outside the US had higher fertility than born in the US. The odd ratio in the model with number of live births was 1,54, whereas in the model with number of biological children was 1,86. These figures suggest that a parenthood was more likely among immigrants.

Table 8. Parameter estimates and standard errors for live births in the ordinary regression model

Variable	Category	Parameter estimates	s.e.	p-value
Number of live births	No children	-0,24	0,02	<0,01
	One child	0,57	0,03	<0,01
	Two children	1,68	0,03	<0,01
	Three children	2,83	0,05	<0,01
	Four children	3,98	0,08	<0,01
Birth place	Outside the US	0,43	0,06	<0,01
	Inside the US	0	-	-

Source: own creation

Table 9. Parameter estimates and standard errors for biological children in the ordinary regression model

Variable	Category	Parameter estimates	s.e.	p-value
Number of biological children	No children	0,73	0,03	<0,01
	One child	1,44	0,04	<0,01
	Two children	2,42	0,05	<0,01
	Three children	3,47	0,08	<0,01
Birth place	Outside the US	0,62	0,07	<0,01
	Inside the US	0	-	-

Source: own creation

Furthermore, the probabilities of having a child with respect to birth order and birth place of mother or father were calculated (table 10 and table 11).

Table 10. Probability of having a child by birth order and birth place of mother

Birth place	Number of live births					
	No children	One child	Two children	Three children	Four children	Five or more children
Outside the US	0,34	0,21	0,23	0,14	0,06	0,03
Inside the US	0,44	0,20	0,20	0,10	0,04	0,02

Source: own creation

Table 11. Probability of having a child by birth order and birth place of father

Birth place	Number of biological children				
	No children	One child	Two children	Three children	Four or more children
Outside the US	0,53	0,17	0,16	0,09	0,06
Inside the US	0,67	0,13	0,11	0,05	0,03

Source: own creation

The probabilities of having a child of any order were higher in the population of immigrants. This pattern was observed regardless of sex of respondent. It is evidence that parenthood will be more likely in the population of immigrants.

6. Control variables analysis level

There are many additional factors which need to be considered in the analyses of fertility. We studied the variables as follows: a period of time being on immigration, marital status, importance of religion, age, Hispanic origin and education. A marital status for the population of women was considered as duration of first marriage. If a respondent was not married before the survey then a zero value was assigned (see Thornton 1979, Mosher *et al.* 1992). We decided to use number of marriages instead of duration of first marriage for the population of men because the initial analyses indicated, that the number of marriages did not determine the fertility significantly. Multiple classification analysis was used and fertility was measured using number of live births (population of women) and number of biological children (population of men).

Table 12. MCA analyses for live births

Birth place	Live births						
	N	Observed means	Adjusted means				
			MS	A	EDU	R	HO
Outside the US	1116	1,51	1,44	1,41	1,48	1,49	1,40
Inside the US	5381	1,21**	1,26	1,23	1,22	1,22	1,23

Notes: A - age, MS - marital status, EDU - education, R - religion, HO - Hispanic origin. ** ≤ o.o5
Source: own creation

Table 12 presents the MCA analyses for the population of immigrant and non-immigrant women. The observed means of number of live births were 1,51 and 1,21 child in the population of immigrant and non-immigrant women, respectively. The difference between means was statistically significant, what suggests that fertility depends on birth place of a woman. Moreover, it indicates that fertility among immigrants were significantly higher than among US-born women. The observed means were adjusted by marital status, age, Hispanic origin, education and religion. The first three control variables explained the difference between the observed means, but to some degree only. When we controlled marital status and age, the difference decreased from 0,30 to 0,18 child for each variable separately, but when Hispanic origin of women was controlled the difference dropped to 0,13 child. As we see, none of the controlled variables explained fully the difference in fertility between two analyzed subgroups of women. It suggests that an immigration process determines the level of fertility.

Table 13. MCA analyses for biological children

Birth place	Biological children						
	N	Observed means	Adjusted means				
			MS	A	EDU	R	L
Outside the US	833	1,06	0,89	0,95	1,06	1,03	0,94
Inside the US	4038	0,64**	0,67	0,67	0,64	0,64	0,66

Notes: A - age, MS - marital status, EDU - education, R - religion, L - Hispanic origin. ** ≤ o.o5
Source: own creation

We repeated these calculations for the populations of immigrant men and non-immigrant men (table 13). Again, marital status, age and Hispanic origin, which played a role in explaining a statistically significant difference between the observed means (1,06 and 0,64). When we controlled marital status, the difference in fertility decreased from 0,42 to 0,22 child, whereas for variables age and Hispanic origin it dropped to 0,28 child. There is evidence that none of variables eliminated the difference between the observed means. It could indicate that a source of differences in fertility is an immigration process.

We would like to point out that changes in the adjusted means compared to the observed mean were larger for immigrants (table 12 and table 13). When we consider marital status, it could be evidence that marriages for immigrant women last longer, whereas for immigrant men, it could prove that they get married frequently. The survey's data shows that the average of all marriages was 6,25 years for immigrant women and 5,63 years for non-immigrant women, whereas number of marriages for immigrant men was about one-third higher in relation to US-born men. When we take into account age, it could suggest that the age distribution differs between the populations of immigrants and non-immigrants. It also could suggest that a maximum of fertility is recorded in the most numerous age group. The data shows that a maximum was observed for women aged 30-34 and 35-39 years for both, immigrants and non-immigrants. On the other side, there were around 21% and 20% of immigrants and around 17% and 16% of US-born women who belonged to these age groups. In case of immigrant men, a maximum of fertility was recorded for men aged 35-39 and 40-44 years. There were 16% and 15%, and 15% and 14% of immigrant men and non-immigrant men who belonged to these age groups, respectively. These figures do not support the initial assumptions about the age distribution and groups with maximum fertility. Finally, when we analyze Hispanic origin, it could indicate that a proportion of Hispanic origin women and men was higher in the population of immigrants. The data shows that there were 57% and 13% of Hispanic origin women among immigrants and non-immigrants, and 60% and 15% of Hispanic origin men among immigrants and non-immigrants, respectively.

The data presented above shows a source of differences in fertility is immigration. This fact confirms our third hypothesis which says that immigration is a crucial determinant of fertility among immigrants.

Duration of residence of immigrants in the United States

A length of time on immigration was the last variable considered in our study. Due to lack of data on biological children for men, we examined this factor solely for the population of immigrant women. We applied the multiple classification analysis to find out how a length of duration of residence determines fertility. The dependent variable was number of live births in the US. We used the control variables such as age, marital status, education and number of children before immigration.

Table 14. MCA analyses for live births among immigrant women

Duration of immigration	Live births in the US						
	N	Observed means	Adjusted means				
			A	MS	EDU	ChBI	A, MS, EDU,ChBI
Up to 10 years	510	0,65	0,67	0,68	0,61	0,64	0,79
11-20 years	340	1,41**	1,36	1,32	1,32	1,35	1,24
21-30 years	162	1,73**	1,48	1,53	1,65	1,56	1,37
31 years and more	67	1,84**	1,67	1,67	1,92	1,80	1,58

Notes: A - age, MS - marital status, EDU - education, ChBI - number of children before immigration. ** ≤ o,o5

Source: own creation

Table 13 shows that none of the control variables determined the fertility distribution among immigrant women in the United States. When all variables were controlled the difference between observed means decreased from 1,18 to 0,79 child. This is evidence that number of live births was determined by the length of time on immigration. Moreover, the level of fertility was directly proportional to duration on immigration. We detailed the MCA analyses by considering live births by age and years of immigration (table 15) (Ford, 1990).

Table 15. Average number of live births by age and time on immigration

Age groups	Years of immigration				N
	Up to 5 years	6-12 years	13-20 years	21 years and more	
15-24	0,35	0,54	0,41	0,71	258
25-34	0,51	1,44	1,54	1,35	442
35-44	0,16	0,88	1,72	1,93	379
15-44	0,38	1,1	1,33	1,65	1079

Source: own creation

We would like to pay special attention to the group of women aged 35-44 years. This group had the lowest fertility level during the first five years after immigration. The average number of live births was 0,16 child. This may indicate that fertility was completed mostly before immigration. Moreover, this could be evidence that mothers e.g. needed some time to raise children who were born before immigration. The data seems to support these assumptions because the average number of live births for women aged 35-44 years before immigration was 0,76 child (not shown in table 15). Despite of these facts, the average number of live births increased 5,5 times for women aged 35-44 years in time of 6-12 years after settled down in the US. This figures support that duration of residence is positively correlated with fertility. Analyzing other age groups we found a similar pattern. The only difference is that fertility among women aged 15-24 and 25-34 years was mostly realized in the United States. It seems to be obvious taking into account age of these women. Finally, we would like to point out the average number of live births throughout the first five years after immigration was 0,51 child for women aged 25-34 years and 0,35 child for women aged 15,24 years, whereas before immigration, 0,34 and 0,10 child, respectively (not shown in table 15). More details present table 15. In general, the presented analyses can be treated as evidence that immigration determines the level of fertility. And, it confirms the last hypothesis saying, that duration of residence is positively correlated with fertility among US immigrants.

7. Conclusions

The paper focused on the fertility analyses between subgroups of US immigrants and US-born individuals, both women and men. The authors verified four hypotheses. The first hypothesis says that higher fertility among immigrants occurs because most of them come from the countries with higher fertility than recorded in the US. Most legal and illegal US immigrants come from Mexico, Salvador, Guatemala and Philippines. The TFR observed in these countries is significantly higher that recorded among the American born citizens. The second hypothesis says

that higher fertility among immigrant women is caused by postponing of motherhood owing to immigration. We showed that immigrant women tend to postpone of motherhood when they plan to immigrate to the US because the average number of live births and number of pregnancies before immigration was significantly lower before immigration compared to after immigration to the US (see Ford 1990). We found that pattern across all 5-year group ages. The third hypothesis was that immigration is a crucial determinant of fertility among immigrants compared to non-immigrants, both women and men. We tested that relation between fertility and place of birth by means of ordinary logistic regression. That relation was statistically significant in model for women subgroup and men subgroup. Respondents who were born outside the US had higher fertility than born in the US (see Kahn, 1994). The final hypothesis says that duration of residence is positively correlated with fertility. The multiple classification analysis was applied to find out how a length of duration of residence determines fertility. We employed the control variables like age, marital status, education and number of children before immigration. When the model controlled all variables the difference between observed average number of live births in the subgroup of immigrant women and US-born women decreased significantly. This stated that fertility was determined by the length of time on immigration (see Rumbaut - Weeks 1986, Ford 1990). Furthermore, we showed that the level of fertility was directly proportional to duration on immigration.

To sum up the research, the following conclusions can be drawn. Firstly, we showed that a relation between fertility and place of birth was statistically significant in both populations, women and men. We proved that, both motherhood and fatherhood is more likely in the populations of immigrants. Secondly, none of the selected controlled variables explained the difference in fertility between immigrants and non-immigrants to the entire extent. It suggests that an immigration process is a major source of variation in fertility between examined populations. What is more, immigration may lead to postponing a maternity or marriage. Finally, we found a positive association between the duration of residence and fertility of the US immigrants. We suggest further detailed research focused on the impact of migration on reproductive behavior including race and Hispanic origin of immigrants to understand and recognize the additional socioeconomic mechanisms which underline this relationship.

References

Agresti, A. 1984: *Analysis of ordinal categorical data.* NY: John Wiley and Sons, New York.

Albert, A.- A. Anderson 1984: "On the existence of maximum likelihood estimates in logistic regression models." *Biometrica*. 71, 1, 1-10.

Andrews, F.- Morgan, J. N.- Sonquist, J. A.-Klem, L. 1973: *Multiple Classification Analysis. A report on a computer program for multiple regression using categorical predictors.* 2nd ed., Survey Research Center, University of Michigan

Armstrong, B. G.-Sloan, M.1989: "Ordinal regression models for epidemiological data." *American Journal of Epidemiology*. 129, 1, 191-204.

Chen, A. Y.-Escarce, J. 2007: "Family structure and the treatment of chilhood asthma." <http://www.nber.org/papers /w13461>.

Ford, K. 1990: "Duration of residence in the United States and the fertility of U.S. immigrants." *International Migration Review*. 24, 1, 34-68.

Goldschneider, C.-Mosher, W. D. 1991: "Patterns of contraceptive use in the United States: the importance of religious factors". *Studies in Family Planning*. 22, 2, 102-115.

Hosmer, D. W.-Lemeshow, S. 2000: *Applied logistic regression*. NY: John Wiley & Sons, New York.

Hwang, S.-Saenz, S. R. 1997: "Fertility of Chinese immigrants in the US: testing a fertility emancipation hypothesis." *Journal of Marriage and the Family*. 59, 1, 50-61.

Jennrich, R. I.-Sampson, P. F. 1976: "Newton-Raphson and related algorithms for maximum likelihood variance component estimation." *Technometrics.* 18, 11-17.

Kahn, J. R. 1994: "Immigrant and native fertility during the 1980's: adaptation and expectations for the future." *International Migration Review.* 28, 3, 501-519.

Lepkowski J. M.-Mosher, W. D.-Davis, K. E.-Groves, R. M.-van Hoewyk, J.-Willem, J. 2006: "National Survey of Family Growth, Cycle 6: Sample design, weighting, imputation, and variance estimation." National Center for Health Statistics, Vital Health Statistics 2, 142.

Long, J. S.-Freess, J. 2006: *Regression Models for Categorical Dependent Variables Using Stata*. 2nd ed. College Station, TX: Stata Press.

McCullagh, P. 1989: "Models for discrete multivariate responses." *Bulletin of the International Statistical Institute*. 53, 407-418.

McCullagh, P. 1989: "What can go wrong with iteratively re-weighted least squares?" In *Multi-level Analysis of Educational Data*, Ed. R. D. Bock. London: Academic Press. Pp. 147-157.

McCullagh, P. 1980: "Regression models for ordinal data (with discussion)." *Journal of the Royal Statistical Society.* Series B, 42, 109-142.

McCullagh, P.-Nelder, J. A. 1989: *Generalized Linear Models*. Chapman & Hall, London.

McNamee, P. 2007: "The effect of economic status on care decisions: results arising from the free personal care policy in Scotland." <http://ssrn.com/abstract=992726>.

Meagher, K. 2008: "Confidentially is not enough: framing effects in students evaluation of economics teaching." <http://ssrn.com/abstract=999261>.

Mosher, W. D.-Williams, L. B.-Johnson, D. P. 1992: "Religion and fertility in the United States: new patterns." *Demography.* 29, 2, 199-214.

Nelder, J. A.-Wedderburn, R. W. M. 1972: "Generalized Linear Models." *Journal of the Royal Statistical Society.* 135, 3, 370-384.

Retherford, R. D.-Choe, M. K. 1993: *Statistical models for causal analysis.* NY: John Wiley and Sons, New York.

Rumbaut, R. G., Weeks, J. R. 1986: "Fertility and adaptation: Indochinese refugees in the United States." *International Migration Review.* 20, 2, 428-466.

Thornton, A. 1979: "Religion and fertility: the case of mormonism." *Journal of Marriage and the Family*. 41, 1, 131-142.

van Ginneken, J.-Razzaque A. 2003: "Supply and Demand Factors in the Fertility Decline in Matlab. Bangladesh in 1977-1999." *European Journal of Population.* 19, 1, 29-45.

Impact of some social, demographic and economic phenomena on sustainability of pension systems in the Arabic world

József Kárpáti dr.[1]

One of the PhD workshops within the Faculty for Economics of the Szeged University designated a series of studies to the problem of aging societies, with a special focus on pension system challenges. The author of this article started participating in this work when the majority of the publication was already edited, so the recently published book of studies (Botos 2009) did contain analyses of pension issues in the EU context, then Germany, Post communist countries, Russia, China and India but the entire Arabic world was missing. This article tries to keep up with the very qualitative series of studies on this issue, introducing a territory that has a very different social, demographic and economic background compared to the other parts of the world.

Keywords: pension systems, MENA countries, women's rights, pension contribution rate

1. Research context and introduction

In Europe relatively little attention is paid on demographic and economic processes of arab societies that currently have a young age structure. By profound analyses we can take such conclusions relating to our pension system which forewarn to the cornerstones of the long-term sustainability of a system in general. In Middle East and North African Arab Countries (MENA countries) the pension systems have several, system level disadvantageous characteristics. The pension systems will be huge and non-fundable in the long run, although currently there is a very narrow coverage of the population. Nowadays there is too little attention turned to that, because these systems based on very young societies still produce surplus, not deficit. This will change soon.

First of all we have to admit that Arab societies' policy makers don't necessarily even want to cover the entire old aged group, since there are several traditions concerning the care about old aged persons, and usually at least two, rather three generations are living together in one single family economy. But this is not

[1] dr. József Kárpáti, Hungarian Central Statistical Office, Head of Planning Department (Budapest), PhD student, University of Szeged (Szeged)

always sufficient to save and treat old aged people – especially women – because a large part of them have no relatives to rely on. This will be introduced later in more detail.

The narrow coverage ratio of the pension systems is not prepared for demographic changes either, which will remarkably reshape the age structure of the society of MENA countries resulting bigger and bigger dependency ratios. In the meantime, the retirement regulations are quite permissive – more to say, exceptional – and large replacement rates can be reached after relatively small vesting period. Moreover, in most of the cases there is no income ceiling on which the pension calculations are based, or it is too high to be long term sustainable. The retirement age limits are low, as well. In addition, on the other side significantly high rate of black and grey economy, the special labour market conditions of women and immigrant workers are typical in most of the examined countries.

The management of the systems is not adequate either. The first objection can be raised regarding the transparency. It is frequent that different pension schemes were established for different labour force groups. The mobility of the labour force between these schemes is difficult and complicated, and sometimes even impossible. The management is weak, the investments are not diversified properly. The characteristics of the labour market result that the majority of the society is uncovered by pension services. First, the pioneer researches of UN and World Bank revealed all these factors between 2002-2007.

Figure 1. MENA countries

Source: OECD

The researchers have started to raise the alarm in MENA countries during the last few years. The first step was the Arab Plan on Ageing to the Year 2012 (ESCWA 2002). The aim of this current paper is to present the economic and social factors behind Arab pension systems and the demographic processes obstructing sustainability in actual regulations of the systems. The paper attempts to reveal those regulating and managing failures which worsen their sustainability further.

2. Special economic and social factors in MENA countries

2.1. Informal economy

Besides the officially registered manufacturing sector in each country there is an informal sector in which not registered labour force does productive work. The size of the informal / illegal / non-regulated labour market and the produced income can be estimated, but significant differences are observed by continent to continent. According to OECD data, in certain African countries black and grey economy generates further 50-55%, in some cases 66-70% of the official GDP. In Latin America this ratio is about 40% in average (with quite important deviation: in Bolivia 68%, in Chile 19%), while in Singapore 13, in Japan 11, in developed

European countries is about 17-18% in average (Schneider 2002). In Hungary according to the calculations of researches referred by the government hidden income produced in black economy is 18% compared to GDP. In Hungary its source is mainly smuggling and concealment of excise tax, tax evasion (VAT), undeclared employment, software product piracy and the violation of the consumer protection act. (Government 2007)

According to a study prepared in Jordan on the Arab world (Saleh 2008), in MENA countries there are two primary causes of the majority of the black economy: first is that the family businesses have a great historical tradition. Services, trade and small-scale output are generally realized by unregistered small enterprises with the help of family members. Other problematic field is the official employment of women which is varying country by country. The second cause is that the labour law regulations are harmonized to the traditions, so in most Arab countries for the internal service sector, the family-owned small businesses and agricultural workers labour law regulations are not valid. So these people don't receive social security services, pension, day-off etc.

From the aspect of the extent of used labour force, in the developing countries the significant part of the available labour force works in the informal sector. In the MENA countries this reaches 48-55% of the total labour force excluding agriculture (Robalino 2005). 51% in Latin America, 78% in Sub Saharan Africa (excluding South African Republic) work illegally (WIEGO 2002).

It's worth mentioning that the measurement and monitoring of the economic performance and labour force usage of the informal sector is highly debated. Although ILO has been attempting to create uniform concepts and methods since 1972, the measurement methods and the comparability of data are still not objective due to the nature of the phenomenon. This field is not part of this analysis, but the phenomenon itself is worthy to note.

Additionally, economies of MENA countries are very heterogeneous. There are ones operating market-type reforms, "civil" states like Jordan or Lebanon. Kingdoms, emirates rich in oil or natural gas form another group. There are „powder kegs" with important row materials like Iran or Iraq, and there are organized military states, like Libya. Depending on the structure of the state, the role and power of the private and public sectors alters in the economy. In the center of this analysis still there are such characteristics which ensure the future warnings and the possibility of taking general consequences.

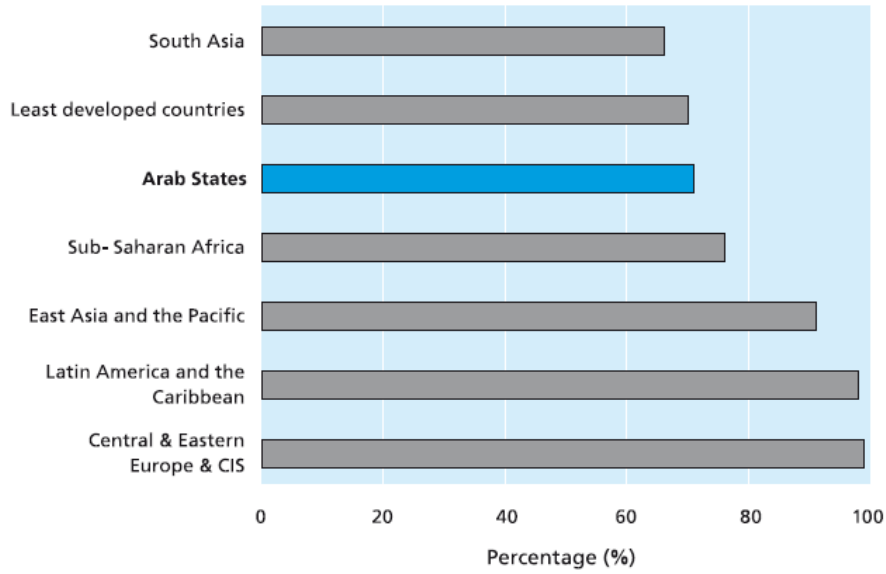
2.2. The position of women in society and their chances

Social state of women is worth special attention as well. It is a worldwide demographic trend that women live longer than men. Actual aged generations represent this phenomenon really sharp. In the world in total, two thirds of the

people at the age of 75 or above are women. In the Arab countries care for older people is mainly based on family connections, the governmental care systems are still not well-prepared for the attending older people, especially old women.

One of the factor of the social isolation of women is that they are deliberately uneducated which can be originated in the previous several decades – the exclusion of women from education. The other fact is that their employment also has obstacles in certain Arab countries (UN 2006).

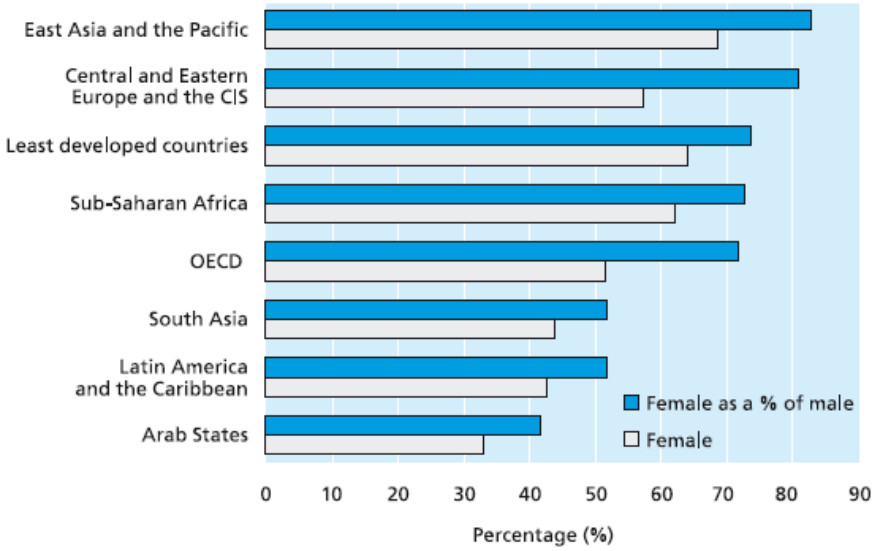
Figure 2. Female literacy rate as a percentage of male literacy rate (age 15+ years), world regions, 2003



Source: UNDP 2005

The major problem of the relationship of these two phenomena is that however the developed countries and international organizations spend high amount of money for the sake of female education programs (and the UN study expresses that these girls usually even score significantly better in school than boys), their entering to the labour market is very difficult anyway.

Figure 3. Female (age 15+ years) economic activity rate (%) and female activity rate as a percentage of male activity rate, world regions, 2003



Source: UNDP 2005

As a majority of women will not be able to enter the official labour market just hidden economy, they also won't be able to make arrangements for their old age. This will increase their dependency from their family members.

Another vivid example of the judgment of women is when they get widowed. For example, in those countries of the examined group where armed conflicts have occurred, 84% of the widowed persons in the area were women in 1996 (UN 2007b). In the most Arab cultures, re-marriage is the exception for men, while women are practically morally stigmatized if they would wish to marry again after their husband had deceased. These factors result in a very negative position for them. In the meantime, the average age of the women increased so much, that their isolation in such situations lasts much longer, locked inside their own society, than before.

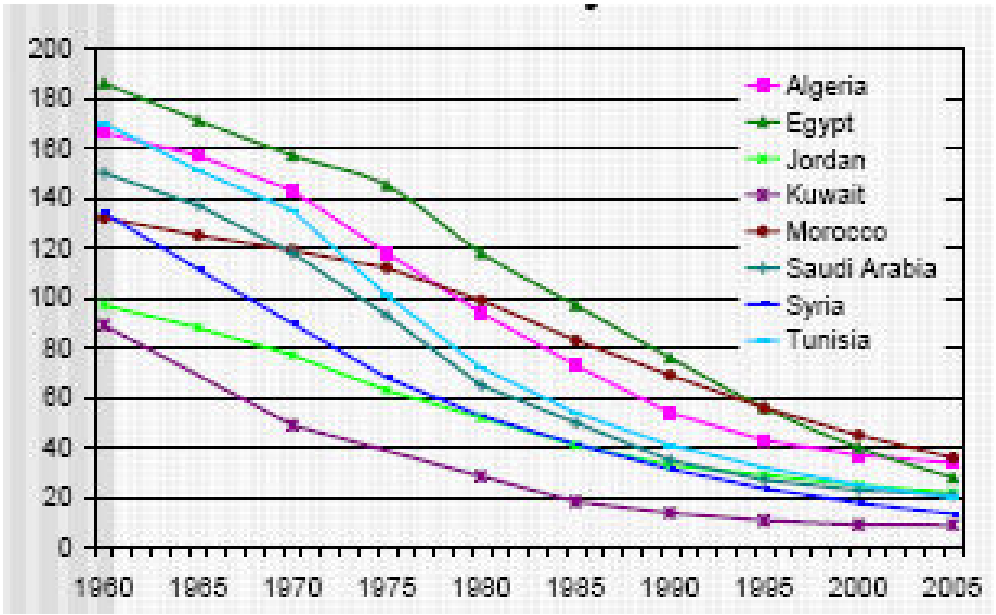
2.3. Demographic factors and labour market interdependencies

One of the most common demographic factors of the MENA countries is the dominance of young aged generations within the total population. In parallel with decreasing, but for a European level still high fertility rates the infant mortality has also radically decreased in these states, while life expectancies at birth are reaching European levels. If we add the already stated growing appearance of highly educated

women to the labour market, we clearly see a warning of dramatic changes within 25-40 years in these nations (ESCWA 2002 IUSSP).

From the end of the fifties very positive changes have started in the MENA countries. Mortality rates decreased, life expectations improved, the quality of medical services improved significantly.

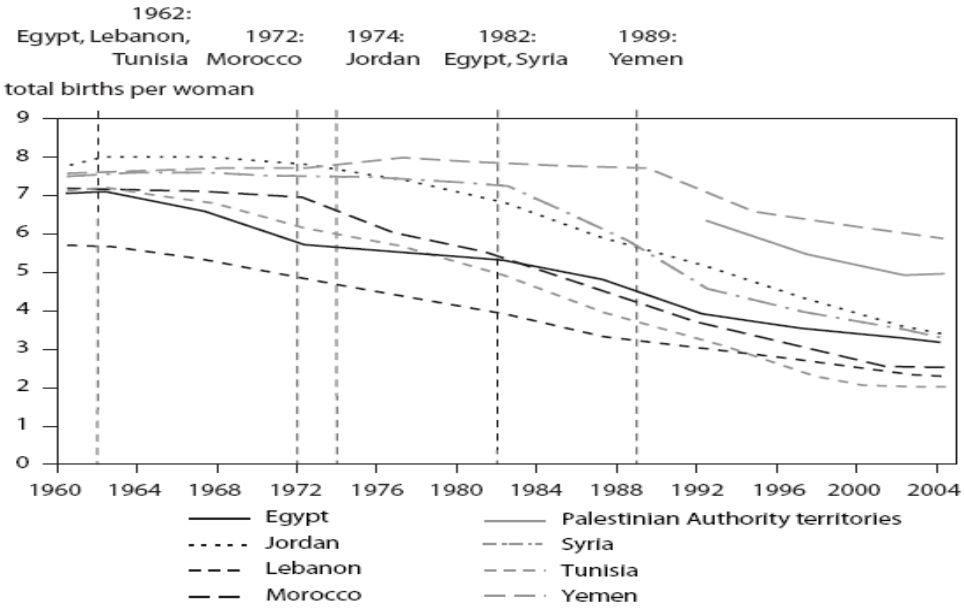
Figure 4. Number infant death cases per 1000 live births (infant mortality rate) in some Arabic countries



Source: Noland-Pack 2007

Although the preciseness of these data concerning Egypt, Saudi-Arabia and Tunisia above is questioned by Jordanian researchers (Saleh 2008), it surely can be stated generally, that thanks to the improvements of the healthcare systems, the infant mortality decreased to its fifth during the last four decades. Fertility rates did not start decreasing at the same time, only about two decades later, int he nineteen-eighties. These fertility rates didn't decrease as strong as the mortality rates, and they are still above European figures.

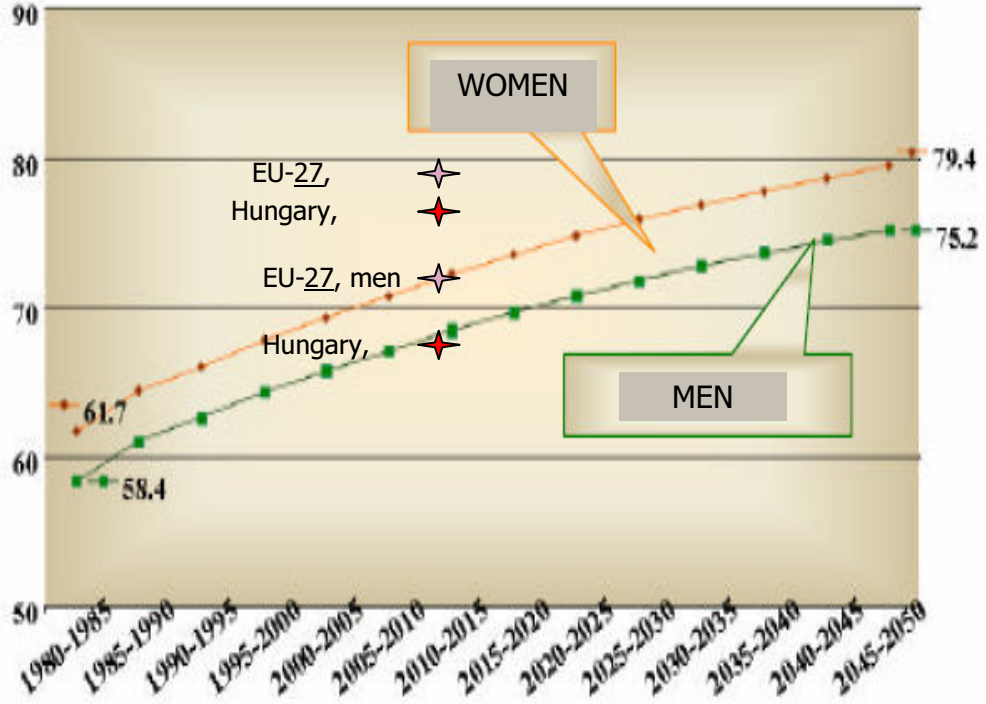
Figure 5. Change of fertility rates in selected countries



Source: World Bank, World Development Indicators, 2004, April 2006

The life expectancies at birth also improved a lot in these countries, resulting in a fact that Arabic countries can proudly offer 75 years for men instead of 58, and almost 80 years for women instead of 62 as it was 25 years earlier.

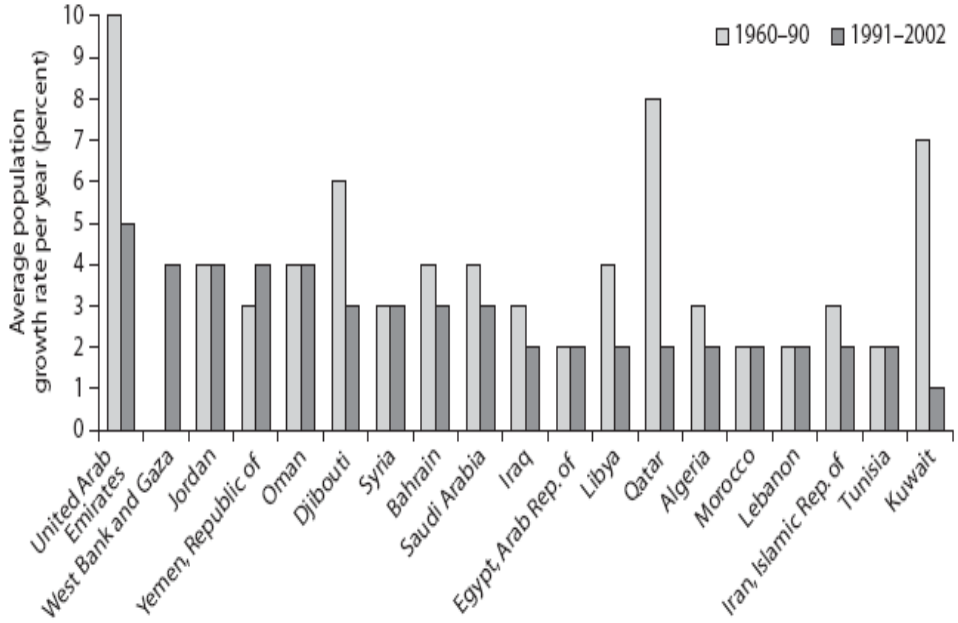
Figure 6. Change of life expectancies at birth in Arabic countries (average)



Source: UN 2007b

These factors bring a high level of rising in the population of the examined countries. The increase is not that intensive anymore, as it was in the sixties or seventies, but the process goes on int he twenty-first century as well. The previously unexpected 6-8 percent of population growth has fallen to 2-3 percent per year, but this growth is becoming very constant. In many Arabic countries we can see a stabile increase of population since forty years. The size of the population in the United Arab Emirates, Qatar and Djibouti is showing a slow down, but is still in the positive range. These dynamics will decrease further between 2015 and 2025, to an average +1,5% per year (Robalino 2005).

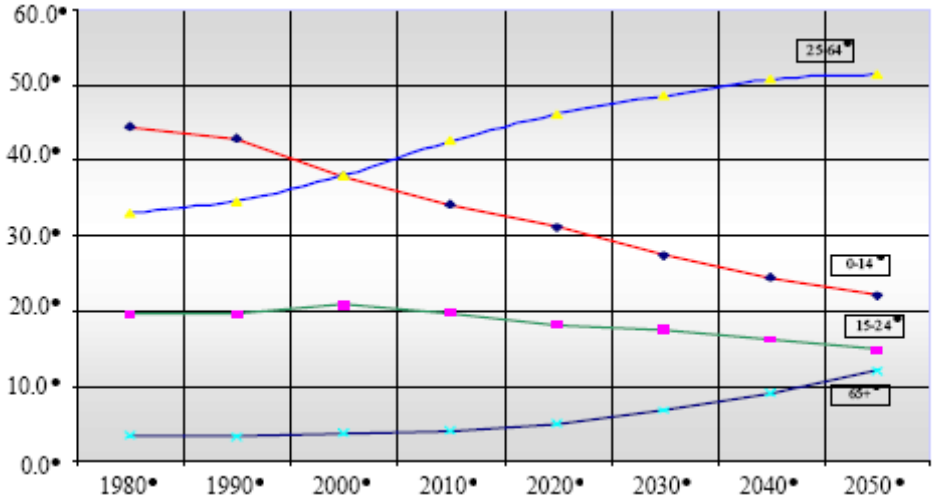
Figure 7. Annual average growth rate of population in MENA countries (percent)



Source: Robalino 2005

It can be stated, that while the age group of economically active cohorts will constantly grow on a small level, the age group of inactives (people above 65 years of age) will suddenly grow from 2020 onwards.

Figure 8. Extrapolated change of the size of age groups in MENA countries (percent of total)

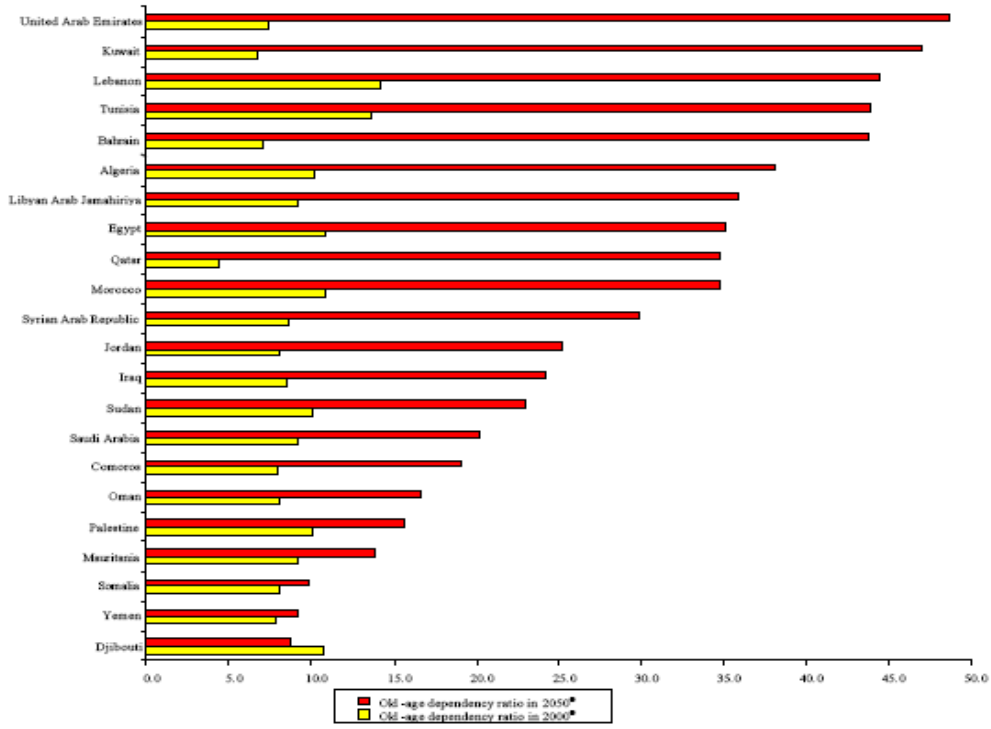


Source: UN 2007b

According to the figure of this UN model from the year 2007 it can be stated that now 55% of the population of the Arabic countries are under 25 years of age and from now onwards this rate will reduce by 20 percentage points (!) by 2050. During the same period of time the rate of the age-group between 25 and 64 years of age will increase from 42 to 51 percent. After the year 2020 the rate of people above 65 will increase dramatically, from the present 5% to 12%, which means more than double share form the population.

Derived from all these the significant increase of old age dependency ratio was predicted [UNO 2007a, 2007b], namely the rate of the 65 and above would increase significantly compared to the age group between 15 and 64. The most stupendous changes are expected in the oil-emirates. In the case of Kuwait, Bahrain, the United Arabian Emirates and Qatar it means that the old age dependency ratio increase from 4-6% to ten times (!), 44-48% in 50 years. The change will concern the South-African Arabian States and Yemen and Palestine less:

Figure 9. The extrapolated change of old age dependency ratios between 2000-2050 (percentage)

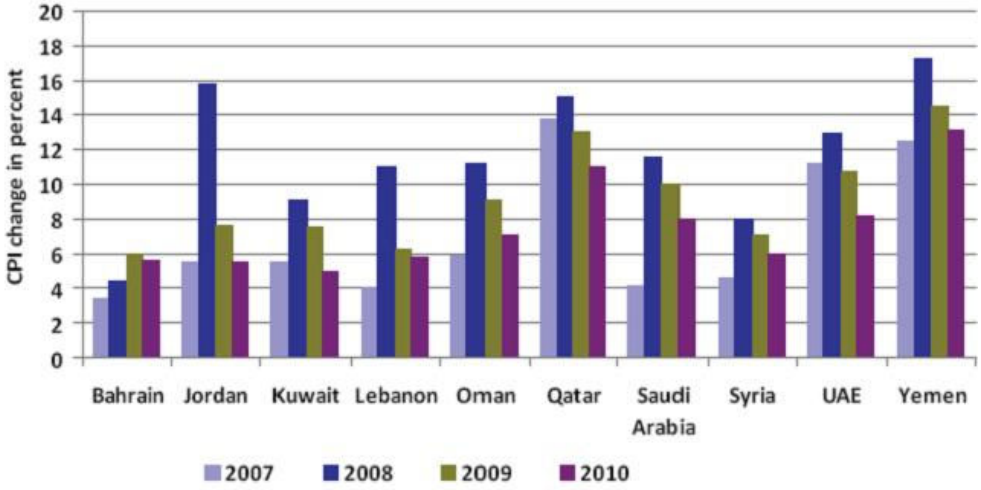


Source: UN, 2007b

The first main question under these circumstances about the effect on labour market will be whether it would be enough legal working places for approximately 50 million more economically active people in the Arabic world until 2015, to absorb the present young generation's labour force. (Here I have to refer again to the subject of the informal economy, which was mentioned earlier.) However even if there would be enough official workplaces under favorable circumstances, the question that how the present high implicit rate of return (see later) that pension systems are offering would remain sustainable. Theoretically if the labour market succeeded in absorbing this mass and the increase remained continuously, the pension systems would be able to roll the present high rate of return to the next generation, and no restrictions, for example income ceiling should be introduced. However this has a really small chance, since the private sector probably wouldn't be able to induce appropriate investments and development for such a growth, while in the public sphere such quantity of workers wouldn't be able to find a job. Among certain Arabic countries we find huge differences in this field, mainly in the number and measure of state enterprises. But it is already frightening that except the states

with rich fossil energy resources, the unemployment rate is getting up officially to between 15 and 30% (Aluwaisheg 2004), while inflation is high as well:

Figure 10. Consumer price indices' changes, including predictions for some MENA countries

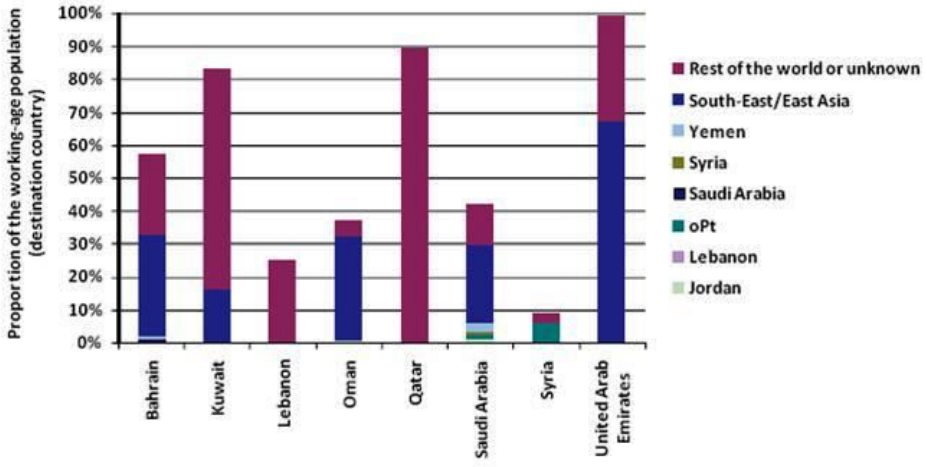


Source: IMF, World Economic Outlook Database, October 2008.

3. The sustainability questions of pension systems

Compared with the relatively large covered share of the population in European (Galasso 2007, Botos ed. 2009) and American pension systems, in the Arabian countries – apart from some exceptions, where the public sphere dominates – on average 25-30% of the active aged population will have a share in pension (SSA/ISSA 2009). The pension systems of certain countries are really young. The longest history of a state operated pension systems in the area can be found in Iraq, where it was founded in 1950. Certain states started to pay pensions only after 10-15 years fundraising periods in the eighties, in an uprising system to their insured clients. We can see many unique cases. It is characteristic that the oil emirates give a huge amount of citizen allowance (mostly to men), then again there is not a compulsory pension system to the private sector, in addition the national pension rules were not extended to the numerous foreign employees of the emirates.

Figure 11. Share of immigrants compared to the total working age population of selected countries



Source: own creation

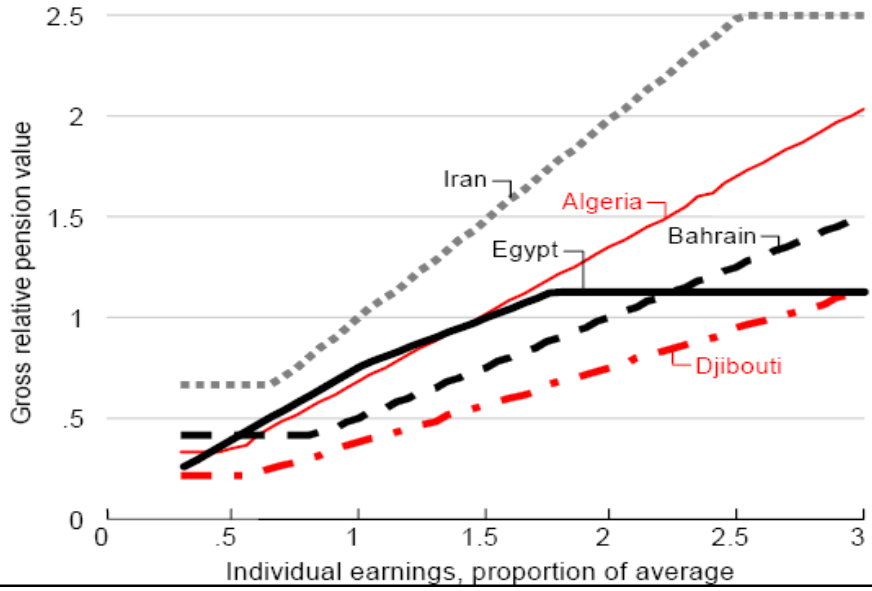
A serious need for jobs will surely force the leaders of the most advanced Arab countries to modify their policies concerning allowances for their citizens and these people will have to start thinking about going to work. But such a decision will have a currently hardly estimated effect on the immigrants' home countries as well.

In the case of the MENA countries with a more serious army (eg. Libya, Gaza Strip or Iran) policy there is a special pension system for this sector, under more favorable conditions of retirement (Country Profile: Iran). The political intention behind that is clear and makes the military a more inviting profession for the employees. To the public employees and state employees the pension system has been evolved everywhere, but the private sphere hasn't. On voluntary ground the joining to the pension system of the private employees is possible just in a small part of the MENA countries. But if someone got into the pension system, can wait his/her inactive years with characteristically good expectations. In addition, the systems are self-financing at the moment, which means nowhere is necessary to draw in additional sources from the state budget: the contributions given from the current active generation produce surplus and cover all the expenses (Rutkowski 2007).

The scale of the granted minimum pension is usually between 35-70% of the average wage, the income ceiling, which is the base of the calculation of pension is missing from the pension systems of the most states. So the pension of those, whose earnings are outstanding, totally act linear upon the higher contributions paid currently by them (see later).

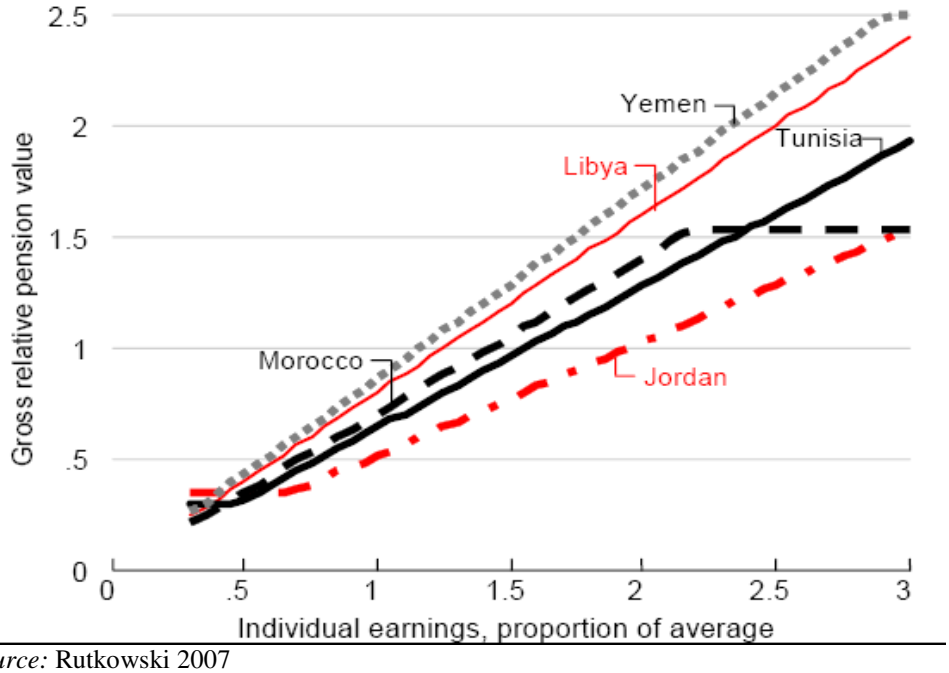
The first important sustainability question is the gross relative value of pensions according to the income before the retirement (rate of income substitution). By the survey of the World Bank it can be seen that the gross relative value of pensions in MENA countries in the case of average wage is its 50-100 percent (Iran pays outstanding high pensions for military), and the amount of pension is usually based on the – mostly – higher income directly before retirement (except from Egypt).

Figure 12. Gross relative value of pensions (y) in the rate of average wage (x) in certain MENA countries (1.)



Source: Rutkowski 2007

Figure 13. The gross relative value of pensions (y) in the rate of average wage (x) in certain MENA countries (2.)



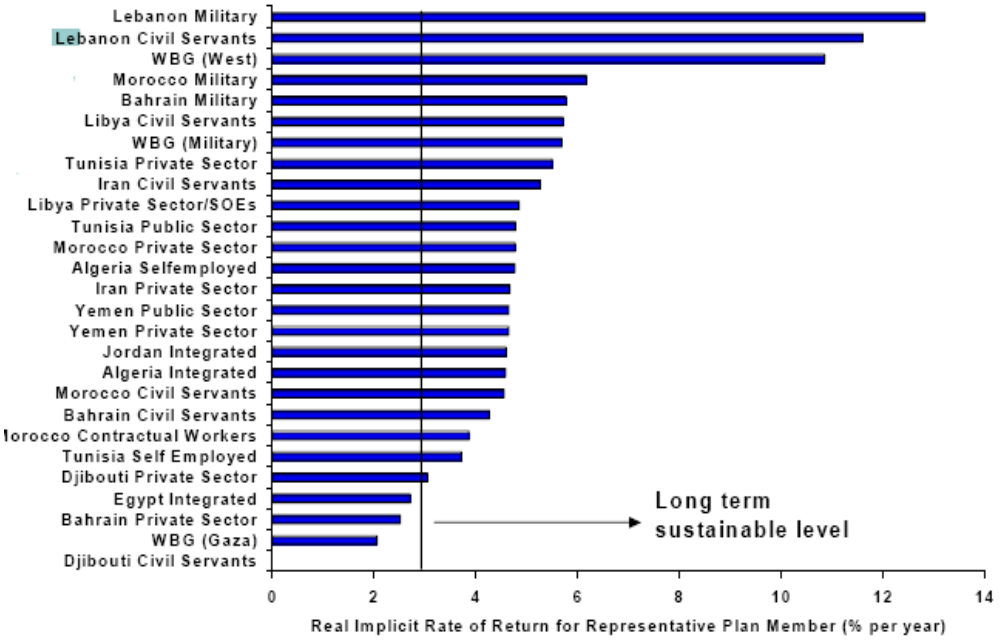
Source: Rutkowski 2007

Especially good services are provided for their retired pensioners by Yemen, Libya and up to the double of the average wage as an individual earning, Morocco.

Besides all that the implicit rate of return paid into the pension cash register is really high in most states, which means paying the pension contribution now is expressly a good investment. This only could sustain besides the certain expansion of the officially reported labour force, whereof there is not too much chance at the moment. So the systems will face the problem that the new entrants finance the result of a good investment for those, who retire now, but by the time the new generation became pensioners, there would not be an adequate coverage for their likely promising investment into their pension.

According to the academic opinion (Robalino 2005, Rutkowski 2007) parallel to the ideal growth of economy, a long-term sustainable system could be operated at about 3 percentage internal rate of return, while primarily the military service twitted state pension systems (Lebanon, Gaza Strip, Morocco, Bahrain) offer such an internal rate of return for their payers, as if their money was put into a commercial bank for 9-12% annual interest rate. The political target of the lately mentioned military systems is obvious again.

Figure 14. Average internal rates of return in the case of a representative (+) considered who retires

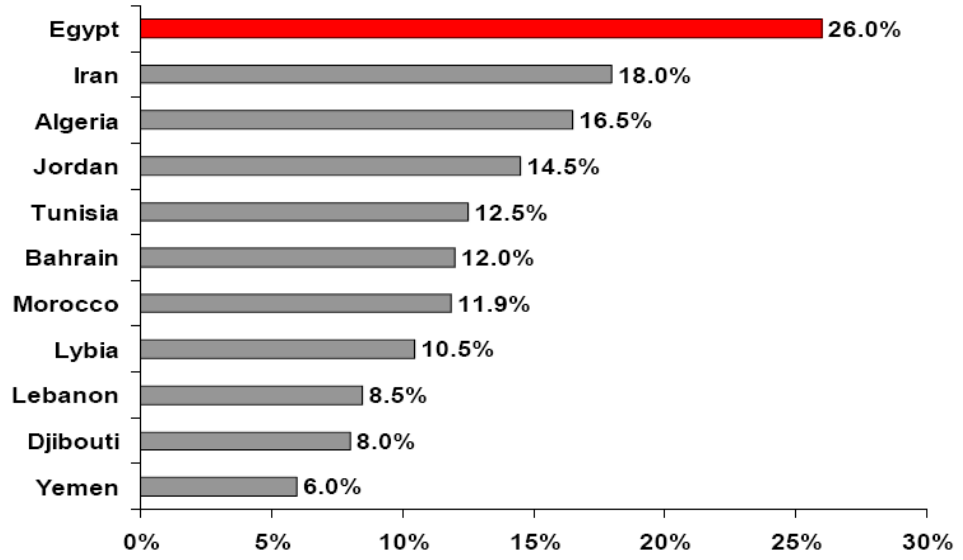


Source: Calculation of World Bank, 2007

(+): representative plan member is someone starting paying a contribution from the age of 25 and working until the legal regular retirement age, having sufficient service years

Offering great services, the contribution rates are not very high for a European eye, except of Egypt maybe.

Figure 15. Pension contribution rates (%) in some Arabic countries



Source: SSA/ISSA

The income that is the base of the pension calculation is usually considered progressively in the most modern pension systems, meaning that the higher for someone the years' income is before retirement, the less they count to the payment sum. In contrary the MENA countries neither don't usually even fix an income ceiling, maximizing the sum of contribution and the achievable service, nor apply a progressive calculation scheme. This brings them currently very high contributions from those who have good jobs right now, but they don't consider how much more they'll have to pay for the same people in few decades, when they retire. So pension systems are facing a high implicit debt, growing this debt from the present 30 to 80-100% of GDP of the given countries. The implicit debt rates of pension systems will grow above the financial debt of the states (Robalino 2005).

A further system-level error is the fact, that these pension systems are not using any income balancing methods for the calculation of the pension sum payable. That means they don't analyze the career progress of the applicant (and the growth of his/her salary) but just take the closing salary before retirement into account and calculate the pension amount with that, without any deflated backwards calculations. In many cases this is unfair and gives the opportunity for people to cheat. With some manipulation the final year's salaries can be "upgraded" heavily, so even a formal raise in the last moth of the salary for example can increase the pension payable for the applicant, while his/her contribution payments during the pervious years did not raise the fundaments for that.

Low retirement ages are also a system-level error. Today in MENA countries the involved women may become pensioners at the age of 55, men at 60 for sure, while special sectors enjoy even more benefits, offering earlier retirement, sometimes at the age of 45.

Management of these pension systems is expensive and not efficient enough. In most cases the infrastructure is not sufficient and the costs of administration are very high. Administrative costs reach up to 5-10% percent of the annual expenses of these pension fund services, while some countries show extreme data (Yemen: 50% of total expenses).

4. Summary statements and recommendations

Pension systems of the MENA countries let the active aged society's minor part (those eligible and covered with the service) retire very early, with a low requirement of service years. The reserves and financial capabilities of these funds are more than sufficient (however, the global financial crisis had an impact on these funds, too). The predictable effects in the near future don't paint a sunny picture about these systems. Sustainability will suffer from serious risk factors, and the structural or regulatory errors of systems will be amplified by the demographic changes. These effects will push the pension systems to unsustainability.

It would be wise starting reform programs for a long term, before crisis will be inevitable. These efforts should focus on the following:

- The possibly largest part of the economically active cohorts of society should pay contributions. An obligatory pension system should be introduced while grey / black economy has to be fought stronger.
- All workers shall pay an affordable contribution but:
- Only the promise of such sums shall be acceptable for pensioners, which can be financed. This means that basic political decision shall be made about the philosophy behind state social support systems in these countries.
- The redistribution of the contribution payment sums shall be transparent. People with larger income shall take larger commitments for the financing of the pensions of people with lower income, but contribution and return should be harmonized.
- Such parametric regulations are required which fortify the sustainability of the systems (e.g. increasing the retirement age).
- The situation of women shall be treated with a high attention.

References

Aluwaisheg, A. A. A. H. 2004: The Need for Integration in the Arab World and the Challenges of Globalization:The GCC Experience, mansucript, 15th Crans Montana Forum, June 2004.

Abu Dhabi Emirate pension law (2000/2.) http://www.pension.gov.ae/PensionsLaw/default.aspx.

Botos, K. (ed.) 2009: Idősödés és globalizáció (Aging and globalization). Tarsoly Kiadó kft. Budapest.

Country Profile: Iran–U.S. Library of Congress, Federal Research Division, 2008, http://memory.loc.gov/frd/cs/profiles/Iran.pdf.

ESCWA The Arab Plan of Action on Ageing to the Year 2012, ENSZ ESCWA, Beirut 2002. http://www.escwa.un.org/popin/docs/ArabPlanofActionAgeing2012Eng.pdf

Galasso, V. 2006: The Political Future of Social Security in Aging Societies, The MIT Press.

Government press release by János Veres on Hungarian black economy http://www.magyarorszag.hu/hirkozpont/hirek/gazdasag/veres20070706.html.

IUSSP-Population Aging in the Arab States-XXV. IUSSP 2005 conference paper, Tours, France.

NOLAND-PACK: The Arab Economies in a Changing World - Peterson Institute 2007, http://books.google.hu/books?id=FzywCP6o0DkC&dq=arab+economies+changing+world&printsec=frontcover&source=bn&hl=hu&ei=2ClOSt2RJNmOsAbqx8nLCA&sa=X&oi=book_result&ct=result&resnum=4.

Robalino, D. A. 2005: *Pensions in the Middle East and North Africa.* World Bank, http://mpra.ub.uni-muenchen.de/10448/.

Rutkowski, M.: MENA Pension Systems and Pension System Objectives, conference presentation, Limassol, December 2007.

Saleh, K. 2008: Informal Sector in Jordan and other MENA countries. manuscript, Department of Statistics, Jordan 2008.

Schneider, F. 2002: Size and Measurement of the Informal Economy in 110 countries around the world. Canberra. http://rru.worldbank.org/Documents/PapersLinks/informal_economy.pdf.

SSA/ISSA, 2009: Social Security Programs Throughout the World: Asia and the Pacific 2008, United States Social Security Administration and International Social Security Association, Geneva. http://www.ssa.gov/policy/docs/progdesc/ssptw/2004-2005/asia/index.html.

UN 2006 - The Arab Human Development Report 2005 (UNDP 2006), http://204.200.211.31/contents/file/ArabHumanDevelopRep2005En.pdf.

UN 2007a. *Gender and Aging.* ENSZ ESCWA Centre for Women, Lebanon 2007,

http://www.escwa.un.org/divisions/ecw_editor/Download.asp?table_name=other&field_name=id&FileID=150.

UN 2007b - *Population ageing in Arab Countries*. ENSZ Economic and Social Commission for Western Asia (ESCWA), manuscript, New York 2007.

WIEGO – *Women in Informal Employment.* Globalizing and Organizing, http://www.wiego.org/about_ie/causes%20and%20consequences.php.

Introducing a Data Manipulation Application for Estimating a Nation's Resources to Manage Cross Cultural Distances

Ismo Koponen[1]

The general purpose of my current research project is to quantify a factor of four elements that seem to affect cultural distances either narrowing or widening gaps between nations in international settings. The A.I.K.A. Factor consists of the following elements: Attitude, Interest, Knowledge and Adaptation Ability.

A more specific aim of this paper is to report on my study in a Nordic setting. A bidirectional survey (total n = 265) was carried out. The resources two Scandinavian nations – Swedes and Finns - and the Russians have for managing their dyadic cultural distances were measured. The surveyed data was processed with The Koponen Manipulator, a two-dimensional data processing tool that converts each determinant to an interdependent one.

Treating the factor with the manipulator results in a method that I tend to call The Dynamic Method.

Keywords: cultural distances, A.I.K.A. factor, Koponen manipulator, Dynamic method

1. A Traditional, static, model …

… for determining Cultural Differences is the one of Geert Hofstede. This chapter illustrates this model with Scandinavian – Russian data. The Scandinavian values of the model's four dimensions are calculations of Swedish and Finnish data, both obtained from the below mentioned source.

[1] Ismo Koponen, M.Sci. (Econ.), senior lecturer, Oulu University of Applied Sciences, School of Business and Information Management, Finland

Acknowledgements: Liikesivistysrahasto (Finland) and The Marcus Wallenberg Foundation (Sweden) have both supported my research project. Thank you, very much!

Table 1. Culture Index Scores of Hofstede's Dimensions

	Power distance	Uncertainty avoidance	Individualism vs Collectivism	Masculinity vs Femininity
Scandinavia	32	44	67	16
Russia	93	95	39	36

Source: Hofstede, 2001: 500, 502. 'Scandinavia' calculated by the author of this paper.

To my understanding, the problem with any traditional Cultural Properties Theory is that they are extremely static of nature. Any such theory illustrates each and one culture in form of a set of fixed dimensions. The quantified dimensions are the same regardless of the viewing point. It can be argued that e.g. Scandinavia be very different from a western than from an eastern viewing point. This is because the viewers are different. Also, it is likely that Scandinavians send different images of themselves to the westerners and to the easterners.

The outcome of the theoretical study is: there is a need for a new – more dynamic - approach. Psychic distances between cultures are both dyadic and unique. With these statements I mean that the distances between any cultures A and B are only between these A and B (a dyad), and that distances A -> B, and B - > A may be very different (unique).

The above became clear to me by 2005. A previous paper of mine reports on the study done in the triad of Sweden, Finland and Russia (Koponen - Tsyvkunov, 2005).

2. The Dynamic Method …

… , on the other hand, consists **of a factor of four elements (determinants), and of a data manipulator**. Together they enable the quantification of cultural distances and resources people have for covering them. The elements of The A.I.K.A. Factor are: Attitude, Interest, Knowledge and Adaptation Ability. The manipulator is a two-dimensional data processing tool that converts each determinant to an interdependent one.

The method makes measuring self attributed psychic distances mathematically possible. The method integrates numeric, and graphic capabilities, making analyzing a problem easy. The two-dimensional manipulator accepts either surveyed or set weights as the Y-dimensions of each XY-projection. Here, I have used surveyed weights. In total, my study is based on 265 filled in questionnaires. The average value of the factor's Y-dimensions is 6.793 (max: 10). It can thus be argued that the relevancy of the factor's elements be 67.93%, on average.

In literature, I can see authors using same concepts as I am doing myself. Someone Edgar Wegman supports my effort by having stated: "Clearly, projection-

based techniques are highly successful and lead to important insights concerning data" (1990). Another, outsider's, impact can be seen in the following: "We want to have easy, flexible, availability of basic or higher level operations, with convenient data manipulation, … We want to be able easily to modify data, … and to do all this and more … adapted to statistical usage" (Chambers 1999). Thank you, both of you!

The scientific assumption on which my theory is based is that the dynamic conditions of The A.I.K.A. Factor indicate people's cultural capabilities.

3. The Empirical Findings …

… in short. The psychic distances within the Nordic Dyad were measured being, firstly: from Scandinavians to Russians 27.08 units, and secondly: from Russians to Scandinavians 23.30 units. Units are 'points' or whatever measures – agreed with the informants – on a scale from zero to 40.

The resources the informants have to cover the distances, are their estimations indicated as the values of The A.I.K.A. Factor. Each determinant has been measured on a scale from zero to 10. The maximum total of resources would, thus, be 40.

Table 2. The Scandinavians' Resources to cover the distance (27.08) to the Russians; n of X = 143, n of Y = 268

	Attitude	Interest	Knowledge	Adaptation ability	Row sums i.e. totals
X - value	4.83	4.84	4.14	4.39	18,20
Y- weight	7.68	6.74	6.41	6.34	
XY - projections	37.11	32.60	26.54	27.85	124.09
Excess	26.47	26.47	26.47	26.47	
Z = X including Y	10.64	6.12	0.06	1.38	18.20

Source: own creation

How to read the above table? The Scandinavians' distance to the Russians was measured by a Scandinavian sample population, and found being 27.08 units from the possible maximum of 40 units. Their resources total in 18.20 units i.e. are approximately 67 % from the maximum of 40 units.

The original X-values do not indicate a specific – more than average – problem. After the data manipulation, however, the Scandinavians low **knowledge** on the Russians' culture, seems to be their most serious problem, in this context. Also, their **adaptation ability** is a problem to be dealt with. The higher readings of **attitude** and **interest** form a good basis for developing the weaker elements of the

factor. At present, the Scandinavians, however, experience being fairly distant from the Russians.

Table 3. The Russians' Resources to cover the distance (23.30) to the Scandinavians; n of X = 125, n of Y = 268

	Attitude	Interest	Knowledge	Adaptation ability	Row sums i.e. totals
X - value	7.51	5.00	3.80	4.97	21.28
Y- weight	7.68	6.74	6.41	6.34	
XY - projections	57.70	33.68	24.36	31.53	147.26
Excess	31.50	31.50	31.50	31.50	
Z = X including Y	26.20	2.18	-7.14	0.03	21.28

Source: own creation

How to read the above table? The Russians' distance to the Scandinavians was measured by a Russian sample population, and found being 23.30 units from the possible maximum of 40 units. Their resources total in 21.28 units i.e. are approximately 91 % from the maximum of 40 units.

The weighted value of the Russians' positive **attitude** towards the Scandinavians is highly emphasized. This, fairly well, 'carries them over' the more problematic elements of their resources. Please, pay attention to **knowledge.** In this context, it has got a negative value, indicating a severe problem! All in all, the Russians feel them being fairly close to the Scandinavians.

4. A final conclusion of the study

The study shows that cultural distances are dyadic and unique. The Dynamic Method makes it possible to measure the distances between cultures. More than that, It also enables measuring the resources cultures experience having for covering the distances.

The Scandinavian and Russian surveys resulted in data that supports my hypothesis that a cultural distance A->B can be different from the cultural distance B->A. In practice, the Scandinavians to Russians -distance is longer than the Russians to Scandinavians -distance.

References

Chambers, J. 1999: Computing With Data: Concepts and Challenges. American Statistician, Feb99, Vol. 53, Issue 1.

http://scholar.google.fi/scholar?hl=fi&q=%22**john+chambers**%22+%22computing+with+data%22&as_ylo=

Hofstede, G. 2001: Cultures Consequences - Comparing Values, Behaviors, Institutions, and Organizations Across Nations. Thousand Oaks: Sage Publications.

http://www.google.com/books?hl=fi&lr=&id=w6z18LJ_1VsC&oi=fnd&pg=PR15&dq=**geert+hofstede**&ots=x4eBD9Krf2&sig=h2yOxSFFzcVrkuzuyjFbZpFlcHY#v=onepage&q=&f=false

Koponen, I. – Tsyvkunov, V. 2005: Psychic Distances to Foreign Cultures Measured within a Nordic Triad - Covering the Distances with the Resources of The A.I.K.A. Factor. A paper presented to the 8th Conference on International Business in Vaasa (Finland), August 22nd - 23rd, 2005.

http://scholar.google.fi/scholar?hl=fi&q=**koponen+tsyvkunov**&as_ylo=

Wegman, E. 1990: Hyperdimensional Data Analysis Using Parallel Coordinates. Journal of the American Statistical Association, Sep90, Vol. 85, Issue 411, 664-675 p.

Application of microsimulation models in the Hungarian Central Statistical Office

Zsolt Kovari[1]

In this article I would like to present a static microsimulation model developed in the Hungarian Central Statistical Office (HCSO), which was used primarily for household's incomes study. This model was intended for a first step of a greater research, thus we worked with the simplest tools. The items of household's incomes were aged, for this the data were gained from the Household Budget Survey (HBS). We were looking for an answer by the method of the microsimulation to the question how exact and reliable result can be produced from a base year for the following year by a static model based on the data of the HBS comparing to the real data. We would like to demonstrate further by some impact analysis whether the models is suitable for analytical presentation of the consequences of economic and political decisions.

The result proved that the method operated well concerning the most important income items. The gross income and its parts, the incomes from work and social incomes were successfully estimated with sufficient accuracy. At those cases where there are considerably differences between the estimated and measured values, however those are interesting from theoretical point of view, in reality they are insignificant small values affecting only few people. The impact analysis is considered also successfully, since its results were as expected, well interpretable and they were in conformity to the measured real values.

Keywords: modelling, microsimulation, income, tax simulation, family allowance

1. Introduction

Modelling is a frequently used and accepted method for the goal of which is to get to know real-life processes more precisely. A model is a simplified picture created by humans. In its creation we do not take into consideration the process of less importance with regard to the examined phenomenon. Theoretical models often appear in a mathematical form for the building of which we use abstract concepts. Simulation is practically a procedure for analysis of mathematical models whose purpose is to make you understand the behaviour of system modelled or to make different experiments with it. In the examination of the economic systems' simulation models are of utmost importance because they are too complex to examine them in another

[1] Zsolt Kovari, senior councillor, Hungarian Central Statistical Office (Budapest)
Author expresses his thanks to Timea Toth for their useful advice.

way and we cannot make experiments with them. Simulation is nearly always carried out using a computer and we draw conclusions on the behaviour of the systems on the basis of the examination.

Microsimulation models are based on so-called micro-units. These are the basic building blocks of the examined system. In case of economic situation these blocks are individuals, households, enterprises and other units. In the process of modelling the characteristics and behaviour are examined on the basis of the statistical data collected. The procedure was developed by G. Orcutt who published his famous article entitled "A new type of socio-economic system" in 1957. Thanks to the extensive research since then, nowadays microsimulation models are made for numerous purposes. These are classified according to many aspects. There are two types: (Molnár 2003)

- Data based models
- Agent based models

Data based models store the micro-units in a database. The database is made by data collection and contains detailed information on micro-units and their economic background. The other component of the model is the rule system which determines changes caused by the examined processes. (Bourguignon–Spadaro 2006) These changes can be events of some probability. Some characteristics of the items, however, can be modified by changing the way of calculation. In the agent-based models the micro-units are usually objects, where they have interrelation with each other and with their environment. Their behaviours are influenced by behavioural rules which are described by algorithms. This article concentrates on data based models.

The data based models are divided into two groups depending on whether we take into consideration the changes in the composition of micro-units in time or not. Thus:

- Static models
- Dynamic models

Changes in time are not taken into consideration in static models, the demographic composition of the micro-units can be left as invariant. Thus these models are used for studying short-term effects of economic or political arrangements. In a household a child can be born, a person may die, or the couple may get a divorced. Moreover new households may be established or ceased. An enterprise may go bankrupt, transform or open new local units. The dynamic models carry out these events with a defined probability and then modify the simulated database properly. Therefore these models are able to analyse the effects of the examined economic decisions in the long run.

Nowadays some microsimulation models appear, which have been developed for special studies, such as spatial models used for regional effects on eco-social political decisions. Microsimulation models were also developed for modelling the effects of crises. In 2008 Robilliard at al. used microsimulation to examine the effects

of poverty and inequality of the financial crisis which happened in Indonesia in 1997. (Robilliard et al. 2008) The effects of the sub-prime crisis which broke out in 2006 in the US and which resulted in the world-wide bank crisis in 2008 will probably be manifested in the data of the Household Budget Survey in 2008 and 2009. This article cannot analyse these questions because at the time of writing this material data on crisis were not readily available. It is an exciting challenge to collect and implement the methodological instruments for the analysis of the effects of the financial crisis.

It can be seen clearly that microsimulation is used very effectively in the process of preparation for decision makings. In the meantime it can be a tool of eco-political experiments, since reliable impact analyses are made with its assistance. The database models show which social groups may be winners or losers of a certain measure with different parameters. You can study the results of the tax rate modifications or imposing of tax on various income components.

2. The former microsimulation work in HCSO

In 1983 at HCSO the Household Statistical Microsimulator (Zafir 1987, 1988) development was started in cooperation with the University of Darmstadt. It was used for income distributions of tax variants. The application was ready for the practical use. (Szívós 1993) In the mid 90's HCSO and Cambridge University developed a simplified microsimulation model, which examined the effects of family supports on the households' incomes. This was useful for the 1996 income survey, but the expedient microsimulation work has not been continued. (Csicsmann-Papp 1998, Éltető-Havasi 1998) In 2006 HCSO in cooperation with Ecostat developed a static model, which was used for the study of tax and imposition of tax on family allowance. The purpose of this article is the presentation of this model.

3. Modelling of the change of the household incomes

The model developed in 2006 was founded on two targets. The one is to be producing some simple IT instruments which can be a starting point for new developments and put the productivity of a rule-based model to the test. The other one to have create a new developer team, which collect the relating special literature, gain developing experience and thus become capable to create more sophisticated models. We set up a static model in 2006, in which the analysis of incomes was emphasized.

3.1. The microsimulation database

The database of the model contained the data of 2004 Household Budget Survey. It was chosen because we would have liked to make the ageing for the next year, and

the 2005 data was available too (but not yet the 2006 data), therefore the calculated and real values could have been correspondent promptly.

The HCSO has been making Household Budget Survey for 50 years. The data were collected up to 1983 yearly, between 1983 and 1993 by every odd year, and after 1993 yearly again. Hungarian citizens living in Hungarian private household were covered by the population. A sample was taken by multiple-stage procedure from the population. The unit of sampling is the dwelling and the observation unit was the household. The sampling framework was the 2001 census. The data collection is continuous all year. Each month the one twelfth of the households in the sample kept a diary, in which making a note

- the expenditures of household's members
- the provisions bought, the quantity of clothing
- the benefits and donations received in kind
- the quantities consumed from own produced foodstuff or given presents

At the process of the questionnaire interrogation the census-takers record the data about

- the most important characteristic of the household's members (sex, age, educational level, economical activity, employment)
- the housing and equipments of the house
- the degree of supply of the consumer durables goods

At the end the households ensure data about

- personal incomes or other revenues
- non-personal (household) incomes
- expenditures on buildings, buying real estate, agricultural activity or significant consumer durable articles.

Consequently the Household Budget Survey contains detailed information about incomes and consuming expenditures of the members of the households in the sample. The data collected are made suitable by proper weighting procedure for representation the total population. After the data quality adjustment and imputation the database is stored in two SAS datasets, the personal and the household file. The two tables can be link by a unique household ID. This is a 15-digit text identifier. The 2004 database contains 23,513 individuals and 8,710 households. 62 variables based on individuals data and 92 variables based on household data were ageing or calculated. These variables can be classified to the next groups:

- Incomes from work (i.e. incomes from all-day employment, different refund of expenses, incomes from secondary employment or casual work, etc.)
- Social incomes (i.e. pensions, unemployment benefits or family allowances, etc.)
- Other incomes (i.e. income of interest payment, dividends, incomes from abroad, etc.)

The first step was the ageing of the incomes of individuals, and then they were summarized for household level. In the third part the household level incomes

(which could not be assigned to persons) were ageing. Household level incomes are for example the amount from insurance companies as life-, pension- or accident- insurance, received life annuity, amount of loan.

3.2. Ageing rules

The model rules were sorted to the following groups:

- Calculation according to regulations (I.e. pensions, different allowances, tax types, tax privileges, social security fees, net incomes are produced on the base of legislation).
- Calculation on the base of a statistical indicator (There were some income items, which were calculated in virtue of inflation or harmonized consumer price index).
- Calculation on the base of some ratio (We figured average growth rate on some data of previous year, to generate the counted values and we hoped that these would be not change significantly in the near future).
- Aggregation (The gross income, the tax base, the household total income, or the total income components were calculated on the summary of the estimated data. In some cases the law had to be taken into account, like the tax regulation asses the content of the base tax.)
- Calculation in virtue of statistical methodological rules (The results for the analysis, like the margins of the income categories, means or dispersion indices has been calculated on the base of usual methodological provisions.)

There are some obvious calculations, which were not sorted to the above mentioned groups, like the production of per capita net income.

When the model was built up we knew the legislation environment and the values of the used statistical pointers were also known. We can't rely on this in the general case, but it is not a contradiction. Naturally a model has some assumptions relating to the certain attributes of the environment for instance the legislation environment. One of the advantages of microsimulation is that we can choose these attributes and rules freely. A hypothetical law can be built in the model because it just means that certain legislation environment will be described. The necessary statistical pointers can be estimated or are assumed some values. Actually we examine the relation between the households and the values. So we always presume the legislation environment and the needful statistical data are known.

4. Computerized implementation

We had to create a proper computer program to carry out the ageing and producing the suitable outputs. Since the starting database stored in SAS datasets, it was a natural choice to elaborate a SAS program to perform the calculations. Those experiences were taken into account, which accumulated on the building of previous sys-

tems, and we used the guides which were found in the special literature or known from general programming methodological notes.

The key parameters were stored in separate parameter tables in order to give or change them easily. We used totally 6 parameter tables. These kept the data for personal income calculations, for defining the family tax privileges or family allowances, moreover further index numbers used in ageing.

The program was broken down to so-called micro-modules. The micro-module – similarly to the modules well-known in program development – is responsible for a certain function. This solution ensured that the modification of the model could be executed with a change or rewriting some micro-modules or parameters. Numerous different models can be formed from suitable system of micro-modules and parameter tables so the construction is extremely flexible.

The final version of our program contained 32 micro-modules, which can be sorted to the following groups:

- Preparatory modules (7 ones has been made. These are responsible for making the copy of the original database, or loading the parameter tables. Accordingly their activity is limited to technical-like tasks.)
- Ageing performing modules (22 ones has been made. These calculate the new values of the income items and produce the aggregated data. Their number is less, then the variables, since the same micro-module operates the variables ageing by the same way.)
- Producing output modules (3 ones has been made. These summarize the individuals and household incomes with and without weights, and create tables, which present different income groups of the households.)

Later other modules have been made too, which produced different publication tables, but these are not built organically into the program. For comparison to the 2005 real data there has to make utility programs, which produce the result from the Household Budget Survey 2005 after the ageing calculation.

5. Results of ageing

The most important result was the comparison between the aged and real income data values of the households. The table below shows the data of the most important income groups:

*Table 1. T*he most important income groups

	Grouping of the income (2004, aged)			
	N	**Total (million Ft)**	**Average**	**Standard deviation**
			(thousand Ft)	
Gross income	**3,775,955**	**10,067,190**	**2,666**	**2,205**
Income from work	3,149,455	7,301,745	2,318	2,392
Social incomes	3,157,745	2,613,833	828	594
Other incomes	3,031,969	151,613	50	326
	Receipts not accounted as income			
Cash from savings	1,181,917	268,780	227	494
Received loans, credits	192,164	114,712	597	1,301
Net income	**3,775,955**	**7,913,930**	**2,096**	**1,370**
	Grouping of the income (2005, real)			
	N	**Total (million Ft)**	**Average**	**Standard deviation**
			(thousand Ft)	
Gross income	3,837,087	9,984,986	2,602	2,045
Income from work	3,187,109	7,083,581	2,223	2,231
Social incomes	3,248,725	2,719,856	837	595
Other incomes	3,007,325	181,549	60	421
	Receipts not accounted as income			
Cash from savings	873,304	166,691	191	365
Received loans, credits	60,709	38,046	627	1,787
Net income	**3,837,087**	**7,986,438**	**2,081**	**1,352**

Source: Own creation

Notes: N means the number of households which have the income from the given type.

It can be seen in the table above that the gross and net incomes were estimated well by the model. The two largest components, the income from work and social incomes, have a good correspondence to their real values. Only in the receipts not accounted as an income has some differences to the fact values. The model significantly overestimated the number of raiser credits and loans, and the money amount from this category. But these are only just a small part of the gross income and touch very few people.

The reason of the differences can be found as the assumption rules are partly realized. The cash from savings was not modified in the model because it was guessed they would not be changed significantly. The government party preparing to the 2006 election did not wanted to deal economic problems, thus the general economic situation has been unchanged. This caused that fewer people must have fewer

cash from previous savings. The estimated value seems to be better, if we take this as cash from savings per capita. While the total amount in 2005 was 62% of the predicted data, the cash from savings per capita estimated 227,000 Fts, however, the real value was 191.000 Fts, as 84% of the estimated in the model.

We estimated the loans and savings on the base of the amount of loans by monetary financial institutions. The largest problem was that the numbers of credit receivers were modified drastically. It can be seen better again if we compare the credit per capita to the aged data. The total received loans in 2005 were 33% of the estimated, the credit per capita was 627,000 Fts, and the aged data was 597,000 Fts. This means only 5% difference. Entirely the receipts not accounted as an income item in 2005 was 53% of the model estimation, but this is not significant concerning the gross income, thus the discrepancy is not considered serious deficiency of the program. The table below shows that the gross income how distributes between the largest income groups.

Table 2. Distribution of the gross income

Gross income share in percentage	
2004 aged	
Gross income	100,0
Income from work	72,5
Social incomes	26,0
Other incomes	1,5
Net income	78,6
2005 real	
Gross income	100,0
Income from work	70,9
Social incomes	27,2
Other incomes	1,8
Net income	80,0

Source: A háztartások jövedelemalakulásának elemzése mikroszimulációs módszerrel (2007)

It's fully visible that the category of other incomes is under 2% both. This insignificant item was underestimated by the model, which is understandable, since the reasons recognizable the changes of laws and changes in questionnaire techniques.

Summarizing the consequences it can be say that the model's assumptions were operated well, since the larger importance income items were estimated by fewer than 5% error. Similarly we received favourable result to the major indicators of the income distribution and the income deciles.

6. Impact analysis

The impact analysis was mentioned among the utilization possibilities of microsimulation. For this demonstration we carried out two impact analysis. The one is related to taxation, in which we calculated the private income tax on the basis of accepted income tax rate of 2005 from the ageing data. This can be kept as an ageing procedure, since the planned tax legislation modifications are know at the previous year autumn, thus it was easy to built into a model. On the other hand it illustrate well, how simple to specify the tax tables, and to study the effects of the planned modifications. The ageing produced immediately the gross and net income, from which it was easily to calculate net income per capita. In addition we examined the shaping of the income deciles, the net income distribution of the different family compositions concerning the sex, or number of children, and we carried out regional examinations too. The incomes were grouping according to the economic activity of the head of household, and the number of active and inactive earners in the household.

The most important part of the modification of the 2005 year tax was that the 3-rated tax became to 2-rated by deleting the middle rate.

Table 3

2004			**2005**	
0–800 000	0	18%	0	18%
800 001–1 500 000	144 000	26%		
1 500 001	326 000	38%	270 000	38%

Source: Hungarian Tax and Financial Control Administration

The tables shows that the new Personal Income Tax has no effect on the incomes under 800,000 Fts, thus there was no foreseen changed for lowest income groups. The income groups between 800,000 and 1,500,000 Fts the system leave more incomes and – as reading from the table – the earners more than 1,500,000 Fts would be richer by 56,000 Fts.

In reality other impacts exists, but summarizing the new tax system lifted the net income per capita in every income deciles. However, the results show that the new system is favourable to the rich. Other investigation results are in harmony with the idea that the households with smaller income did not feel the effect of tax changes. For instance an examination was carried out for the regional impact analysis on the changes of net incomes of the households. Relevant difference was not demonstrable, but the smallest growth was estimated in Northern-Hungary and Northern-Great-Plain (1.3%), which are sorted to the poorest regions of the country. The biggest growth (1.9%) can be declared to the richest part, Central-Hungary.

Similarly explanation is when the households with active earner head of household could be rely on 2.3 % growth, meanwhile the pensioner households

could get 1.3% net income growth. As more active earners were presenting in a household, as larger net income growth happened, this also lead back to the features of the tax system detailed already. Where there was no active earner, the ageing net income surpassed only by 0.1% of the 2005 data, namely significant change was not predicted. Where 3 or more active earners lived in the house, we waited 2.6% growth in net income. More detailed examination would need to reveal the link of the net income growth and the number of children in the household. In our estimation the biggest decrease can be experienced by the couple with one child or with more children in 2005 (2% and 1.9%). Here the interpretation can not refer only to the tax system, since these families can receive family allowances or tax privileges. Simultaneously the result illustrated that the new system does not penalize the family with children.

The other illustrative impact analysis was the family allowance taking into tax base. The original idea was to increase the family allowance by 30%, and calculated into the tax base with taxation on the base of 2005 year rules.

In order to do this we had to link the family allowance to a member of the household. The member must not be child-aged or pensioner and we chose the one who possessed the lowest income. Household Budget Survey considered family allowance household level income, and we also did this during the simulation process.

The result was examined by the similar theory as the taxation. The net incomes were experienced to have increased to a small degree. The biggest growth (2,3%) happened in the first income decile. It is understandable, as the incomes are the lowest there, thus the taxed family allowance could not lifted them. This is considered to be as a positive impact. The other results have similar interpretation. We predicted 2,6% net income growth to the households which had not earner members. Insignificant effect presented by the types of settlement groupings. Here 0,4% net income growth could have been predicted to the household in villages.

Greater differences would have been revealed if the family allowances could have been lifted by 30%. We calculated this amount as the incomes in forms of taxes back to the budget has not been changed. The gross-up of net wages by more than 30% would have needed budget approval, which was not realistic in that economic and political situation.

The government in 2009 submitted a proposal, which involved the family allowance to be taken to the tax base. however, this proposal shared the family allowance among the members of the household in other way, than our way. According to the planned orders the half of the family allowance has to be calculated to the tax base of the parents who bring up together the children. But this proposal was annulled by the Constitutional Court of Hungary, because it trespasses on the theory of the proportional sharing on taxation. Namely in case when the state does not make payable the family allowance to the parent, this means that there is a taxation of not acquired income.

In our version the only one member would have to be under taxation and – assumption a rational activity of human beings – we supposed that this person would be who has the smallest tax base. The theory is rather unfeasible, as according to the Constitutional Court the family allowance to be under taxation can infringe the theory of taxation by ability. Theoretically a case can occur, when a parent brings up more children, she/he has to pay more tax. Because of these decisions – in spite of this can be a manageable by microsimulation – the result executed by the examination carried out can not be usable. But this does not diminish the productivity of the model and the elaboration of the model was not redundant, since we worked several algorithm out which can be useful for similar tasks.

7. Consequences

The microsimulation model developed in 2006 aimed to realize a tool with simple methods in order to carry out income ageing and impact analysis. It was a trial to compare the real and calculated data, and to find that the model rules and the starting database capable for microsimulation examinations. In this article I summarized some results, which proved that the model in its framework operated well. It was important to write it down, because it focuses the methodology of microsimulation and the fact that microsimulation can be happened by simple tools, raising interesting questions. For instance those cases, when the rule based model has also differences between the real and predicted data. Their explanation may lead to the same simple reasons, however, it can reveal a new idea to the model, and it encourages to a better understand of a part.

We planned the model showed to age in more dimensions. It would be important from economical or statistical point of view to build in demographical modules, or model of the activity rate or build in consumer data. It would make it capable to build sufficient and sensitive tax models. By informatics point of view it would be great to solve the easily handling of the parameters, or the automation of data supply.

A new tax simulator is described in XXXVII issue of recurrent publication which has been developed by private persons recently (Belyó et al. 2009). The designers produced a complex model oriented microsimulation system used open source components that can be used for simulation of personal income tax, corporation tax and value added tax. The authors made simulated personal income tax data from tax return data 2007. Corporation data was estimated from fact data 2008. The ageing data of VAT was calculated by consumption patterns of households which were constructed based on Household Budget Survey 2007. The authors could use information on 25,573 firms, 8,547 households and 22,305 individual.

The calculation of personnel income tax and corporation tax carried out by the similar ageing rules as of the model of HCSO[2]. The developers created a modern IT background and designed an effective framework system to help the modelling. The application basis on Web service technologies is a serious step towards a general microsimulation system which does not need collaboration of IT expert or computer man for the building of the model. The main importance of this part of the work, however, that it points to the necessity of a framework system by which the preparatory work can be carried out for the simulation and can be set of the model parameters.

References

Molnár, I. 2003: *A mikroszimuláció alkalmazása.* Ecostat, Budapest.

Bourguignon, F. – Spadaro, A. 2006: *Microsimulation as a Tool for Evaluating Redistribution Policies.* ECINEQ Working Paper 2006-20

Robilliard, A. – Bourguignon, F. – Robinsor, S. 2008: *Crisis and Income Distribution: a micro-macro model for Indonesia* The impact of macroeconomic policies on poverty and income distribution: macro-micro evaluation techniques and tools (Editors: François Bourguignon, Maurizio Bussolo Luiz A. Pereira Da Silva) Palgrave-Macmillan Publishers 93-118p.

Zafir, M. 1988: A háztartási mikroszimuláció felhasználása a statisztikai munkában. *Statisztikai Szemle*, 66, 4, 367-372. p.

Zafír, M. 1987: A *háztartási mikroszimuláció.* KSH, Budapest.

Szívós, P. 1993: A mikroszimulációs rendszer és alkalmazása. *Statisztikai Szemle*, 71, 4, 316-325. p.

Csicsman, J. – Papp, P. 1998: A családtámogatási rendszerek hatásvizsgálata mikroszimulációval. *Statisztikai Szemle*, 76, 3, 238-249. p.

Éltető, Ö. – Havasi, É. 1998: Mikroszimulációs kísérlet a családtámogatások hatásvizsgálatára. *Statisztikai Szemle*, 76, 4, 324-340. p.

Cserháti, I. – Dobszayné – Havasi, É. – Keresztély, T. – Kővári, Zs. – Szép, K. – Takács,. T – Tallér, A. – Tamási, B. – Varga, Zs. 2007: *Háztartások jövedelemalakulásának elemzése mikroszimulációs modellel.* Ecostat, Budapest.

Bátori, D. – Belyó, P. – Molnár, I. – Sinka, I. – Szathmári, B. – Tóth, G. Cs. 2009: *Adó-szimulátor.* Ecostat, Budapest.

[2] The model of HCSO only focused on the personal income tax.

Investigating Dental Care Status with Multilevel Modelling

Elif Çoker [1] *– Meral Yay* [2] *– Ömer Uysal* [3]

The aim of this study is to explore the status of dental care. For this purpose, the data set is taken from a real data set which the survey was carried out in 2003 at Istanbul University Cerrahpaşa School of Medicine, Turkey. The population is defined to be people who are over the age of 18, living in Istanbul. According to the Turkish Statistical Institute indicators, Istanbul is divided into three regions. In these regions, there are 25 towns in total and all of them are included. From these 25 towns, 285 districts are selected randomly. In total, the analyses are performed with the participation of 931 individuals. Taking into account the nested structure of the data set (individuals are nested within districts, districts within towns and towns within regions), multilevel modelling approaches are investigated.

Keywords: Dental care, Multilevel modelling, Nested data, Gifi

1. Introduction

Statistical methods are commonly used in social research. The usage of statisical methods in social research can be explained in two parts: the first half of it introduces descriptive statistics and inferential methods (confidence intervals and significance tests), the second half of it introduces bivariate methods (contingency table analysis, regression) and advanced regresion methods (multiple regression, analysis of variance, logistic regression and its extensions). In this study we focused on an extension of logistic regression called multilevel multinomial logistic regression and its application in social research.

Generally social research is based on individuals. However, since individuals live in social groups, they can not be considered independently from them. Despite

[1] Elif Çoker, PhD student, Mimar Sinan Fine Arts University Faculty of Science and Letters, Department of Statistics (Istanbul)
[2] Meral Yay, PhD student, Mimar Sinan Fine Arts University Faculty of Science and Letters, Department of Statistics (Istanbul)
[3] Ömer uysal, PhD student, Istanbul University Cerrahpaşa School of Medicine, Department of Biostatistics (Istanbul)

this structure, many social researchers aim to explain variability in behaviour and attitudes of individuals, here we are specifically interested in the use of statistical models to analyze quantitative data. Multilevel modellling which has an important place in statistical methods aims to redress the balance, by emphasizing both individulas and their social contexs.

The aim of this study is to describe the Multilevel Multinomial Logistic Regression Models and to apply these models to a data set collected in Turkey on Dental Care Status and then use the Gifi transformation to re-analyze the data set.

2. Logistic Regression Models

A great many variables in social sciences are categorical. It is hardly surprising that social scientist frequently wants to estimate regression models in which the response variable is categorical. In this context logistic regression is a statistical modeling method that can be useful. It describes the relationship between the categorical response variable and one or more continuous and/or categorical explanatory variables. Logistic regression is used when explanatory variables are either continuous or categorical and response variables are categorical. Categorical variables have two main types of measurement scales. These are nomainal and ordinal scales. Nominal categorical variables have categories that have no natural order to them. Ordinal categorical variables have a natural order. The goal of logistic regression is to correctly predict the category of response for individual cases using the most parsimonious model.

Early uses were in biomedical studies, but the past 20 years have also seen much use in social science research and marketing (Agresti 2002). In this context there are two main uses of logistic regression. The first is the prediction of group membership. Since logistic regression calculates the probability or success over the probability of failure, the results of the analysis are in the form of an odds ratio. Logistic regression also provides knowledge of the relationships and strengths among the variables.

Logistic regression does not make the typical assumptions: the responses, conditional on the explanatory variables do not have to be normally distributed, they don't have to be linearly related, and we don't require equal variance within each group. Suppose there is a single quantitative explanatory variable X. For a binary response variable Y, recall $\pi(x)$ denotes the "success" probability at value x. This probability is the parameter for the binomial distribution. Logistic regression has a linear form for the logit of this probability,

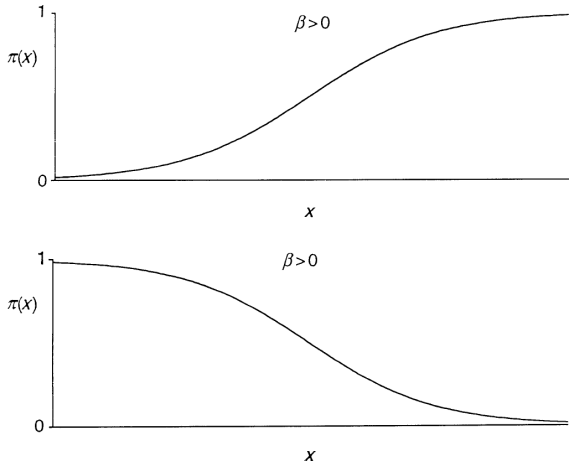
$$\log it\left[\pi(x)\right] = \log\left[\frac{\pi(x)}{1-\pi(x)}\right] = \alpha + \beta x \qquad (1)$$

This is called the logistic regression function (logit function). Equation (1) implies equation (2) for the probability $\pi(x)$, using the exponential function $\exp(\alpha+\beta x)=e^{\alpha+\beta x}$,

$$P(Y=1/X=x)=\pi(x)=\frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}=\frac{1}{1+e^{-(\alpha+\beta x)}} \quad (2)$$

In this function the parameter β indicates the rate of decrease or increase of the curve in the Figure 1. When $\beta>0$, $\pi(x)$ increases as x increases (Agresti 2007).

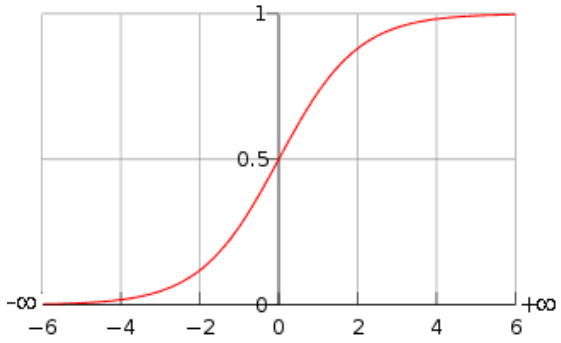
Figure 1. Logistic regression functions



Source: Agresti 2007

The shape of the logistic regression function, which describes the mathematical form of the logistic model can be seen in Figure 2. This shows that $\pi(x)$ increases or decreases as an S shaped function of x. The change in the $P(Y/X=x)$ per unit change in x becomes progressively lower as the conditional mean gets closer to zero or 1 (Hosmer-Lemeshow 2000).

Figure 2. Linear approximation to logistic regression curve



Source: Hosmer-Lemeshow 2000

3. Multinomial Logistic Regression Models

Multinomial logistic regression can be used with a categorical response variable that has more than two categories, but the categories can be ordered or unordered. It compares multiple groups through a series of binary logistic regressions. The group comparisons are equivalent to the comparisons for a dummy-coded response variable, with the category with the highest numeric score used as the baseline category. Like logistic regression, multinomial logistic regression does not make any assumptions of normality, linearity, or variance homogeneity. The multinomial logit model also assumes that the response variable cannot be perfectly predicted from the explanatory variables for any case.

Suppose Y to be a categorical response with C categories and $\{\pi_1, \pi_2, ..., \pi_C\}$ denotes the response probabilities, satisfying $\sum_{c=1}^{C} \pi_c = 1$.

With n independent observations, the probability distribution for the number of outcomes of the C types is multinomial. This distribution defines the probability for n observations into C categories. Multinomial logit models simultaneously use all pairs of categories by specifying the odds of outcome in one category instead of another (Agresti 2002). Logit models pair each response category with the baseline category. The model can be expressed as in equation (3):

$$\log \frac{\pi_c(x)}{\pi_C(x)} = \alpha_c + \beta_c x, \qquad c = 1, 2, ..., C \tag{3}$$

The model has $C-1$ equations, with separate parameters for each. The effects vary according to the category paired with the baseline. If $C=2$, then this model has a single equation, reducing to ordinary logistic regression for binary responses.

4. Multilevel Multinomial Logistic Regression Model

Multilevel multinomial logistic regression models (MM-LRM) are developed for data sets which have a nested structure. These models are also known as mixed-effects multinomial logistic regression models or multilevel logistic regression models for polytomous data (Hedeker 2003; Skrondal-Rabe-Hesketh 2003).

For the terminology of multilevel analysis, let i denote the level-1 units (individuals) and j denote the level-2 units (clusters). Suppose that there are $j=1,...,N$ level-2 units and $i=1,...,\ n_j$ level-1 units nested within each level-2 unit. Thus the total number of level-1 units across level-2 units is $n = \sum_{j=1}^{N} n_j$.

If the nominal response variable has c categories, the multilevel multinomial logit model can be defined in terms of a mixed Generalized Linear Model (Grilli-Rampichini, 2007):

$$\eta_{ij}^{(c)} = \alpha^{(c)} + \beta^{(c)'} x_{ij} + \xi_j^{(c)} + \delta_{ij}^{(c)}, \quad c = 2,...,C \tag{4}$$

It should be noted that there are no category-specific explanatory variables in equation (4), although this is possible. Each equation in this model may have a different intercept ($\alpha^{(c)}$) and regression coefficients ($\beta^{(c)}$). Also ξ_j and δ_{ij} are vectors of random error terms which show unobserved heterogeneity at the cluster and individual level, respectively. We assume the errors are distributed normally ($\xi_j \sim N(0,\Sigma_\xi)$ and $\delta_{ij} \sim N(0,\Sigma_\delta)$), the errors for different levels are assumed to be independent from each other.

The multinomial logit link is defined as in equation (5):

$$P(y_{ij} = c / x_{ij}, \xi_j, \delta_{ij}) = \frac{\exp\left(\eta_{ij}^{(c)}\right)}{1 + \sum_{c=2}^{C} \exp(\eta_{ij}^{(c)})} \tag{5}$$

We consider the response variable y_{ij} to follow a multinomial distribution spanning the set of categories $c = 2,...,C$. We use $c = 1$ as the baseline category for which all the parameters and the random error terms are set to zero. Thus, the conditional probability of $y_{ij} = 1$ is $\left[1 + \sum_{c=2}^{C} \exp(\eta_{ij}^{(c)})\right]^{-1}$.

The likelihood function of the multilevel multinomial logistic regression model is given in equation (6):

$$L(\theta) = \prod_{j=1}^{N} \int \prod_{i=1}^{n_j} \left\{ \int P(y_{ij} / x_{ij}, \xi_j, \delta_{ij}) f(\delta_{ij}) d\delta_{ij} \right\} f(\xi_j) d\xi_j \qquad (6)$$

where $\theta' = (\alpha^{(2)},...,\alpha^{(C)}, \beta^{(2)},...,\beta^{(C)}, \Sigma_{\xi}, \Sigma_{\delta})$. We must use integral approximations to maximize the likelihood, since the integrals do not have closed-form solutions (Grilli-Rampichini 2007). Thus, several methods are proposed and implemented in various software packages for the estimation of these models. But the most frequently used methods are Marginal quasi-likelihood (MQL), Predictive or Penalized quasi-likelihood (PQL) and Full Information Maximum Likelihood (FIML). MQL involves expansion around the fixed part of the model and tends to underestimate the values of both the fixed and random parameters. PQL involves expansion around both the fixed and random part of the model and is more accurate than MQL but computationally less stable (Hedeker, 2008; Pickery-Loosveldt 2002). PQL and MQL are used in MLwiN (Rasbash et al. 2005). FIML uses Gauss-Hermite quadrature for the approximation of the likelihood function's integral and is avaliable in Supermix, SAS PROC NLMIXED (SAS/Stat 2004), Stata (StataCorp 2005), LIMDEP (Greene 2002) etc.

5. Gifi

Gifi is a transformation proposed by Albert Gifi (1989). For a data set which is a combination of continous and categorical variables (called mixed), the Gifi transformation converts the non-linear categorical variables to a linear scale. Once the non-linear variables are transformed to a linear scale, several classical multivariate techniques can be applied to the transformed continous data.

Although Albert Gifi wrote a book about the Gifi transformation, it did not recieve much interest for a long time. Michailidis and de Leeuw (1996) applied this transformation on a pure categorical data set and then used the classical multivariate techniques on the transformed scale to determine the patterns in the data set. Following this study, Suman Katragadda (2009) used the Gifi transformation in a mixed data set which is more complex than a pure categorical data set. After implementing the transformation, the data set was composed of only continous

variables. Thus, he applied classical multivariate techniques in the transformed continous space and identifed useful patterns.

As a brief overview of the Gifi transformation, suppose we have *m* categorical variables and denote these variables as h_j, $j = 1,...,m$. Each variable is assumed to have k_j categories. Suppose that there are *n* observations obtained from these *m* variables. As a result, an *n x m* dimensional information matrix H can be defined. Since the transformation process will lead to some information loss, this loss is expressed in a loss function (Gifi 1989):

$$\sigma(X;Y_1,...,Y_m) = m^{-1}\sum_{j=1}^{m} SS(X - G_j Y_j) = m^{-1}\, tr\left[(X - G_j Y_j)'(X - G_j Y_j)\right] \quad (7)$$

In equation (7), SS is the sum of squares of the H matrix. For each categorical variable, k_j dummy variables can be composed. Thus $G_j(i,t) = 0$ or 1 can be defined as a $G = [G_1,...,G_m]$ vector with $n\,x\sum k_j$ dimensions.

The lost function given in equation (7) is the heart of the Gifi system (Michailidis-de Leeuw 1996). The goal is to minimize the function simultaneously over the X and Y_j's. In this minimizing problem, several restrictions can be imposed. In order to avoid improper solutions corresponding to $X = 0$ and $Y_j = 0$, Gifi (1989) imposed the restrictions given in equation (8) and equation (9).

$$X'X = n\,I_p \quad (8)$$

$$u'X = 0 \quad (9)$$

In equation (9), u is a $p\,x1$ dimensional vector consisting of all 1's. The first restriction given in equation (8) standardizes the squared length of the observed scores to be equal to n and in addition for two and more dimensions, it requires the columns to be orthogonal. The second restriction given in equation (9) requires the graph plot to be centered around the origin. We can use the Alternating Least Squares algorithm to minimize the loss function.

6. Application

In the twenty-first century, the considerable part of the health services will contain studies about reducing both the extensiveness and the volume of a group of diseases starting cardiovascular system diseases, respiratory diseases, cancer, diabetes and tooth diseases. In this study dental care status is examined in particular.

The data set used in the application is taken from a real data set which the survey was carried out in Istanbul University Cerrahpaşa School of Medicine, Turkey in 2003. The aim of the survey was to examine the dental health of adults. To carry out the research, the target group of the survey was selected as people who

are over the age of 18, living in Istanbul. According to the Turkish Statistical Institute indicators, Istanbul is divided into three regions. In these regions, there are 25 towns in total and all of them are included. Our main interest is on Dental Care Status (DCS) which is measured in a nominal scale from 1 to 6. Here 1 is coded as 'all teeth are present', 2 as 'most teeth are present, no dentures', 3 as 'some dentures are present', 4 as 'all teeth are dentures', 5 as 'neither teeth nor dentures are present' and 6 as 'other'.

The data set includes 1000 individuals, but the analyses are performed with non-missing 931 individuals. The data set has a nested structure: 931 individuals (1st-level) are nested within 285 districts (2nd-level) which are nested within 25 towns (3rd-level) and 3 regions (4th-level) of Istanbul. Since there are so many variables to predict DCS in the survey, as a pre-analysis factor analysis is used for the purpose of data reduction. For the prediction of DCS, explanatory variables gender, age, tooth brushing, me, doctor, chance and environment are used. The last four explanatory variables are composed using factor analysis results. The factor 'me' can be defined as the individual considers herself/himself responsible, 'doctor' as the individual considers the doctor responsible for his/her dental care, 'chance' as the individual thinks the dental care status of his/her is like that by chance and 'environment' as the individual thinks the environment is responsible for his/her dental care. Gender is coded as zero for women and one for men. Tooth brushing variable has a ordinal scale from 1 to 7. For example 1 encodes 'I brush my teeth once a day' and 7 codes 'I never brush my teeth'.

Since we have a nested data structure and our response variable is measured on a nominal scale, the first part of the application is about multilevel multinomial logistic regression models. The application is performed with the Supermix software. To begin with modelling, first of all we have to check that the data set really has a 4-level data structure.

To answer this question two models are composed: a 4-level and a single-level multinomial logistic regression model. Since these models are nested, the deviance statistics is used for comparison. Of course, here, 'nested' indicates that a specific model can be derived from a more general model by removing parameters. For nested models, the difference in the deviances has a chi-square distribution with degrees of freedom equal to the difference in the number of parameters estimated in the two models. The deviance test can be used to perform a formal chi-square test, in order to test whether the more general model fits significantly better than the simpler model (Hox 2002). The results suggested that the 4-level multinomial logistic regression model is statistically significant compared to the single-level model ($p<0.001$). Thus, it is sensible to go on with the multilevel models. Considering that we want a good, but at the same time parsimonious model, we thought to reduce this 4-level model to a 3-level model by including an explanatory variable describing region instead of a level. Besides, it should be kept in mind that four regions is a very low number to have variation for DCS. The comparison of these models

suggested that a 3-level model including the explanatory variable describing the region is better than the 4-level multinomial logistic regression model ($p<0.001$). Next, the first-level explanatory variables are added to the model and the results can be seen in Table 1.

Table 1. Three-level multinomial logistic regression model including the first-level explanatory variables

Term	1 vs. 2		1 vs. 3		1 vs. 4		1 vs. 5		1 vs. 6	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	**-2,6312**	0,3779	**-4,593**	0,4629	**-11,4619**	1,4688	**-15,6171**	3,8366	**-4,8639**	0,7164
Region 2	-0,2296	0,3038	-0,2734	0,2823	-0,1943	0,5447	-1,1929	1,5509	-0,6153	0,4758
Region 3	-0,0203	0,2512	-0,4276	0,2897	*-1,0607*	0,6419	-0,1751	1,2643	0,5513	0,5501
Gender	-0,0081	0,1874	-0,3259	0,2216	*-0,8629*	0,4623	-1,2613	1,1103	-0,0535	0,3381
Age	**0,061**	0,0088	**0,1147**	0,0102	**0,2097**	0,0229	**0,2133**	0,0507	**0,041**	0,0149
Tooth brushing	**0,3177**	0,0727	**0,1761**	0,0848	**0,3505**	0,1356	**0,8159**	0,2508	**0,3419**	0,1147
Me	0,0595	0,0999	**0,237**	0,1169	0,006	0,2162	0,4997	0,6008	**0,6339**	0,2267
Doctor	-0,0837	0,096	*0,2176*	0,1143	0,1221	0,2378	*1,3671*	0,6976	*0,2582*	0,1628
Chance	0,0837	0,0974	0,1243	0,1136	0,0089	0,2274	0,0531	0,5255	-0,1869	0,1909
Environment	**0,22**	0,1003	**0,3579**	0,1115	**0,576**	0,2407	-0,3497	0,5001	-0,0211	0,1545
Level-2 variances	0,0425	0,1027	0,0717	0,1669	1,1376	1,0861	1,1506	2,8945	0,435	0,4306
Level-3 variances	0,3654	0,3834	0,1646	0,1354	0,0742	0,1574	0,8209	1,7352	2,616	0,7196
-2log L	2096,04									
AIC	2216,04									
bold indicates $p < 0.05$ *italic* indicates $0.05<p<0.10$										

Source: own creation

Looking at Table 1, the estimates of the model suggest that dummy-coded region variables, gender and doctor variables are non-significant for all five equations. After removing these variables, the final model is obtained and the estimates can be seen in Table 2. What we can see from the final model is we have a two-level multinomial logistic regression model where individuals are nested within towns. This means after removing the non-significant variables from the model, the district level also became non-significant. At the end of the table, the ICC values, which indicated how much of the variation of DCS lies within the district level, are also given. Since all of them are higher than 0.05, it can be considered enough for multilevel models (Muthén & Satorra, 1995).

Table 2. Final Model: Two-level multinomial logistic regression model including the significant first-level explanatory variables

Term	1 vs. 2		1 vs. 3		1 vs. 4		1 vs. 5		1 vs. 6	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	**-2.7117**	0.3293	**-4.8958**	0.4044	**-10.926**	0.9257	**-14.816**	2.3016	**-4.6028**	0.5531
Age	**0.0626**	0.0088	**0.1138**	0.0099	**0.1932**	0.0173	**0.1925**	0.0352	**0.0439**	0.014
Tooth brushing	**0.3144**	0.0709	*0.1596*	0.0818	**0.2897**	0.1197	**0.7327**	0.2257	**0.309**	0.1079
Me	0.0775	0.0939	**0.2968**	0.1105	0.0274	0.1823	0.5896	0.5422	**0.6149**	0.2079
Doctor	-0.0542	0.0922	*0.2003*	0.1073	0.0999	0.2156	**1.3055**	0.6265	**0.3024**	0.1486
Environment	**0.2118**	0.0961	**0.3324**	0.1075	**0.4862**	0.2108	-0.2719	0.4254	0.0065	0.1436
Level-2 variances	**0.4892**	0.3095	**0.1804**	0.124	**0.1855**	0.2113	1.0641	1.9177	2.2127	0.4766
-2log L	2130.14									
AIC	2200.14									
ICC	**0.129**		**0.052**		**0.053**		**0.244**		**0.402**	

Source: own creation

The second part of the application is about using the Gifi transformation. The Gifi transformation is employed for the response variable DCS and the explanatory variable tooth brushing which has an ordinal scale. Since the response variable becomes continous after the transformation, an ordinary multilevel regression modelling approach is used to predict DCS. Hox's (2002) 5-step modelling approach is used throughout modelling the multilevel regression model. According to this 5-step approach, a four-level random-intercept model (M1) is composed and compared with a single-level random-intercept model (M2) as a first step in order to check if multilevel modelling approach is appropriate. A random-intercept model is a model which has no explanatory variables. The estimates of these models can be found in Table 3.

Table 3. First step comparing several models to identify optimal structure

Model	**M1**: Four-level random-intercept model		**M2**: Single-level random-intercept model		**M3**: Three-level random-intercept model	
Fixed part						
Predictor	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	-0.010	0.040	0.000	0.026	-0.010	0.040
Random part						
σ_e^2	0.000	0.000	0.632	0.029	0.016	0.011
σ_u^2	0.016	0.011			0.049	0.021
σ_v^2	0.049	0.021			0.573	0.030
σ_f^2	0.573	0.030				
Deviance	2202.867		2214.904		2202.867	

Source: own creation

For the comparison of the 4-level and the single-level models, the difference between the deviances is 12.04 with 3 degrees of freedom following a chi-square distribution and the related significance level p=0.003<0.01. The null hypothesis is rejected here, which means that there is some error variance in the 4-level model for regions. So now we are sure that we're dealing with a 4-level multilevel model in which individuals are nested within districts, districts are nested within towns and towns are nested within regions. Next, we estimated a 3-level regression model (M3) and compared it with the 4-level regression model. Since the deviance statistics are exactly the same, the 3-level model, which has one fewer parameters is prefferred over the 4-level model.

As a second step, the first-level explanatory variables are added to the model and thus (M4) is composed and the results are given in Table 4, which shows that all the factors (me, doctor, chance and environment) are non-significant. After removing them, the the model is re-estimated and (M5), which is the final model, is obtained.

Table 4: Second step comparing several models to identify optimal structure

Model	**M4**: + First-level predictors		**M5**: The final model	
Fixed part				
Predictor	**Coefficient**	**S.E.**	**Coefficient**	**S.E.**
Intercept	0.498	0.081	0.500	0.080
Gender	0.064	0.050	0.064	0.049
Age	-0.015	0.002	-0.015	0.002
Tooth Brushing	0.213	0.031	0.211	0.031
Me	-0.021	0.027		
Doctor	-0.027	0.026		
Chance	0.012	0.026		
Environment	0.019	0.025		
Random part				
σ_e^2	0.505	0.026	0.500	0.026
σ_u^2	0.032	0.017	0.037	0.017
σ_v^2	0.010	0.008	0.015	0.010
Deviance	2066.693		2068.781	

Source: own creation

The third step would be to add higher-level explanatory variables. Since there wasn't any higher-level explanatory variables, the first-level variables are aggregated to the second level but seen that none of them were significant. The fourth step, which is assessing whether any of the slopes of explanatory variables have a significant variance component between the groups, is also evaluated. Again, there wasn't any significant results. Since we couldn't find any significant variable which has a

random-slope, in order to explain this random slope, adding cross-level interactions would be final step. Thus, this step is passed.

After finding the final models for both multilevel multinomial logistic regression and ordinary multilevel regression models, Akaike Information Criterion (AIC) is used for comparison. The AIC for the two-level multinomial logistic regression model is 2200,14. For the ordinary three-level regression model, the AIC is calculated with the formula AIC = Deviance + 2p where p is the number of estimated parameter in the model and is calculated as 2082,78. Since the lower value of AIC is better, we can conclude that the ordinary three-level regression model is preferred.

7. Discussion

If the economical conditions of the investment of treatment servives are taken into consideration, the minimization of the health budgets is noteworthy. One of the most important parts of the health budgets is left to dental health. From this point of view, dental care status is explored. For this reason, a multilevel multinomial logistic regression and an ordinary multilevel regression model through Gifi transformation is examined for dental care status and compared. In the multilevel multinomial logistic regression model, we ended up with a two-level model in which individuals are nested within towns. And for the prediction of DCS, the final model includes age, tooth brushing, me, doctor and environment explanatory variables. The most fundamental variable of the model is 'tooth brushing'. This means, the lower the frequency of tooth brushing, the worse the status of dental care is, which is reasonable. In the ordinary multilevel regression model, we ended up with a three-level model in which individuals are nested within districts and districts are nested within towns. Here, the final model includes gender, age and tooth brushing explanatory variables. And the most important variable of this models seems to be gender. With regards to AIC, the ordinal three-level regression model is selected as a final model over the two-level multinomial logistic regression model.

References

Agresti, A. 2007: *Introduction to Categorical Data Analysis*. Wiley, New Jersey.

Agresti, A. 2002: *Categorical Data Analysis*. Wiley, New Jersey.

Gifi, A. 1989: *Nonlinear Multivariate Analysis*. Wiley Series in Probability and Mathematical Statistics, New Jersey.

Greene, W. H. 2002: *LIMDEP Version 8.0 User's Manual. 4th edition*, Econometric Software, Plainview, New York.

Grilli, L. – Rampichini, C. 2007: A Multilevel Multinomial Logit Model for the Analysis of Graduates' Skills. *Stat. Methods&Applications*, 16, 381-393. p.

Hedeker, D. 2008: Multilevel models for ordinal and nominal variables. In J. de Leeuw & E. Meijer (Eds.): *Handbook of Multilevel Analysis*. Springer, New York.

Hedeker, D. 2003: A mixed-effects multinomial logistic regression model. *Statistics in Medicine*, 22, 1433-1446. p.

Hosmer, D. W. - Lemeshow, S. 2000: *Applied Logistic Regression*. John Wiley & Sons, Canada.

Hox, J. 2002: *Multilevel Analysis: Techniques and Applications*. Lawrence Erlbaum Associates, New Jersey.

Katragadda, S. 2008: Multivariate Mixed Data Mining with Gifi System using Genetic Algorithm and Information Complexity. *Unpublished Doctorate Thesis*, University of Tennessee, Knoxville.

Michailidis, G. - de Leeuw, J. 1996: *The Gifi System of Descriptive Multivariate Analysis,* UCLA Statistical Series Preprints 204. Univ. California, Los Angeles.

Muthén, B. O. - Satorra, A. 1995: Complex sample data in structural equation modeling. In P.V. Marsden (Ed.): *Sociological methodology.* Blackwell Publishing, Oxford, 267-316. p.

Pickery, J. – Loosveldt, G. 2002: A Multilevel Multinomial Analysis of Interviewer Effects on Various Components on Unit Nonresponse. *Quality&Quantity*, 36, 427-437. p.

Rasbash, J. - Charlton, C. - Browne, W.J. - Healy, M. - Cameron, B. 2009: *MLwiN Version 2.1*. Centre for Multilevel Modelling, University of Bristol.

SAS/Stat, 2004: *SAS/Stat User's Guide. version 9.1*, SAS Institute, Cary, NC.

Skrondal, A. – Rabe-Hesketh, S. 2003: Multilevel Logistic Regression for Polytomous Data and Rankings. *Psychometrika*, 68, 267-287. p.

StataCorp, 2005: *Stata Statistical Software: Release 9*. College Station, TX.

Supermix, http://www.ssicentral.com/supermix/index.html

The prospects of using diaries in a research relating medical services

Beáta Vajda[1]

Diaries have been used in several branches of social sciences for constructing pictures of respondents' (authors') reality. This method is less frequently used than other, „classic" ones, but is gaining ground recently in the field of market research as well as in topics like time allocation, consumer expenditures, travelling habits, social networks, health, diseases and behaviour in connection with them, or criminality. Diary researches are often complemented by other methods, most frequently by in-depth interviews to gain additional information about the topic and the circumstances of using the diary. The most important advantages of using diaries are that they may provide alternative to an interview, especially in cases when events are difficult to recall, or the topic is sensitive. They may also help to overcome problems associated with inaccurate or biased recall. Semi-structured diaries were used to test the usability of this method on the on hand, and to explore the most typical communication processes, feelings and thoughts in a consulting room. Results indicate that diaries are suitable to use in this field and suggest several specific features of medical services.

Keywords: medical services, doctor-patient communication, diary method

1. Introduction

The healthcare system is one of the pillars of the non-profit sector, and nowadays, even developed countries have to struggle with financing difficulties of the healthcare system – it became a worldwide problem. The system of medical services concerns all citizens and determines the whole society, where every single person is a potential consumer. In Hungary, not only the general mortality rate is high, but that of the middle-aged managers as well, and this problem concerns the economically active group to a greater extent. That is why not only the reproduction, but the maintenance of workforce is needed as well (Kincses 2000). It is important because of the fact that in developed countries, knowledge-intensive sectors become dominant, and, as Hungary lacks important natural resources, we have to appreciate our human resources.

According to the scientific literature (Molnár-Csabai 1994, Street et al. 2009), and our preliminary research results (Málovics et al. 2009) as well, communication

[1] Beáta Vajda, PhD student, University of Szeged, Faculty of Economics and Business Administration (Szeged)

style can affect the quality of interaction between the participants, or the willingness for cooperation – thus, the quality and the outcomes of medical services. Getting to know the perceived risk and the efficient communication of risk are important tools for achieving the satisfaction of the patient. The proper risk communication and the involvement of the patient into the decision process are more and more widely considered as factors which improve the quality of medical services. Therefore, it would be important and worth exploring the Hungarian features of this area.
In the process of trying to obtain information about the processes described above, we have decided to use not only focus group interviews, in-depth interviews and questionnaire, but a special type (semi-qualitative, semi-quantitative) of a less frequently used method: diaries. In this paper, the most important characteristics of this method are described, as well as how we have used it in our research on doctor patient relationship and the most important results.

2. About the diary method

Recording a diary originally is a qualitative research method, of which there are several types to be distinguished. Diaries as personal documents or journal types of accounts have been used for a long while by historians and recently by sociologists (Corti 1993) – but the diary technique in scientific research rather refers to records that are kept by a research worker's request, with the intention to "obtain, as accurately as possible, a daily or even hourly record of certain of the respondent's activities" (Oppenheim 2005).

The most popular topics of investigation where diaries are frequently used are time allocation problems (accounts of time telling much about quality of life, social and economic well-being as well as patterns of leisure and work), consumer expenditures, transport planning and health behaviour. Other topics covered using diary methods are social networks, social policy, clinical psychology, crime behaviour or alcohol consumption; and they are increasingly being used in market research (Corti 1993, Jones 2000).

The different types of diaries are illustrated in Table 1. As described, the most important difference between solicited and unsolicited diaries is that while the former is kept at the request of the researcher, the latter have a rather spontaneous nature. Their usability characteristics (from a research point of view, e.g. aspects and methods of analysis, getting answers to research questions etc.) are a result of the attributes arising from this basic difference. "Contemporary" refers to one of the most important feature of all diaries, namely that activities are recorded close in time to their happening, being able to ensure a more accurate picture about the behaviour – this is one of the advantages of using diaries.

Table 1. Types of diaries

Solicited	Unsolicited
Kept at the request of / induced by the researcher	Spontaneous nature
Most often externally structured	Can be strctured or unstructured
Can be rewarded	Unrewarded
Contemporary	Contemporary

Source: Own construction on the basis of Jones 2000

In connection with the acceptance (or usability) of diaries in the scientific research domain, there is no concordance among researchers. While according to one of them, "self-completion diaries have a number of advantages over other data collections methods…" (Corti 1993), others consider it less beneficial: „since the technique is expensive, difficult to design and to place, and hard to analyse, it should only be used, when the necessary estimates cannot be obtained in another way…" (Oppenheim 2005). Naturally, both sides have significant arguments – that is why it is worthwhile reviewing the most important advantages and disadvantages of using diaries.

- Advantages of using diaries:
- They can be more informative – people are often more willing to tell things to a paper than to another person. That is how the problem of sensitive information – which can occur in case of personal interviews – can be eliminated.
- They are less constrained than a questionnaire, and more like an interview. However, with an enough number of diaries, not only qualitative, but also quantitative analysis (e.g. frequencies) may be possible.
- They are respondent-led in emphasis, feelings and expression.
- They can be beneficial to respondents as well, who often state the personal value of recording feelings and experiences in this way.
- They can be useful in checking other methods, such as interviews or questionnaires.
- As filling in the diaries (recording activities, feelings or experiences) happens close in time (certainly the same day, or even right after a happening), they are potentially less constrained in recalling; inaccurate recall, biased recall or the use of heuristics are less feasible.
- Disadvantages of diaries:

- Limitations in the number of them to be collected.
- Interesting or important problems that arise cannot be examined immediately (like with an interview).
- Accurate recordings depend from respondents' commitment to write about events as they occur.
- Much energy have to be applied to respondent preparation and getting diaries in.
- Selectivity – diaries may not include all aspects of the examined area.
- Dependency on potential respondents' willingness to participate (as more effort from them is needed, compared to e. g. a questionnaire).
- Errors can arise from inappropriate respondent conditioning, insufficient cooperation and sample selection bias.
- They require a higher level of literacy of respondents.
- The question of representativeness: how typical was the respondents' behaviour in the particular event?
- (Carlow 2002, Oppenheim 2005)

After considering all features of this method – convinced eventually by the possibility to be able to record feelings close after a consultation, avoiding as much as possible inaccurate recall, we have decided to prepare a pilot diary-module of doctor-patient communication research. Considering this special situation, we have planned to take a special form of diaries: one-occasion diaries. Therefore, we accepted the definition of Oppenheim (2005) for the diary method ("a daily record kept by the respondent at the research worker's request to obtain, as accurately, as possible, a record of the respondent's activities"), with one alteration: we have omitted the word "daily" that originally was a feature of the definition.

3. Research method

As mentioned above, we have planned a special type of diary: a one-occasion one, to be recorded by both patients and doctors (different ones by this two groups, naturally), right after a consultation. With this, our aim was to be able to capture thoughts and feelings of respondents close to the interaction. For this, a questionnaire would have been too restrained (it would not have been able to record feelings with the respondents own words, for example), while interviews would have been very difficult to record in a consulting or a waiting room. Therefore, we have planned a semi-structured diary, in which participants could write their thoughts down.

3.1. The development of the one-occasion diaries

When forming the specific diaries, we had to take both our research requirements and the general method requirements into consideration.

To be able to capture the most important features of doctor-patient communication, we have decided not to ask respondents simply about what had happened during the consultation, but to create a semi-structured diary, which leads the way of telling the experiences and feelings, but still leans on the respondents' own focuses and phrasings – to reflect what the respondent feels or thinks is of personal importance.

From a formal and functional point of view, we have taken the literature's findings and observations into account. According to Carlow (2002), diaries should be tailor-made for each speciality (for example for information, explanation and reassurance received, or feelings that something went wrong), while general instructions should include the need to fill the diary in at the requested time. Oppenheim (2005) confirms the need to pay lots of attention to definitions and written instructions (including the need for accuracy and frankness), and to make the diary format clear, providing adequate space and making it look interesting and important. Corti (1993) also suggests that the inside cover page should contain a set of instructions on how to complete the diary. Naturally, we have tried to consider all the suitable suggestions when developing the diaries. One "unit" in our investigation was a consultation.

3.2. The structure of patient and doctor diaries

Both types of diaries had a cover page with the title and a logo of our university. The second page contained the aim and description of the research, the assurance of anonymity and detailed instructions about how to fill the diary in – these instructions naturally were different for the two groups (besides of asking the respondents to write their experiences down as soon as possible after the consultation and telling that at each question, they are expected to answer in several sentences, using their own words and telling what is important for them).

Doctor diaries instructions included the need to touch upon:

- the style and procession of communication
- uncertainties
- whether psychic support or professional consultation was expected/given
- trust/distrust
- expectations and their fulfilment
- patient competence
- satisfaction.

Patient diaries instructions included the need to touch upon:

- fears and anxieties
- reassurance or increase in anxiety
- understanding and listening
- support
- clear explanations, communication style
- trust/distrust
- expectations and their fulfilment.

After the instruction page, two pages contained the actual questions.

Doctor diaries contained the following questions:

- data on the consultation (its length, those present, the length of the relationship with the patient)
- the runoff of the consultation
- how the consultation went (communication style, who was the initiative, whether there were questions)
- feelings during the consultation
- what was the patient's competence like (whether he/she understood what had been told)
- how decisions about the therapy were made (whether the patient wanted/could participate)
- any other issue the doctor considers important in connection with this topic.

Patient diaries contained the following questions:

- data on the consultation (its length, those present, the length of the relationship with the doctor)
- the runoff of the consultation
- feelings during the consultation
- feelings after the consultation
- decision making about the therapy (the doctor's role, the patient's involvement, satisfaction)
- whether there is anything the doctor should have done differently

- whether there is anything the patient should have done differently

3.3. Participants and data collection

Data collection was realized by the help of a paediatrician, who (by snowball method) managed to call several practicing paediatricians in the research – they constituted the doctor respondents, therefore, one doctor filed several diaries, after several patients. These doctors or their assistants gave the diaries to their patients after the consultations, asking them to participate in a research independent from themselves. Arisen from this specific area (paediatrics), not the actual patients (children) filled the diaries, but their parents accompanying them. One patient therefore filled only one diary.

Altogether, 56 doctor diaries and 70 patient diaries could have been collected and used. The motivations of the visits were: 4 of the patients went to the doctor for counselling with a baby, 4 of them for a control because of some kind of disease, 16 of them had to go to hospital because of the gravity of the situation or because the family doctor had not been available, but it was emergency, and most of them (43) went to the family doctor with common problems or diseases, and in 3 cases, motivation couldn't been identified because of insufficient responses. Although there are diaries, which haven't been filled completely – supposedly because the anxiety of the parent, or the lack of time -, in our opinion, valuable information can be retrieved of them too.

4. Results

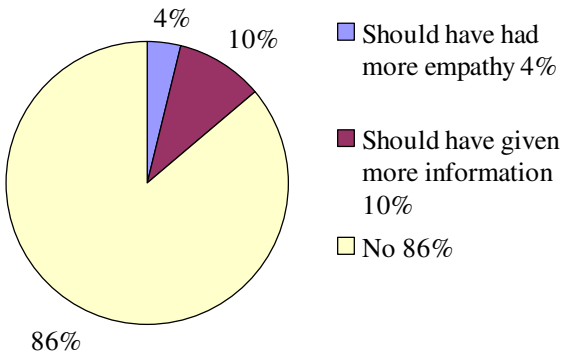
The content analysis of diaries has been realized by the researchers; in this phase of the research, no software was used for the content analysis. As a relatively high amount of diaries could have been collected, frequency analysis of data was realised as well.

As for the description of the happenings during the consultation, naturally, the answers were quite different in some ways, as there can be innumerable problems, and respondents can have diverse styles. But, the description of a process of complaints and questioning – examination – diagnosis – advices and prescriptions is apparent from both types of diaries. Taking the contexts into consideration as well, we can tell, that a paternalistic communication and decision making had been generally used in the cases we examined; and, concluded from other answers, especially from the answers to the question concerning unsatisfied expectations of patients („Is there anything the doctor should have done in a different way?"), we can tell, that most of the parents are satisfied with this style of consultation, or rather do not complaint about it (this result can be eventuated by other factors as well – for further explana-

tion, see later). This result reinforces our preliminary expectations, based on results of in-depth interviews in another module of the research.

Concerning the proportion of parents who thought the physician should have done something differently, it turned out that about 14 per cent of them were not satisfied with the performance of the doctor; and all of them had some kinds of problems with communication (Figure 1).

Figure 1. Whether there is anything the doctor should have done in a different way



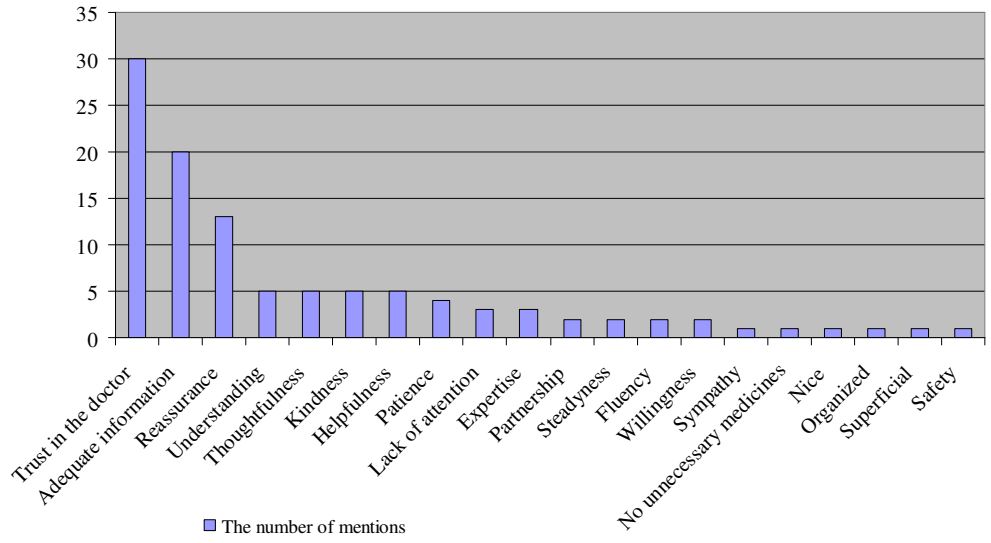
Source: Own creation

It may confirm our assumption of information asymmetry, namely that patients cannot assess the professional competence of the physicians – therefore, they do not necessarily realize if something has gone wrong –, only their communication style and manner – so, if there is some sort of complaint, it is connected to the area he/she can perceive and offer an opinion about – as one of the respondents laid down: „*I cannot criticize the decision of the doctor, as I am incompetent.*" However, even more than a half of those parents who had problems or criticism, left feeling reassured (and all the other respondents stated that too). Interestingly, most of the parents who were dissatisfied, have known and have been seeing the given doctor for years, or even more than a decade – despite the fact that (as we can conclude from their answers) they consider it as a general problem about their doctor. Still, changing doctor (supplier) does not occur. Self-criticism was a peculiarity of 7 of the patients; admitted mistakes were: not being calm enough, not taking the child to doctor in time, not asking the physician for explanation, and trusting the physician in attendance despite the inadequate examination. The results form this question indicated that diaries really had the advantage of being beneficial to respondents – several of them have realized what he/she would do differently another time.

The most common feelings and thoughts respondents had during the consultation were the feeling of trust, reassurance; and adequate disclosure of information

and answers to questions. Figure 2 and Figure 3 show more detailed results of both types of diaries.

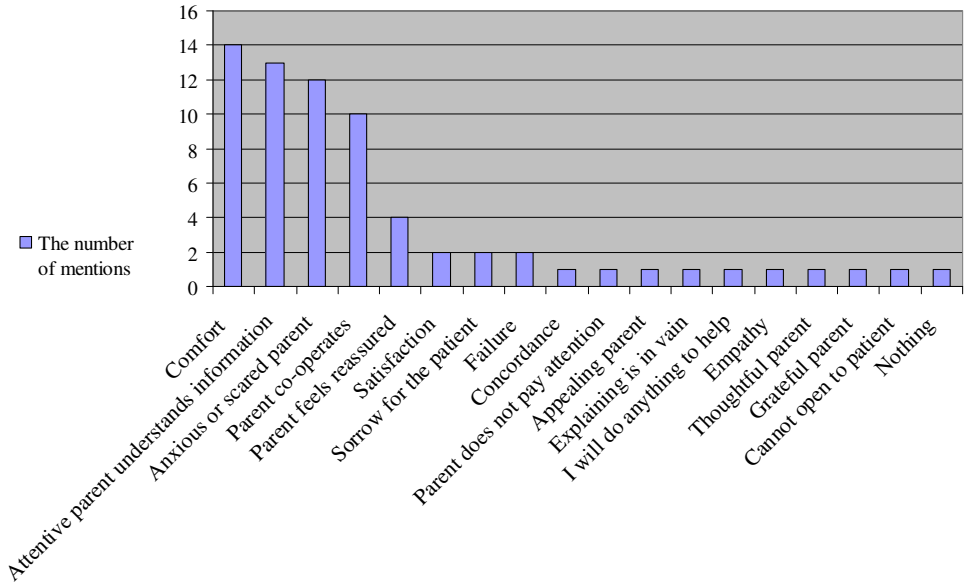
*Figure 2.*Feelings and thoughts of parents during the consultation



Source: Own creation

Patients place most emphasis on social – and communication – competences. If we compare the most typically mentioned feelings and thoughts of doctors and parents of patients, we can see that factors connected to the transmission of information are significant with both groups of respondents – and its positive side is emphasised: many parents think they received adequate information and feel it important enough to mention, and many physicians are satisfied with the way they can transfer information to parents.

*Figure 3.*Feelings and thoughts of doctors during the consultation



Source: Own creation

Nevertheless, phrasings of the doctors refer to the fact that this flow of information is rather one-way; they expect parents to be attentive and understand the information they give. *Signs of two-way communication are missing* from the parents' answers as well. Comfort felt by physicians is the most frequently mentioned feeling – this can be connected to trust, being the most frequent answer of patients: doctors may feel comfortable because they perceive trust oriented at them.

The doctors had the possibility to tell their thoughts about the topic of doctor–patient relationship and communication at the end of the diary. It was a peculiarity that many of them had something to add. The answers were quite informative, and included the suggestions as follows:

- a special consultation room would be desirable,
- staff is incomplete – more doctors are needed,
- much more time is needed for appropriate communication,
- it is substantial to communicate placidly and moderately even in case of a serious illness,
- it is important to get a feedback from to parent if he/she has totally understood the instructions, especially in case of a new medicine,
- attention, empathy and enough time is needed from the physician,

- in many cases, the measurement of whether the parent (who has scant health–related knowledge) managed to understood all the information should be more accurate
- communication trainings should be compulsory for general practitioners.

On the base of these comments of the respondents, we can state that physicians recognise and raise the most important issues about improving the quality and effectiveness of risk communication. They are aware of details that have to be revised and that education and training in communication is missing overmuch (none of the respondents have participated in any communication training). Young graduate physicians have to encounter the task of adequately informing a patient who is not an expert, and additionally emotionally affected by the situation.

5. Limitations and conclusions

Naturally, certain limitations turned out in the application of diaries in our research, especially in case of the patient diaries.. They were filled in the waiting hall of the consulting room, received by a physician or his/her assistant, and several of them given back to the assistant. Reading the answers (although anonymity naturally was emphasised) suggests that respondents could have had „fears" about their doctor – from whom they are dependent. These feelings could have affect their answers, possibly negative opinions or critics were missed out of the answers. Another problem, as diagnosed by the quality of the completion of the diaries, that using a one-event diary, as ours, carries the risk of distortion, as respondents are not suitably prepared to understand what the researchers would like to read about. Further features to improve in the future can be: exploring other areas of therapy (paediatrics is a very special one), data analysis with content analysis software.

However, despite the limitations mentioned above, we think that it is worthwhile setting an improved module of diary research up in the topic of doctor-patient communication, as this stage of research have shown that several valuable information are possible to retrieve, exploiting the advantages of diaries and thoughtfully planning the data collection procedure and analysis.

References:

Carlow, U. 2002: A Guide to Patient Involvement in ABC Programmes. *Ready Steady Book!* National Booking Programme, UK.

Corti, L. 1993: Using diaries in social research. *Social Research Update,* 1993, 2, available at: http://www.soc.surrey.ac.uk/sru/SRU2.html, visited 15 09, 2009.

Jones, R. K. 2000: The Unsolicited Diary as a Qualitative Research Tool for Advanced Research Capacity in the Field of Health and Illness. *Qualitative Health Research*, 10, 4, 555-567. p.

Kincses, Gy. 2000: Az egészségügy új szemlélete. *Egészségügyi menedzsment*, 2000, 5, 91-94. p.

Málovics, É. – Vajda, B. – Kuba, P. 2009: Paternalizmus vagy közös döntés? Páciensek az orvos-beteg kommunikációról. In Hetesi E. – Majó Z. – Lukovics M. (eds): *A szolgáltatások világa.* SZTE Gazdaságtudományi Kar Közleményei, JATE Press, Szeged. 332-343. p.

Molnár, P. – Csabai, M. 1994: *A gyógyítás pszichológiája*. Springer Hungarica, Budapest.

Oppenheim, A. N. 2005: *Questionnaire Design, Interviewing and Attitude Measurement.* Continuum, New York.

Street, R. L. – Makoul, G. – Arora, N. K. – Epstein, R. M. 2009: How does communication heal? Pathways linking clinician-patient communication to health outcomes. *Patient Education and Counseling*, 74, 295-301. p.

Double View at the Quality of Life in districts (LAU1) of the Czech Republic[1]

Renata Klufová[2] *– Ivana Faltová Leitmanová*[3]

The economic development, observed implicitly by the level, respectively by the development of the quality of life, is characterized more or less by heterogeneous economic, social, demographic, ecologic and other indicators. Variable approach to the definition of the quality of life and its indicators is further modified in case of an investigation at the lower than the national level. The paper focuses in detail on economic and social aspects of the quality of life at the level of districts (LAU1) of the Czech Republic. Problems stem for example from unavailability of data on a given or lower hierarchy levels. In some cases, the obtained data comes from sample surveys and they can be insufficiently representative considering a lower hierarchy level. In other cases, it is not possible to analyze data obtained at higher levels. However, the information about the quality of life and its aspects at the districts and lower levels is essential, at least from the point of view their functioning and sustainable development, whose strategy is hard to define if there is a lack of data and information.
The aim of the paper is to assess the possibility of using quality of life indicators at the various hierarchy levels, especially at the level of Czech districts (LAU1).

Keywords: quality of life, regional development, cluster analysis, GIS

1. Introduction

Districts and municipalities present the basic units of the administrative classification in the Czech Republic. According to the law, they are obligated to take care of fulfilling needs of their inhabitants. We can see the administrative units as

[1] * The research reported in this paper was supported by the Ministry of Education of the Czech Republic (grant FRVŠ 237/09) and Ministry of Agriculture of the Czech Republic (project NAZV QI92A023)

[2] RNDr. Renata Klufová, PhD, senior lecturer, University of South Bohemia, Faculty of Economics, Czech Republic
[3] doc. Ing. Ivana Faltová Leitmanová, CSc, associate professor, University of South Bohemia, Faculty of Economics, Czech Republic

Acknowledgement:

environment in which its inhabitants naturally make their considerations about the quality of life. However, the relationship between wide spectrum of needs and the degree of their satisfaction on one hand, and the level of the quality of life on the other hand, is not unambiguous. The quality of life can be examined by various types of indicators (economic, demographic, social etc.) in particular administrative units. The changes of the quality of life in time and space are also important for the assessment and regional policy decision making. We tried to create a typology of the Czech districts according their level of the quality of life on the base of chosen indicators and its changes in the last decade.

2. Surveys

Canadian system "Quality of Life Reporting System" offers an inspirable look at the quality of life on the level of municipalities on the whole. It enables scenario definition of indicators according an aim, users and problematic areas (Burrett 2009). These considerations are contextualized competitiveness of large towns and urbanized units in a global world, where it is needed to make smaller towns or less urbanized complexes more attractive for living. Measuring of the quality of life is in the condition of municipalities connected with many problems which than complicate a comparison, as it can be seen on the methodology applied in Spanish 314 province of Barcelona (Royuela et al. 2003). The analysis of more than 60 municipalities in Austria concluded that the situation in agriculture is one of the critical factors for the quality of life in the countryside (Baaske et al. 2009). The research of municipalities of the Silesian province showed that the effort to balance the development needs a change of a model of their management (Skotnicka-Zasadzien – Midor 2009).

A synthetic index of 19 indices was constructed for the measuring of the quality of life in 643 Spanish villages. The indices describe consumption, social services, housing, transport, environment, labour market, health, culture and leisure, education and security. These characteristics in the frame of Value Efficiency Analysis (VEA), which is a more detailed specification of the Data Envelopment Analysis (DEA) (Gonzáles et al. 2009), provided for the construction of the order of municipalities according to their quality of life. The quality of life (as a component and a precondition of the sustainability of the development in the context of their three parts – economic, social and environmental) served also as a base of a methodology for performance evaluation (Scipioni et al. 2004). Thus, it allowed overcome the usual contrast between model usefulness and its simplicity relatively easily. Only demographic characteristics, especially a family size, served as a starting point for the research in Ghana (Arthur 2008).

Canadian municipalities (Ontario) define for the evaluation of the quality of their inhabitants priorities in four areas: economy (Business and Employment Variety, Image, Advanced Technology, Information Process, Promoting Excellence, Fi-

nancials Tools, Partnerships, Education and Research), environment (Growth and Infrastructure, Gross-Boundary Coordination, Green Space, Natural Heritage Protection, Waste Diversion, Air, Water and Soil Protection, Resources Conversation, Health), social (Volunteerism, Community Participation, Inclusiveness, Leadership, Community Identity, Vitality, Cultural Heritage, Arts and Culture) and land-use planning (Housing Choice, Access, Community Design, Development Form, Land-Use Planning Tools, Reducing Barriers, Staff Capacity, Collaboration) (Municipal 2005). It is one of examples of conceptual framework of the regional development policy, as it is perceived by OECD. Just in case, it is also advisable to consider two main "profits" of urbanization (OECD 2007), which are the decrease of commuting times and the possibility of cheaper households living. Above mentioned processes can be perceived as positive externalities which do not possess the market characteristic.

In the Czech Republic, the attention is primarily aimed, regarding the disposable database, to analyses on the level of districts. One example is the analysis of the quality of life using the socio-economic deprivation indices which were constructed only for Moravian districts (Tomášková et al. 2007). Two methods were selected for creating the SESDI based on (1) Townsend index that is calculated as a sum of Z-scores of specific factors (INDEX1); (2) MATDEP and SOCDEP that use standardized score for each of the factors for maximal value of all areas; the final index reaches values in the interval from 0 to 8 (INDEX2). The GIS was used for presentation of deprivation of enumeration districts.

Mederly et al. (2004) covered with indicators of the quality of life and sustainable development. They tried using the tools of multivariate statistical analysis to develop an alternative indicator of the quality of life and sustainable development to commonly used indicators (HDP or Human Development Index). The so called Regional Index of the Quality of Life was constructed on the regional level which is based on three parts of life: a) precondition for long and healthy life, b) preconditions for creative life with sufficient education and c) preconditions for adequate living standard. The second indicator Index quality and sustainability of life characterized the development in the Czech Republic on the national level in the period 1990-2000. This indicator was compiled for four main areas of life: a) socio-political, b) social, c) economic and d) environmental area. On the global level, there has been constructed the Index of sustainable development for 179 countries in the world. Svobodová (2008) was interested in the general determination of the term of quality of life from views of different institutions, its relation to the standard of living and to the sustainable development. She categorized and measured the quality of life on the national and regional level. The aim of Charvát´s work (2009) "Quality of life in the border regions of the Czech Republic" is to reveal differences in the level of quality of life in different parts of the borderland. The main content category consists of issues of broad groups of indicators, indicator programs, a brief socioeconomic analysis and in particular to measure the quality of life. Measurements are carried out at

the level of districts (LAU1) using alternative quality of life measure. From the results, it is obvious that the areas with the best quality of life are České Budějovice, Zlín and Liberec. Conversely, very low quality of life can be found in the districts of Chomutov, Most and Teplice. The importance of the transport accessibility, especially in the countryside and peripheral regions, is described in the work of Květoň and Marada (2006).

Another view on the quality of life presents the work of the authors Janský and Létalová (2009), they have analysed life quality in the regions of the Czech Republic from the view of natural resources. Applied indicators of quality of life by the view of natural resources evaluate natural conditions, agricultural land areas and forested areas. In the area of protection of natural resources there are evaluated investments in protection of environment in total and also per 1 inhabitant, further are evaluated noninvestment protection of environment expenses, economic gains on the activities of the protection of environment, coefficient of ecological stability, production of communal (municipal) waste in total and per 1 inhabitant.

The quality of life and its appraisal is not complete without consideration of ecologic indicators. The availability of ecologic data on the level of regions is limited in conditions of the Czech Republic because the monitoring of emissions of CO^2, NO^2 and the other for example is realized only on the level of regions (NUTS3).

3. Methods

The level of economic and social development of Czech districts was assessed in two periods: in 2001 and after five years, in the period 2007 – 2008. The second period is longer with respect to the availability of comparable data. Following indicators were chosen for a comparison both the periods:

- average salary per capita in 2001 and 2005 (CZK)[4]
- crime rate (number of crimes per 100 inhabitants) in 2001 and 2007
- dependency ratio (number people above 65 to children to 15 years) 2001 and 2008
- road transport density (km/km^2) – 2001 and 2007
- net migration 2001 and 2008
- rate of unemployment in 2001 and 2007
- hospitals bed per 10 000 inhabitants in 2001 and 2007
- completed dwellings per 1 000 inhabitants in 2001 and 2008
- average amount of pensions (CZK) in 2001 and 2008

[4] Average salary per capita was available only for the year 2005.

The transport density in 2001 was derived by spatial tools in ArcGIS from the vector layer of Czech road net, while the density for the second period was calculated from the data in Statistical yearbooks of regions in 2007. With respect to absolute dissimilarity of these sources, it is necessary to take this fact into consideration in the results interpretation.

Cluster analysis (Aldenderfer - Blashfield 2006, Everitt - Dunn 2001, Romesburg 2004) was used to obtain rough typology of Czech districts in 2001 and in the second period 2007 - 2008. We also tried to classify the Czech districts according the changes of the used variables between the two studies periods. Various algorithms and metrics were tested and their results compared to find the best solution. Hierarchical clustering served as the first step of the classification. The aim of this step was to find a proper count of clusters. The clusters were identified by the *k*-means method. The computations were realized with the programs Statistica, R, and the results visualized in ArcGIS.

Prague was excluded from the analysis because it differs too much from other districts.

4. Results

The variables used for the classification in 2001 are correlated, as it can be seen in Table 1. For that reason, we used only the non correlated variables in cluster analysis but we studied the potential influence all of the variables on the socio-economic development of regions. We also tried to realize the classification by the use of Mahalonobis´ metrics which removes the influence of oversized correlation of variables. The appropriate number of clusters is five according the results of hierarchical clustering. Wards method gave us the best results.

Table 1. Correlation matrix 2001

	Dependency ratio	Average salary	Road density	Crime rate	Net migration	Rate of unemployment	Average pension	Hospitals beds	Dwellings completed
Dependency ratio	1,00	-0,00	**0,33**	-0,16	-0,15	-0,09	**0,27**	**0,24**	-0,05
Average salary		1,00	**0,59**	**0,71**	**0,25**	-0,06	0,03	**0,28**	0,06
Road density			1,00	**0,50**	**0,29**	-0,09	0,08	**0,27**	0,11
Crime rate				1,00	0,16	0,15	-0,21	0,29	-0,13
Net migration					1,00	**-0,30**	-0,14	**-0,39**	**0,59**
Rate of unemployment						1,00	**-0,34**	0,22	**-0,41**
Average pension							1,00	0,09	-0,01
Hospitals beds								1,00	**-0,25**
Dwellings completed									1,00

Marked correlations are statistically significant at $p<0,05$.
Source: own creation

The results of our attempt to create a typology of Czech districts according the level of their economic and social development in 2001 are apparent from Table 2. Variables, statistically significant for the inclusion of districts into clusters are: average salary, transport density and crime rate.

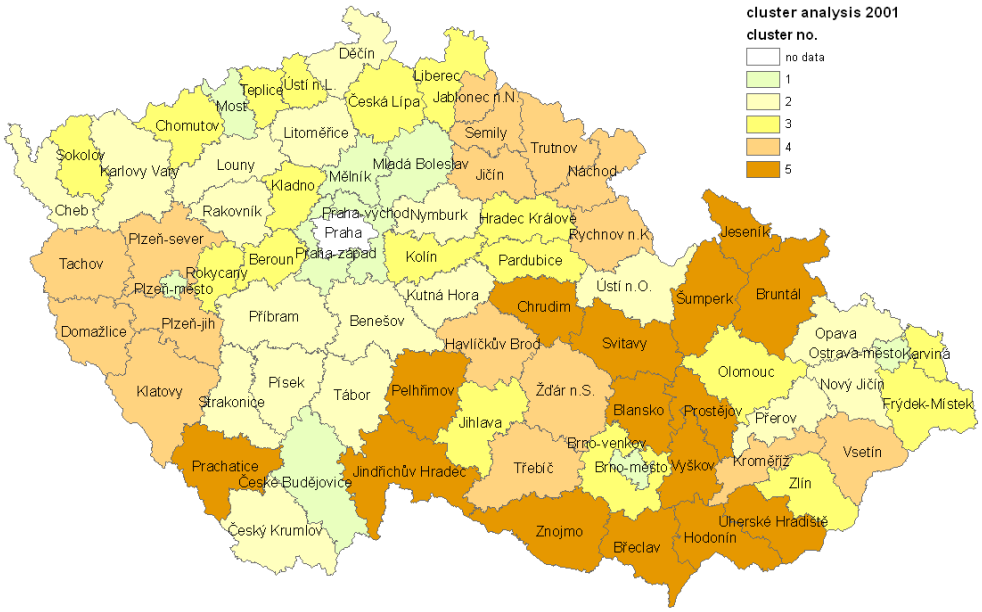
The inclusion of districts into clusters in 2001 was determined by the main forces of regional development: the average salary, transport density, the crime rate, the average pension and the level of health care. It is obvious that there exists a dependency among these variables (Table 1): transport density determines a development of activities in a region and consequently high average salaries, pensions, housing construction and also the quality of health care. An interesting finding is the positive correlation between the crime rate and the transport density. Is it possible to judge that the criminality (as partially technical-technologic phenomenon) would significantly eliminated without mobility in the 21th century? The membership of the districts into clusters is visible in Figure 1.

Table 2. Analysis of variance – 2001

variable	between group variability	df	within-group variability	df	F	p-value
Dependency ratio	563	4	13482	71	0,7408	0,567302
Average salary	84777160	4	7730815	71	194,6489	0,000000
Transport density	1	4	1	71	11,2952	0,000000
Crime rate	30	4	30	71	17,9829	0,000000
Net migration	75	4	1166	71	1,1416	0,344082
Rate of unemploy-ment	10	4	1028	71	0,1698	0,953130
Average pension	10391060	4	399471,2	71	461,7135	0,000000
Hospitals beds	11044	4	44555,1	71	4,3996	0,003104
Dwellings completed	1	4	89,0	71	0,2176	0,927804

Source: own creation

Figure 1. Districts clusters – 2001



Source: own creation

Table 3. cluster comparison (means) – 2001

Variable	cluster				
	1	2	3	4	5
Dependency ratio	96,40	93,57	96,12	96,56	96,23
Average salary	15541,67	12871,89	13829,94	12752,44	12028,33
Road density	0,91	0,60	0,70	0,61	0,56
Crime rate	4,23	2,84	3,02	2,26	1,96
Net migration	2,88	0,15	0,26	-0,06	-0,25
Rate of unemployment	8,98	9,68	10,18	6,94	9,81
Average pension	6355,33	6083,28	6475,56	6859,88	6269,87
Hospitals beds	77,05	50,05	63,96	42,10	52,62
Dwellings completed	2,52	2,10	2,14	2,17	2,63

Source: own creation

Table 3 confirms mild differences among the Czech districts in the variables which are not significant for including into cluster – variables which cannot be marked as the main driving forces for the regional development in 2001: population age structure, housing construction and rate of unemployment.

Cluster 1 consists of the metropolitan districts concentrating a wide range of activities with the largest cities in the Czech Republic. They can by characterized by the highest salaries, high transport accessibility, high level of crime, high migration gains, relatively lower rates of unemployment, average pensions, high level of health care and above-average housing construction. We can mark these districts as the development cores. The districts with under-average dependency ratio, under-average salaries, worse transport accessibility, average criminality, moderate migrations increases, slightly above-average unemployment, lower pensions and weaker health facilities constitute cluster 2. Cluster 3 consists of the districts which are distinguished by moderately above-average salaries and pensions, relatively high road density, above-average criminality, small migration gains, high level of unemployment and relatively good health care.

Relatively low salaries, average transport accessibility, lower criminality, migration decreases, under-average unemployment and also worse accessibility of health care are the features of the districts in the cluster 4. Cluster 5 is formed by the districts with low salaries and lower pensions, bad transport accessibility, low crime rates, the highest migration decreases, the highest rate of unemployment, above-average unemployment, the average level of health care.

Correlation analysis confirmed that the level or the rate of mutual influencing among studied variables demonstrated the high of average salaries in the relationship with net migration and also with crime rate. Another relatively important depend-

ency has been manifested between average salaries and the transport density. Considering the crime rate, it is not fully unique manifestation of a synergic effect which results from a sense of security of inhabitants in particular districts and the level of their public order, experience in the work of the Police of the Czech Republic and also in negative phenomena in its work, and last (but) not least in the state of prevention. The migration oneself is then, according the Strategy of Regional Development of the Czech Republic for the period 2007 – 2013 (Strategy of the Regional Development 2009) influenced by (up to present) not much developed housing market which creates a barrier to a potential development of general migration mobility. The migration flows are primarily realized as a temporary migration and commuting, changes of permanent residences are less often.

The values of the variables for the period 2007 – 2008 were much more correlated. The data in Table 4 confirm a similar pattern of mutually influencing the variables, contrary to the period of 2001, the influence of unemployment on the other variables has been increased.

Table 4. Correlation matrix 2007 – 2008

	Dependency ratio	Average salary	Road density	Crime r	Net migration	Rate of unemployment	Average pension	Hospitals beds	Dwellings completed
Dependency ratio	1,00	**-0,46**	**0,28**	**-0,25**	-0,18	-0,22	**0,36**	0,20	**-0,51**
Average salary		1,00	**0,26**	**-0,25**	**0,50**	**0,44**	**-0,45**	**-0,28**	**0,65**
Road density			1,00	**-0,27**	**0,63**	**0,40**	0,12	**0,37**	0,08
Crime rate				1,00	-0,08	0,21	0,13	**0,36**	**-0,26**
Net migration					1,00	**0,58**	-0,10	0,22	**0,38**
Rate of unemployment						1,00	-0,10	0,21	**0,24**
Average pension							1,00	**0,38**	-0,16
Hospitals beds								1,00	**-0,23**
Dwellings completed									1,00

Marked correlations are statistically significant at p<0,05.
Source: own creation

Again, using the Mahalanobis´ metrics in cluster analysis, we have found that the differentiation of economic and social development is statistically significantly determined by average salary, road density, criminality, the average pension and the

level of health care and also the more and more increasing importance of migration flows and the unemployment in the period 2007 – 2008. The allocation of the districts into clusters denotes Figure 2.

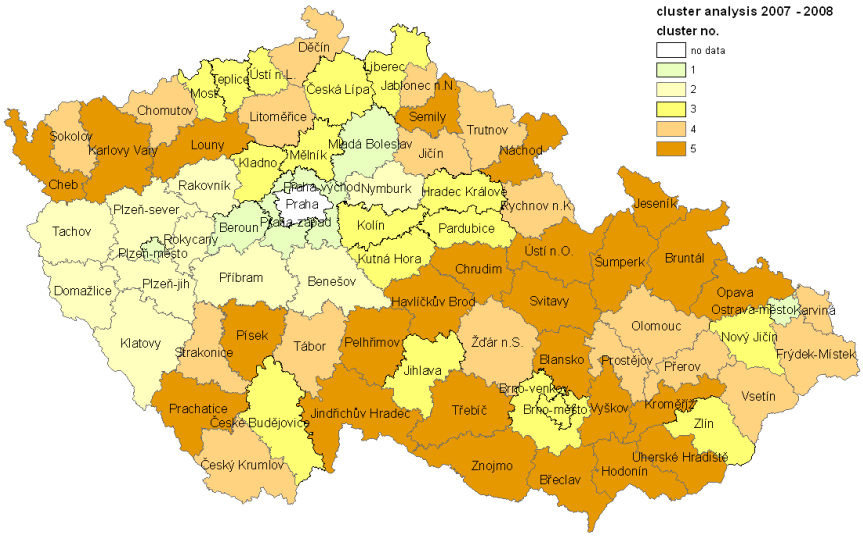
Table 5. Analysis of variance – 2007 – 2008

	between group variability	df	within-group variability	df	F	p-value
Rate of unemployment	91	4	592	71	2,7255	0,035935
Net migration	2605	4	4240	71	10,9071	0,000001
Crime rate	26	4	43	71	10,6880	0,000001
Dependency ratio	231	4	11454	71	0,3585	0,837281
Average salary	136436600	4	12483620	71	193,9941	0,000000
Road density	2	4	2	71	14,3283	0,000000
Average pension	5107522	4	4138119	71	21,9081	0,000000
Hospitals beds	9566	4	41609	71	4,0809	0,004925
Dwellings completed	60	4	171	71	6,2367	0,000233

Source: own creation

The increasing importance of the net migration and unemployment proves the increasing differentiation among Czech districts.

Figure 2. Districts clusters – 2007 – 2008



Source: own creation

Table 6. cluster comparison (means) – 2007 – 2008

Variable	cluster				
	1	2	3	4	5
Rate of unemployment	4,33	5,21	7,32	7,79	7,36
Net migration	24,12	12,30	7,78	4,52	2,76
Crime rate	3,83	2,27	3,17	2,53	1,96
Dependency ratio	97,00	101,80	102,24	99,94	103,00
Average salary	19867,17	16080,10	17829,12	16232,33	15307,80
Road density	1,28	0,70	0,81	0,73	0,66
Average pension	8857,50	8488,70	9192,06	9303,33	9172,96
Hospitals beds	53,25	27,11	65,96	50,12	52,17
Dwellings completed	5,08	1,74	2,03	1,80	1,70

Source: own creation

Keeping with the increasing importance of migration and unemployment, Table 6 confirms the increasing differentiation of the Czech districts not in the values of these variables. The typology of the districts has changed during the studied period in the following way: cluster 1 consists of the districts influenced directly by Prague or some other large cities (Plzeň, Ostrava) – the centres of various activities with the tradition of industry and tertiary sector. The typical characteristics of are: relatively convenient age structure, the highest salaries, the best transport accessibility, high criminality, the highest migration increases and low rate of unemployment. The districts distinguishing by demographic ageing, slightly under-average salaries, average road density, under-average crime rates and also large migration increases, and low rate of unemployment constitute the cluster 2. Cluster 3 consists of the districts with verifiable demographic ageing, above-average salaries, relatively good transport accessibility and higher criminality, slightly above-average migration increases and above-average rate of unemployment. The remaining clusters 4 and 5 can be characterized by under-average salaries, lower transport accessibility, lower criminality and low migration gains. While the cluster 4 has a relatively good age structure, the cluster 5 suffers from the demographic ageing. An interesting finding is that whilst these two clusters show low average salaries, they show concurrently the highest average pensions.

The final step of our analysis represents an attempt to classify the Czech districts according the changes of the followed variables in the period 2001 – 2008. In this case, the districts were classified into four groups.

Table 7. Analysis of variance – rates of change 2001 – 2008

variable	between group variability	df	within-group variability	df	F	p-value
Average salary	0,001	3	0,1644	72	0,1406	0,935331
Road density	0,026	3	1,6493	72	0,3834	0,765258
Crime rate	0,048	3	1,2432	72	0,9283	0,431612
Average pension	0,048	3	0,5652	72	2,0259	0,117864
Dependency ratio	0,060	3	0,7182	72	1,9897	0,123152
Rate of unemployment	0,154	3	1,3333	72	2,7725	0,047588
Net migration	3306,516	3	375,7849	72	211,1750	0,000000

Source: own creation

As can be seen in the Table 7, the typology of the Czech districts on the base of the changes of variables mostly determining economic and social development of the regions and so the quality of life (average salary and pensions, transport accessibility, crime rates, age structure, rate of unemployment and migration), the Czech districts can be divided into five distinct groups. Only two variables determined significantly the groups in this case: rate of unemployment and migration.

Table 8. cluster comparison (means) – rate of change 2001 – 2008

variable	cluster			
	1	2	3	4
Average salary	1,25	1,25	1,26	1,25
Road density	1,18	1,19	1,15	1,15
Crime rate	0,86	0,91	0,95	0,92
Average pension	1,41	1,40	1,42	1,46
Dependency ratio	1,01	1,07	1,05	1,10
Rate of unemployment	0,61	0,76	0,78	0,77
Net migration[5]	24,67	10,71	3,88	0,61

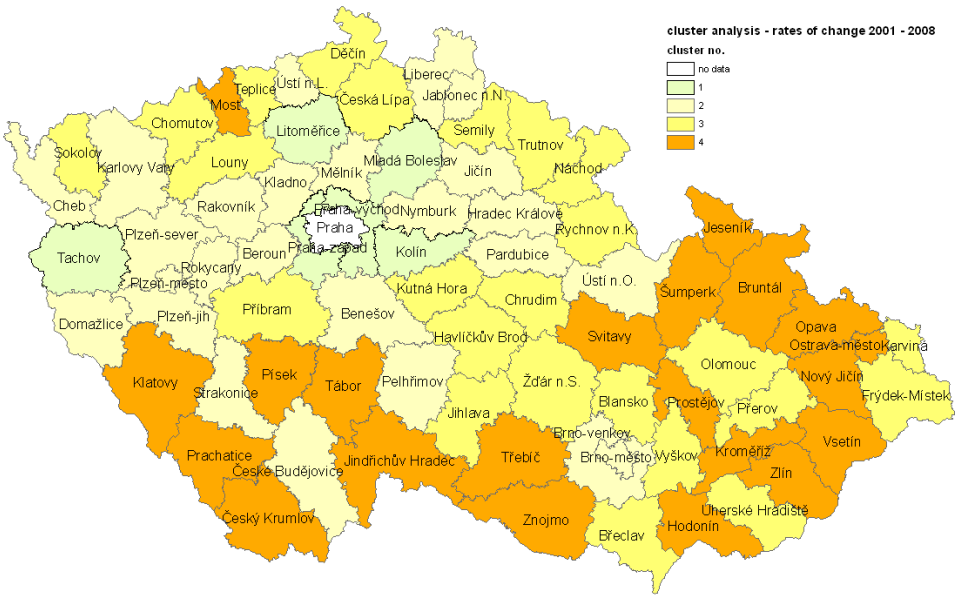
Source: own creation

Cluster 1 consists of the districts which can be characterized by the increase of transport accessibility, reduction of the criminality, low rate of unemployment and high migration gains – some adjacent districts influenced by Prague. Surprising fact is the including of district Tachov into this cluster. The districts forming the cluster 2 show similar characteristics, they differ only from the amount of migration.

[5] The change of migration was computed as the difference between the net migration in 2008 and 2001.

Most of these districts create a whorl around Prague, the other represent regions with various traditional activities. Cluster 3 consists of districts which noticed the smallest decrease of unemployment and also the migration gains are small here. The last group of districts suffers from the unemployment and also migration decreases. The peripheral or transformed former industrial regions constitute this cluster.

Figure 3. District clusters – rate of change 2001 – 2008



Source: own creation

The comparison of the values of the followed variables in the two periods (tables 2 and 4) shows the increasing differentiation of the socio-economic conditions and thus the quality of life in the Czech districts. It is also obvious that almost a half of districts changed their position in groups in the followed period. The assessment of this change is problematic. The peripheral or restructured regions experienced the biggest changes in the studied period.

5. Conclusions

In both the studied periods, the differentiation of the quality of life in the Czech districts was significantly determined by the average salary, transport accessibility and

criminality. The increasing role of migration flows follows from the analysis. Migration can be perceived as the mediating influence of economic and social factors. The Czech Republic suffers from the demographic ageing (like most of the developed European countries) which is partially eliminated by migration.

The realized comparison of the Czech regions (LAU1) points out changes of all the examined regions in time. It is obvious that the generalization of the results from the analysis is not possible because of the limited extent of the used variables. In spite of this, further to the foreign experience, the extension of the spectrum of examined variables enables us to obtain information which is needed for decision processes and which will become the starting points for the management aimed at the harmonic development of particular regions followed by a positive impact on the quality of life in municipalities.

The analysis described above should be perceived as a starting point to a deeper investigation of the situation of the regional development processes in the Czech Republic on the various levels of hierarchy relating to the quality of life. The future considerations will relate to the choice of better indicators which is influenced by the availability of data from various sources. The use of more appropriate statistical methods and solving some methodological problems seems to be the main task for the further analysis.

References

Aldenderfer, M., S. - Blashifield, R., K. 2006: *Cluster Analysis*, SAGE University Press, 88. p.

Arthur, J. L. 2008: Family Size and Quality of Life Nexus, Case of The Sunyani Municipality, 3rd National Conference on Population, Health and Environment, Conservation International, Tagaytay City, Philippines, Ghana.

Baaske, W. E. - Filzmoser, P. - Mader, W. - Wieser, R. 2009: *Agriculture as a success factor for municipalities*. Jahrbuch der Österreichischen Gesellschaft für Agrarökonomie, 18/1., 21-30. p.

Burrett, J. 2009: *Quality of Life in Canadian Municipalities*. In Sirgy, M. J. - Phillips, R. - Rahtz, D. R.: Communnity Quality-of-Life Indicators: Best Cases III. Springer, Netherlands, pp. 155-165.

Charvát, O. 2009: *Quality of life in border regions of the Czech Republic*. Faculty of Economics and Administration, Masaryk University, Czech Republic.

Everitt, B. S. – Dunn, G. 2001: *Cluster Analysis*. Edward Arnold a member of the Hodder Headline Group, Bristol, 237 p.

González, E. - Cárcaba, A.. - Ventura, J. - Garcia, J. 2009: *Measuring duality of life in Spanish municipalities*. Governing Regional Development Policy. OECD 2009.

International Transport, Urban Form and Economic Growth 2007. OECD/ECMT.

Jánský, J. - Létalová, P. 2009: *Analysis of life quality in the regions of the Czech Republic*. In ICABR. s. 1. (http://www.icabr.com/fullpapers/L%E9talov%E1%20Petra,%20Jaroslav%20J%E1nsk%FD%20(Analysis%20of%20life%20quality%20in%20the%20regions%20of%20the%20Czech%20Republic).pdf)

Květoň, V. - Marada, M. 2006: *Význam dopravní obslužnosti v rozvoji venkovských oblastí*. Sborník z konference Venkov je náš svět, Sociologická laboratoř ČZU Praha.

Mederly, P. - Topercer, J. - Nováček, P. 2004: *Indikátory kvality života a udržitelného rozvoje. Kvantitativní, vícerozměrný a variantní přístup*. Praha, CESES, 117s.

Municipal Quality of Life for Economic Prosperity. 2005: Ministry of Municipal Affairs and Housing, Ontario.

Romesburg, H. C. 2004: *Cluster Analysis for Researchers*, North Carolina, Lulu, 340. p.

Royuela, V. - Suriňach, J. - Reyes, M. 2003: *Measuring Quality of Life in Small Areas Over Different Periods of Time*. Social Indicators Research, Vol. 64. Issue 1., October 2003, 51- 74. p.

Scipioni, A. - Mazzi, A. - Fornasiero, L. - Schiavon, A. 2004: *The duality of life in Padova municipality and the ISO 14031 model – an experience to measuring urban sustainability by context indicators and performance indicators*. Engineering Management Conference, 2004, Proceedings, 2004 IEEE International, Vol. 1, pp. 139-143.

Skotnicka-Zasadzien, B. - Midor, K. 2009: *The effect of balanced development idea on the duality of life of inhabitants in selected municipalities o the Silesian province*. Qual Quant.

Strategy of the Regional Development of the Czech Republic 2007-2013. [Cit 2009-11-10] on-line http://www.dhv.cz/regstrat/

Svobodová, A. 2008: *Regional evaluation of quality of life*. Thesis, Faculty of Economics and Administration, Masaryk University, Czech Republic.

Tomášková, H. - Šlachtová, H. - Polaufová, P. - Šplíchalová, A**.** 2007: *Socioeconomic deprivation index of the Czech Republic - construction*. Central European Journal of Public Health : International Conference and Central and Eastern European Chapter Meeting of International Society for Environmental Epidemiology. vol. 15/51, Suppl., 30-31. p.

Competitiveness of Human Capital in the countries of the Visegrád Group (V4)

Lampertné Akócsi Ildikó[1]

The analysis of competitiveness has become an irreplaceable tool of economic studies by now. Professional literature has several approaches to the definition of competitiveness. The most accepted of these is the pyramid model of competitiveness. Each of the success factors featured in the model (including human capital) plays an important role in influencing the future competitiveness of an economy. In my essay I look at the role of human resources in shaping territorial competitiveness in the group of the Visegrád countries.

Keywords: competitiveness, human capital, territorial competitiveness, cluster analysis, factor analysis, typifying of regions, competitiveness cluster

1. Introduction

The essay focuses on the NUTS2 level spatial units of the Visegrád Countries. The name comes from the north Hungarian town of Visegrád, which was the place of a historical meeting of the Central European kings in the 14th century. In the past the primary objective of the countries of the group was the accession (integration) of the Visegrád countries into the Euro-Atlantic structure. After achieving this goal, the foreign policy of the V4 countries significantly expanded and now covers other fields as well. The Visegrád group strives for the strengthening of the identity of Central Europe within the European Union and supports the regional cooperation of the countries of Central Europe.

The aim of my research was to explore correlations between the competitiveness of this group of countries and the development level of human resources living there. These countries have been cooperating for centuries in the field of economy, culture, politics and trade. The research was meant to answer the following questions:

- Do the multi-dimensional relations within the group of the countries have common features determining their competitiveness?
- Are human resources just as dominant for all regions with different culture and mother language?

[1] assistant professor of College of Dunaújváros, lampaki@mail.duf.hu

- Are there cross-border similarities or maybe differences despite the belonging together?

2. Competition of spatial units

The concept of competitiveness defined in the 6th regional report of the EU, which is the most widespread and most accepted definition, says the following: "the ability of companies, industries, regions, nations and supranational regions to generate, while being exposed to international competition, relatively high income and employment levels" (Lengyel 2003) In other words, the goal of the respective spatial units is to acquire "abilities" by which they are able to increase the amount of incomes generated in the given territory, their employment level and thereby the living standards of the inhabitants living there.

It means that there is competition among the territories (countries, regions, micro-regions and settlements). There is competition for the tenderable resources, for investors, and for all available resources, including not last the human resources (Tóth et al. 2008).

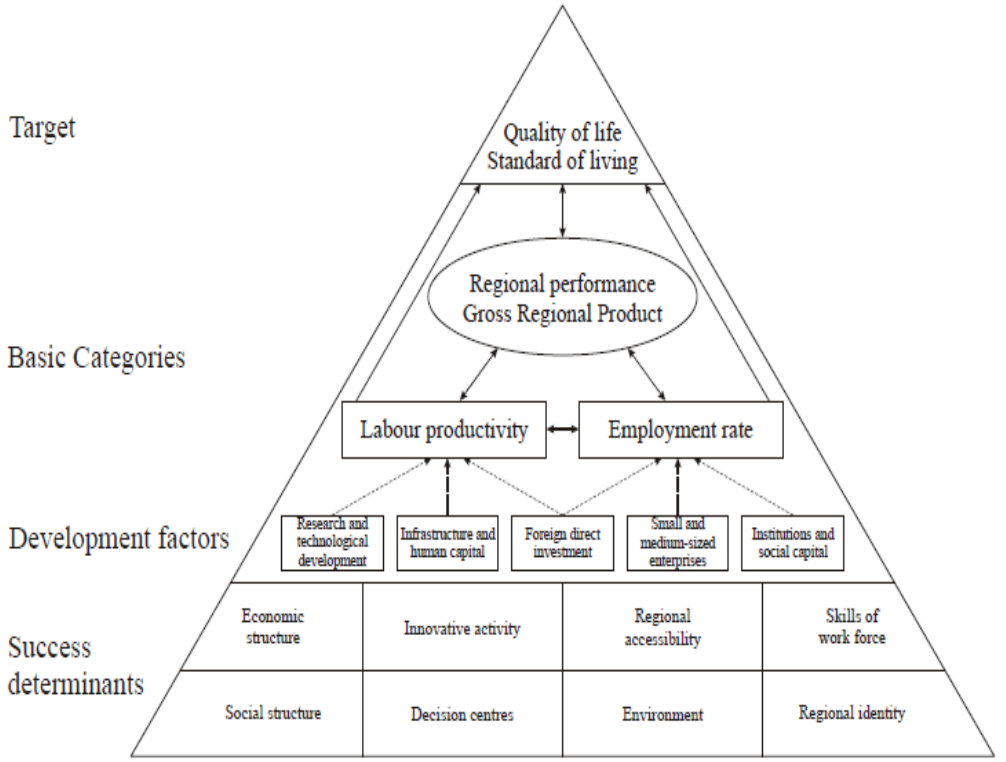
Regional competitiveness can be measured by three basic categories: work productivity, employment rate and incomes generated (GDP). The factors influencing these indices can be divided into two groups. One group is made by the basic factors that directly and usually in the short run influence incomes, employment and productivity. The development of these factors can strikingly improve the competitiveness of a region (Lengyel 2006). Basic factors are research and development, infrastructure and human capital, external direct investments, small and medium-sized enterprises, institutions and social capital. The indirectly working factors of success improve competitiveness in the long run, through indirect impacts. These factors are economic structure, innovation culture, regional accessibility, the preparedness of the labour force, social structure, decision-making centres, quality of the environment and the social cohesion of the region. The transformation of the basic factors, including then structure of human capital, is a direct tool for the improvement of competitiveness in the short-run.

3. The role of human capital in the development of the competitiveness

A basic question of regional science is the classification of regions. The typifying of regions has several methods, but each includes the development level of human capital as a basic criterion. In other words, one of the main powers of spatial organisation is knowledge. (Tóth 2009) The ability for the creation, acquisition and adaptation of knowledge determines the innovation capacities of the regions, and thereby competitiveness. Especially in the developed countries knowledge is an

outstanding power of spatial organisation. Knowledge is the main feature of the human resources. The neoclassical theory already recognised human capital as a key factor in economic development that explains the disparities among the regions. The theory differentiates between materialised and non-materialised technical development. Materialised technical development is an innovation investment itself, while non-materialised development is the provision of qualified workforce able to apply the technology. It is non-materialised technical development that is the basic source of disparities among the regions. The regions that have a high level of knowledge specialise themselves in activities such as R & D, scientific research and innovation. Regions lass abundant in human resources will specialise themselves in routine activities, the technology of which is globally available. This process further deepens the disparities among the regions. In the pyramid model of competitiveness (Lengyel 2006) human capital is among the basic factors, i.e. among the driving forces outside the economy that explain the competitiveness of a region in the short run. (Figure 1)

Figure 1. The pyramid model of regional competitiveness



Source: Lengyel (2003)

The basic hypothesis of the research is that human capital is a dominant factor in determining the competitiveness of regions. Global competition in the 21st century is not for goods or services or capital; it is for humans, for intellectual capital. It means that those regions will have a competitive advantage and develop that concentrate the advanced human resources.

4. The data

In my essay I look at the NUTS2 level territorial units of the 4 Visegrád countries, i.e. 8 regions in the Czech Republic, 7 regions in Hungary, 16 regions in Poland and 4 regions in Slovakia (Appendix 1.).

The research is built on the analysis of seven indices. The first index is the most frequently used measure of competitiveness, Gross Domestic Product per capita, the remaining indices focus on some aspects of human resources.

The indices examined were as follows:

- GDP per capity in Euro
- Number of higher education students in per cent of the total population
- Number of R & D employment in per cent of total employment
- Unemployment rate (in per cent)
- Number of employment in the field of science and technology in per cent of the total active population
- Employment rate (in per cent)
- Number of pupils and students in per cent of the total population

The data are so-called hard data that come from the databank available on the website of the Eurostat. I calculated relative indicators from the data of 2006, so that the differences coming from the size of the respective regions should not influence the research findings. The potential mistakes caused by the different units and magnitude of the data were handled by the method of standardisation.

5. The methods

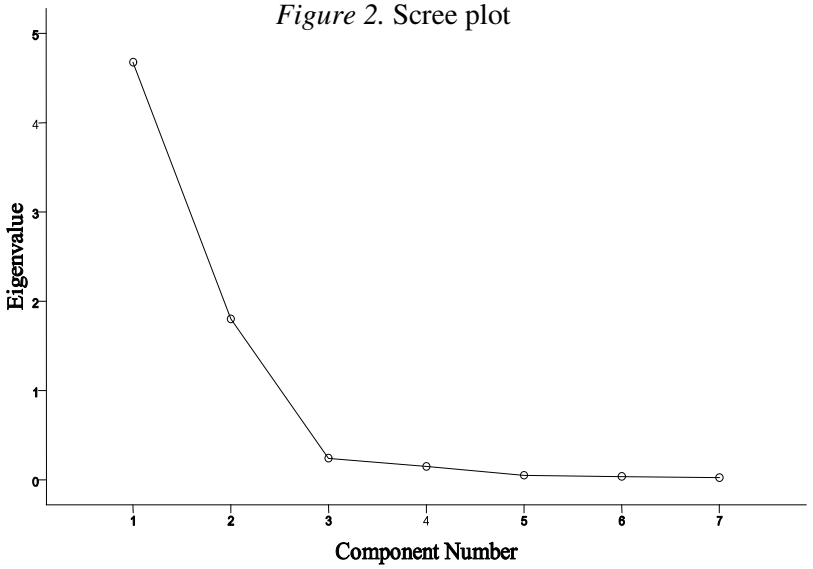
The research operated, in addition to the tools of descriptive statistics, with the methods of principal component analysis and cluster analysis. Principal component analysis is a special case of factor analysis. Factor analysis is a tool that can be successfully used for the explanation of a large number of variables with a smaller number of uncorrelated latent variables. Principal component analysis can be used for the preservation of the mass of information accumulated in the variables without a major loss, in a smaller number of uncorrelated variables, principal components (Kovács – Lukovics 2006). This method is good for making a statistical analysis in a

transformed smaller dimension space without losing important information. It can be used successfully if we have a large number of variables, strongly interrelated stochastically, that carry redundant information (Ketskeméty - Izsó 2005). Cluster analysis is a multi-variable statistical method, a segmentation of data by which blocks of data can be ordered into homogenous groups. These groups are called clusters. The main objective of cluster analysis is to classify the examined cases into relatively homogeneous groups, using of selected data, in a way that the observation units in a respective cluster should resemble each other but be different from the members of other groups.

6. The research findings

By using principal component analysis I ordered the listed indices into two well separable factors. The reasonability of ordering the indices into two principal components is justified by the values in Figure 2 and of the communality. The principal component analysis computed communality for each variable, which is actually a multiple coefficient of determination. The multiple of coefficient correlation that can be calculated from this shows the closeness of the correlation between the principal components as explanatory variables and original components as dependent variables. The communality of all seven indices is over 0.8. This means that the two principal components quite well aggregate the information content of the seven indices.

Figure 2. Scree plot

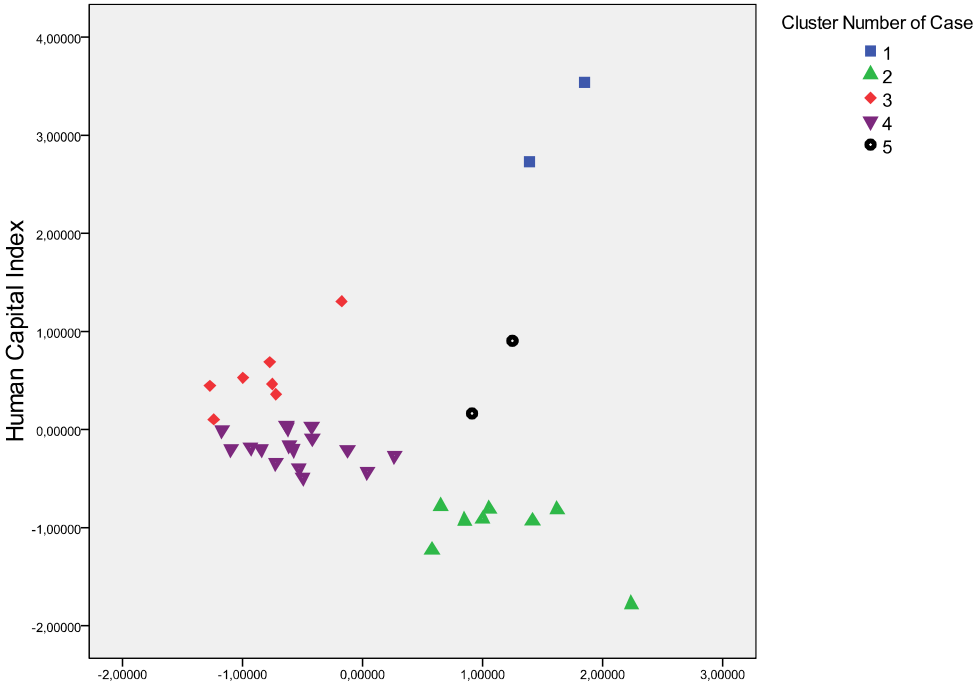


Source: Own calculation based on data from Eurostat's with statistical softver SPSS 17

The Eigenvalue of the first principal component is 4.678 and explains 66.822% of all variances of the original variables. This factor contains the proportion of higher education students, the proportion of R & D employees, the number of pupils and students in per cent of the total population, and the number of those employed in science and technology in per cent of the total active population. I called this factor Human Capital Development Index. The Eigenvalue of the second principal component is 1.803 and it explains 25.763% of the total variance of the original variables. Thus these two factors together determine 92.585 of the variance of the original variables. The indicators of the second principal component are employment and unemployment rate and the value of GDP per capita. I named this factor the Competitiveness index, because two of its components are measurable base categories of competitiveness.

In the next step I put the 35 territorial units into clusters based on the two principal components. The method I used was k-means clustering. On the basis of the analyses run, the creation of two clusters or five clusters is also a meaningful solution. First I looked at the five-cluster solution (see Figure 3.).

Figure 3. The Clusters



Source: Own calculation based on data from Eurostat's with statistical sotver SPSS 17

Cluster one includes the capital city of the Czech Republic and the region around the capital city of Slovakia. The average GDP per capita in the cluster is € 21,250. The two regions are characterised by low unemployment and a high employment rate. Twenty-nine per cent of their population takes part in some form of full time education, 3.05% of them in higher education. A high proportion of the economically active population works in the fields of science and technology, and research and development. In all indices examined, the values of Prague are better, i.e. incomes generated are higher, as is employment level, and the human resources of this area are more qualified.

Cluster two involves eight regions, which are characterised by a low development level of human resources and relatively high competitiveness. Among the members of the cluster we find two Hungarian, one Slovakian and five Czech regions. It is interesting to observe the location of the regions. The eight regions surround in a semi-circle, in an arch, Austria and the developed regions of South Moravia and Bratislava. The average GDP per capita in this cluster is € 9,050. The relatively high employment level has a standard deviation in a no more than 6.3% interval. The differences of the unemployment rates are much bigger among the regions that make this cluster. The proportion of higher education students is very low, as is the percentage of those employed in the field of R & D. Among the members of the cluster, the competitiveness of Middle Bohemia (Stredni Cechy) is outstanding, but the proportion of pupils and students is very low, so it is the region with the least developed human resources. The reason for this is that the region is situated around Prague. The capital city as a knowledge centre concentrates the educational and research functions, and the highly skilled labour force living here provides one of the most important factors of competitiveness, human capital, for the neighbour region as well.

Cluster three has low competitiveness regions with medium developed human resources. This cluster concentrates seven regions from Poland. The average GDP per capita in this cluster is € 7,371; unemployment rate on the average is 14.56%, with a standard deviation at an only 5% interval. The number of pupils is high, but only 1.07% of the population are higher education students. The proportion of those employed in R & D sector and in science and technology is almost the same as the average of the regions in cluster five.

Cluster four has regions with weak competitiveness, in which the development level of human capital is low. This cluster involves one Czech, two Slovakian, four Hungarian and nine Polish regions. The regions that belong to this cluster lag in all examined indices behind the average values of the other clusters. The cluster is generally characterised by a high unemployment level, but there are significant differences across the individual regions in this respect. The lowest unemployment rates are recorded in the South Great Plain region in Hungary, while the highest unemployment strikes East Slovakia (Vychodne Slovensko). The proportion of higher

education students is low; the smallest number of students study in East Slovakia and the Lublin Voivodship in Poland (Lubelskie).

Cluster five has two regions as well, the region of Central Hungary and South Moravia in the Czech Republic. These two regions are characterised by a medium competitiveness and medium developed human resources. In this cluster the average GDP per capita is 49.1% lower, and average unemployment rate is 64.9% higher than in the previous cluster; also, the average rate of employment lags somewhat behind that. The rate of pupils and students is almost the same as in the previous cluster, but only 1.4% of the population is students in higher education. The proportion of those employed in science and technology, and research and development is approximately half of the values of the previous cluster. (Table 1.)

At the designation of two clusters, the capital cities and the regions surrounding them were put into the same group. In this cluster there is a positive correlation between the development level of human capital and competitiveness, while this correlation is negative in all other regions. The first cluster features an average GDP per capita value of € 17,175. Of the total population, 27.10% are pupils and students, of which the proportion of higher education students from all inhabitants is 2.28% on the average; 3.06% of all employees work in research and development, 46.15% of the active population are employed in high-tech sectors, unemployment rate is 6.20% and employment rate is 65.73%. The regions in the other cluster have an average GDP per capita of € 7,087.10. Of the total population of the cluster, 22.75% are pupils and students, of which the proportion of higher education students is only 0.81%. In R & D sector 0.88% of the population, in science and technology 27.75% of the population is employed. Unemployment rate is high (11.78%), while the average value of employment rate is 56.85% (Figure 4.).

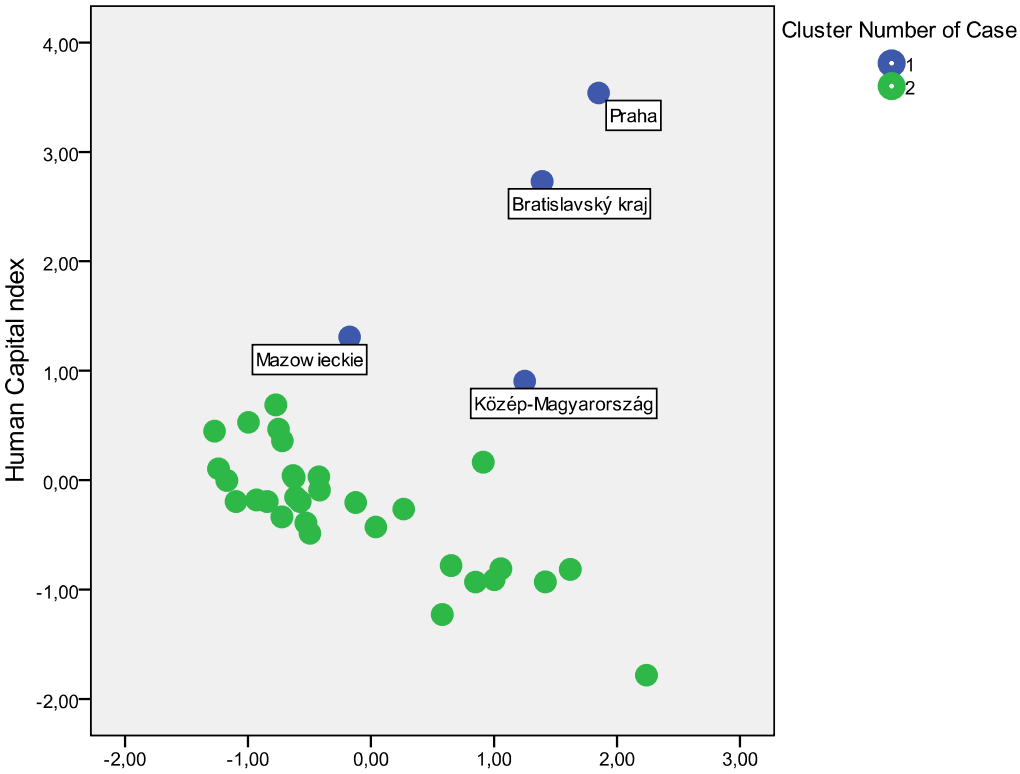
Table 1. Cluster membership of the regions

NUTS 2	Region	Cluster
CZ01	Praha	1
SK01	Bratislavský kraj	1
CZ02	Střední Cechy	2
CZ03	Jihozápad	2
CZ04	Severozápad	2
CZ05	Severovýchod	2
CZ07	Střední Morava	2
HU21	Közép-Dunántúl	2
HU22	Nyugat-Dunántúl	2
SK02	Západné Slovensko	2
PL12	Mazowieckie	3
PL21	Malopolskie	3
PL41	Wielkopolskie	3
PL42	Zachodniopomorskie	3
PL51	Dolnoslaskie	3
PL62	Warminsko-Mazurskie	3
PL63	Pomorskie	3
CZ08	Moravskoslezsko	4
HU23	Dél-Dunántúl	4
HU31	Észak-Magyarország	4
HU32	Észak-Alföld	4
HU33	Dél-Alföld	4
PL11	Lódzkie	4
PL22	Slaskie	4
PL31	Lubelskie	4
PL32	Podkarpackie	4
PL33	Swietokrzyskie	4
PL34	Podlaskie	4
PL43	Lubuskie	4
PL52	Opolskie	4

PL61	Kujawsko-Pomorskie	4
SK03	Stredné Slovensko	4
SK04	Východné Slovensko	4
CZ06	Jihovýchod	5
HU10	Közép-Magyarország	5

Source: Own calculation

Figure 4. Two clusters int eh countries of Visegrád



Source: Own calculation based on data from Eurostat's with statistical sotver SPSS 17

7. Summary

When classifying the regions of the Visegrád countries into five clusters we can see an axis along which the competitiveness and the development level of human capital of the regions improves from east to west. This classification justifies that the human resources do not determine the competitiveness of regions to the same extent, but this

fact is irrespective of the individual countries. In all four countries we find highly competitive regions and regions lagging in competitiveness, with human resources of different endowments.

If we order the regions into two clusters, we can see a basic identity. The regions around the capital cities, with their advanced economies, entertainment, cultural, career and other possibilities attract and keep intellectual capital, gaining thereby a considerable competitive advantage. The biggest difference between the two clusters can be seen in R & D sector and in the proportion of those employed in the sector of science and technology. It is clear that in regions where knowledge is concentrated, high-tech industry appears and the regional incomes increase. Between the school education indices of the two clusters there are no such significant disparities as in the case of employment in research and development, and science and technology. This justifies the assumption that in these regions one of the most significant competitiveness factors is human capital. The negative correlation between the development level of human capital and competitiveness is an interesting phenomenon that requires further researches.

Appendix 1. Regions in the countries of Visegrád

Code NUTS 2	**Name of Region**
CZ01	Praha
CZ02	Střední Cechy
CZ03	Jihozápad
CZ04	Severozápad
CZ05	Severovýchod
CZ06	Jihovýchod
CZ07	Střední Morava
CZ08	Moravskoslezsko
HU10	Közép-Magyarország
HU21	Közép-Dunántúl
HU22	Nyugat-Dunántúl
HU23	Dél-Dunántúl
HU31	Észak-Magyarország
HU32	Észak-Alföld
HU33	Dél-Alföld
PL11	Lódzkie
PL12	Mazowieckie
PL21	Malopolskie
PL22	Slaskie
PL31	Lubelskie
PL32	Podkarpackie
PL33	Swietokrzyskie
PL34	Podlaskie
PL41	Wielkopolskie
PL42	Zachodniopomorskie
PL43	Lubuskie
PL51	Dolnoslaskie
PL52	Opolskie
PL61	Kujawsko-Pomorskie
PL62	Warminsko-Mazurskie
PL63	Pomorskie
SK01	Bratislavský kraj
SK02	Západné Slovensko
SK03	Stredné Slovensko
SK04	Východné Slovensko

Source: Own calculation

References

KETSKEMÉTY, L. – IZSÓ, L. 2005: *Bevezetés az SPSS programrendszerbe, Módszertani útmutató és feladatgyűjtemény statisztikai elemzésekhez.* ELTE Eötvös Kiadó, Budapest

KOVÁCS, P. – LUKOVICS, M. 2006: Classifying Hungarian Sub-regions by their Competitiveness.“Globalization Impact on Regional and Urban Statistics” SCORUS 25th Conference on Urban and Regional Statistics and Research, Wroclaw.

LENGYEL, I. 2006: A regionális versenyképesség értelmezése és piramismodellje, *Területi Statisztika 9.* 46, 2, 131-147.

LENGYEL, I. 2003: A regionális versenyképesség értelmezése és piramismodellje, Verseny és területi fejlődés, JATEPress, Szeged.

TÓTH, T. – PUPOS, T. – GÖRÖG, M. 2008: Területi tervezés és programozás. Szerk.: Tóth T. Szent István Egyetemi Kiadó Gödöllő.

TÓTH, T. 2009 szerk.: *Regionális gazdaságfejlesztés és menedzsment.* SzIE Gödöllő Gazdaság- és Társadalomtudományi Kar, Gödöllő, 37.-50., 79.-126.

Bibliography

COOKE, P. 2001: Knowledge Economies. Clusters, learning and cooperative advantage. Routledge, London.

DTI: UK Productivity and Competitiveness Indicators 2003: DTI Economics Paper. 2003, 6.

LENGYEL, I. – RECHNITZER, J. 2004: *Regionális gazdaságtan.* Dialóg Campus Kiadó, Budapest-Pécs.

LUKOVICS, M. 2004: Regionális gazdaságfejlesztés: eltérő fejlettségű megyék versenyképességének összehasonlító elemzése. *Tér és Társadalom.* 2004, 4.

MALIZIA, E. – FESER, E. J. 1999: Understanding Local Economic Development. Center for Urban Policy Research, New Jersey.

MARTIN, R. L. 2003: A Study on the Factors of Regional Competitiveness. A final report for the European Commission DG Regional Policy. University of Cambridge, Cambridge.

PORTER, M. E. 1998: Cluster and the new economics of competition. *Harvard Business Review*. 6, Nov-Dec, 77-90.

SAJTOS, L. – MITEV, A. 2007: *SPSS kutatási és adatelemzési kézikönyv.* Alinea Kiadó, Budapest.

SIKOS, T. T. 1987: *Matematikai-statisztikai módszerek alkalmazásának lehetőségei a területi kutatásban.* Akadémia Kiadó, Budapest, 301.p.

SZEGVÁRI, P. 2006: Versenyképesség az új országos területfejlesztési koncepcióban. *Területi Statisztika*. 2006, 2.

SZÉKELYI, M. – BARNA, I. 2003: *Túlélőkészlet az SPSS-hez. Többváltozós elemzési technikákról társadalomkutatók számára.* Typotex Kiadó, Budapest

SZILÁGYI, GY. 2008: A versenyképesség mérése a nemzetközi összehasonlítások módszertanának tükrében. *Statisztikai Szemle.* 86, 1.

Innovation Performance and Competitiveness in the Hungarian Sub-regions

Miklós Lukovics[1] *– Péter Kovács*[2] *– Imola Rittgasszer*[3]

Besides globalization, or rather, parallel with it, knowledge-based economy seems to represent a highly important ground-gaining force – quasi becoming a trendy buzzword – that attracts increasing attention in developed countries, although its forms vary in different regions owing to the differing situation and set of conditions of the given area and the new type of international specialization emerging as a result of global competition. All this is a fundamental factor at the level of sub-regions, since competitiveness is determined by knowledge base on the local level.

The present paper aims to prepare the development of an indicator system and a complex method to measure the connection between the innovation performance and competitiveness of local units. We try to demonstrate the determining role of the innovation performance on the regional disparities measured by the competitiveness on sub-regional level with the help of multi-variable data analyzing methods based on a determined system of viewpoints, a correctly chosen theoretical models and statistical data.

Keywords: regional competitiveness, innovation performance, creativity, spatial analysis

1. Introduction

Today the permanent competitive advantage in developed regions derives from creativity, and the introduction of new products, services and processes which have not been replicated by the fellow competitors. We can observe that services having high added value and the intangible assets are gaining ground, which also may cover the high labour costs. For these products the quality, creativity and specialization of workforce on a given field become essentially important thus the quality of production factors are also greatly emphasized instead of their quantity. Based on Bajmócy's (2008) definition these knowledge-based economies are characterized by

[1] Dr. Miklós Lukovics, PhD, senior lecturer, University of Szeged, Faculty of Economics and Business Administration, Institute of Economics and Economic Development (Szeged)
[2] Dr. Péter Kovács, PhD, senior lecturer, University of Szeged, Faculty of Economics and Business Administration, Institute of Business Studies (Szeged)
[3] Imola Rittgasszer, PhD student, University of Szeged, Faculty of Economics and Business Administration, Institute of Business Studies (Szeged)

the growing dependence on information and knowledge as well as the variety of technological change and innovations, which lead to the increase in productivity.
The effective and fast learning becomes very important for enterprises and small regions in order to steadily keep their advantage. "The rapid introduction of innovations and new technology **means competitive advantage**." (Lengyel 2000, p. 980.). The presence of innovations crucially determines the competitiveness of the regions.

The process never sets short-term partial objectives but a social political aim, that is, the increase of local inhabitants' welfare. This – a little high – aim can be achieved by improving competitiveness, which is equivalent to the improvement of productivity according to Porter. However the rate of growth of productivity is primarily dependent on the **innovations**, that is, on new products, but mainly new technologies which enables the enterprise to obtain and strengthen permanent competitive advantages.

2. Theoretical background of the analysis

The concept of competitiveness that, due to the special attributes of global competition, has become one of the central terms in economics, offers an opportunity for the analysis of local units. International literature obviously ties analyzing the spatiality of economic influences to **competitiveness** and thoroughly designed models are available especially for the analysis of countries' competitiveness. The European Union's 2007-2013 programming period also devotes special attention to competitiveness as well as improving its influencing factors in order to facilitate cohesion and catching up (EC 2004, 2006a, 2006b).

Excellent competitiveness reports are completed each year at country level, however, in the case of studying regional competitiveness, focus must fall on smaller and smaller spatial units. Towns and town areas constitute the obvious basic units of such analyses, since the competitiveness of a country or region is mostly determined by towns, whose competitiveness tends to significantly exceed the competitiveness of the areas situated among them. International surveys dealing with the competitiveness of towns have also pointed out that the competitiveness of towns is also defined by the agglomeration area surrounding the town core that can be regarded as a nodal region, and therefore, is difficult to handle in the case of empirical analyses (Parkinson et al 2004, 2005, 2006). Sub-regions as administrative-statistical spatial units mostly correspond to the category of local unit as an economic criterion; however, the boundaries of these obviously somewhat differ from the actual economic catchment areas.

Beyond taking a position, it is also significant to **introduce the definitions** that constitute the basis of empirical analysis, since the selected approach is also accompanied by the methodology applicable in the course of empirical analysis. In the case of any empirical analysis, it is especially important to define the concepts that

the analyst intends to rely on in his or her research. This statement is especially true for competitiveness analyses, since the concept of regional competitiveness constituting the object of the analysis is a controversial term – as I demonstrated it in my paper –, and, on the other hand, it can be interpreted in various ways. Since regionalists also tend to accept approaches of regional competitiveness with highly different content, in competitiveness analyses it is really important to precisely express the definition, based on which analysis is carried out. In fact, the selected concept strictly determines the further logic of the analysis as well as its applicable method.

There are several, well known definitions of regional competitiveness, which interpret the approach of competitiveness on territorial units variously. Perhaps, the approach of regional competitiveness, published in the Sixth Periodic Report of the EU is based on the widest consensus: "*The ability of companies, industries, regions, nations and supra-national regions to generate, while being exposed to international competition, relatively high income and employment levels"* (EC 1999. p. 75.). In our research we depend on this standard definition of competitiveness, which is increasingly used in the regional policy of the European Union (Lengyel-Rechnitzer 2000, EC 2004).

3. The theoretical relations between competitiveness and innovation

The above presented standard competitiveness concept have already included the effect of innovation and research development on competitiveness between the lines. Based on Lengyel's (2003) deduction if the wages do not decrease and also not low in an economy, in addition the products are competitive, that is, they are not more expensive than other products and also marketable, this all can be implemented provided there is a constant innovation and technology change in the economy. Thus the productivity is increased by the innovations. Its essential condition is the research development activity and the flow of knowledge.

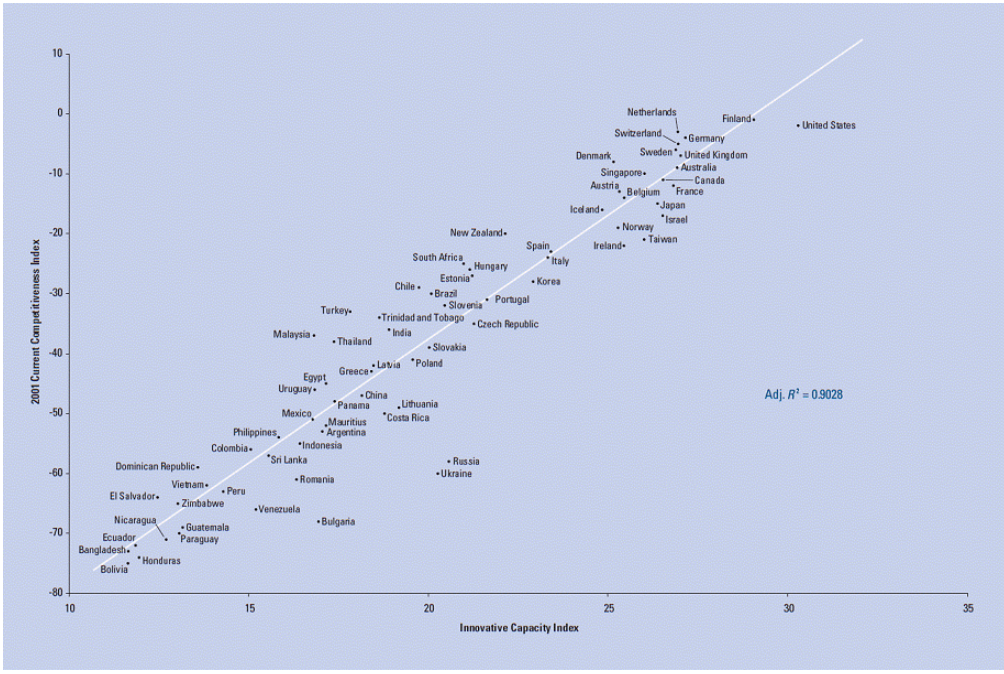
In terms of our research Porter and Stern's research in 2001 has major importance. The authors undertake to quantify the relationship of innovation and competitiveness with the help of multiple-variable data analysis methods. It is methodologically carried out by that on the basis of the research four subindexes an Innovation Capacity Index is calculated, which is compared by country – obviously by Porter's influence – to the competitiveness index of WEF. From this regression and correlation relationship is quantified.

In terms of our research the most significant part of Porter and Stern's work is that they examined the relationship between the Innovation Capacity Index (ICI), the Competitiveness Index (CI) and the GDP per person. The Competitiveness Index is used in the 2001 analysis of the World Economic Forum. It is not surprising that a strong correlation was shown between the innovation capacity and the competitiveness. They emphasize that utilizing and developing the innovation capacity, to

achieve the high level of productivity it is necessary for a given country to have growing and sustainable competitiveness.

The fact that the significant proportion of the countries are placed along the regression line between the two indexes (ICI and CI) indicates that the more innovative the country, the more competitive it is (Figure 1). (In addition, the R^2 index reveals a quite strong relationship: R^2=0,9028).

Figure 1. Relationship between competitiveness and innovation



Source: Porter-Stern (2001)

Based on these studies we can state that the innovation becomes an important source of the competitive advantage of the countries but in the case of many countries the many good innovation capacities are in vain if they cannot utilize and turn them into economic value, from which they could increase their income. Here within the frame of the Innovation Capacity Index the bases of innovation are comprehensively examined and they also took the factors into consideration that are needed for the good business utilization of innovation for the enterprises. The development of the innovation capacity has positive relation with the competitiveness and the prosperity of the country. The developed countries have an important role in creating innovations and it is necessary for gaining their competitive advantage, while the innovation strategy of undeveloped countries will be important to connect them to the

global knowledge base and take over the developed new solutions (Porter-Stern 2001).

Lederman and Maloney (2003) examined that how it is possible to qualify the effect of R&D costs on GDP growth. In their research they conducted regression calculations based on the base data of 53 countries. To conduct the regression calculations in order to smooth out the occasional outliers of certain years they used five year averages, namely the period between 1975 and 2000. The main finding of the research is quantifying the relationship according to which increasing the GDP proportional R&D expenditure with 1% causes a 0.78 % point increase in the rate of growth of the GDP. Their result is particularly important in terms of our research since it shows that one of the priority indicators of innovation potential, the GDP proportional R&D expenditure can be quantified with and has close relationship with one of the priority indicators of competitiveness, the GDP.

4. The role of innovation in some demonstration models of competitiveness

In this section we review the demonstration models which make the wide range of factors influencing regional competitiveness graphic through the extending the concept of competitiveness and including the *development side* in the examination. It is important that of all the demonstration models to be presented the competitiveness top hat, the competitiveness tree and the pyramid-model are all rely on the standard definition[4] of competitiveness, thus it is expected to form an organic whole with the ideas written in the previous section.

4.1. The competitiveness hat

A research group of the University of Cambridge attempts to synthesize the various definitions of competitiveness (Martin et al 2003). They find that despite the diverse approaches, some issues which have to be considered as distinctive feature of regional competitiveness can be taken out from the international literature. These factors are systematized by the regional competitiveness hat, which consists of four different levels: regional result, regional output, regional intermediate output and the factors of regional competitiveness[5].

[4] The standard concept of competitiveness is used by, for example, the Department of Trade and Industry (DTI) in the United Kingdom, in addition this approach appears in the study titled Benchmarking the Competitiveness of European Industry.

[5] The experts of World Bank in working out the urban and local economic strategies in 2000 separated three types of the examined indicators, which approach have become widely accepted in the European Union: the input indicators, the output indicators and the outcomes (Worldbank 2000). This three way subdivision or its further consideration can be observed in quite many approaches circling the concept of regional competitiveness with the use of indicators.

The factors that can be associated with innovation are placed on the brim of the top hat, between the secondary factors. Three such factors can be named, however, their exact role are not detailed by the model (Martin et al 2003):

1. Technology:
2. Innovativity
3. Knowledge infrastructure

4.2. The competitiveness tree

The research group of ECORYS under Jan Maarten de Vet's (2004) direction created a graphic demonstration model of the regional competitiveness similarly to the competitiveness top hat (Figure 3.). According to the authors' ideas, the tree as an organic whole characterizes the cyclicness of the competitiveness concept. This metaphor is filled with content by that the quality of the soil and the efficient functioning of the roots, the trunk and the branches determine the strength of the tree and the quality of the fruits.

This metaphor is accordance with our findings to a great extent, according to which innovation is one of the particularly important influential factors of competitiveness, since it is placed in the root of the competitiveness tree, which determines the strength of the tree and the quality of the fruits, which may in itself be corresponded with competitiveness.

This is a dynamic process, since the fertility of the soil is partly determined by the extent to which the falling fruits revitalize it. As a result, the "fruit" of competitiveness (welfare, sustainability, etc.) is unambiguously determined by the categories, such as innovation, which affect the competitiveness on different levels and to different extent.

4.3. The pyramid-model of competitiveness

The **pyramid-model** of the competitiveness of regions is based on the standard definition of competitiveness, which is a convenient conceptual approach, since it highlights two measurable economic categories: the levels of income and employment (Lengyel 2000).

To improve the competitiveness the *simultaneous* development of research, innovation, education and training, in addition, the spread of scientific and technological knowledge and its appearance in the competitive advantages of the enterprises operating in the region are essential. Besides the public research institutes, the enterprise research institutes have a crucial effect on the competitiveness of a region. Generally speaking, there is no significant companies' R&D activity in the relatively undeveloped regions. Provided that there is no public research institute which could provide the knowledge necessary for technological catching up, taking over the technology becomes emphasized, that is the technology transfer. As a result, the in-

novation results may also arrive from outside the region (for example technology transfer), but the competitiveness of the region is decisively promoted by the effective R&D activity, the creation of innovations and their extensive, **rapid spreading** in the region[6]. This can have decisive impact on the region and the competitiveness of its enterprises.

The innovation not only has a distinguished position among the basic factors, but also among the long-term influential success factors. That is, as a result of the logic of the Pyramid-model the *innovation affects* a welfare of a region *in two ways*: on the one hand, strengthening the development of research and technology directly improves the competitiveness of a region, while the existence of innovation culture indirectly contributes to all this.

One success factor is formed by the "**Innovation culture and capacity**". In the successful regions many patents are obtained, there are extensive innovation activities and the spread/diffusion of innovations is efficient. Only few regions are able to perform high technology R&D activities, but the majority can also be successful in traditional manufacturing industries by the use and rapid adaptation of new technologies. The innovation culture and skill are necessary for a region to respond with adequate adaptation strategy to any kind of technical, business, environmental, etc. change, and to be able to turn the innovation challenges to its own advantage. The innovation ability does not only mean the capacity of research and development institute or university but especially that of the enterprise, the prepared, innovative small and medium enterprises in the new market sectors and prospering branches.

4.4. Competitiveness and innovation in the operational programme of EU 246

On behalf of the European Union, the report that examines the potential of the regional policy instruments of the European Union in the framework of a quite large empirical research and focusing on the objectives of Lisbon and Göteborg, was completed on 22nd June 2009. the analysis examines the policy instruments of cohesion in 27 EU member states, through analysing 246 operational programmes which are supported by either the European Regional Development Fund or the Cohesion Fund or the National Strategic Reference Frameworks of the given countries.

The analysis includes several important findings that highlight important connections in terms of our research. One of the most important overall connection is that the Convergence as well as the Competitiveness and employment programmes promoting the first two objectives of the Structural Funds both contribute to achieving the objectives of Lisbon and Göteborg. It is accomplished by promoting innova-

[6] The WEF uses the technology index to calculate the economic creativity index. However, in this the emphasis in placed on the *existence* of innovations, regardless if it is from own development or external source (Lengyel – Deák 2001).

tion, supporting the enterprises and strengthening the synergy between growth and environmental protection.

According to the other quite significant metaphorical finding of the document, six roads lead to Lisbon and Göteborg. These roads are arranged into two main groups, the competitiveness main group and the convergence main group, within which there are 3 roads respectively. These roads summarize the focus of the reviewed operational programmes based on the budget and the programmes of regional policy.

1. Competitiveness roads:

- The first of the competitiveness roads is innovation, R&D and entrepreneurial skill. The writers of the study identify this strategy primarily with the relatively small countries, which can be characterized by small regional differences and their GDP per person exceeds the EU average.
- The second competitiveness road is focused on the increase of employment, the urban renewal and the energy, also with R&D and innovation relationship. This strategy was found mainly in the case of countries having GDP more than the EU average, but the regional differences of these countries are bigger than that of the countries of the first road.
- The third competitiveness road focuses on the economic and environmental synergies through priorities such as renewable energy, urban and rural developments and tourism. This strategy was found primarily in the case of larger countries divided by regional differences and having average GDP level.

2. Convergence roads:

- The road focused on growth and employment: recognizing the importance of transport accessibility and employment. This road includes the stimulation of innovation and entrepreneurial skill in the operational programmes, but with less intensity than in the competitiveness roads. This strategy can be experienced in the relatively developed cohesion countries having strong central region, where the GDP per person is slightly under the EU average. The analysis ranks Hungary, along with Portugal, the Czech Republic, Slovenia, Malta and Cyprus in this category.
- Developing human resource and institution capacity through the stimulation of knowledge, in addition, providing better access to innovation and the programmes improving entrepreneurial skill in the interest of improving the quality of labour force and attaining sustainable development. This strategy dominates mainly in smaller cohesion countries where the GDP/person is lower than the EU average.
- Regional cohesion: mostly the development of infrastructure dominates in the interest of bridging the urban-rural gap. This strategy primarily dominates in large, polycentric countries characterized by regional inequalities with low GDP per person.

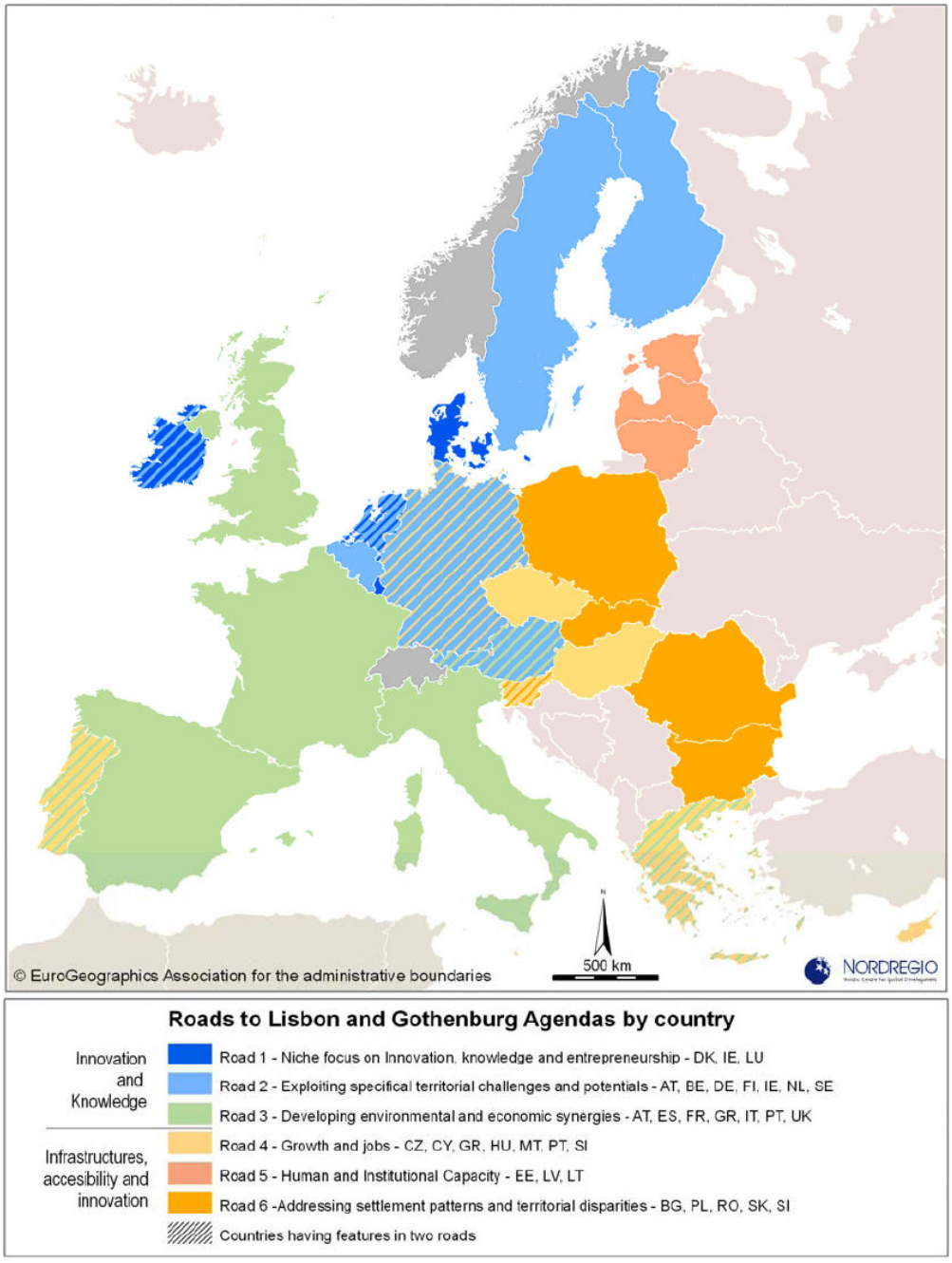
The consequence can be drawn from the logic of the 6 roads that the innovation as a condition can be assigned to both the competitiveness (objective 2) and the convergence (objective 1) priorities, but with different meaning and intensity.

In the *Regional competitiveness and employment* operational programmes strong Lisbon-orientation can be observed, the emphasis is on the innovation, knowledge and technology. The member states where there are predominantly objective 2 regions the priorities related with innovation and entrepreneurial skill clearly dominate in the competitiveness and employment operational programmes (R&D expenditure, innovation grants).

In the *Convergence* operational programmes the transport infrastructure objective is present in the largest rate, but the CSG priority of innovation and entrepreneurial skill does not suffer significant disadvantage compared to this objective.

It is quite important to emphasize a very significant difference: it is unequivocally the creation of knowledge that is meant by the innovation priorities in the Competitiveness and employment operational programmes, which increases the competitiveness of the regions. By contrast, the innovation priorities of the Convergence operational programmes primarily mean the preparation for receiving the innovation results from outside the region, which is mainly necessary to stop further dropping behind, in addition, in the case of its success it is necessary to start catching up. Based on this logic it can be stated that developing new results is necessary but not sufficient condition for competitiveness. The member states according to the ways are spatially concentrated (figure 2).

Figure 2. The spatial position of the six roads in the European Union



Source: EC (2009)

5. The key factors determining innovation potential: knowledge and creativity

The previous sections show clearly that science have already achieved quite fine results in the process of revealing the relationship of competitiveness and innovation. It was clearly justified that the two concepts are in fairly close relationship with each other. Many of the analyses revealing the relationship between innovation and competitiveness find that the examination of the relationship between the two concepts has to be extended: they highlight the knowledge or the creativity (or both) as the input of innovation. This section aims to carry out this extension towards knowledge and creativity, and to extend the connections revealed in the previous sections to the definition of competitiveness-innovation-creativity-knowledge.

According to De Bono (1990) the creativity is the ability to reach several innovative solutions with using unusual methods. Csikszentmihályi (1997) defines creativity as the ability to communicate irregular thoughts, which forms new ideas and helps to create complex solutions through examining reality from different approaches. Overall, according to the most accepted characteristics the creativity is an ability that is necessary for problem solving, seeking different alternatives, utilizing resources as effectively as possible and distinction. The owner of creativity as a characteristic, the creative person notice the problems sooner due to the features listed earlier, and his/her freedom of thought enables him/her to develop more efficient and new solutions (O'Rafferty–O'Connor–Curtis 2009).

In our opinion, creativity is a special human attribute that can be mostly characterized by curiosity, desire for knowledge, tendency for self-realization and "not thinking in patterns". It is important to underline that the efficient solutions promoting development are formed if the creativity is accompanied by sufficient knowledge (Florida 2002).

Several studies show the positive impact of the R&D activity as the activity promoting innovation on the economic growth and productivity. Guellec and Pottelsberghe (2001) examined the long-term effect of the different types of the R&D related to the productivity based on the data of the period between 1980 and 1998 for the certain countries of OECD. They concluded that the R&D of both the public sector and the private sector and foreign countries has significantly positive effect on the increase of productivity.

Kakko and Inkinen (2009) show the close relationship of the "homo creativus" and the innovation through comparing the features of innovation with the characteristics of creative people, thereby they demonstrate the clear relation according to which creativity, the creative people's added value is one of the crucial factors of developing innovations.

Several researches proved the positive impact of innovation and R&D activities on influencing incomes. In Porter's theory the innovation means the basis of productivity and it will become the determiner of competitiveness, the aim of which is increasing welfare. At the same time it has to be highlighted that in order for the

increase of productivity as a result of innovation not to be at the expense of employment, the human resource have to continuously train itself to be creative and able to exploit opportunities. This is particularly important in knowledge-based economies.

In Huggins and Izushi (2008) model the base of knowledge is essentially important for the economy to develop innovations. They call it the base and the recipe of innovation. In developed economies utilizing knowledge to create some kind of novelty or innovation greatly contributes to the competitive advantage of the enterprises. Thus in the modern economies the knowledge becomes the key factor of the innovation. However, they emphasize that economic seizure of knowledge is very difficult; it cannot be characterized by a single index.

The World Bank's examination of knowledge-based economies the innovation and knowledge also appear as the main influences of competitiveness. The aim of their survey is to help the countries to exploit knowledge more in order to become knowledge-based economy. The 4 factors examined (ICT infrastructure, economic-institutional structure, innovation system, human resource development) are taken into account with equal weight, emphasizing that all of them are equally important to exploit knowledge. The World Bank does not highlight the innovation, but put emphasis on the utilization of knowledge as a key determinant. Since a strong positive relationship was shown between the development, the rate of growth of development and the knowledge-base of an economy, the knowledge and its utilization have become the key to future development.

With the special focus on knowledge-based economies the role of knowledge and creativity also appreciate. Taking Florida's work (Florida 2002), Huggins-Izushi's theory, Huggins and Davies's ideas (Huggins-Davies 2006), and Swann and Birke's model (Swann-Birke 2005) into consideration, we can observe that more and more people, going beyond innovation, place great emphasis also on its input factors, knowledge and creativity in terms of competitiveness.

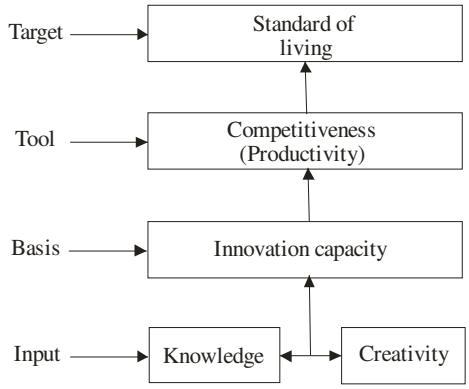
It supports the important role of knowledge and innovation that the EU also puts great emphasis on them. As early as in 2000 it emphasized the highlight of knowledge-based economy in the framework of Lisbon Strategy. One of the Community Strategy Guidelines the development of knowledge and innovation serving growth, and VIK Program the innovation efforts receive the most sources. The importance of innovation performance is shown by that it is annually surveyed in the framework of EIS. Recognizing the significance of knowledge needed for innovation the year of 2009 was named the European Year of Creativity and they also conducted its survey, which was summarized in the document titled Design, Creativity and Competitiveness.

Based on the above presented findings of international literature it can be clearly stated that **there is a clear and strong relationship between competitiveness and innovation**, in addition, that **the relationship is two-directional** (Carayannis – Gonzalez 2003). The dominant direction is that the existence of innovation

results strengthens the competitiveness of the region (that is the innovation affects competitiveness), however – as the competitiveness tree also demonstrated – **there is a feedback** in the structure as well, since in the competitive regions the local business environment further intensifies, which provides further drive for the innovation.

Furthermore, it can be stated that the most recent international research findings are directed towards that **the input side of the innovation capacity is increasingly highlighted**, and the research findings extend the examination clearly towards **knowledge and creativity**. In our opinion, to create new findings both knowledge and creativity are necessary. The knowledge is necessary but not sufficient condition for the creation of innovation results, because the static way of thinking built on patterns is not adequate for creating innovation results. The same can be said about creativity: the creativity is also necessary but not sufficient condition for the creation of innovation results, behind the way of thinking searching for new directions there has to be serious knowledge and professional competence to realize it in some kind of result. Based on this, completing Porter's line of thought with knowledge and creativity we receive the logical structure of knowledge-based economy development (Figure 3).

Figure 3. The logical structure of knowledge-based economy development



Source: Own construction based on Lengyel (2002)

6. The role of innovation in the set of indicators of the most significant competitiveness reports

The objective of the next stage of our research was to review the set of indicators and methodology of competitiveness reports and researches from two aspects: on the one hand, we examined the role and extent of innovation in competitiveness researches (whether its temporal change possibly carries significant information), and

the extent of benchmarking reasonably allowed when we intend to conduct an indicator-based analysis on the small regions in the Southern Great Plain region. In the study **the set of indicators of 17 international competitiveness reports and competitiveness researches was reviewed** (ACRC (2008), BERR (2008), BERR (2009a), BERR (2009b), BHI (2008), Couto et al (2004), Forfás (2009), IMD (2009), Huggins, R. – Davies, W. (2006), InnoMetrics (2008), Kronthaler, F. (2003), Piech, K. (2008), PSRC (2008), RDC (2003), Snieska, V. – Bruneckiene, J. (2009), WEF (2009), World Bank (2008))

It was found that in all the examined reports dealing with competitiveness the innovation was present. Although the concepts used to define competitiveness are not identical, the role of the growth of productivity is emphasized throughout and the associated innovation also has a great role. At the same time, its weight in competitiveness is different in each analysis.

It can be concluded that various competitiveness analyses have been created in the international literature on national and regional level. The majority of these work with quite a **sophisticated methodology, consistent and established use of concepts and a set of indicators having been refined** for a number of years, the majority of which may also be adapted to national, innovation-centred competitiveness analyses conducted on local regions. In many cases the adaptation is made more difficult by that certain indicators of the competitiveness analyses conducted on the level countries and regions are not available or not interpreted on small region level.

The reviewed competitiveness analyses worked with not only different region concept but also with **different competitiveness concept**. Of the reviewed competitiveness studies the IMD, the WEF, the Forfás, the DTI and the BHI give an **own competitiveness definition**, on which they **consistently build** competitiveness analysis, while the examinations of the European Union – also under consistent use of concepts – draw on the standard competitiveness definition as appropriate.

It is instructive that in addition to exactly defining the definition and using the concepts consistently, the WEF and the Forfás also apply a **model** forming the base of the indicator definition and **matching the uttered definition of competitiveness** (diamond-model, Forfás-pyramid) in the course of analysis, which significantly facilitates the clarity of the logical structure of the analysis, therefore its expected acceptability. Consequently, in our view in building the model to be worked out for the purpose of the complex analysis of competitiveness it is appropriate to take this logical structure into account.

After reviewing the set of indicators of the examined analyses we can also conclude that the majority of the applied scorecards use **indexes with both ex post and ex ante character**. The studied analyses work with quite **heterogeneous set of indicators**, which are often based on different logic. However, the appreciation of the role of the **soft data** in the certain set of indicators is definitely remarkable. The greater proportion of the examined analyses in terms of their methodology rather

undertook **comparing** and evaluating with the use of **simple statistical methods**, but in the most recent analyses the multi-variable analyses and the pursuit of index formation dominate.

In connection with the reports that are published annually (or at certain intervals) updated we had the opportunity to observe temporal tendencies as well. It outlines the European Union's intention of increasingly placing the basis of competitiveness on innovation. It can be observed on the change of the set of indicators, which can be clearly detected on the increase of the weight of factor groups and indexes related to the innovation.

Reviewing the set of indicators it can be found that despite the above mentioned heterogeneity certain lines can be formulated (Chart 1.):

1. One of these lines is indicated by undeniably the frequency of occurrence of certain indicators. The number of specific patents can be considered as a dominant index, which appears in 13 of the 17 reviewed set of indicators.
2. The rate of the entire R&D expenditure calculated in GDP % can also be regarded as a highlighted index, however, the demand on breaking the total R&D expenditure down to the R&D expenditures of government, business and higher education sector appears in more and more places.
3. The increase in the weight of output indicators compared to the input indicators also can be interpreted as a tendency. Due to the characteristic of innovation activity there is not under any circumstances a deterministic relationship between the input data and the output data (perhaps not even stochastic), thus the producers of the set of indicators increasingly try to move towards the output indexes, but this data is quite difficult to be produced.
4. The requirement according to which the innovation results should also be utilized in industry appears in more and more competitiveness reports. The indexes which try to quantify the university-industry cooperation have been involved in the set of indicators as a consequence of this.
5. A certain sectoral delimitation appears as a novelty is the most recent competitiveness reports: the high-tech and/or ICT sector is separated in a number of places.

7. The ideas related to the possibility of measuring the competitiveness and innovation performance in Hungarian small regions

Since both the competitiveness and the innovation performance are complex concepts, neither can be characterized by a single index. Instead, indicator systems as well as factors and indicators formed from basic indicators by multi-variable data

analysis techniques have to be used to measure and characterize them. To produce them a complex, high quality database is needed based on which these indicators and indicator systems can be created. However, the existing databases leave a lot to be desired. Since the success of data analyses and drawing the right conclusions are greatly dependent on the quality of the basic database, prior to the indicators and the multi-variable data analysis **firstly** we have to **create a proper basic database**, which is suitable for characterizing the competitiveness and innovation performance. The experience of the reviewed sets of indicators gives an appropriate help with this (chart).

From the basic database given by the study of literature, **secondly** we have to select the indicators which indeed have a **relevant impact** on competitiveness and innovation performance. For this we will characterize the categories of the applicable models with maximum two or three indicators artificially created by principal components analysis in a way that the principal components convey the 75-80% of the information content of the original variables. We will not include in the model the original variables the communality of which will be low and their omission would not mean a distorted model in terms of the examination.

After forming the circle of relevant variables, the **clustering** of small regions comes next. We plan to conduct it on two lines. We complete the clustering, on the one hand, based on the whole pyramid model, on the other hand, based on only the innovation indicators. From the received results and their comparison we try to type the small regions based on their competitiveness and innovation performance, furthermore we check that in what kind of pattern and combination the type categories according to the competitiveness and innovation performance occur.

Finally, we establish a **ranking** of the small regions **based on their competitiveness and innovation performance**. The multi-dimensional procedure is a suitable procedure for this. From the results and their comparison we try to conclude the character of relationship that may be between the level of competitiveness and the innovation performance in the case of small regions. We also examine whether the existing relationship is the same character as if we examined country or regional data.

This study aimed at reviewing the set of indicators and methodology of competitiveness reports and researches from two aspects: on the one hand, we examined the weight and role of innovation in the competitiveness researches (whether its temporal change possibly carries significant information), and the extent of benchmarking reasonably allowed when we intend to conduct an indicator-based analysis on the small regions. The study reviewed the set of indicators of **17 international competitiveness reports and competitiveness researches** (table 1).

Table 1. The occurrence of certain indicators in the reviewed sets of indicators

Name of indicator	WEF GCR 2009-2010	IMD WCY 2009-2010	ECI 2006-2007	BERR 2008 PCI	Forfás ACR 2009	WB KEI 2008	European Inn. Sc.	BERR REPI 2009	BHI 2008	ACR Croatia 2008	Estonian Comp. R.	Puget Sound 2008	Lithuanian RCI 2009	German regions 2003	Portugal regions 2004	Poland	World Bank 2008
1. Innovation capacity for enterprises	X									X							
2. The quality of scientific research institutes	X									X							
3. University-industry research cooperations	X									X			X				
4. The public purchase of advanced technology products	X																
5. Availability of researchers, engineers	X								X								
6. The number of patents per unit	X	X		X	X	X	X		X	X	X	X	X			X	X
7. Intellectual property	X																
8. The number of technological cooperations between enterprises		X								X						X	
9. Supporting technological development		X								X							
10. Technological regulation, legal background		X								X							
11. High-tech export		X															
12. Total R&D expenditure in all (million US $)		X	X														
13. Total R&D expenditure in all (per unit)		X															
14. Total R&D expenditure in GDP %		X		X	X		X			X	X	X			X	X	
15. R&D expenditure of business sector in total (million US $)	X	X	X											X	X	X	
16. R&D expenditure of business sector in GDP %	X	X		X	X	X					X	X					
17. The number of R&D employees		X									X					X	
18. The number of R&D employees per inhabitant		X	X				X			X	X						
19. The number of R&D employees in enterprises		X															
20. The number of R&D employees in enterprises per inhabitant		X	X														
21. The number of basic researches		X								X							
22. Science and engineering degrees in proportion to the first university degrees		X															
23. The number of published scientific articles		X				X	X				X						X
24. The focus of sciences in schools		X															
25. The number of Nobel Prizes in total (chemistry, physics, economics, life science)		X															
26. The number of Nobel Prizes in total per unit (chemistry, physics, economics, life science)		X															
27. The number of patent applications		X	X									X		X		X	
28. Intellectual property rights		X															
29. Patent productivity (patent number/the number of R&D persons working in the business sector (1000)		X															
30. Scientific research		X															
31. The rate of innovative enterprises within all the enterprises				X													
32. The number of trademarks and designs per unit				X	X												
33. The income from research of higher education institutes and publicly financed research institutes				X													
34. The rate of turnover form innovation activity in turnover				X													

Name of indicator	WEF GCR 2009-2010	IMD WCY 2009-2010	ECI 2006-2007	BERR 2008 PCI	Forfás ACR 2009	WB KEI 2008	European Inn. Sc.	BERR REPI 2009	BHI 2008	ACR Croatia 2008	Estonian Comp. R.	Puget Sound 2008	Lithuanian RCI 2009	German regions 2003	Portugal regions 2004	Poland	World Bank 2008
35. The supply and demand side of innovation				X													
36. Employment in R&D and high-technology, in addition, in technology demanding industries.								X									
37. Innovation index of European Innovation Scoreboard					X												
38. The rate of firms doing innovation activity within all the enterprises					X												
39. The part of income that the enterprises spend on innovation activity					X											X	
40. R&D expenditure of higher education institutes in GDP %					X		X										
41. The number of researchers in proportion to employees (1000)					X												
42. BERD by enterprise type (with foreign or domestic ownership)					X												
43. The rate of people having PhD degree in relation to the population (1000)					X						X						
44. Amount of royalties per capita						X											X
45. The rate of labour force with higher education degree in relation to the total workforce							X										
46. The number of participants in scientific training, the number of PhD students							X									X	
47. ICT expenditure per capita							X										
48. Broadband access in proportion to the population							X										
49. HERD (R&D expenditure of higher education) in proportion of GDP (%)			X				X		X		X						
50. Knowledge transfer between universities and enterprises										X							
51. The size of venture capital per capita									X	X		X					
52. The number of university students per 1000 inhabitants													X				
53. The number of college students per 1000 inhabitants													X				
54. The number of technologies produced by universities													X				
55. The cooperations of universities of different regions													X				
56. The participation of universities in international research													X				
57. The rate of R&D financed by foreign countries within GERD																X	
58. The number of high-tech patents per unit																X	
59. The rate of medium-high-tech and high-tech researches in the total number of researches																X	
60. The number of employees in ICT sector			X														
61. The R&D expenditures of higher education per capita			X														

References

ACRC 2008: *Annual Competitiveness Report Croatia*. National Competitiveness Council, The Institute of International Relations, Zagreb.

Bajmócy, Z. 2008: A regionális innovációs képesség értelmezése és számbavétele a tanulás-alapú gazdaságban. In Lengyel I. – Lukovics M. (szerk.): *Kérdőjelek a régiók gazdasági fejlődésében*. SZTE Gazdaságtudományi Kar Közleményei, JATEPress, Szeged. pp. 26-46.

BERR 2008: *The 2008 Productivity and Competitiveness Indicators*. Department for Business Enterprise and Regulatory Reform, London.

BERR 2009a: *The 2009 Productivity and Competitiveness Indicators*. Department for Business Enterprise and Regulatory Reform, London.

BERR 2009b: *Regional Economic Performance Indicators*. Department for Business Enterprise and Regulatory Reform, London.

BHI 2008: *Eighth Annual State Competitiveness Report 2008*. The Beacon Hell Institute at Suffolk University, Boston.

Bono, E. D. 1990: *Lateral Thinking: A Textbook of Creativity*, Penguin. United Kingdom.

Carayannis, E. G. – Gonzalez, E. (2003): Creativity and Innovation= Competitiveness? When, Hoe, and Why, *The International Handbook on Innovation*, 587-605. o. Pergamon, Amsterdam.

Couto et al 2004: *Competitiveness of Portuguise Regions. A Comparative Analysis*. University of Azores, Azores.

Csikszentmihályi M. 1997: *Creativity: Flow and the Psychology of Discovery and Invention*, Harper Collins, New York.

DTI 2004: Regional Competitiveness and State of the Regions. Department of Trade and Industry, London.

EC 1999: *Sixth Periodic Report on the Social and Economic Situation and Development of Regions in the European Union*. European Commission, Luxembourg.

EC 2004: *A new partnership for cohesion - Third Report on Economic and Social Cohesion*. European Commission, Brussels.

EC 2006a: *The Growth and Jobs Strategy and the Reform of European Cohesion Policy. Fourth Progress Report on Cohesion*. European Commission, Brussels.

EC 2006b: Council Decision of 6 October 2006 on Community strategic guidelines on cohesion (2006/702/EC). *Official Journal,* L 291/11, Luxembourg.

EC 2009: *The Potential for regional Policy Instruments, 2007-2013, to contribute to the Lisbon and Göteborg objectives for growth, jobs and sustainable development. Final report*. European Commission, Brussels.

Florida, R. 2002: *The rise of the creative class. And How It's Transforming Work, Leisure, Community, and Everyday Life.* Basic Books, New York.

Forfás 2009: *Annual Competitiveness Report 2009.* National Competitiveness Council and FORFÁS, Dublin.

Guellec, D. – Pottelsberghe, B. 2001: *R&D and Productivity Growth: Panel Data Analysis of 16 OECD Countries.* OECD STI Working Papers, 2001/3, Paris.

Huggins, R. – Davies 2006: *European Competitiveness Index 2006-07*. Robert Huggins Associates Ltd, United Kingdom.

Huggins, R. – Izushi, H. 2008: Benchmarking the knowledge competitiveness of the globe's high-performing regions. A review of the World Knowledge Competitiveness Index. *Intenational Business Journal,*1-2, 70-86. o.

IMD 2009: *World Competitiveness Yearbook 2009*. International Institute for Management and Development, Lausanne.

InnoMetrics 2008: *Rethinking the European Innovation Scoreboard: A New Methodology for 2008-2010.* Pro Inno Europe, Inno Metrics.

Kakko, I. – Inkinen, S. 2009: *Homo creativus*: creativity and serendity management in third generation science and technology parks. Science and Public Policy, 36, 537-548. o.

Kronthaler, F. 2003: A Study of the Competitiveness of Regions based on a Cluster Analysis: The Example of East Germany. Institut für Wirtschaftsforschung, Halle.

Krugman, P. 1994: Competitiveness: A dangerous obsession. Foreign Affairs, 2, 28-44. o.

Lederman, D. – Maloney, W. F. 2003: *R&D and Developement.* Policy Research Working Paper, 3024, World Bank, Washington.

Lengyel, I. 2000: A regionális versenyképességről. *Közgazdasági Szemle*, 12, 962-987. o.

Lengyel, I. 2002: A regionális gazdaság- és vállalkozásfejlesztés alapvető szempontjai. In Buzás N.-Lengyel I. (szerk.): *Ipari parkok fejlődési lehetőségei: regionális gazdaságfejlesztés, innovációs folyamatok és klaszterek.* JATEPress, Szeged, 24-54. o.

Lengyel, I. 2003: Verseny és területi fejlődés: térségek versenyképessége Magyarországon. JATEPress, Szeged.

Lengyel, I. – Deák Sz. 2001: A magyar régiók és települések versenyképessége az európai gazdasági térben (elméleti és fogalmi háttér, főbb irányzatok). *Kézirat*. SZTE GTK, Regionális és Alkalmazott Gazdaságtani Tanszék, Szeged, p. 98.

Lengyel, I. – Rechnitzer J. 2000: A városok versenyképessége. In Horváth Gy. – Rechnitzer J. (szerk.): *Magyarország területi szerkezete és folyamatai az ezredfordulón*. MTA RKK, Pécs, 130-152. o.

Marteen de Vet, J. et al 2004: *The Competitiveness of Places and Spaces. A Position Paper*. ECORYS, Rotterdam/Leeds/Birmingham/Brussels.

Martin, R. L. et al (2003): *A Study on the Factors of Regional Competitiveness*. A final report for The European Commission DG Regional Policy. University of Cambridge, Cambridge.

O'Rafferty, S. – O'Connor, F. – Curtis, H. (2009): *The creativity gap? – bridging creativity, design and sustainable innovation.* Joint Actions on Climate Change Conference, Aalborg, Denmark.

Parkinson, M. et al (2004): *Competitive European Cities: Where do the Core Cities Stand?* Office of the Deputy Prime Minister, London.

Parkinson, M. et al (2005): *State of the Cities. A Progress Report to the Delivering Sustainable Communities Summit.* Office of the Deputy Prime Minister, London.

Parkinson, M. et al (2006): *State of the English Cities.* Office of the Deputy Prime Minister, London.

Piech, K. (2008): *Knowledge- and Innovation-based Competitiveness – Assessing the Position of Poland in Europe*. The Knowledge and Innovation Institute, Warsaw.

PSRC (2008): *Puget Sound Regional Competitiveness Indicators 2008-2009*. Puget Sound Regional Council, Washington.

RDC (2003): *Competitiveness and Future Outlooks of the Estonian Economy*. Research and Development Council, Tallin.

Porter, M. E. – Stern, S. (2001): *National Innovative Capacity*. In *The Global Competitiveness Report 2001-2002.* Oxford University Press, New York.

Snieska, V. – Bruneckiene, J. (2009): *Measurement of Lithuanian Regions by Regional Competitiveness Index*. Engineering Economics, Kauno technologijos universitetas, Kaunas.

Swann, P. – Birke, B. (2005): *How do creativity and design enhance business performance?* DTI, London.

Török, Á. (1999): Verseny a versenyképességért. MeH Integrációs Stratégiai Munkacsoport, Budapest.

WEF (2009): *The Global Competitiveness Report 2009-2010.* World Ecomonic Forum, Geneva.

World Bank (2008): Measuring Knowledge in the World's Economies. World Bank Institute, Washington.

The spatiality of the creative micro-regions in Hungary

Imola Rittgasszer[1]

Today the economic environment is influenced and transformed by other background processes besides globalisation, which worth being and have to be paid attention to as area organiser powers. Nowadays the development of globalisation processes unambiguously shows that human factors and the knowledge based on it may become the future key factors to development. We live in a world where knowledge, human creativity and the ability to process infinite information are also one of the prime movers of development.
This study aims to analyse how Hungarian micro-regions can be classified according to the position of the creative class. The starting question is whether the micro-regions as regional units are suitable to examine the spatial position of the creative class.
In the analysis of the Hungarian micro-regions I take Richard Florida's study dealing with the creative class as a starting point. I examine the database and its indicators developed for Hungarian micro-regions with multivariable statistical methods, like Multidimensional Scaling (MDS) and Hierarchical cluster analysis.

Keywords: creative class, knowledge-based economy, micro-regions, regional analysis

1. Introduction

In the global contest *not the material possessions* but instead the knowledge and the relationship capital have become factors of vital importance, the most important movers of economic development (Enyedi 2000). In regional science, knowledge is identified as a *decisive factor of regional development* (Lengyel 2003; Rechnitzer – Smahó 2005; Varga 2005). Considering the changes of factors determining regional growth and development we can find also in Hungary that in the years after the change of regime those areas stood out where the human factors were concentratedly at present. Nowadays the quality of *human capital* – according to several recognised academic trends, its innovativity and creativity – *and economic development are in very close connection with each other*, furthermore, the success of a region is among others determined by the available labour force's ability to innovative economic performance. (Lengyel–Rechnitzer 2004; Varga 2009). Naturally, besides skilled

[1] Imola Ritgasszer, PhD student, University of Szeged, Faculty of Economics and Business Administration, Institute of Business Studies (Szeged)

workforce several factors play a role in the economic performance of a region, but today knowledge, human creativity and the ability to process information have become one of the most important factors of development. Consequently, knowledge-based economic development strategy can be applied successfully in regions where the human resource of a required quality is available (Lengyel 2007).

The theory worked out by American economist-geographer Richard Florida stands out from the trends tracing regional development back to the quality of labour force. Not only did Florida introduce the concept of creative class but he also worked out the method suitable to examine the creative class of the city regions in the USA. (Florida 2000a).

2. Knowledge-based economy and the creative class

It turns out from the most widely accepted definitions of knowledge-based economy that it is quite a complex concept and it has wide-ranging properties, which can be approached from several aspects. (OECD 1996; OECD 2005, ESRC 2005; Kok 2003; DTI Competitiveness White Paper 1998; Leadbeater 1999; Huggins–Izushi 2008; Leydesdorff 2006; WBI 2007; Lengyel B.–Leydesdorff, L. 2008). There is only one common segment of the definitions, which emphasises the outstanding role of the creation of knowledge. It can also lead to the conclusion that mapping knowledge-based economy with indicators cannot happen with some selected ones but a complex, multivariable analysis has to be applied.

The explanations of global economy linked to knowledge can be classified according to basically two kinds of logic:

1. The "knowledge-based economy" as a programme has rather an *economic political character*. The economic political public opinion considers the R&D ability as the pledge of succeeding in knowledge-based economy (Bajmócy 2008; Lengyel B. 2008).
2. The other explanation of knowledge-based economy is connected to the *human capital side*; it plays an important role in creating and spreading knowledge (Varga 2009).

The latter forms the subject of the study, that is *the special segment of the human capital side of knowledge-based economy, the analysis of the position of the creative class*, and it does not deal with the economic political side of knowledge-based economy. János Rechnitzer and Melinda Smahó (2005, p14) highlight that *"national regional research has so far neglected the analysis of knowledge as a new component of regional development"*, so – although in this field there have been significant steps forward recently – I intend to move in this direction with my research. Naturally, the classification of the explanations of global economy linked to knowledge according to two kinds of logic cannot be separated, since the quality of human capital essentially influences whether it is possible to use knowledge-

based economic development strategy in the given area. I go back to this briefly in the analytic part of the study.

In the analysis of the character of knowledge-based economy a basic question is *what moves knowledge-based economy*? According to one of the accepted answers, which can be approached from the human capital segment of knowledge-based economy, the development possibilities of certain regions are greatly determined by the quantity and quality of the human resources in the area. (Rechnitzer et al. 2004). According to Florida (2002a) in the 21st century *not simply knowledge-based economy,* but rather *a creative economy was created*, which *is moved by a special segment of human recourses, the creative class.*

3. Richard Florida's basic model

Florida explains economic development with the so-called *3T model*, that is *Technology*, *Talent* and *Tolerance*. The point of this is that besides technology, as one of the – generally accepted – keys to economic growth, talent and tolerance are also considered as forwarders of the creative economic growth. According to Florida these three factors are necessary but in themselves not sufficient conditions of economic development. Florida claims that if a region has these three factors and they can also work together well then the region is able to show development, otherwise it falls behind (Florida 2002a):

1. *Technology, the first T*: About the first component of the 3T, technology, neoclassical Robert Solow says in his work awarded Nobel Prize that boosting economic development rather depends on the increase of the extent of technological progress than on income accumulation or market extension studied until then (Lengyel–Rechnitzer 2004). By technology, Florida means *the economic-technological development of the given area, which means the existence of high-tech industries*. He finds their presence a necessary but not sufficient condition for a region to be able to become a creative centre and to attract further creative people and businesses, which can generate dynamic economic growth and higher employment.
2. *Talent, the second T*: Talent as the second variable can be traced back to Nobel Prize winner Robert Lucas's theses emphasising the role of human capital. This key factor means *the quality of the human resources available*. Highly skilled people can create new knowledge. They can efficiently and creatively use their already existing and new knowledge, and make values by this. Higher education is not necessarily needed for creative work; however, most creative workers have degree. The members of the creative class usually work in knowledge-intensive industries and have extensive creative capital, which they use in their work as well as in

other fields of their lives. Universities as catalysts have an important part in producing the creative class.

3. *Tolerance, the third T*: The third, new factor of Florida's 3T model *may be its most important, decisive variable.* Tolerance itself can be approached in several ways, and this factor is one of the cornerstones of the creative class's way of life.

Florida regards tolerance as a factor that essentially influences the model, because technology and talent are variables which have already been known and mentioned as key factors, which were earlier used to describe the economic development of some regions as well. Compared to them, the appearance of tolerance in the model is a new feature. Since the aim for individualism is a common characteristic of the creative class members. They like being their own masters and defining their identity themselves in a way that it expresses their creativity. They are open-minded to diversity as a thing that is needed to display and enrich their creative potential (Florida 2004). *In creative economy the competitive regions recognise the creative workforce*, they accept that it has to be judged by expertise and not by appearance. "*Cities and also regions lose competitiveness if they do not reflect the new culture"* (Florida 2002a). According to the theory, *economic development depends on creative people's decisions on settlement*. And these people look for places which are colourful, tolerant and open to new ideas. This realisation makes it necessary to have tolerance appeared as a variable in the 3T model of the growth of the new and creative economy. Since tolerance helps unfold talent, which forwards economic development. Florida claims with this that there is a close connection between receptive and open cultural environment and economic development (Florida 2005).

Florida gives a detailed description about his method in none of his works, the reader has to make it out and interpret. Besides the method, I found deficiencies in presenting the indicators and the indicators mapping them. However, it has advantages and disadvantages as well. The advantage of the not completely defined method that studying and taking it as a starting point we can make our own 3T model of an examined country and region. Its disadvantage definitely is that mapping the method and its indicators step by step is impossible. Florida uses *one or two indicators to define each index* in most cases in his 3T model. The own model later contains the indicators that are relevant, but it is necessary to involve further indicators according to the characteristic features of the examined regions.

4. The international adaptations of the model

Florida's method and indicators can hardly be adapted in Hungary because of the American specialities, on the other hand, the aggregation level analysed by Florida

is too high compared to the microregional level I intend to analyse. I expect going through the benchmark examples to make answering these hypotheses easier.

The 10 international adaptations surveyed during the research involved extremely important experience relating the adaptability of the method in Hungarian micro-regions. Several authors found that while analysing European regional levels it is impossible to adapt the method totally, because *most of the indicators* belonging to the indicators which form Florida's 3T model *are unattainable from the database of different countries* (Andersen–Lorenzen 2005; Lengyel–Ságvári 2008; Hackler–Mayer 2008; Houston 2008; Zimmerman 2008). Several studies used much more indicators and complete indicator systems than Florida to achieve genuine results (Andersen–Lorenzen 2005; Clifton 2008; Hara 2008; Mellander–Florida 2007; Lengyel–Ságvári 2008) and used multivariable statistical analyses (Lengyel–Ságvári 2008; Sharp–Joslyn 2008). The other important experience is that in the analysis it is practical to make the overall examination of the selected regional level first, then after drawing the lessons and conclusions, *narrowing the involved regions down and making a further study (*Lengyel–Ságvári 2008; Andersen–Lorenzen 2005). This method can be entirely used for the regional level I have chosen, on the basis of the quite different development and characteristics of the certain domestic micro-regions.

All the studied international works are founded on Richard Florida's basic methodology and basic model, but we can state that each study without exception has something new which, after considering Florida's methodology carefully, is due to the characteristic features of the given country and the chosen regional level. After all this we can state that the international works approached the study of the creative class's regional position with *identical set of concepts, identical analytic methods and similar set of indicators*.

I can only use the experience of international adaptations based on Florida's methodology indirectly in my analysis, since the authors studied *cities, regions, city regions or workforce catchment areas* in international relation – except of Lengyel and Ságvári – which are not suitable for the regional level I intend to examine. From the international studies, only the city regions, the regional units used by Andersen and Lorenzen (2005), are similar to the microregional level which I have chosen as a basic unit for my analysis.

Naturally, the method I developed and intend to present is only a possible version of the way to interpret and analyse the spatial position of the creative class and the regional development in Hungarian micro-regions.

5. Analysis of Hungarian micro-regions

The majority of the reviewed analyses studied city region, workforce catchment area and microregion levels. The first two can correspond to the *intersectional region*

from the region concepts (Lengyel–Rechnitzer 2004), which is quite important and well interpretable from economic aspect, but which is very difficult to be mapped with statistical data, because its boundaries cannot be given accurately.

According to international practice, I would also apply workforce catchment areas or city regions as the basic unit of the analysis. In domestic data collection, micro-regions can be corresponded to workforce catchment areas most of all (Lukovics 2007; Lukovics 2008).

6. Used data

For developing the data, I collected basically the data belonging to the indicators that form Richard Florida's 3T model, and I founded on international and national studies. It is important to mention that the final data the analysis is based on *reached its final form after repeated refining.*

My work was made more difficult by *the limited amount of data available on microregional level* and the lack of *data* used in the basic model but *not collected in Hungary*. The overwhelming majority of the data regarding the 174 micro-regions is from the National Region Development and Region Organisation Information System. In addition, the number of patents are collected from the Hungarian Patent Office "PIPACS" industrial property rights database, the number of the public body members are from the Hungarian Academy of Sciences Public Body database, and I received the R&D data from the Central Statistics Office after personal data request.

While making the database I took several aspects into consideration, but I mostly aimed to use the latest data available in 2008 in the analysing part of the study. The database is built up from the *microregional data of 2006* and the *national census data of 2001*. It is true that the census data of 2001 shows the state 7 years before, but I could use only that because this is the latest complete data source which is available.

The database formed in this way contained *93 basic data*[2], from which the database that the analysis is based on was developed by making specific and rate indicators. In the next step I arranged the data *according to the 3T model*, thus Florida's Technology dimension was mapped by 11, Talent dimension by 26 and Tolerance dimension by 16 indicators. *I started the analysis with altogether 53 indicators*. It is to be noted that the disproportionate distribution of the indicator numbers belonging to the different dimensions does not mean the overrepresentation of the dimension mapped with more indicators. This is because rankings are made separately within each dimension and the final T index is made from the unweighted arithmetic average, so the number of indicators the dimensions consist of becomes indifferent during the calculation the final T index.

[2] Including the data used for the calculation of final T index correlation.

The analysis is not completely the adaptation of Florida's methodology and model, but based on the experience of the application of the basic model and its foreign use, an analysis which is supported by appropriate database and also provides methodological novelties compared to the basic model. In fact, I take the *main idea, the mental framework and the methodological milestones* and I adjust them to the characteristics of the regional unit I study. I mean by methodological novelties that *I use much more variables* to develop the model than Florida and the other authors, and I form the analysis to suit Hungarian characteristics.

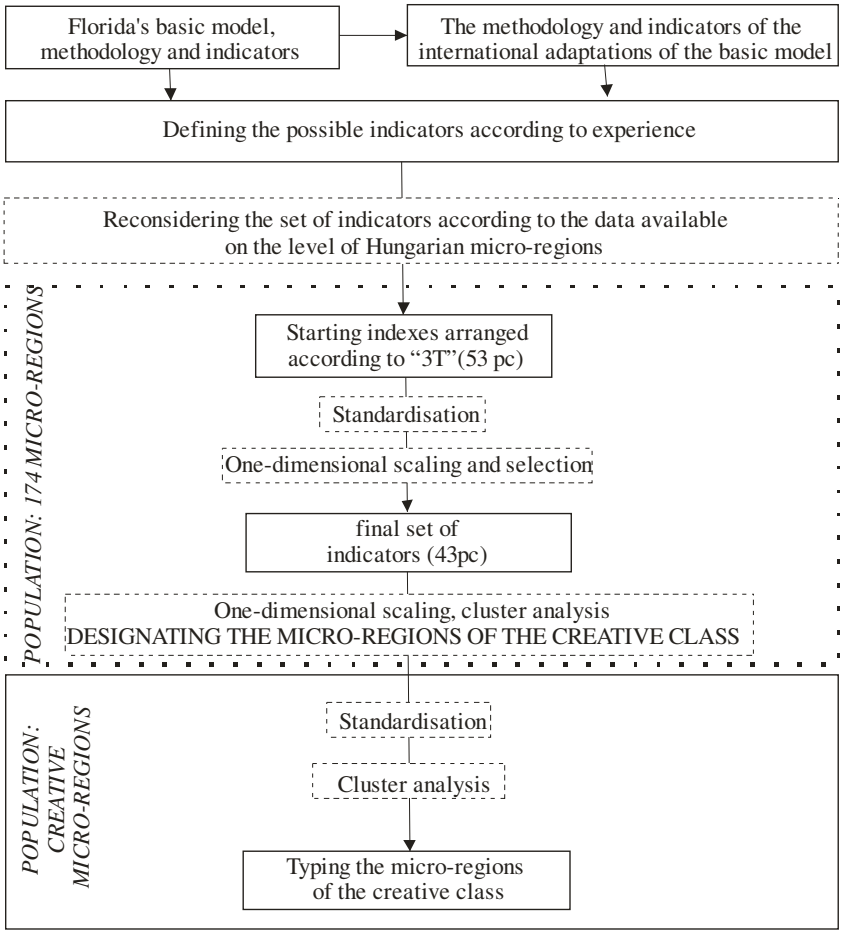
7. The methodology of the analysis

After standardising the 53 starting indicators of the database we intend to make the rankings of the micro-regions based on each T, then the final rankings based on it. On the basis of the reviewed analyses it became obvious that *each region of the whole area structure cannot be regarded as a place of where the creative class appears*, so several people attempted to sophisticate the analysis in a way that *they continued to make it with regions* which were *detached* according to certain aspects *as sampling population.*

The former idea, according to which the 174 micro-regions considered as sampling population should be narrowed down after examination to micro-regions where the creative class can more probably be found, follows also from Florida's works. According to him, it does not make sense even theoretically to examine regions with extremely different development level, thus *it is practical to detach and continue to examine the regions where the creative class is more concentratedly at present* than in the other ones. (Florida 2002b). The method which is suitable to narrow down is making a ranking with the help of a one-dimensional scale, and as a result the best performing micro-regions stand in the first places and the worst performing ones in the last positions. However, the main habitat of the creative class according to the one-dimensional scale can be chosen in quite a subjective way. In the interest of detaching as objectively as possible I use cluster analysis.

After this I expect that *the circle of micro-regions in which the creative class very probably appears can be selected from the 174 micro-regions*, thus in the following part of the analysis I consider the *n* pieces of micro-regions as sampling population. I continue differentiating the *n* pieces of micro-regions defined as the habitat of the creative class according to the 53 indicators with the help of cluster analysis (Figure 1). Finally, I refine the results with correlation calculation.

Figure 1. The logical system of the methodology of the analysis



Source: own creation

8. Final set of indicators

As I wrote in the methodological part, first I ran[3] a one-dimensional scaling on each T dimensions, as a result of which each microregion got a coordinate separately in Technology, Talent and Tolerance dimensions, on the basis of which they could be ranked. In the course of this examination all variables remained with appropriate S-stress value in the case of the es mapping Technology (S-Stress value: 0,03) and Talent (S-Stress value: 0,11) , but I had to select from the indicators in the case of Tolerance (S-Stress value: 0,2). After the selection with mathematical-statistical methods the *final database consisted of 11 indicators mapping technology, 25 indicators mapping talent and 6 indicators mapping tolerance, that is 43 indicators in total* (Chart 1).

The *final ranking number* based on the 3 dimensions together was formed by averaging the ranking numbers of the three one-dimensional scaling.

Chart 1. The final set of indicators

	The name of the indicators
TECHNOOGY	1. Number of patents per 10000 inhabitants in the microregion from 2000 to 2006 (pc)
	2. Number of R&D places per 10000 inhabitants, 2006 (pc)
	3. Investments of R&D places per 1000 inhabitants, 2006 (thousand HUF)
	4. Expenses of R&D places per 1000 inhabitants, 2006 (thousand HUF)
	5. Expenditure of R&D places per 1000 inhabitants, 2006 (thousand HUF)
	6. Actual total number of R&D places per 1000 inhabitants, 2006 (person)
	7. Actual number of scientific researchers in R&D places per 1000 inhabitants , 2006 (person)
	8. Number of public body members of the Hungarian Academy of Sciences per 10000 inhabitants, 2006 (person)
	9. Number of computers in public educational institutions per 1000 inhabitants, 2006 (pc)
	10. Number of work places with Internet supply in public educational institutions per 10000 inhabitants, 2006 (pc)
	11. Number of ISDN lines per 1000 inhabitants, 2006 (pc)
TALENT	12. Number of regular cultural activities per 1000 inhabitants, 2006 (person)
	13. Number of participants in regular cultural activities per 1000 inhabitants, 2006 (person)
	14. Number of students in higher education in all departments according to the seat of the institution per 1000 inhabitants, 2006 (person)
	15. Number of teachers working in higher education according to the seat of the institution per 10000 inhabitants, 2006 (person)
	16. Number of visitors to permanent theatres per 1000 inhabitants, 2006 (person)
	17. Number of registered businesses in health and social supply national economy sector per 10000 inhabitants, 2006 (pc)
	18. Number of registered businesses in education national economy sector per 10000 inhabitants, 2006 (pc)
	19. Number of registered businesses in financial transmission national economy sector per 10000 inhabitants, 2006 (pc)
	20. Number of registered readers in libraries of workplaces, higher education and other institutions per 1000 inhabitants, 2006 (person)

[3] I used the SPSS version 13.0 in my analyses.

	21. Number of cultural events per 1000 inhabitants, 2006 (pc)
	22. Number of participants on cultural events per 1 inhabitant, 2006 (person)
	23. Number of visitors to museums per 1000 inhabitants, 2006 (person)
	24. Number of other intellectual workers per 1000 inhabitants, 2001 (person)
	25. Number of workers in service sector per 1000 inhabitants, 2001 (person)
	26. Number of workers in service character industries per 1000 inhabitants, 2001 (person)
	27. Number of managerial and intellectual workers per 1000 inhabitants, 2001 (person)
	28. Number of people having qualifications in higher education per 1000 inhabitants, 2001 (person)
	29. Number of employees in financial activities per 1000 inhabitants, 2001 (person)
	30. Number of employees in health and social supply per 1000 inhabitants, 2001 (person)
	31. Number of employees in estate business and economic service per 1000 inhabitants, 2001 (person)
	32. Number of employees in education per 1000 inhabitants, 2001 (person)
	33. Number of employees with service occupations per 1000 inhabitants, 2001 (person)
	34. Number of employees with managerial and intellectual occupations per 1000 inhabitants, 2001 (person)
	35. Number of employees with other intellectual occupations per 1000 inhabitants, 2001 (person)
	36. Income of intellectual works per 1 inhabitant, 2006 (thousand HUF)
	37. Number of medical specialists per 1000 inhabitants, 2006 (person)
TOLERANCE	38. Number of immigrations per 1000 inhabitants, 2006 (person)
	39. Number of dependent men per 1000 inhabitants, 2001 (person)
	40. Number of single men and women over 15 per 1000 inhabitants, 2001 (person)
	41. Number of divorced people over 15 per 1000 inhabitants, 2001 (person)
	42. Number of ethnic minority members per 1000 inhabitants, 2001 (person)
	43. Number of families based on common-law relationship per 1000 inhabitants, 2001 (pc)

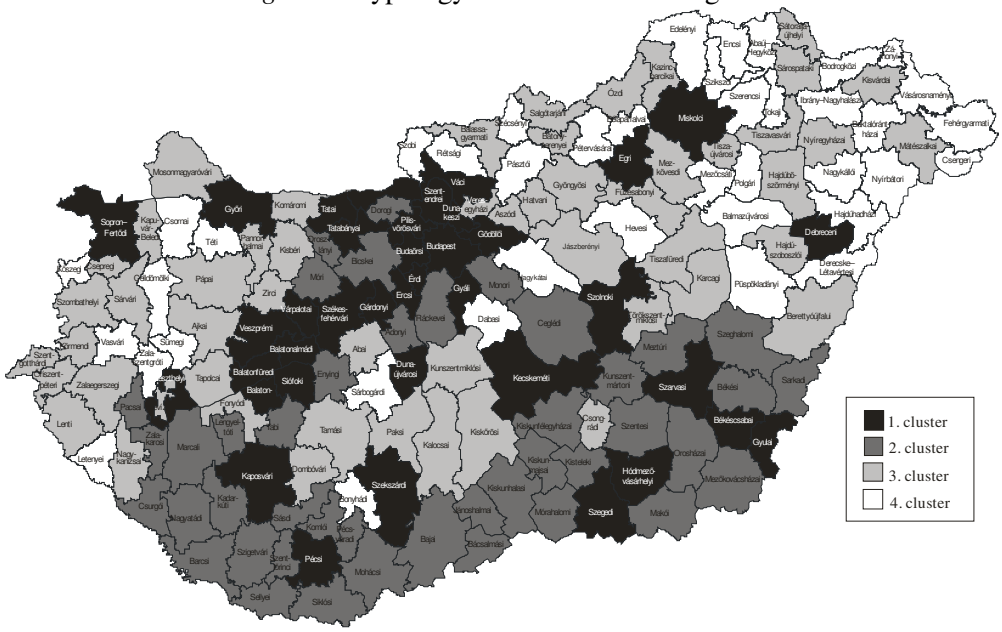
Source: own creation

9. Designating the creative micro-regions

As I have already mentioned, we cannot unambiguously determine the micro-regions considered to be the most probable habitat of the creative class with the help of one-dimensional scaling, since we cannot draw arbitrarily the limit after one microregion of the rankings. To be able to separate the creative and less prospering micro-regions, it is practical to use cluster analysis.

I chose *hierarchical clustering* for my analysis, because there had not been any former guiding regarding the number of clusters to create. Based on the structural chart of the hierarchical clustering procedure four clusters could be identified (Figure 2).

Figure 2. Typology of the 174 Micro-regions



Source: own creation

Those micro-regions got into the first cluster which are the most outstanding from the aspect of technology, talent and tolerance as well. The second, third and fourth clusters are built up from micro-regions that have lower value than the average according to at least one of the T dimensions, thus these micro-regions – as non-creative ones – are left out of the further part of the analysis. So in the following I continue analysing exclusively the micro-regions that form the first cluster.

10. Typing the creative micro-regions

The new sampling population to be studied is the 38 micro-regions belonging to the first cluster. The indexes used to analyse the new sampling population are identical with the ones used in analysing the 174 micro-regions, so I study the 38 micro-regions with the help of 43 indicators.

After standardising the indicators belonging to the 38 micro-regions once more, I create additional groups by means of cluster analysis, through the analysis of which I expect to be able to differentiate further and make the position of the creative class more accurate.

Since the number of clusters to be created – similarly to the analysis on the 174 micro-regions – was not definable beforehand, I ran *hierarchical clustering program*

again. It can be gathered from the merging chart of the clusters that designating 4 clusters is justified this time as well. The 1st cluster includes 5 micro-regions, the 2nd cluster includes 5 micro-regions, the 3rd cluster includes 23 micro-regions and the 4th cluster includes 5 micro-regions (Chart 2).

Chart 2. The Typology of the Creative Micro-regions

Super creative micro-regions	**„Spill-over" driven micro-regions**	**Potentially creative micro-regions**	**Moderately creative micro-regions**
– Budapest	– Budaörsi	– Békéscsabai	– Balatonalmádi
– Debreceni	– Dunakeszi	– Dunaújvárosi	– Balatonföldvári
– Pécsi	– Érdi	– Egri	– Balatonfüredi
– Szegedi	– Pilisvörösvári	– Ercsi	– Keszthelyi
– Veszprémi	– Szentendrei	– Esztergomi	– Siófoki
		– Gárdonyi	
		– Gödöllői	
		– Gyáli	
		– Győri	
		– Gyulai	
		– Hódmezővásárhelyi	
		– Kaposvári	
		– Kecskeméti	
		– Miskolci	
		– Sopron-Fertődi	
		– Szarvasi	
		– Székesfehérvári	
		– Szekszárdi	
		– Szolnoki	
		– Tatabányai	
		– Tatai	
		– Váci	
		– Várpalotai	

Source: own creation

I put a great emphasis on finding the main characteristics of each cluster, so after examining the values belonging to each cluster of the charts in the output of the cluster analysis thoroughly and one by one, I attempt to name the four clusters in a way that the name can reflect the relation of the micro-regions in certain clusters to the 3T dimension (Chart 3).

Chart 3. The Typology of the 38 Creative Micro-regions

Cluster	Technology (T1)	Talent (T2)	Tolerance (T3)
Super creative micro-regions	high	high	high
„Spill-over" driven micro-regions	relatively low	relatively high	relatively high
Potentially creative micro-regions	relatively high	medium	medium
Moderately creative micro-regions	medium	relatively low	relatively low

Source: own creation

After running the cluster analysis, it is remarkably important for interpreting the results to *define the most important characteristics of the certain clusters* according to criteria based on which the certain micro-regions were arranged into clusters. Technically, this can be realised by listing the average value of each variable in the given cluster into one of the outputs of the SSPS. After analysing these values we can unambiguously define what characteristics the micro-regions have to get in the certain clusters, in this way the four basic types of creative Hungarian micro-regions could be created (Figure 3):

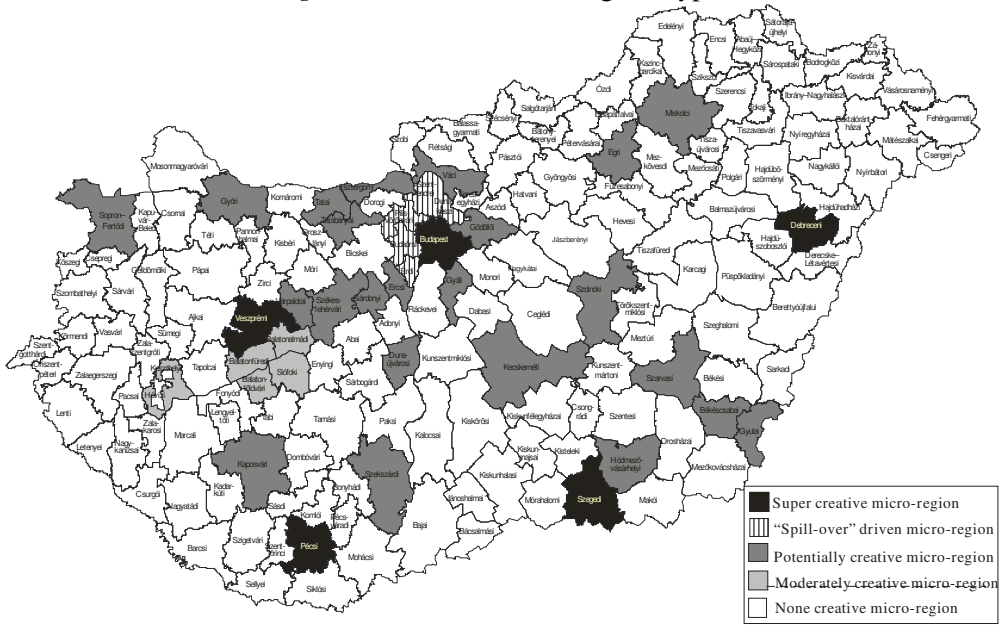
1. *Super creative micro-region*[4]: the 5 micro-regions where the variables of all the three T dimensions have the highest value. The micro-regions of Budapest, Debrecen, Pécs, Szeged and Veszprém stand out from the aspects of Technology, Talent and Tolerance. These micro-regions also stand in the first 5 places of the final rankings the one-dimensional scaling resulted in. The variables of all the three Ts show higher values than the average in this cluster than in the other three.
2. *„Spill-over" driven micro-region:* from the variables grouped according to the 3T those ones show a relatively high value which belong to the circle of Tolerance and Talent, while the variables of Technology have a relatively low value. This cluster includes the micro-regions belonging to the agglomeration ring of Budapest, bordering from north, north-west, west and south-west. Studying the social-economic processes nowadays we find that as a response to the urbanisation drawbacks being present in Budapest, on the one hand, the economic and political elite concentrated in Budapest moves to the agglomeration more frequently and commutes to Budapest to work, on the other hand, more and more businesses chose premises in settlements in the much more liveable and very close

[4] The super creative compound is from Florida's works. In addition, as to be noted, the micro-regions in this cluster compared to developed countries cannot be termed super creative.

agglomeration instead of Budapest. The creativity of the micro-regions of this cluster can be mainly attributed to the spill-over of the knowledge developed in Budapest.

3. *Potentially creative micro-region:* the variables in the Technology dimension of these micro-regions have a relatively high value; in addition, the variables of Talent are averagely high as well as in the case of the Tolerance dimension. The high value of the Technology variables is due to that in 90 per cent of the micro-regions in this cluster there is a seat or an affiliated department of some kind of higher educational institution. If we take a look at the final rankings of the one-dimensional scaling, the micro-regions belonging to the Potentially creative cluster are in the first third of the rankings on the basis of their ranking number.
4. *Moderately creative micro-region:* it represents micro-regions which have an average value in Technology dimension, while the variables of Tolerance and Talent dimensions are relatively low in value from the variables grouped according to the 3T dimensions. There are only 5 micro-regions in the cluster of the Moderately creative micro-region; however, they are close to each other geographically near Lake Balaton.

Figure 3. Creative Micro-region Types



Source: own creation

11. Conclusions

The study looked for the answer to the basic question that how it is possible to classify the Hungarian micro-regions according to the position of the creative class. The first statement based on the examination is that the *creative class cannot be mapped with one or two selected indicators; a complex indicator system is needed to be used*. The concept of knowledge-based economy is so complicated and complex that the analyses based on and started from it have to have a complex set of indicators so that the results drawn from the analysis can be correct.

Based on international studies and my own examination it can be stated that not all the 174 micro-regions have the "critical mass" of the creative class, because of this it is necessary to examine and type the so-called creative micro-regions separately. Based on the four groups which were formed after making the examination of the technology, talent and tolerance variables with one-dimensional scaling and the cluster analysis, the micro-regions that can be defined as creative micro-regions become unambiguously detachable. These detached, 38 creative micro-regions are the most outstanding in all the three dimensions.

As a result of the analysis, the 38 creative micro-regions form four groups according to the average extent of values of technology, talent and tolerance taken in each cluster and after evaluating these properties, named Super creative micro-regions, "Spill-over" driven micro-regions, Potentially creative micro-regions and Moderately creative micro-regions. The micro-regions involving cities with county rights are creative micro-regions with five exceptions (Nagykanizsa, Nyíregyháza, Salgótarján, Szombathely and Zalaegerszeg). These five micro-regions can be ranked among the non-creative micro-regions on the basis of technology, talent and tolerance dimensions together.

References

2007. évi CVII. törvény a települési önkormányzatok többcélú kistérségi társulásáról szóló 2004. évi CVII. törvény módosításáról.

Andersen, K. V. – Lorenzen, M. 2005: *The geography of the Danish Creative Class. A Mapping and Analysis*. Copenhagen Business School, Frederiksberg.

Bajmócy Z. :2008: A regionális innovációs képesség értelmezése és számbavétele a tanulás-alapú gazdaságban. In Lengyel I. – Lukovics M. (szerk.) *Kérdőjelek a régiók gazdasági fejlődésében*. JATEPress, Szeged, 26-46. p.

Clifton, N. 2008: The „creative class" in the UK: an initial analysis. *Geografiska Annaler Series B: Human Geography*, 1, 63-82. p.

DTI 1998: *Competitiveness White Paper*. Department of Trade and Industry, London.

Enyedi, Gy. 2001: Globalizáció és magyar területi fejlődés. *Tér és Társadalom*, 1, 1-10. p.

ESRC :2005: *ESRC Strategic Plan 2005-2010*, Swindon.

Florida, R. 2002a: *The rise of the creative class.* Basic Books, New York.

Florida, R. 2002b: The economic geography of talent. *Annals of the Association of American Geographers,* 92, 743–755. p.

Florida, R. 2004: *Cities and the creative class*. Routledge, New York.

Florida, R. 2005: *The Flight of the Creative Class: The New Global Competition for Talent*. HarperCollins, New York.

Hackler, D. – Mayer, H. 2008: Diversity, entrepreneurship, and the urban environment**.** *Journal of Urban Affairs*, 3, 273-307. p.

Hara, I. 2008: *A Study of Regional Creative Development and Policy*. Paper presented at the "Culture, Cohesion and Competitiveness: Regional Perspectives" 48th Congress of the European Regional Science Association, Liverpool.

Houston, D. 2008: Will attracting the „creative class" boost economic growth in old industrial regions? A case study of Scotland. *Human Geography*, 2, 133-149. p.

Huggins, R. – Izushi, H. 2008: Benchmarking the knowledge competitiveness of the globe's high-performing regions. A review of the World Knowledge Competitiveness Index. *International Business Journal,* 1-2, 70-86. p.

Kok, W. 2003: *Enlarging the European Union. Achievements and Challenges.* Report of Wim Kok to the European Commission, European University Institute Robert Schuman Centre for Advanced Studies, Florence.

KSH 2007: *Területi Statisztikai Évkönyv 2006*. Központi Statisztikai Hivatal, Budapest.

Leadbeater, C. 1999: *New measures for the New Economy Report*. Demos, London.

Lengyel, B. 2008: Tudásteremtés és ko-evolúció: az egyetem–gazdaság–kormányzat kapcsolatok globális és lokális vetületei. In Lengyel I. – Lukovics M. (szerk.): *Kérdőjelek a régiók gazdasági fejlődésében*. JATEPress, Szeged, 47-61. p.

Lengyel, B. – Leydesdorff, L. 2008: A magyar gazdaság tudásalapú szerveződésének mérése: az innovációs rendszerek szinergiáinak térbelisége. *Közgazdasági Szemle*, 55, 6, 522-547. p.

Lengyel, B. – Ságvári, B. 2008: A kreatív munkaerő feltérképezése Magyarországon. A „Tehetség, Technológia, Tolerancia, a kreatív gazdaság lehetőségei Magyarországon" *Kézirat.* Demos, Budapest.

Lengyel, I. 2003: *Verseny és területi fejlődés: térségek versenyképessége Magyarországon.* JATEPress, Szeged.

Lengyel, I. – Rechnitzer, J. 2004: *Regionális gazdaságtan*. Dialóg Campus, Budapest-Pécs.

Lengyel, I. 2007: Fejlesztési pólusok, mint a tudásalapú gazdaság kapuvárosai, *Magyar Tudomány*, 6, 749-758. p.

Leydesdorff, L. 2006: *The Knowledge-Based Economy: Modeled, Measured, Simulated*. Universal Publishers, Boca Rota, Florida.

Lukovics, M. 2007: A lokális térségek versenyképességének elemzése. *Doktori értekezés,* Szegedi Tudományegyetem Gazdaságtudományi Kar, Szeged.

Lukovics, M. 2008: *Térségek versenyképességének mérése*. JATEPress, Szeged.

Mellander, C. – Florida, R. 2007: *The Creative Class or Human Capital? Explaining regional development in Sweden*. CESIS, Stockholm.

OECD 1996: *The Knowledge-Based Economy. Science, Technology and Industry* Outlook 1996. Organisation for Economic Co-operation and Development, Paris.

OECD 2005: *Oslo Manual. Guidelines for Collecting and Interpreting Innovation Data.* Third edition. Organisation for Economic Co-operation and Development, Paris.

Rechnitzer, J. – Csizmadia, Z. – Grosz, A. 2004: A magyar városhálózat tudásalapú megújítóképessége az ezredfordulón. *Tér és Társadalom*, 2, 117-156. p.

Rechnitzer, J. – Smahó, M. 2005: *A humán erőforrások sajátosságai az átmenetben*. MTA Közgazdaságtudományi Intézet, Budapest.

Sharp, E. – Joslyn, M. 2008: Culture, Segregation, and Tolerance in Urban America. *Social Science Quarterly*, 3, 573-591. p.

TeIR 2008: Országos Területfejlesztési és Területrendezési Információs Rendszer. VÁTI, Budapest.

Varga, A. 2005: *Agglomeráció, technológiai haladás és gazdasági növekedés: A K+F térszerkezet makrogazdasági hatásainak vizsgálata*. MTA Doktori Értekezés, Pécs.

Varga, A. :2009: *Térszerkezet és gazdasági növekedés*. Akadémiai Kiadó, Budapest.

WBI 2007: *Building knowledge economies: advanced strategies for development*. World Bank, Washington, D.C.

Zimmerman, J. 2008: From brew town to cool town: Neoliberalism and the creative city development strategy in Milwaukee. *Cities*, 4, 230-242. p.

Spatial differences in inflation persistence in Hungary

Zsuzsanna Zsibók[1] *– Balázs Varga*[2]

On the basis of a disaggregated data set, we study inflation persistence in Hungary by focusing on regional cross-sectional variation. To this end, we use regional inflation series constructed from individual store-level price quotes. The price observations were collected for the CPI database at a monthly frequency and were provided by the Central Statistical Office of Hungary. In order to estimate inflation persistence, we assume time-varying-coefficient autoregressive models as described in Darvas - Varga (2007). The aim of the study is to describe the spatial patterns of Hungarian inflation persistence on the NUTS-3 level by using various exploratory spatial data analysis (ESDA) techniques. Previous researches found that while an apparent co-movement exists between the inflation rates in different regions, the decomposed inflation rates are quite disperse. We show that the overall level of inflation persistence decreased during the sample period, however, there are notable differences between the local patterns.

Keywords: inflation persistence, time-varying coefficient models, exploratory spatial data analysis

1. Introduction

Understanding the determinants and dynamic patterns of inflation is a crucial issue for modern economies since it can have implications for economic efficiency and wealth. There is a general consensus in the literature that inflation is a monetary phenomenon in the medium and the long run. Nonetheless, over shorter horizons various macroeconomic shocks temporarily move inflation away from its long-run trend. These shocks or their effects on inflation can be persistent and lead to persistent deviations of the level of inflation from its mean representing price stability. Inflation persistence is defined as the tendency of inflation to converge slowly towards the target level in response to shocks, therefore inflation persistence is a measure of

[1] Zsuzsanna Zsibók, junior research fellow, Centre for Regional Studies of the Hungarian Academy of Sciences, Hungary
[2] Balázs Varga, PhD candidate, Corvinus University of Budapest, Hungary, his work concerning this paper was supported by the OTKA grant No. 76868

Acknowledgement: The authors would like to gratefully acknowledge the Central Statistical Office of Hungary and the National Bank of Hungary for providing access to the data used in this study.

the speed of adjustment (Dossche - Everaert 2005). It is the backward-looking component of inflation and it indicates that to what extent inflation is anchored to past rates.

The most common measure of inflation persistence suggested in the literature is based on a univariate time-series model that assumes an autoregressive process, then, inflation persistence is measured as the sum of the autoregressive coefficients. Alternative measures are discussed in e.g. Pivetta - Reis (2007).

Franta et al. (2007) conducted a comparative analysis of Central and Eastern European countries including Hungary, however, their estimates are constrained due to data limitations. Lendvai (2005) found evidence of a high inflation persistence in Hungary while the weights of the lagged and forward-looking components are roughly equal. Estimations for the degree of persistencie in Hungary by the National Bank of Hungary (Magyar Nemzeti Bank 2008) suggested a high level of persistence, since agents assume permanent effects of monetary shocks.

A large-scale research project has been carried out by the Eurosystem central banks called the Inflation Persistence Network (IPN) which aimed to analyse micro price data in the euro area. The findings of the IPN were summarized by Dhyne et al. (2005), and a survey of the IPN evidence concerning price setting in the euro area were conducted by Altissimo et al. (2006). The latter found that inflation persistence had been moderate and had decreased in the respective period while the measured persistence had been different at the different levels of aggregation (euro area, national level, sectoral level, individual prices), namely, at higher levels of aggregation the adjustment of inflation had been slower. Further implications of aggregation were dealt with by Altissimo et al. (2007) in details.

Concerning spatial differences, the above mentioned studies focussed on cross-country comparisons, especially on comparative analyses of the eurozone countries or the US and the euro area. Our opinion is that investigating spatial differences in monetary (inflationary) processes is a relevant issue within a single country, as well. Such questions were discussed in the United Kingdom (Hayes 2005), in Italy (Fabiani et al. 2004, Veronese et al. 2005), in Spain (Alberola-Marqués 2001) and in the United States (Schunk 2005) so far. Broda - Weinstein (2008) compared cross-border and within border price differences in the US and Canada, and found that the degree of market segmentation across the border is similar to that within borders.

Beck et al. (2006, 2009) took a novel approach to the analysis of inflation dynamics within and across euro area countries since their comparative researches were conducted at the regional level. They employed a model where regional inflation dynamics were explained by common euro area and country specific factors as well as an idiosyncratic regional component. They found that while the area wide factors are strongly significant and have a high explanatory power, their loadings are different across different regions, which suggests that differences in regional inflation developments are partly due to area wide phenomena.

Micro-level analyses of inflation dynamics and price-setting behaviour have been conducted in Hungary, as well. The major studies in this field available in English language include Rátfai (2006, 2007), and Dusek - Budaházy (2008), however, these analyses cover only a fraction of the CPI basket. A detailed analysis of Hungarian data with a much higher coverage was conducted by Gábriel - Reiff (2008).

The aim of our study is to analyse spatial differences in inflation persistence in Hungary with micro price data. We use the store-level individual price quotes of the Central Statistical Office of Hungary. A detailed description of the database can be found in Gábriel - Reiff (2008) while the spatial differences were analysed by Reiff - Zsibók (2007). The rest of our paper is structured as follows: after introducing the database in section 2 and describing the methodology in section 3, section 4 presents the findings. Finally, section 5 concludes.

2. Data

We use a data set collected by the Central Statistical Office of Hungary (KSH), containing micro Consumer Price Index (CPI) data from December 2001 to June 2007 (altogether 67 consecutive months) where the observations were made at store-level. The CPI coverage changes each year. For example, in 2006 we have data about 770 representative items of the 896 on the CPI item list, while in 2007, we have 747 out of 876. The total weight of these representative items in the consumer basket is 70.122% in 2006 and 69.272% in 2007, indicating a reasonable coverage. The total number of observations is approximately 4.7 million.

Our original database contains 5 variables: product codes (5-digit representative item codes), prices, store codes, the dates of observations (year and month) and change codes. The "change code" indicates sales, normal price increases/decreases, price imputations, forced store and/or product replacements, changes of suppliers, changes in product outfits, and mistakes in previous months' quotes. The store codes include a store identifier and a location identifier in terms of county which allows us to apply county-level investigations.

Monthly inflation rates were computed as the weighted averages of the item-specific monthly changes in the average price, based on the log differences of average prices. The CPI figures were computed, differently from KSH's method, with unchanged weights based on year 2006 which allows us to eliminate the methodological biases since changing weights themselves otherwise could lead to breaks in the time series. We selected 2006 as a basis of the weights because coverage was the largest in that year. A disadvantage of using 2006 weights instead of changing weights is that it decreases comparability with KSH data.

Regional disaggregation was made at two levels: at county-level, including 19 counties and the capital city, Budapest; and at regional level, including the seven NUTS-2 regions of Hungary.

3. Methodology

3.1. Time-varying coefficient methods

There are various methods to measure inflation persistence (Dossche - Everaert 2005, Altissimo et al. 2006), nevertheless, estimations suggest a certain degree of uncertainty since the use of different methods resulted different degrees of inflation persistence. Using univariate autoregressive time series models, authors found very high degree of persistence, while those who studied the estimated autoregressive models for sub-periods identified by the break points found significantly smaller persistence (Darvas - Varga 2007). There is a growing empirical literature in support of the notion that inflation persistence could be changing in time (Gadzinski - Orlandi 2004) which implies that it is reasonable to use time-varying coefficient methods to study the persistence of inflation. Darvas - Varga (2007) argue that the new member states of the EU (including Hungary) went through substantial structural changes during the transition period and are still changing in a faster pace than mature economies. For this reason, they stress that one cannot assume constant parameters concerning inflationary processes.

In this paper, we apply the methods of Darvas - Varga (2007) for the data set described in Section 2. [3] The standard approaches in the literature measure inflation persistence in a single number which holds for the whole period and cannot say anything about the dynamics of the parameters. The authors compared the properties of two time-varying coefficient methods, (i) the maximum likelihood estimation of a state-space representation of the unobserved components models where the likelihood function was evaluated with the Kalman filter (Kalman 1960), and (ii) the distribution-free estimator of the same model via Flexible Least Squares (FLS) introduced by Kalaba - Tesfatsion (1988, 1989, 1990).

According to the widely-used procedures in the literature, we assume a first-order autoregressive process for inflation persistence:

$$\pi_{t,i} = \beta_{0,t,i} + \beta_{1,t,i} \cdot \pi_{t-1,i} + u_{t,i} \qquad t = 1, ..., T \text{ and } i=1, ..., n, \qquad (1)$$

where $\pi_{t,i}$ is the rate of inflation at time t and region (county) i, $\beta_{0,t,i}$ and $\beta_{1,t,i}$ are region (county) specific time-varying coefficients and $u_{t,i}$ denotes error terms. $T = 65$ and n can be 7 or 20 depending on the territorial units we choose to analyse (regions or counties). Our task is to give estimations for the sequences of the time-varying constant parameters and the time-varying first-order autoregressive parameters of the model. Traditional ordinary least squares (OLS) techniques are not suitable for estimating this model since we have to find time-varying coefficient sequences.

[3] Since the time series exhibit seasonality, first we carried out a seasonal adjustment procedure where seasonal effect is represented by the month averages (month-dummies) which were extracted from the individual observations and then the average of the 12 months' values were added.

To apply the FLS algorithm, two main assumptions have to be made. The first one states that the residual errors of the estimation are small, that is,

$$\pi_{t,i} - (\beta_{0,t,i} + \beta_{1,t,i} \cdot \pi_{t-1,i}) \approx 0 \qquad t = 1, ..., T \text{ and } i = 1, ..., n. \tag{2}$$

The second one states that the coefficients evolve slowly over time:

$$\beta_{o,t+1,i} - \beta_{o,t,i} \approx 0 \qquad t = 1, ..., T\text{-}1 \text{ and } i = 1, ..., n, \text{ for both } o = 0 \text{ and } 1 \tag{3}$$

We have to find coefficient sequence estimates which satisfy both assumptions in an acceptable manner. The idea of the FLS method is to assign two types of residual error to each possible coefficient sequence estimate. One consists of the sum of squared residual measurement errors, $r_M^2(\beta_i, T)$, regarding the first assumption (similarly to the ordinary least squares method), and the other is the sum of squared residual dynamic errors, $r_D^2(\beta_i, T)$, regarding the second, smoothness assumption, where β_i is the 2×1 vector of the region (county) specific coefficients to be estimated for each region *i*.

Our objective is to simultaneously minimize the two types of errors for each possible coefficient sequences. As described in Kalaba - Tesfatsion (1988), an incompatibility cost is specified for any β_i coefficient sequence, which is the weighted sum of the two kinds of errors:

$$C(\beta_i, \mu, T) = \mu \cdot r_D^2(\beta_i, T) + r_M^2(\beta_i, T). \tag{4}$$

By minimizing the incompatibility cost for β_i, given any weighting parameter $\mu > 0$, we will have the unique FLS estimate for β_i. Consequently, this procedure results a continuum number of solutions for a given set of observations, depending on the μ parameter. If μ approaches zero, the smoothness assumption will not be taken into consideration which results that while the sum of squared measurement errors will be brought down close to zero, the sequence of estimates will be volatile. If μ becomes arbitrarily large, we consider only the dynamic error. This is equivalent to the ordinary least squares solution, since r_M^2 is minimized subject to $r_D^2 = 0$, and the parameters will be constant. Consequently, the selection of the μ parameter is a rather important issue in this procedure. In this study, we compute FLS smoothed estimations with two specific μ parameters: $\mu = 10^{1.5}$ and $\mu = 10^{2.5}$, as suggested by the simulation results of Darvas and Varga (2007). The problem is a conditional minimization problem which can be solved with a dynamic programming algorithm. Kalaba and Tesfatsion suggest an estimation mechanism which is based on a recursion to get the members of the unknown sequence of parameters.[4]

[4] Practically, standard statistical software packages not allow us to perform this, therefore we use an own Gauss code.

3.2. *Exploratory spatial data analysis*

For the analysis of spatial differences, we use exploratory spatial data analysis techniques.[5] ESDA comprises techniques for exploring spatial data: summarizing spatial properties of the data, detecting spatial patterns in data, formulating hypotheses which refer to the geography of the data, identifying cases or subsets of cases that are unusual (outliers) concerning their location (Haining 2007). These methods provide measures of global and local spatial autocorrelation as well as spatial heterogeneity. The concept of spatial autocorrelation refers to the coincidence of value similarity with locational similarity (Anselin 2001). There is positive spatial autocorrelation when high or low values of a random variable tend to cluster in space and there is negative spatial autocorrelation when geographical areas tend to be surrounded by areas with very dissimilar values (Le Gallo - Ertur 2003). Spatial heterogeneity refers to the lack of structural stability of the various phenomena over space which may generate characteristic spatial patterns of economic development under the form of spatial regimes (core and periphery) (Anselin 1988, Le Gallo – Ertur 2003). We measure spatial autocorrelation with the Moran's I statistic (Cliff - Ord 1973) which is formulated as follows:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - \bar{x})(x_j - \bar{x})w_{ij}}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{5}$$

where $(x_i - \bar{x})(x_j - \bar{x})$ is the product of the variables' deviation from the mean, w_{ij} is the general element of the spatial weight matrix and n is the number of spatial units. The expected value of I is $-1/(n-1)$, which is in our case -0,0526 since the number of spatial units is 19. Values of I larger than E[I] indicate positive spatial autocorrelation and values smaller than E[I] indicate negative spatial autocorrelation. The significance test of this statistic is carried out by a randomization procedure.

The spatial weight matrix is the means of expressing spatial dependence which is an $n \times n$ matrix where w_{ij} is higher than 0 when location i and j are considered to be neighbours and is 0 otherwise, and the diagonal elements are also 0. At this stage of the analysis, we assume two kinds of contiguity matrices. The first one is a binary contiguity matrix based on 1st order neighbourhood which means that if two spatial units share a common border they considered to be contiguous and a positive value is assigned in the matrix. The other one is a binary spatial weight matrix with a distance-based critical cut-off where the neighbourhood is defined on the basis of a fixed distance[6] (Euclidean distance) which is the same for all spatial units.

[5] The software package used in the analysis is GeoDa 0.9.5-i.
[6] Distances are calculated between region centroids.

Since the selection of the weight matrix is a central decision in the analysis, we present the results with both types of spatial weights in each case. The most important difference between the two weight matrices is that the capital city, Budapest has only one neighbour with the 1st order contiguity-based neighbourhood definition: Pest county, and it has 6 neighbours with the distance-based neighbourhood definition.

The Moran's I statistic is a global measure of spatial autocorrelation and does not inform about the spatial structure of the neighbourhood effects. For this reason, we use „local indicators of spatial autocorrelation" (Anselin 1995) which help us identify spatial clusters with low or high values, as well as regions with low values surrounded by high values and regions with high values surrounded by low values (i.e. spatial instability). These spatial structures can be detected by the Moran scatterplot which plots the variable's values (x_t) against the spatially lagged values ($\boldsymbol{W}x_t$). Each point in the Moran scatterplot corresponds to a given region, and there are four quadrants of the Moran scatterplot. In the upper right quadrant we find locations with high values surrounded by neighbours with high values ("high-high"). In the lower left quadrant we find locations with low values surrounded by neighbours with low values ("low-low"). In these two quadrants spatial association is positive. In the upper left quadrant we find locations with low values surrounded by neighbours with high values ("low-high"), while in the lower right quadrants we find locations with high values surrounded by neighbours with low values ("high-low"). In these two quadrants spatial association is negative. The significance of the spatial clustering can be investigated by the local version of Moran's I statistic for each region i and year t ($I_{i,t}$). A positive value for $I_{i,t}$ indicates spatial clustering of similar values (high or low) whereas a negative value indicates spatial clustering of dissimilar values between a region and its neighbours (Anselin 1995, Le Gallo - Ertur 2003). Here, significance means that similar values are not randomly centered in certain locations, but there is a system behind the given spatial structure (Anselin 1995).

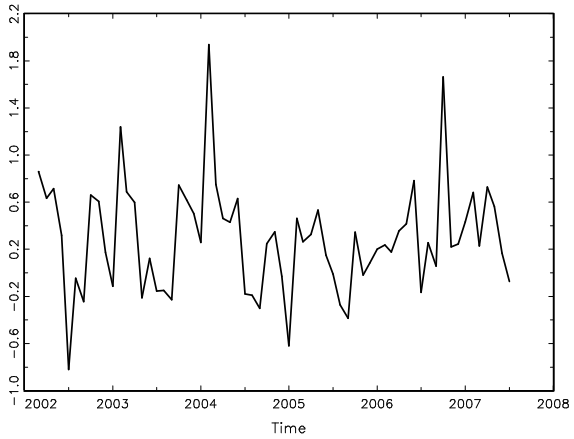
4. Results

In this section, we present the region-specific inflation persistence parameter series ($\beta_{1,i}$) and analyse their deviations from the county-level parameter series which is used as benchmark. To draw a more complete picture, we compare the dynamics of inflation persistence to the results presented by Reiff - Zsibók (2007) for the spatial differences of inflation and price levels. Additionally, we carry out a descriptive analysis of the local and global indicators of spatial autocorrelation.

4.1. Inflation rates

Monthly inflation rates were highly volatile in the sample period at the country-level. As depicted in Figure 1, it varied between –0.8 percent and 1.94 percent with a standard deviation of 0.48 percentage points. The data show seasonal fluctuations where local maximum points were in January each year, excepting 2006 (the highest monthly inflation rates were measured in May and June this year).

Figure 1. Country-level monthly inflation rates between January 2002 and June 2007 (percentages)



Source: own creation from KSH data

We can take a look at spatial differences by studying the cross-sectional standard deviation of regional-level monthly inflation rates (Figure 10).

The highest differences are measured in December and January each year, while the variation is relatively high in June and July each year, as well as in September 2006. The cross-sectional standard deviation varied between 0.05 and 0.46 percentage points in the sample period.

For average monthly inflation rates on the NUTS-2 level, global spatial autocorrelation, measured by the Moran's *I* statistic, did not prove significant for the whole period, however, in 11 out of 66 months (16.67%) it was significant at the 5 percent pseudo-significance level with the distance-based weights (in 8 out of 66 months, 12.12%, with the 1st order contiguity-based weights). The values of Moran's *I* did not show a clear tendency and were highly volatile if we used monthly inflation rates. Nevertheless, they had similar values with the two kinds of weight matrices, excepting June 2003 where this statistic indicated a positive global spatial autocorrelation ($p = 0.09$) with the 1st order contiguity-based weight matrix, and a negative global spatial autocorrelation ($p = 0.06$) with the distance-based weight matrix. The highest negative spatial autocorrelation was measured in August 2004 (Moran's $I =$

–0.2986 with the 1st order contiguity weight matrix, or -0.2917 with the Euclidean distance-based weight matrix) and the highest positive spatial autocorrelation was measured in May 2006 (Moran's $I = 0.2627$ with the 1st order contiguity weight matrix, or 0.2785 with the Euclidean distance-based weight matrix). On Figure 11, we compared the values of the Moran's *I* statistic to the cross-sectional standard deviation of monthly inflation rates.

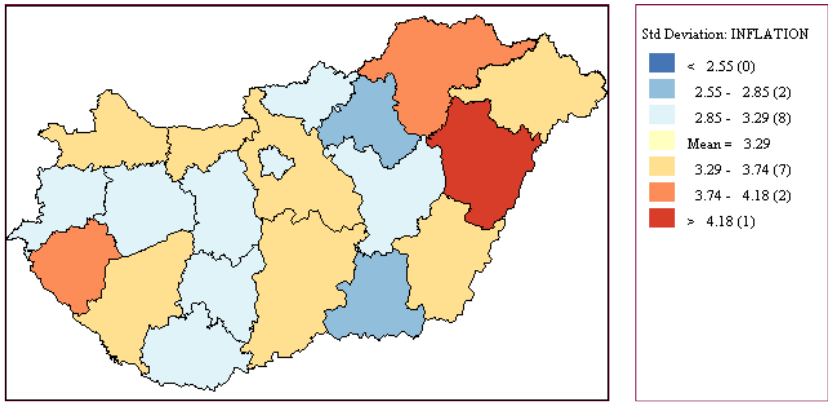
Tests of the local spatial autocorrelation for the whole period were significant in two cases: in Szabolcs-Szatmár-Bereg county at 1 percent level (low-high quadrant), and in Bács-Kiskun County at 5 percent level (high-low quadrant). We found significant spatial clustering of low values in 41 cases and significant spatial clustering of high values in 31 cases with the distance-based weights (in 30 and 33 cases, respectively, with the 1st order contiguity-based weights). Spatial instability was more common, we found significant values in the low-high quadrant in 42 cases and significant values in the high-low quadrant in 28 cases with the distance-based weights (in 43 and 39 cases, respectively, with the 1st order contiguity-based weights). For monthly inflation rates, the indicators of local spatial autocorrelation did not show a clear picture, therefore it would be hard to interpret these results at this stage of the analysis.

Moran's *I* statistics are less volatile for yearly inflation rates (see Figure 12) and the values are rather similar for the two different weight matrices. These statistics were significant at the 5 percent pseudo-significance level in only 5 and 4 out of 54 months (0.093% and 0.074%) with the distance-based and the 1st order contiguity-based weight matrices, respectively. The highest negative spatial autocorrelation was measured in February 2007 ($I = –0.2672$, with distance-based contiguity) and the highest positive spatial autocorrelation was measured in November 2005 ($I = 0.2360$, with distance-based contiguity).

For yearly inflation rates, local spatial autocorrelation was significant for Szabolcs-Szatmár-Bereg county, however, with an opposite sign as in the case of monthly inflation rates (high-high quadrant), and also for Bács-Kiskun and Pest counties (high-low quadrant). The results are much clearer than in the case of monthly frequency data. We found significant spatial clustering of low values in 46 cases and significant spatial clustering of high values in 34 cases with the distance-based weights (in 31 and 37 cases, respectively, with the 1st order contiguity-based weights). We found values in the low-high quadrant in 13 cases, and in the high-low quadrant in 22 cases with the distance-based weights (in 20 and 21 cases, respectively, with the 1st order contiguity-based weights). In more details, Budapest and Pest County had similar results to each other: they were at the low-low quadrant at the first 6 months of the year 2003, after it, they moved to the high-low quadrant at the end of this year. Later, much less significant values were found for these two counties. For Baranya County we found significant values in the first half of 2003 in the high-high quadrant. Significant clustering of low values were found for Bács-Kiskun County, Borsod-Abaúj-Zemplén County, Csongrád County and Nógrád

County. Szabolcs-Szatmár-Bereg County was in the high-high quadrant until March 2005, after it, it moved to the low-high quadrant of the Moran scatterplot. See Figure 2 in support of these findings.

Figure 2. Average county-level yearly inflation rates between January 2003 and June 2007 (percentages)

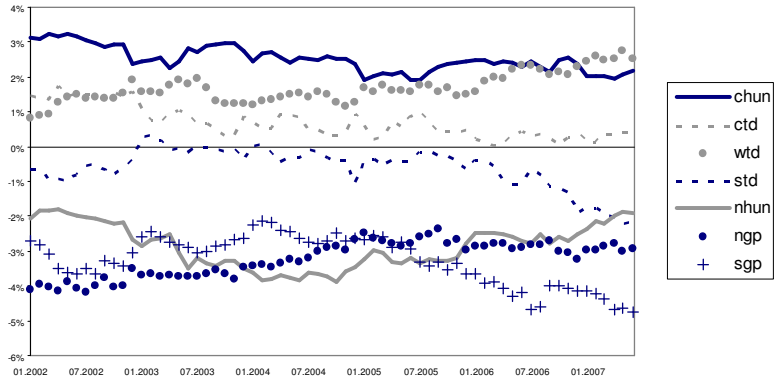


Source: own creation from KSH data

4.2. Price levels

The presented findings raise the question whether spatial differences in inflation rates can be explained by spatial price level convergence. If this were the case, inflation were higher in regions with relatively lower initial price levels than in regions with relatively higher initial price levels. This would cause price levels to approximate each other during the period. To investigate this question, we depict the deviation of regional-level prices from the country-level where country level is represented by 0 (Figure 3).

Figure 3. Regional-level relative prices (percentages)



Source: own creation from KSH data
Notes: chun: Central Hungary, ctd: Central Transdanubia, wtd: West Transdanubia, std: South Transdanubia, nhun: Northern Hungary, ngp: Northern Great Plain, sgp: Southern Great plain

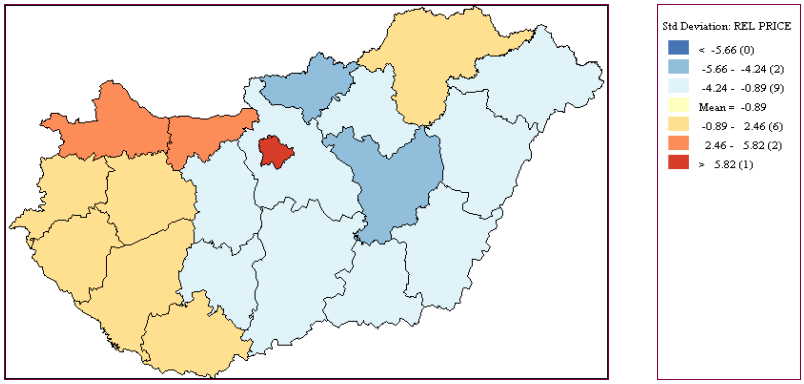
It is clear from Figure 3 that the relative price levels and the spatial differences can be regarded more or less stable during the sample period. The cross-sectional standard deviation of the relative prices support this finding (Figure 13), since it does not show a clear decreasing tendency which implies that the spatial differences have not decreased during the period.

The Moran's *I* statistic indicates positive spatial autocorrelation of the relative prices which was significant at the 5 percent pseudo-significance level in 49 out of 66 months (74.24%) with the 1st order contiguity-based weight matrix, however, in only 11 months (16.67%) with the distance-based weight matrix. Though Moran's *I* statistics are similar concerning the direction of their change, the values computed with the 1st order contiguity-based weight matrix exceed the values computed with the distance-based weights in each month. The differences were relatively high in 2003 and 2004 and decreased after that period. The highest level of global spatial autocorrelation was measured in July 2006 (Moran's $I = 0.2987$ with the 1st order contiguity-based weights) and in June 2007 (Moran's $I = 0.2641$ with the distance-based weights). The values were relatively high until August 2003, then stabilised at a relatively lower level until April 2005 while after that period they began to increase (see Figure 13 in the Appendix).

By investigating the local indicators of spatial autocorrelation we found significant clustering of high values in the case of Komárom which was significant in each month and in Győr-Moson-Sopron, Vas and Veszprém. Significant clustering of low values was found in Békés, Csongrád and Hajdú-Bihar counties. In Budapest, we found significang values in the high-low quadrant in 22 months. These results

indicate that high relative price levels are significantly clustered in the northwestern part of Hungary while low relative price levels are significantly clustered in the southeastern part of Hungary. Budapest, with a high relative price level, is surrounded by counties with generally low relative price levels (Figure 4).

Figure 4. Average relative price levels between 2002-2007 (percentages)



Source: own creation from KSH data

4.3. Inflation persistence

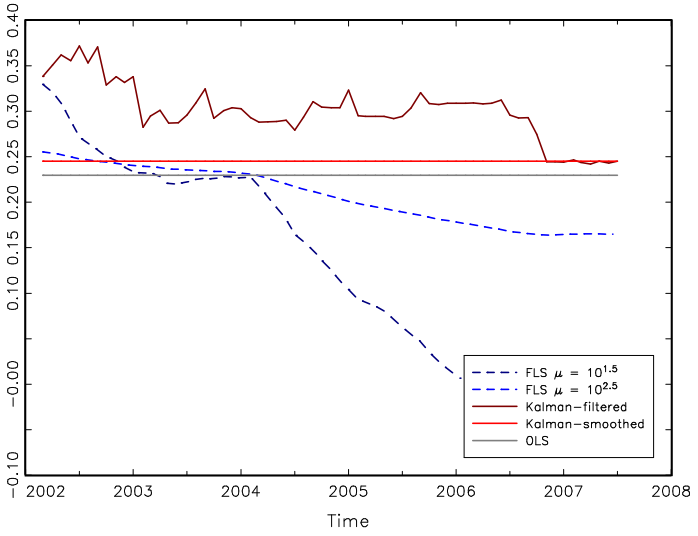
Figure 5 shows the country-level inflation persistence computed from monthly data with different estimation methods. OLS estimates yielded a parameter value of 0.23. This method assigns a single parameter value to the whole sample period and cannot reflect the shifts in inflation persistence. The results of each time-varying coefficient estimation methods indicate that inflation persistence have decreased in the observation period.[7] This means, that although inflation have generally risen, the influence of previous periods declined. This may refer to the phenomenon that agents are basically forward-looking, and generally anticipate future changes, or can be a result of the VAT changes of the years 2004 and 2006.

As described in Section 3, we compute FLS estimations with two different weighting parameters. With $\mu = 10^{1.5}$, the country-level estimated values declined from 0.33 to –0.06, so this method, in contrast to other methods, yielded even negative AR(1) parameters in the period after January 2006. Setting $\mu = 10^{2.5}$ gives a much smoother, though continuously declining trend, where inflation persistence declined from 0.26 to 0.16.

[7] In this study, we focus on the results of the FLS estimation, while results yielded from Kalman filtering and Kalman smoothing are presented only for comparison. Note that with Kalman filtering, we can estimate confidence intervals around the parameter values. These confidence intervals are relatively wide which suggests that we cannot state that persistence declined significantly.

The dynamic pattern of inflation persistence changed during the period, since it declined in a relatively fast pace before April 2003, then it declined much more slowly and it declined steeply until June 2006 which is followed by a more stable period.

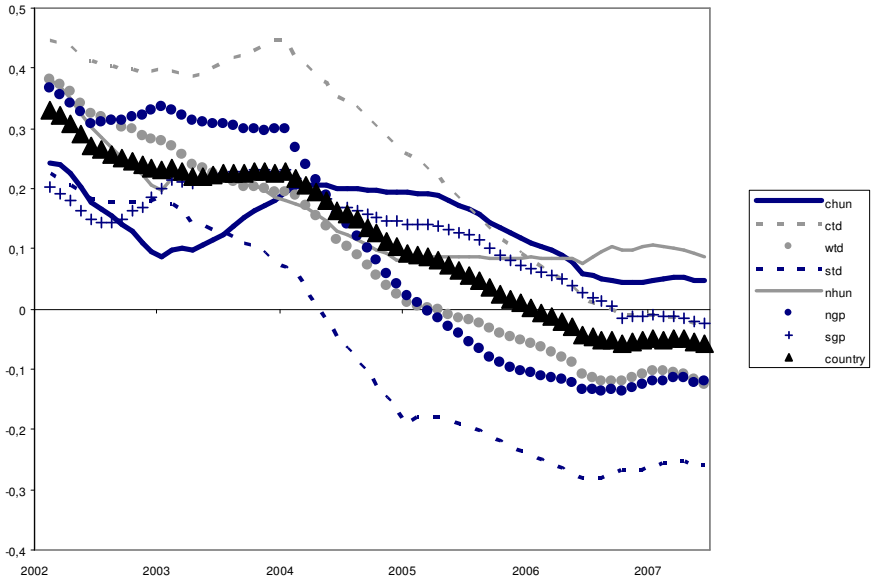
Figure 5. Estimated AR(1) parameters of the country-level monthly inflation series



Source: own creation from KSH data

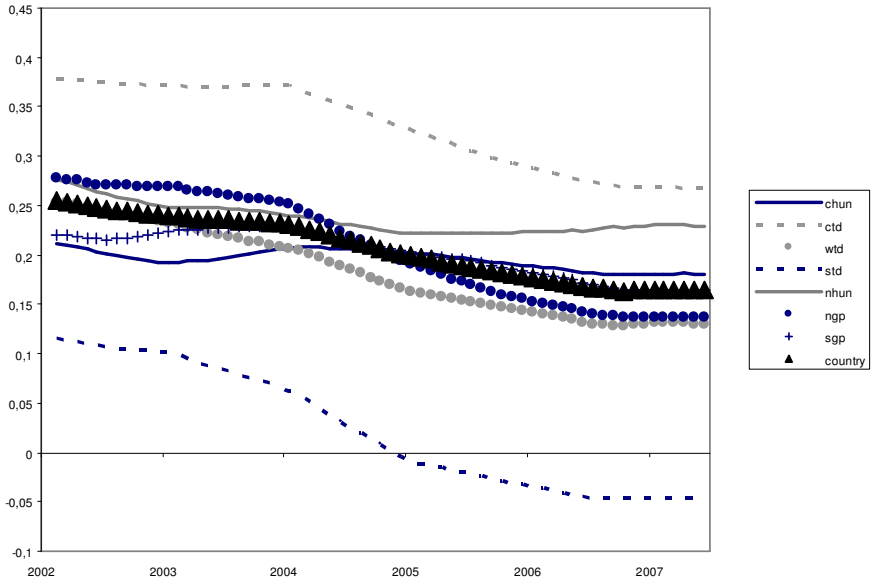
In what follows, we investigate spatial differences in inflation persistence (Figures 6a and 6b). Similarly to the patterns of price levels, spatial differences are relatively large and stable in time in Hungary. This is even more characteristic of the FLS estimation with $\mu = 10^{2.5}$, since parameter estimations are less volatile in time. Our general findings are more or less reflected by the results obtained at the level of the individual regions, too, however, temporal differences are high. In Central Transdanubia, inflation persistence exceeds the country-level persistence during the whole period, and with $\mu = 10^{2.5}$, it is the highest between the regions. With $\mu = 10^{1.5}$, it is also the highest in the first 42 months of the period. Similar tendencies can be found in South Transdanubia, though with an opposite sign, since persistence is below the country level during the whole period. With $\mu = 10^{2.5}$, persistence is the lowest between the regions, and with $\mu = 10^{1.5}$, it is the lowest in the last 48 months of the period. South Transdanubia is the only region where we obtained negative AR(1) parameters with $\mu = 10^{2.5}$, and, with $\mu = 10^{1.5}$, persistence becomes negative in this region long before the other regions (in March 2004).

Figure 6a. Estimated AR(1) parameters of the regional-level monthly inflation series with FLS method ($\mu = 10^{1.5}$)



Source: own creation from KSH data
Notes: chun: Central Hungary, ctd: Central Transdanubia, wtd: West Transdanubia, std: South Transdanubia, nhun: Northern Hungary, ngp: Northern Great Plain, sgp: Southern Great plain

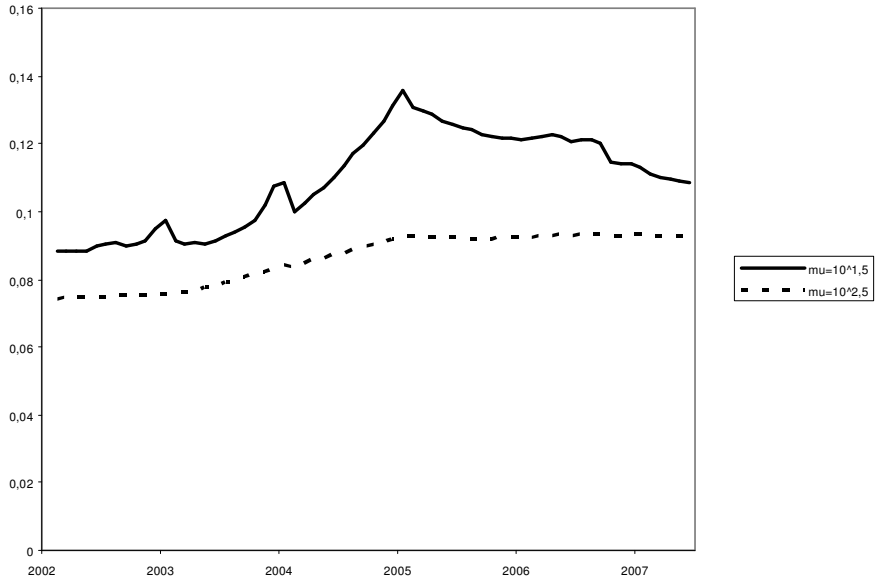
Figure 6b. Estimated AR(1) parameters of the regional-level monthly inflation series with FLS method ($\mu = 10^{2.5}$)



Source: own creation from KSH data
Notes: chun: Central Hungary, ctd: Central Transdanubia, wtd: West Transdanubia, std: South Transdanubia, nhun: Northern Hungary, ngp: Northern Great Plain, sgp: Southern Great plain

Figure 7 indicates the cross-sectional standard deviation of the estimated inflation persistence parameters. With higher μ, the variation has lower values and, similarly to the parameter values, it is less volatile in time. The lowest cross-sectional standard deviation values were measured at the beginning of the sample period: 0.088 (with $\mu = 10^{1.5}$) and 0.074 ($\mu = 10^{2.5}$). The highest values of the cross-sectional standard deviation were 0.136 (with $\mu = 10^{1.5}$), measured in January 2005, and 0.093 (with $\mu = 10^{2.5}$), measured in August 2006. Increased spatial differences can be assumed by the end of the sample period, even though a somewhat decreasing tendency can be seen after January 2005. We could not find such clear tendencies neither in the case of relative price levels, nor in the case of inflation concerning temporal changes of spatial differences. Relative prices, with some exceptions, seem to be relatively constant in time, as well as spatial differences. Concerning inflation, its rate and the spatial differences are not stable in time, however, we cannot discover any clear tendency, therefore we cannot claim anything about their long-run direction.

Figure 7. Cross-sectional standard deviation of the AR (1) parameter of the regional-level monthly inflation series estimated with FLS method



Source: own creation from KSH data

Global spatial autocorrelation was significant at the 5% pseudo-significance level only in 6 out of 65 months (9.23%) with the distance-based weights, and was not significant in any cases with the 1st order contiguity-based weights concerning the AR(1) parameter of the inflation series with $\mu = 10^{1.5}$. The values of the Moran's *I* statistic computed with the distance-based weights were, in almost all cases, above the values computed with the 1st order contiguity-based weights. The highest values were measured in January 2003 ($I = 0.201$) and between August 2004 and December 2004 ($I = 0.2069$). It continuously increased after the beginning of the sample period until January 2003, as well as between January 2004 and December 2004, while it decreased between February 2003 and December 2003, and from January 2005 (Figure 14).

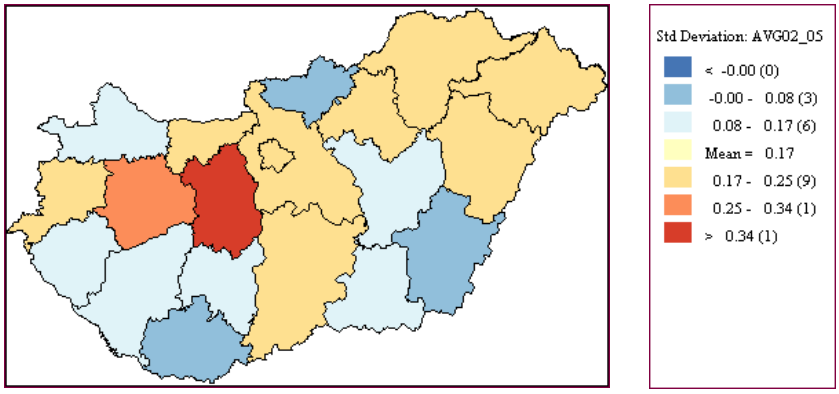
For the AR(1) parameter of the inflation series computed with $\mu = 10^{2.5}$, the differences between the results with the two types of weight matrices were even more pronounced, and the deviations seem quite stable. The Moran's *I* statistic was not significant at the 5% and at the 10% pseudo-significance level, at all. With the distance-based weights, it was positive for the whole period and varied between 0.047 and 0.11. The dynamic patterns are similar in the case of the 1st order contiguity-based weights, however, at lower values: Moran's *I* statistic varied between

–0.017 and 0.055. Global spatial autocorrelation tended to continuously decline during the whole sample period, excepting the period between January 2004 and January 2005 in which it rose (Figure 15).

For the FLS estimation of the AR(1) parameters of the inflation series with $\mu = 10^{1.5}$, significant local spatial autocorrelation was measured for 9 out of the 20 counties. Until February 2003, significant spatial clustering of high persistence values was present in Vas, Veszprém and Zala counties (the western part of Hungary). In the second half of the year 2002, Jász-Nagykun-Szolnok County was in the high-low quadrant of the Moran scatterplot, then, in January and February 2003 it moved to the low-low quadrant. Significant clustering of high values were measured in Komárom-Esztergom County from November 2003 until October 2005, in Fejér County from April 2004 to July 2006, in Nógrád County from November 2005 till the end of the sample period and in Győr-Moson-Sopron County in the first 3 months of 2003 (these counties are quite near to the capital city and/or to the western border of Hungary). As inflation persistence started to sink, Komárom-Esztergom County moved to the low-high quadrant between November 2005 and March 2006. From June 2006 till the end of the sample period, Jász-Nagykun-Szolnok County was in the low-high quadrant.

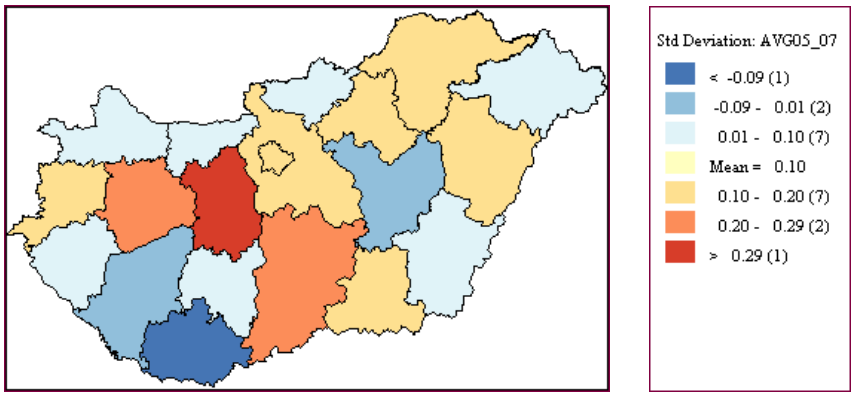
If we used $\mu = 10^{2.5}$, local spatial autocorrelation was significant only in two counties. Significant local clustering of high inflation persistence was measured in Komárom-Esztergom County between the beginning of the sample period and July 2005, and in Fejér County between November 2003 and July 2005, and again from May 2006 till the end of the sample period. Komárom-Esztergom County moved to the low-high quadrant after May 2006. (See Figure 8a and 8b.)

Figure 8a. Average values of the AR(1) parameter of the inflation series ($\mu = 10^{2.5}$) between February 2002 and July 2005



Source: own creation from KSH data

Figure 8b. Average values of the AR(1) parameter of the inflation series ($\mu = 10^{2.5}$) between August 2005 and June 2007



Source: own creation from KSH data

We examined whether there is a co-movement between the regional-level series and the country-level series by using correlation coefficients. The results are presented in Table 1.

Table 1. Correlation coefficients between the trend filtered regional-level series and the trend filtered country-level series of inflation persistence estimated with FLS method (*t* values are in parenthesis)

Region	**Central Hungary**	**Central Transd.**	**West Transd.**	**South Transd.**	**North Hungary**	**North Great P.**	**South Great P.**
$\mu = 10^{1.5}$	0.5285 (4.9415)	0.9233 (19.0818)	0.7720 (9.6408)	0.5449 (5.1574)	0.4892 (4.4525)	0.7223 (8.2901)	0.6747 (7.2562)
$\mu = 10^{2.5}$	0.4101 (3.5687)	0.9589 (26.8327)	0.8777 (14.5347)	0.7613 (9.3207)	0.4106 (3.5739)	0.9242 (19.2086)	0.8567 (13.1836)

Source: own creation from KSH data

Correlation coefficients were computed for the trend filtered series where trend filtering was conducted with the Hodrick-Prescott filter (Hodrick - Prescott 1997). The coefficients proved to be significant at conventional significance levels. The data suggests that in Central Hungary and North Hungary, inflation persistence was relatively far from the overall trend, similarly, though to a smaller extent, to South Transdanubia. A very high correlation is reported in the case of Central Transdanubia.

5. Conclusion

In this study, we investigated the dynamics of inflation persistence between January 2002 and June 2007 in Hungary by using micro price data. We intended to analyse not only temporal changes but also spatial differences at the level of the 20 NUTS-3 regions of Hungary. We took a novel approach by using time-varying coefficient methods in order to investigate temporal changes. We found that inflation persistence decreased between 2002 and 2007 at the country-level and also at the regional-level. This means that inflation is less anchored in the past at the end of the sample period than before, therefore, inflation is determined by the actual innovations rather than past values. In contrast to other empirical investigations in the literature, our results indicate a quite moderate level of persistence (far below 1). The spatial aspects of these phenomena were described by measures of global and local spatial autocorrelation which allowed us to investigate the neighbourhood effects. The interpretation of the results was highly dependent on the underlying series itself. For the highly erratic series of monthly inflation rates, we could not find a clear spatial structure, but for yearly inflation series, the indicator of global spatial autocorrelation evolved more smoothly. This is even more characteristic to the spatial autocorrelation indicators of the relative price levels and the inflation persistence coefficients, though the differences between the results obtained with the two kinds of weight matrices were higher.

A central element of these analyses is the selection of the spatial weight matrix. We aim to refine our research in the future by generating more sophisticated weight matrices and check the robustness of our results to the use of various different weight structures. We think that a reasonable definition of neighbourhood connections would be determined by accessibility weights which incorporate the influence of communication between regions such as transportation links (Anselin 1988). The estimations of the model presented in Section 3 provided information about the constant parameters which were not analysed so far and worth future investigations. Exploratory spatial data analysis techniques serve as a preliminary descriptive tool to look at the data prior to a more sophisticated spatial econometric formalization. This would be an obvious further direction of our research, nevertheless, a prerequisite for this is the extension of our data set with appropriate economic indicators. Due to data limitations, the highest level of regional disaggregation is the NUTS-3 level, therefore we cannot improve the results in this respect.

References

Alberola, E. - Marqués, J.M. 2001: On the evolution of relative prices and its nature at the regional level: The case of Spain. Journal of Regional Science, 41/3, 451-474. p.

Altissimo, F. – Ehrmann, M. - Smets, F. 2006: Inflation Persistence and Price Setting Behaviour in the Euro Area – a Summary of the IPN Evidence. European Central Bank Occasional Paper Series No. 46 (June, 2006), ECB, Frankfurt am Main.

Altissimo, F. – Mojon, B. - Zaffaroni, P. 2007: Fast Micro and Slow Macro: Can Aggregation Explain the Persistence of Inflation? Federal Reserve Bank of Chicago Working Paper No. 2007-02.

Anselin, L. 1988: Spatial Econometrics, Methods and Models. Kluwer Academic, Boston.

Anselin, L. 1995: Local indicators of spatial association-LISA. Geographical Analysis 27, 93–115. p.

Anselin, L. 2001: Spatial econometrics. In Baltagi B. (ed.): Companion to econometrics. Basil Blackwell, Oxford, 310-330. p.

Beck, G. W. - Hubrich, K. - Marcellino, M. 2006: Regional Inflation Dynamics Within and Across Euro Area Countries and a Comparison with the US. European Central Bank Working Paper Series No. 681., ECB, Frankfurt am Main.

Beck, G. W. - Hubrich, K. - Marcellino, M. 2009: Regional Inflation. Economic Policy, January, 141-184. p.

Broda, C. - Weinstein, D. E. 2008: Understanding International Price Differences Using Barcode Data. NBER Working Paper No. 14017, May 2008.

Cliff, A.D. - Ord, J.K. 1973: Spatial Autocorrelation. Pion, London.

Darvas, Zs. - Varga, B. 2007: Inflation persistence in the new members of the EU. manuscript.

Dhyne, E. et al. 2005: Price Setting in the Euro Area: Some Stylised Facts from Individual Consumer Price Data, European Central Bank Working Paper Series No. 524 (September 2005), ECB, Frankfurt am Main.

Dossche, M. - Everaert, G. 2005: Measuring Inflation Persistence – A Structural Time Series Approach. European Central Bank Working Paper Series No. 495 (June 2005), ECB, Frankfurt am Main.

Dusek, T. - Budaházy, Gy. 2008: Regional price differences in Hungary. In: 48th Congress of the European Regional Science Association, Liverpool, 2008, CD-ROM.

Fabiani, S. - Gattulli, A. - Sabbatini, R. 2004. The pricing behaviour of Italian firms: New survey evidence on price stickiness. European Central Bank Working Paper Series No. 333 (April 2004), ECB, Frankfurt am Main.

Franta, M. - Saxa, B. - Smidkova, K. 2007: Inflation persistence. Euro area and new member states. European Central Bank Working Paper Series No. 810 (September, 2007), European Central Bank, Frankfurt-am-Main.

Gábriel, P. - Reiff, Á. 2008: Price setting in Hungary – A store level analysis. manuscript, MNB.

Gadzinski, G. - Orlandi, F. 2004: Inflation persistence in the European Union, the Euro Area, and the United States. European Central Bank Working Paper Series No. 414 (November, 2004), ECB, Frankfurt am Main.

Haining, R. P. 2007: Spatial Data Analysis – Theory and Practice. Cambridge University Press, Cambridge.

Hayes, P. 2005: Estimating UK Regional Price Indices, 1974-96, Regional Studies, Vol. 39/3, 333-344. p.

Hodrick, R. - Prescott, E. C. 1997: Postwar U.S. business cycles: An empirical investigation. Journal of Money, Credit and Banking, Vol. 291/, February, 1-16. p.

Kalaba, R. - Tesfatsion, L. 1988: The Flexible Least Squares Approach to Time-varying Linear Regression. Journal of Economic Dynamics and Control, Vol. 12, 43-48. p.

Kalaba, R. - Tesfatsion, L. 1989: Time-Varying Linear Regression Via Flexible Least Squares. Computers and Mathematics with Applications, Vol. 17, 1215-1245.

Kalaba, R. - Tesfatsion, L. 1990: Flexible Least Squares for Approximately Linear Systems, IEEE Transactions on Systems. Man and Cybernetics, Vol. 20, 978-989. p.

Kalman, R.E. 1960: A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82/1. 35-45. p.

Le Gallo, J. - Ertur, C. 2003: Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe 1980–1995. Papers in Regional Science 82, 175–201. p.

Lendvai, J. 2005: Hungarian Inflation Dynamics. Occasional Papers No. 46, November 2005, Magyar Nemzeti Bank.

Magyar Nemzeti Bank 2008: Elemzés a konvergenciafolyamatokról. Magyar Nemzeti Bank, March 2008.

Pivetta, F. - Reis, R. 2007: The persistence of inflation in the United States. Journal of Economic Dynamics and Control, Vol. 31, 1326-1358. p.

Rátfai, A. 2006: How Fast Is Convergence to the Law of One Price? Very. Economics Bulletin, Vol. 6, Issue 10, 1-12. p.

Rátfai, A. 2007: The frequency and size of price adjustment: microeconomic evidence. Managerial and Decision Economics, 2007, Vol. 28/7, 751-762. p.

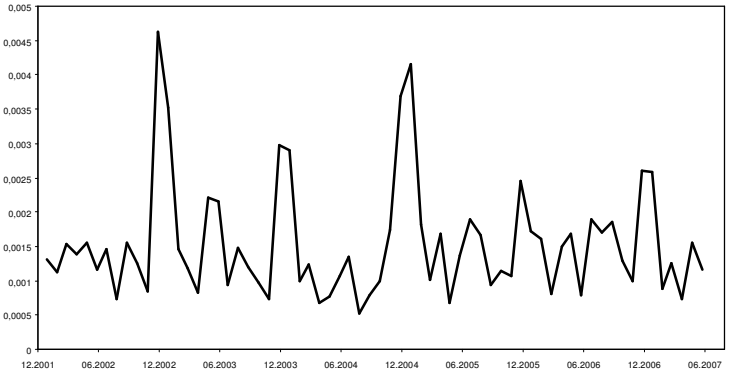
Reiff, Á. - Zsibók, Zs. 2007: Regional characteristics of price-setting and inflation in Hungary. A micro-level analysis, manuscript.

Schunk, D. L. 2005: The differential impacts of monetary policy: Are the differences diminishing? Papers in Regional Science, Vol. 84/1, 127–136. p.

Veronese, G. - Fabiani, S. - Gattulli, A. - Sabbatini, R. 2005: Consumer Price Behaviour In Italy: Evidence From Micro CPI Data. European Central Bank Working Paper Series No. 449, ECB, Frankfurt am Main.

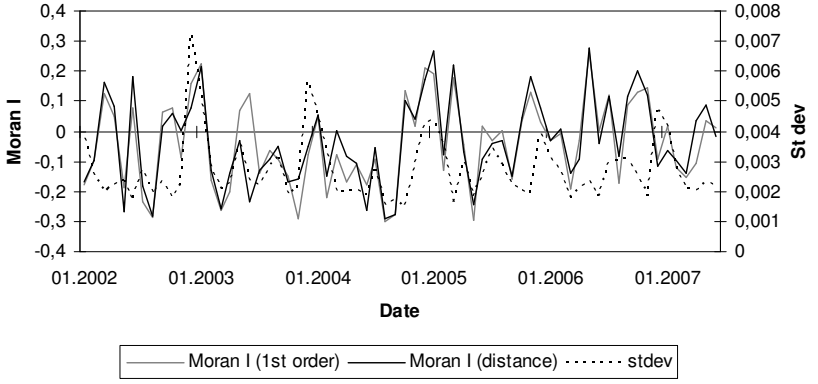
Appendix

Figure 10. Cross-sectional standard deviation of regional-level monthly inflation rates (percentages)



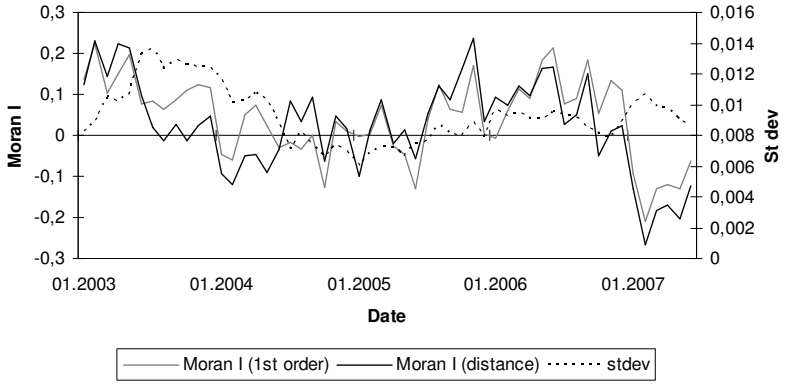
Source: own creation from KSH data

Figure 11. The co-movement of the Moran's I statistics and the standard deviation of monthly inflation rates



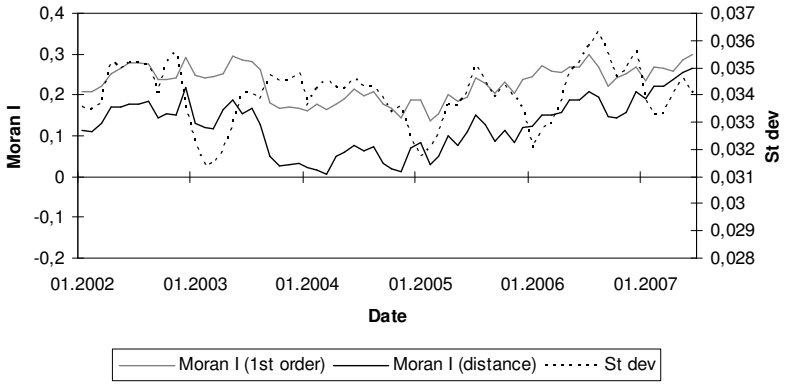
Source: own creation from KSH data

Figure 12. The co-movement of the Moran's I statistics and the standard deviation of yearly inflation rates



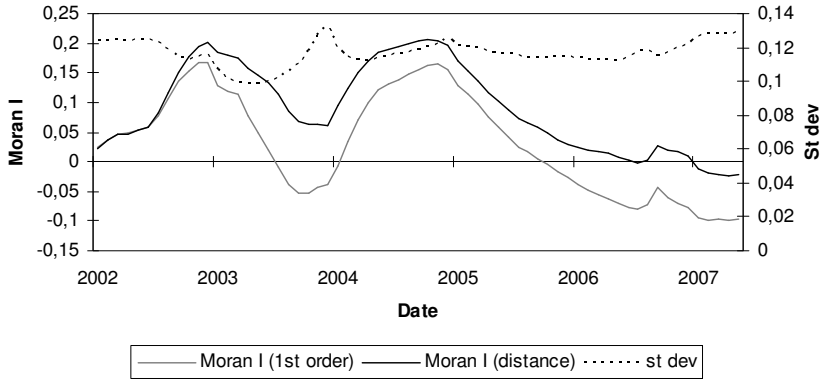
Source: own creation from KSH data

Figure 13. The co-movement of the Moran's I statistics and the standard deviation of relative price levels



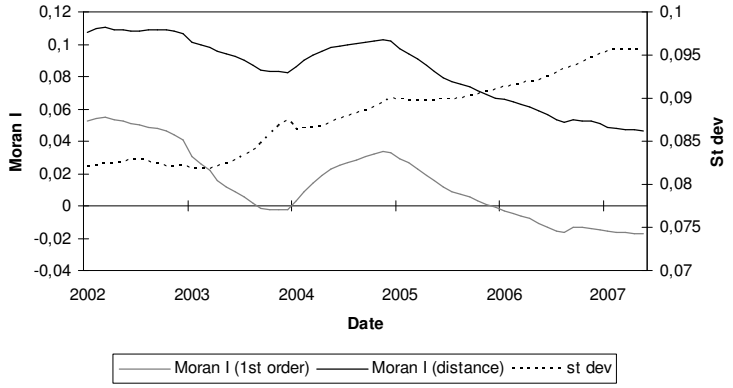
Source: own creation from KSH data

Figure 14. The co-movement of the Moran's I statistics and the standard deviation of the monthly inflation series' AR(1) parameter with $\mu = 10^{1.5}$



Source: own creation from KSH data

Figure 15. The co-movement of the Moran's I statistics and the standard deviation of the monthly inflation series' AR(1) parameter with $\mu = 10^{2.5}$



Source: own creation from KSH data

Where is delinquency nowadays? (Measuring the regional features of delinquency)

Katalalin Kovacsicsné Nagy[1]

The significance of regional aspects of criminality was already recognised in the early examinations of crime statistics. Also at the beginning of the 20th century the statistic data collection covered the place of commitment. The statistical process of these data restricted to the classification by counties. In the first part of the 20th century the task of law enforcement was relatively easy because the place of commitment and the criminal's domicile was the same settlement, the local policemen knew the potential criminals.
As a result of the development of motorisation crimes became more mobilised. By the end of the 20th century and the beginning of the 21st century not only the place of commitment and the domicile had got separated from each other but also the so-called moving crime occurred. In respect of crime prevention it means important information to show which settlements are the most endangered ones regarding crimes and which ones serve as the criminals' domiciles.
The system of indicators described in the study may provide help in the examination of the reasons for crimes because observation covers many factors that may be brought into connection with the crimes. I wish to illustrate the methods used to examine the regional features of crimes and the results through the data of a county (villages, towns, subregions).

Keywords: crime, crime prevention, crime-mapping

1. Introduction

As it is suggested by the subheading, in my lecture I want to deal with measuring the regional features of delinquency and - considering the venue of our conference - I am going to demonstrate the methods using the figures relating to Csongrád county. In the introduction I would like to give a short summary on why we deal with delinquency and how this negative mass phenomenon can be measured.

Why do we deal with delinquency? Nowadays we can hear a lot about delinquency. The newspapers, the radio and the television flood their readers, listeners and viewers with hair-raisers. Is the situation really as tragic as we are informed by the media? In the next few minutes I would like to present the objective reality as far as it is possible on the basis of scientifically grounded measurements.

[1] Katalin Kovacsicsné Nagy, Professor, Eötvös Loránt University, Faculty of Law, Budapest

Basic terms. How to measure? Before projecting lot of numbers onto the screen I must explain how we can measure delinquency, this negative social mass phenomenon. We are aware of the fact that length can be measured in cm, m, area can be measured in square meters, hectares or acres. But what measure can be used for delinquency? The development of the present measuring methods was preceded by long professional debates but now these methods have been being used for several decades. Measuring is carried out using three equally important units of measurement:

- Number of publicly indicted crimes become known
- Number of perpetrators

These two measures and the figures measured by them are provided by the United Crimestatistics Database of the Police (Investigating Authority) and the Public Prosecutor's Office.

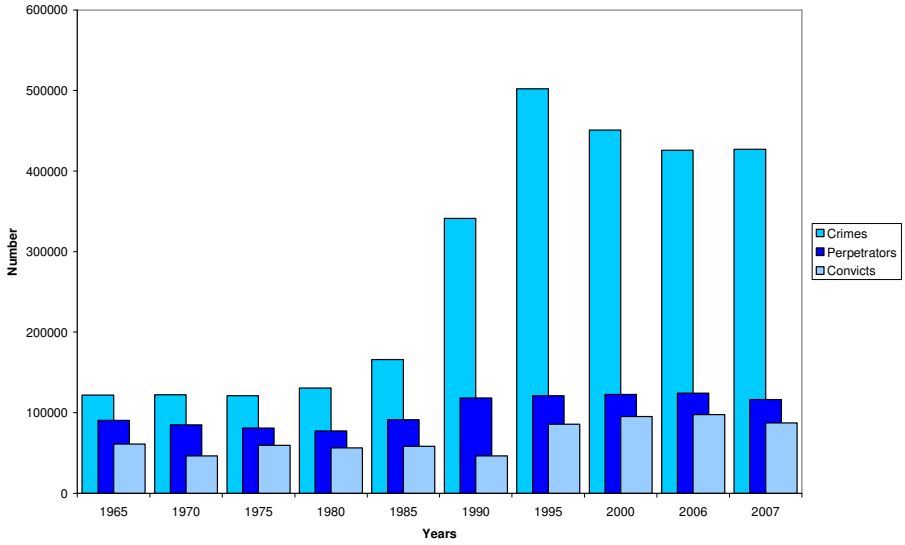
- The number of persons with definitive sentence

This latter one is the measure of the Judicial Statistics. Let me note that the former one was established in 1963, i.e. about 50 years ago and the latter one in 1880, i.e. about 130 yeas ago.

2. Time series and regional data

What does this long time sequence show? If we want to examine the delinquency in Csongrád county we have to start the examination from an earlier stage so that we can make comparisons. Let's have a look at a relatively long time sequence, from 1965 till today. It can be seen that the number of the committed crimes steadily fluctuated around one hundred and twenty thousand per year until 1980, the number of perpetrators and the persons with definitive sentence was approximately the same. It started to increase after 1980 and in 1990 it exceeded three hundred thousand and in 1995 it was more than five hundred thousand. It was at its peak in 1998 when the number of the publicly indicted crimes known exceeded even 600 thousand. It has been slowly decreasing since then and it has been fluctuating between 400 and 450 thousand during the past years. Thus we can say that - even if slowly - the criminal situation has been getting better since 1998. (Fig. 1.)

Figure 1. Number of crimes, perpetrators and persons with definitive sentence in Hungary, 1965-2007



Source: own creation

Measuring the regional features of delinquency. The significance of the territorial aspects of criminality was already realised in the early criminal statistical studies. As early as at the beginning of the 20th century the statistical data collection covered the place of commitment and after 1909 it recorded also the criminal's place of birth. Bud (1910) called these data and the indicators calculated from them criminal geography. He emphasized the importance of criminal geography in searching the causes of delinquency.

The statistical process of these data was restricted to the classification by counties and besides it the classification according to the type of the place of commitment was also considered (Budapest, towns, villages). This system of data collection was in use until the World War II and also the classification covered the counties and the characteristics of the places of commitment.

After World War II until the 60's the criminal statistics were secret. Afterwards and also today the judicial statistics process the data by counties. The classification according to the types of the place of commitment has been terminated because - due to the lots of changes in public administration - the data became incomparable. The number of towns and villages frequently varies and as such also the number of their population, thus comparison becomes pointless[0].

Year 1964 brought a new event in criminal statistics, ERÜBS (United Crimestatistics Database of the Police and the Public Prosecutor's Office) was introduced that year. The system observed the committed crimes including the place

of commitment and the criminals with their domiciles. Although the collected data ensured more varied possibilities for processing, ERÜBS made classifications only according to counties until lately also with regards to the place of commitment and the domicile.

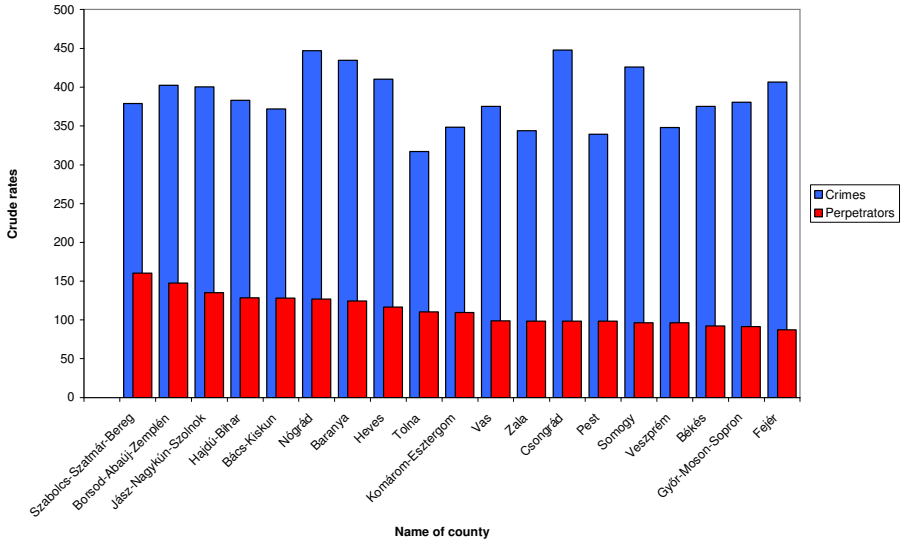
In the second part of the 20th century the composition of criminals considerably changed regarding the connection between the place of commitment and the domicile. As a result of the development of motorisation crimes became more mobilised. By the end of the 20th century and the beginning of the 21st century not only the place of commitment and the domicile had got separated from each other but also the so-called moving crime had occurred, then it was rare to find that a criminal committed the crime at their place of residence. By opening the boarders the mobility of criminals crossing the boarders became widespread, especially in case of the most severe crimes.

From the above it follows that practically it was not possible to examine the regional features of delinquency earlier because the various statistics (police, public prosecutor's office, judicial) observed and processed the data of the region only up to county level. The improvement of methods in computer science made it possible to extend observation and recording also to the level of villages and towns. Since 2001 ENYÜBS (United Crimestatistics Database of the Crime Investigation Service and the Public Prosecutor's Office) has been collecting data up to the level of settlements making it possible to establish "criminal geography".

Examination of the regional factors today. Let's have a look if there is any difference between the delinquencies of the various counties? Here we have to face another problem of measuring, because if we compare the above mentioned measuring data we might easily come to misleading consequences because both the area and the number of population of the counties are different and it might influence the formation of delinquency. Therefore hereinafter we use the number of crimes (perpetrators, sentenced persons) fallen to one thousand (or ten thousand) inhabitants in our calculations. These indicators (called crude criminality rates) can be realistically compared in case of comparing counties with different numbers of population or in case of comparing the delinquency of towns and villages. You can see in the figure that the delinquency in Csongrád county is higher than the national average, it is the highest in comparison with all other counties in 2007 regarding the committed crimes.

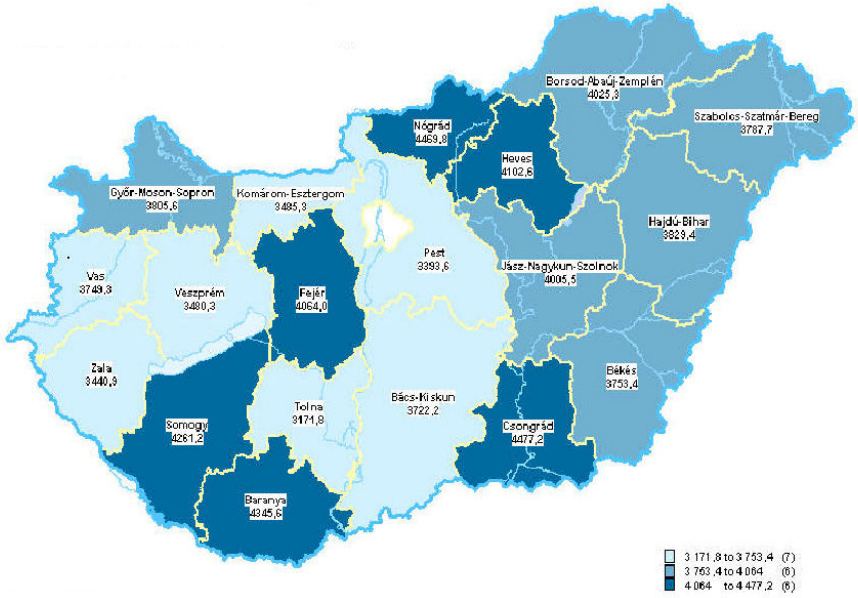
If we perform the comparison on the basis of the number of perpetrators we receive a better picture, the county will be ranked on the 13th place, the rate is lower only in five counties. (Fig 2., 3., 4.)

Figure 1. Number of crimes and perpetrators per 10000 inhabitants in each county, 2007



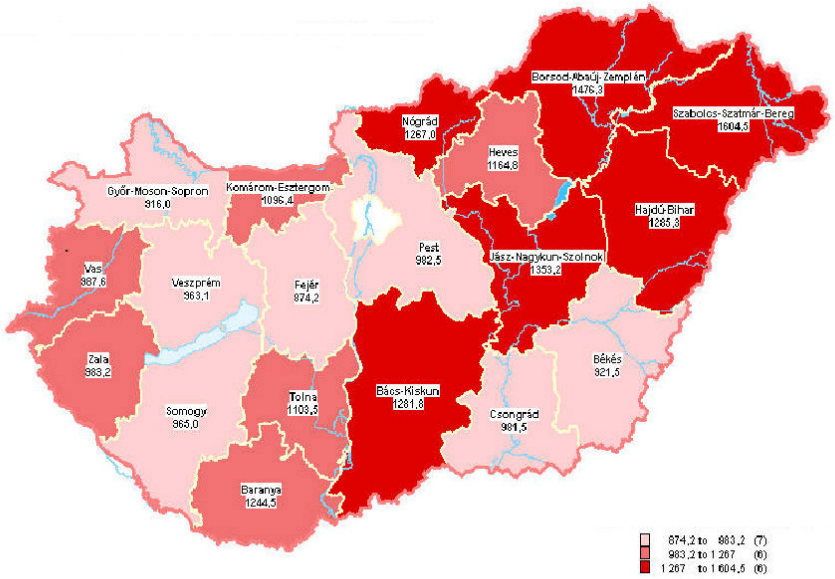
Source: own creation

Figure 2. The frequency of publicly indicted, known crimes, in each county, 2007



Source: IRM BBFO

Figure 3. The frequency of known perpetrators, in each county, 2007



Source: IRM BPFO

3. Analysis of data of Csongrád county

The delinquency even within Csongrád county is not homogenous. Let's compare the delinquency in the subregions of the county. (Table 1.)

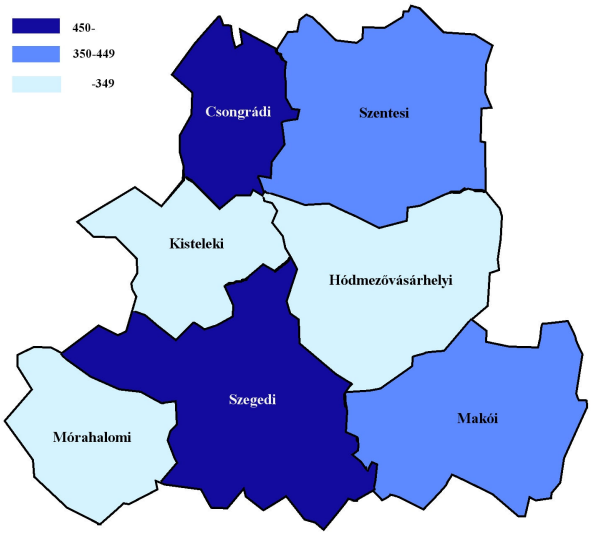
Table 1. Crude criminality rates and number of crimes committed by 100 perpetrators

Subregions	Perpetrators		Discovered crimes		Discovered crimes per 100 perpetrators	
	per 10000 inhabitants					
	2001	2007	2001	2007	2001	2007
Csongrádi	128,7	118,5	490,8	485,7	381,4	410,0
Hódmezővásárhelyi	110,5	110,1	242,9	329,3	219,8	299,2
Kisteleki	120,1	81,2	267,4	211,1	222,6	260,0
Makói	91,7	108,6	402,9	419,5	439,4	386,3
Mórahalomi	60,8	65,6	200,9	223,3	330,4	340,1
Szegedi	118,5	92,8	497,7	531,4	419,9	572,4
Szentesi	132,3	106,0	407,8	444,3	308,2	419,3
Average by counties	112,9	97,7	412,2	445,6	365,2	456,2

Source: own creation

The number of publicly indicted crimes become known and fallen to 10.000 inhabitants in the subregion of Csongrád, Szeged and Szentes is the highest in both observed years and the lowest in the subregions of Kistelek and Mórahalom. (Fig. 5.)

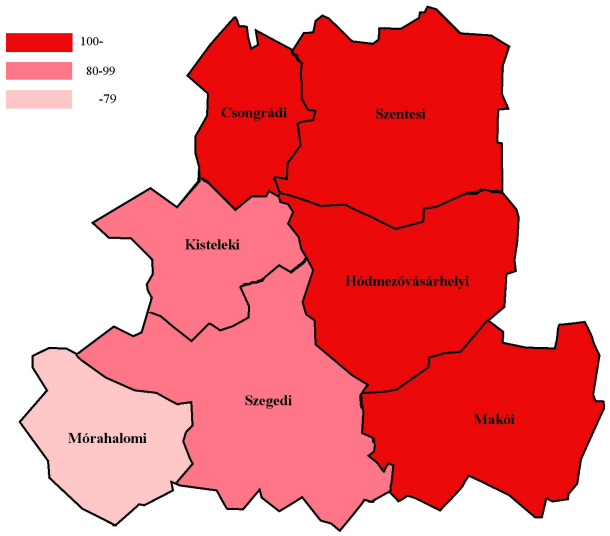
Figure 4. Number of publicly indicted crimes become known per 10000 inhabitants in subregions of Csongrád county, 2007



Source: own creation

The number of perpetrators per 10.000 inhabitants is the highest in the subregions of Csongrád and Szentes and the lowest in the subregions of Kistelek and Mórahalom. (Fig. 6.)

Figure 5. Number of known perpetrators per 10000 inhabitants in subregions of Csongrád county, 2007



Source: own creation

The significant numerical difference between the two types of indicators can be attributed to three factors:

- The perpetrator may commit more than one crime
- If the crime has become known it is not certain that the perpetrator who committed it has become known too
- In case of crimes the regional categorization is carried out in accordance with the place of commitment, in case of the perpetrator's place is the residence

The system of indicators applied during the research might be help in measuring the causes of delinquency because the observation includes many factors that can be connected to delinquency, as such, to the types of settlements, the demographic characteristics of the population living there, their education, employment relations, the rate of unemployment, the state of development of the settlement, tourism, financial situation, budget. Comparison with the categories of the types and the structures of settlements, and the size of population might also reveal important correlative conjunctions.

Below you can find the indicators calculated according to categories by the legal status of the settlements. (Table 2.)

Table 2. Crude criminality rates according to the legal status of the settlement, 2001, 2007

Legal status	Discovered publicly indicted crimes		Discovered perpetrators	
	per 10000 inhabitants			
	2001	2007	2001	2007
Towns of county rank	488,7	541,2	124,9	99,8
Other towns	448,5	430,2	123,2	109,0
Villages	233,4	277,2	80,5	83,2

Source: own creation

Delinquency has obvious connection with the legal status of the settlement, – regarding the committed crimes, – in both periods under examination the rate is the highest in the town of county rank, Szeged, and the lowest in the villages. However, the indicators are not as unambiguous in case of the perpetrators. The rate is not the highest in Szeged in 2007, which might indicate the success of law enforcement in Szeged or the fact that the number of perpetrators not living there has increased.

It seems that the Regional Clerk's Office has no influence on the tendency of delinquency, there is no significant difference between the delinquencies of the seat of the Regional Clerk's Office and other villages.

All indicators are the lowest on the settlements which attract few tourists, they are higher in the National Park and the highest on the settlements with medical tourism. This statement is in line with my examinations carried out in other counties (Veszprém, Nógrád, Heves, Borsod). (Table 3.)

Table 3. Crude criminality rates on the territory of the National Parks and on settlements with medical tourism, 2001, 2007

National Parks, medical tourism	Discovered publicly indicted crimes		Discovered perpetrators	
	per 10000 inhabitants			
	2001	2007	2001	2007
Körös-Maros National Park	387,0	405,7	99,5	115,9
Medical tourism	482,1	477,9	113,5	105,8
National Park & medical tourism	489,1	536,6	128,2	102,2
Neither National Park nor medical tourism	246,4	263,8	85,5	81,1

Source: own creation

If we make a classification according to the size of the population of the settlements at first sight the regularity cannot be noticed, the category of 500-999 inhabitants shows a very high value.(Table 4.)

Table 4. Crude criminality rates according to the size of the population of the settlements, 2001, 2007

Size of the population	Discovered publicly indicted crimes		Discovered perpetrators	
	per 10000 inhabitants			
	2001	2007	2001	2007
-499	162,3	188,8	103,9	97,9
500-999	751,9	738,7	81,4	79,7
1000-1999	180,8	182,0	90,4	81,4
2000-4999	211,0	269,9	81,5	83,6
5000-9999	232,4	214,7	87,1	83,8
10000-99999	417,4	449,5	124,7	116,4
100000-	555,8	595,8	128,4	96,1

Source: own creation

But if we examine which villages can be classified into this category we can see that the population of the border station, Nagylak, belongs to this group too. If we remove the data of Nagylak from the group then the rate of criminality increasing along with the number of population can be recognized. The rate of the group with a population of 2000-4999 persons is increased by the data of Röszke. There is no similar regularity in case of the perpetrators.

It is interesting to see that there is no connection either between the data relating to the crimes or the data of the perpetrators and the rate of the unemployed people. (Table 5.)

Table 5. Crude criminality rates according to the ratio of unemployment, 2001, 2007

Unemployed(%)	**Discovered publicly indicted crimes**		**Discovered perpetrators**	
	per 10000 inhabitants			
	2001	**2007**	**2001**	**2007**
3,0-3,9	512,8	556,5	119,8	93,8
4,0-4,9	396,7	390,4	120,6	102,4
5,0-5,9	291,4	335,2	98,6	100,1
6,0-6,9	530,0	458,4	111,3	97,2
7,0-7,9	167,1	272,7	107,9	108,0
8,0-9,9	233,1	312,9	86,2	80,7
10,0-	259,1	207,6	73,6	133,3

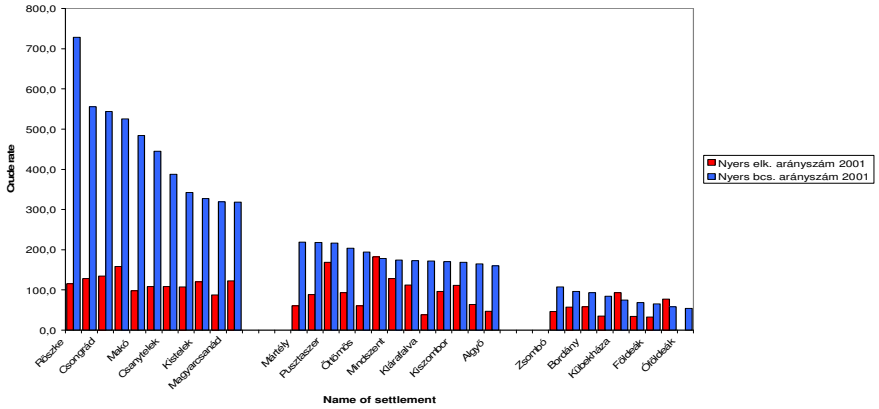
Source: own creation

Because the action data include the place of commitment and the personal data include the perpetrator's place of residence the detailed regional data show the perpetrators' mobility. With regard to crime prevention important information is provided by revealing which settlements are the most endangered in respect of crimes and which of them serve as places of residence for the perpetrators.

Now I will introduce the examination of the criminality of each settlement. (Fig. 7., 8.) In the figure Röszke has the highest rate, Nagylak could not be indicated because of the small size of the figure, its graph is eight times bigger than that of Röszke. Compared to the highest numbers of crimes the number of the perpetrators is much lower. In case of the settlements located in the central part of the diagram in

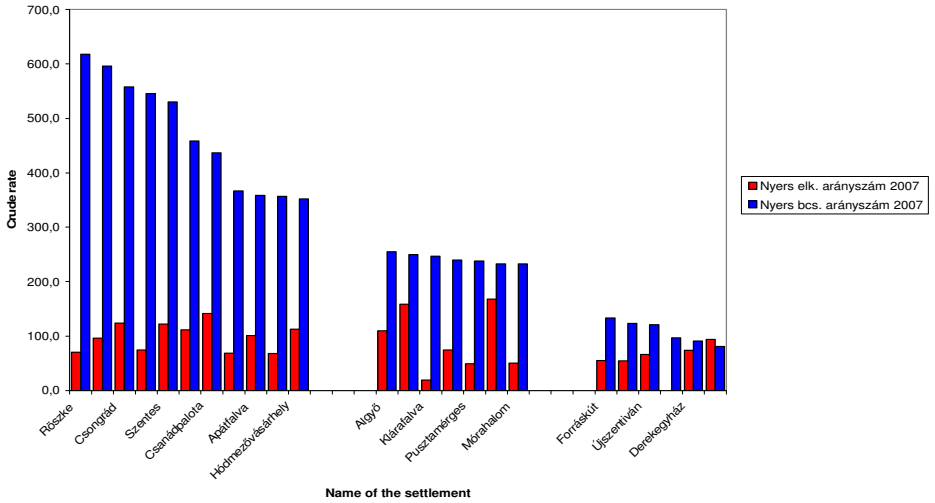
many cases the numbers of the crimes and that of the perpetrators is more or less the same. On some settlements next to the smallest blue columns the number of the perpetrators is higher than that of the crimes.(Kübekháza, Újszentiván)

Figure 6. Discovered publicly indicted crimes and their perpetrators in the settlements of Csongrád county, 2001



Source: own creation

Figure 7. Discovered publicly indicted crimes and their perpetrators in the settlements of Csongrád county, 2007



Source: own creation

In 2007 it was Nagylak and Röszke, Szeged, Csongrád again which had the highest rates calculated from the crimes. The rates of the perpetrators are low in every place. It is only Kübekháza where the number of perpetrators is higher than that of the crimes.

Using the modern technology we can already prepare crime maps. Since we have sufficient information on the delinquency features of each settlement the map can provide information on their regional connections.

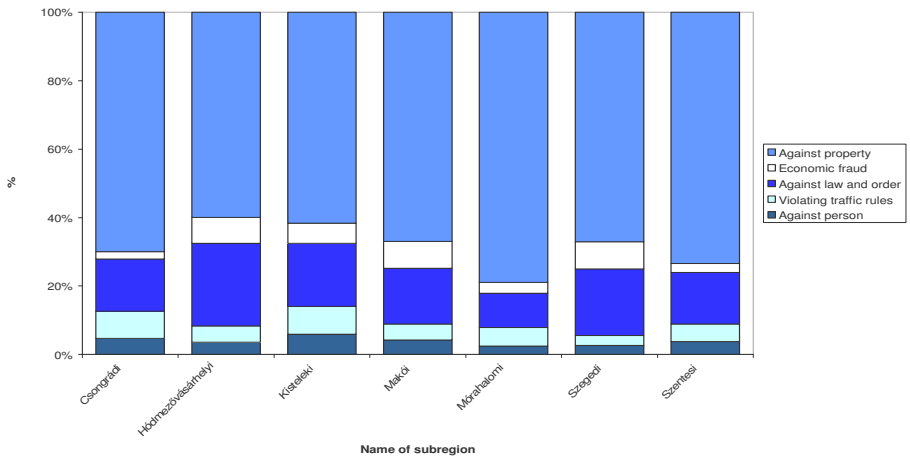
Distribution of types of crimes. The subregions are not identical considering the distribution of the crimes. The highest rate relates to crimes against property in all subregions and it has the highest rate in the subregions of Mórahalom and Szentes. It is followed by the rate of the crimes against law and order and it is the highest in the subregions of Hódmezővásárhely and Szeged, and the lowest in the subregion of Mórahalom. The ratio of crimes against people, the traffic and economic crimes is below 8% in all subregions. (Table 6., Fig. 9.)

Table 6. The distribution of crimes according to the main groups of the Penal Code 2007

Subregions	**Against person**	**Violating traffic rules**	**Against law and order**	**Economic fraud**	**Against property**
Csongrádi	4,6	7,7	15,0	2,1	68,5
Hódmezővásárhelyi	3,5	4,7	23,6	7,4	58,6
Kisteleki	5,7	7,9	17,9	5,7	59,8
Makói	4,1	4,5	15,6	7,6	64,2
Mórahalomi	2,4	5,3	9,7	3,1	76,9
Szegedi	2,6	2,8	19,1	7,7	65,7
Szentesi	3,6	5,1	14,8	2,6	72,1

Source: own creation

Fig. 8. The distribution of crimes according to the main groups of the Penal Code 2007



Source: own creation

4. Definition of the centre of delinquency

The objectives of the research include the determination of the centre of delinquency on both county and subregion levels. In prevention of crime it is important to answer such questions of organisation which are based on the determination of the regional centre of delinquency[0].

The centre point is an indicator produced by averages it eliminates the distorting influence of the accidental cases, and the dynamic examination of these indicators may be a very useful tool when analysing the regional distribution of delinquency.

The centre of delinquency can be defined following the analogy of centre of population applied in demography and the interpretation of this latter was made using the concept of the centre of gravity known as a model in physics. The precondition of generating this indicators is to set a coordinate system on the map of a specific area (governmental district, county or country) and to define the place of the settlements located on the area (villages, towns) using the coordinates of this coordinate system. Today the most practical space coordinates of the settlements can be the GPS coordinates serving for defining the places.

An important indicator of the regional distribution of delinquency is the centre of delinquency which is a point, the coordinates of which is the arithmetic average of the coordinates of each settlement (villages, towns) weighted by the number of the crimes committed on the specific settlement. In other words the two coordinates

of the centre, if t_i means the number of crimes committed on settlement No i., and n is the number of settlements:

$$x_s = \frac{x_1 t_1 + x_2 t_2 + \ldots + x_n t_n}{t_1 + t_2 + \ldots + t_n} = \frac{\sum_{i=1}^{n} x_i t_i}{\sum_{i=1}^{n} t_i},$$

$$y_s = \frac{y_1 t_1 + y_2 t_2 + \ldots + y_n t_n}{t_1 + t_2 + \ldots + t_n} = \frac{\sum_{i=1}^{n} y_i t_i}{\sum_{i=1}^{n} t_i}$$

5. Conclusions

To sum it up, the research is very timely, it is of stopgap character, its realization is possible with the new technical facilities, partly using the methods of accessing the necessary data, partly by using the methods provided by the information technology. In addition to the theoretical significance of the research it serves success in the field of both law enforcement and prevention of crimes.

References

Bud, J 1910.: Általános jelentés. In: A magyar szent Korona országainak Bűnügyi statisztikája az 1904-1908 évekről, p. 61-65. *(In English: General Report; Criminal Statistics of the Countries of the Hungarian Saint Crown in 1904-1908)*

Vavró, I. 1985: A bűnözés területi különbségei, Statisztikai Szemle, 1985 Issue 10. *(In English: Regional Differences of Delinquency, Statistical Review).*

Kovacsicsné Nagy, K. 1973: A bűnözés középpontja és súlypontja, Statisztikai Szemle, Budapest, Vol. 51. Issue 4. p. 371-382. (*In English: Centre and Focus of Delinquency, Statistical Review)*

Telecottages: Hopes and Results in Rural Development[1]

Prof. Dr. László Kulcsár [2] - Ferenc Hohl[3] - Dr. Csilla Obádovics[4]

Rural development is a constant challenge in every region, all around the world. To various extents and in different forms, but rural areas are coping with significant social and economic disadvantages both in the developed countries and the former socialist countries of Europe, as well as of other continents. Many researcher and politician believed that the digital revolution, the headway of the information society would greatly contribute to the decrease of social and regional inequalities. But these expectations have not been validated by any data.

Location of telecottages in Hungary supports the fact that the digital "revolution", the results of the information society do not increase the probability of underdeveloped rural areas' closing up, rather they sustain, and, in some cases, enhance the disparities.

Keywords: telecottage, rural development, regional inequalities, human resource

1. Introduction

The notion of the information society has emerged in literature in the 1960s. It quickly gained popularity, partly because it was expected to significantly diminish social and economic disparities in access to and disposing of information. All the new technologies ensuring the circulation of information enhanced the perception of the globalization and supported the approach that they help to achieve the decrease of the inequalities. So the gap between developed and underdeveloped countries, regions and areas can close.

From time to time, people's mind gets infected with the belief in the all-might

---

[1] A former and shorter version of this study was published in the journal *Gazdaság és Társadalom* (Economy and Society) (2009. 1.)

[2] Prof. Dr. László Kulcsár, *University of West Hungary, Institute of Social Sciences and Pedagogy (Sopron)*

[3] Ferenc Hohl, PhD student, *Szent István University (Gödölől)*

[4] *Dr. Csilla Obádovics, Phd student, Szent István University, Faculty of Economy and Social Sciences (Gödöllő)*

of technological development. The emerge of new mass communication tools – from wide-spreading printed press through the start of the radio and the television (from the last half of the 19th century) up until now – has always created such an illusion. There are other approaches, which treat the effect of mass communication as merely technical or technological issue, and these are similar to those "illusionist" approaches that treat the problem of social and regional inequalities as a simple technical or technological issue, and expect the broadband Internet penetration to decrease these disparities. Regarding the effects of mass communication, by now, at least the researchers and the literature have managed – more or less – to overcome these "teething-troubles". However, regarding the Internet and other digital technologies, the public opinion and the political thinking still follows "schoolboy logic".

The effects of the new information technologies, the information society, and the phenomena of the digital divide can be examined through several approaches and at many different levels. One part of the studies represents a digital horizon expecting the penetration of the digital technology and the Internet to diminish the differences between developed and developing countries. Other part of the studies, thinking over the chances of closing up certain regions, expects the same, but analyse regional inequalities within one country. Studies, representing a third kind of approach, emphasize the various situations of the different social levels, mainly by examining the entrepreneurs or the poor. In this paper, within this topic, we focus particularly on the regional correlations and inequalities.

2. Literature

2.1. Dreams and the reality

Early sources of the literature are full of great, optimistic expectations. For example Edwin (1978) believed that achievements of the information technology would greatly contribute to the close-up of the underprivileged rural areas, especially concerning the development of business ventures. James (2002), in order to close the digital divide, primarily emphasized the necessity of access to cheap IT tools. But James hardly paid attention to the cultural and sociological barriers, just like that IT conference held in India, in 2000. Permkumar and Roberts (1999) went further, and simply ignored the social and cultural factors when they examined the application of the information technology. Beside many other studies, Cavill's paper (1997) also represents this optimistic approach regarding the Australian situation. Studies published in Hungary also expressed similarly optimistic expectations, even recently (Varga 2000). For example Forgács (2008) strongly believes in the penetration of telework (formerly known as distance work), and he tries to present the economic benefits of the telework houses. There are governmental papers and presentations on

telework penetration, but there are no data on the regional distribution of the telework. We can only assume that telework cannot help to overcome regional disparities, since more than half of Hungarian teleworkers are highly educated seniors.[5]

After a while, of course, the critics had arrived and the tone of the literature has changed. The enthusiasm fell off, and those issues came to the fore that – beyond the technology change – prevent closing the digital divide. Gillespie (1987) showed that access to and use of new IT tools is much more restricted in the underdeveloped regions of Western Europe. Pick and Azari (2008), after analyzing more than 70 countries, have concluded that the digital divide between developed and developing countries is apparently not closing. Grimes (2000) criticizing the EU policy of information society points out that the European Union follows an explicitly technological approach toward the issues of the information society, and as a result of ignoring the social and cultural factors, huge amount of money and facility run to waste. He claims that to utilize the new IT tools, components of the human dimension have to be taken into consideration. This statement is the conclusion of that Hungarian survey in which researchers examined the operation of the so-called eHungary-points. These e-points as technical solutions slightly decreased the digital divide between urban and rural areas (at least concerning the technical possibility of access), but, more important, the usage is quite limited, and the demographics of the users (social level, occupation, qualification, age) are almost totally unknown (Pándi–Takács 2006). Social inequalities in Internet usage generally show that the more privileged social groups are in the more favourable situation, so the decrease of the digital divide is not proven in this sense either (Albert et al. 2007). Angelusz and Tardos (2004) indicate serious social inequalities in Internet usage, regarding not only the intensity, but also the quality of the usage. Homoki (2002) claims that the most significant selective factor is the type of settlement and indicates that smaller settlements fall behind.

Previously international organizations had put pressure on developing countries to remove legal barriers that prevent the penetration of new information and communication technologies (like Internet). Deregulation was expected to jump-start the economic growth, thanks to the fact that these countries join into the international trade. Ngwenyama and Morawczynski (2009) recently have criticized this approach and pointed out that deregulation of the early 1990s was insufficient, since political efforts ignored social factors such as conditions of the human resource, cultural characteristics or the weakness of the civil sector. According to the data of Rodriguez and Wilson (2000) although developing countries reached some results concerning information technology investments, but the digital divide has widened in the last few years of the 20th century.

---

[5] http://www.tavmunka.org/content/view/656/236/

Advocates of the knowledge society have also expected much from this paradigm, for example in accordance with the decrease of regional disparities. In recent years in Hungary a political version of this technological approach has emerged, in which there are expressions like "intelligent town", "intelligent region", "intelligent settlement", etc. It is accompanied by certain information technology innovations, like establishment of networks and planned services, but without thoroughly analyzing the social and geographical characteristics of employment and usage. (Pintér 1999, Molnár et al. 2007). Apart from the fact that knowledge society presently exists only in some of its elements, though its virtual establishment progresses forcefully and its language spreads well (Varga), one of Work Foundation's studies demonstrates that in Europe the indicators of the knowledge society created a situation that enhances the sustenance (increase) of the regional inequalities, and regions of Eastern and Southern Europe are in a devastating disadvantage even in this field (Rüdiger-McVerry 2007).

2.2. Information society and rural contexts

Cleevely and Walsham (1980) have relatively early set out their problems concerning the possibilities provided by the new tools of telecommunication. There is no clear correlation between the development of underprivileged regions and the new communication tools, they wrote, and these circumstances should warn the creators of development policy. In the midd-1980s, Mowlana (1984) warned that the consequences and effects of the information revolution are uncertain in many aspects, and we should carefully form our opinions about how it changes the lives of rural areas and the people living there, and to what extent it can be the engine and facilitator of the development.

Malecki (2003) warned that information technologies do not provide quick solution to the problems of underdeveloped rural areas, since those are influenced by long term processes. Penetration of digital technologies is just a small piece of rural development. It opens up certain "doors", but it is insufficient to have anyone to step in. Berkeley et al. (1996) drew attention to that rural areas cannot be treated unitedly. Concerning access to and usage of information and communication tools politics have to create different goals and assign different tools to underdeveloped regions and regions being in a more favourable situation. There are two important packages of measures for underprivileged rural areas: on the one hand, the number of telecottages has to be increased, and, on the other hand, there is need to be an educating activity involving and contribution with local governments and educational institutions. According to Mehta and Kalra (2006) new information technologies, theoretically, are able to contribute to diminishing the poverty, developing human resources and services. But, practically, by analyzing Indian examples, they experienced that the lower levels of the social pyramid have a more difficult access to these services and they utilize them less effectively. Similarly,

they see several problems in the fact that the most underdeveloped regions are mostly crowded out from access to telecottages (in India: kiosks). They think that during planning the usage of the technology it is particularly important to focus on that the conditions of the operation should fit to the social and cultural standards prevailing in the region, and the services should be offered free or affordable prices even for the poorer levels, to avoid deepening the social inequalities. The study of Mehta and Kalra, drawing away from the simple technical approach, demonstrates the importance of cultural context. Many Hungarian researchers share the idea that purely the technical tools are insufficient to bridge the digital divide (Lengyel et al. 2003, 2004).

Technological changes and new technologies – the technological foundation of information society among other things – are double-edged weapons in the rural America, writes the editor's letter in the 1st issue of *Technology in Society* in 2001. Peter Korsching (2001) proposing an essential dilemma, indicates that despite several forms of help and support certain rural areas in America keep lagging behind. Technological changes of information revolution and the digital tools were considered by many people as a possibility, and they hoped that underprivileged rural areas can be taken back to the mainstream of the American social and economic life. Korsching, in accordance with the new technologies, observed the following main issues: on the one hand, the access to these technologies is difficult due to the unequal development of infrastructure; and, on the other hand, people living in the country lack the required knowledge and skills. There is also a third, increasing issues. Internet usage by content is getting fragmented, claims Cavanagh (2009), and this condition is exclusive, since it requires special communication. People with lower level of qualification, living in underprivileged regions (civils, entrepreneurs) hardly can meet this requirement. In other words, the condition of human resources is a major barrier in utilizing the new tools. Corea (2007) also emphasized the effect of human resources' condition when studied how the innovations of information technology can serve the development. Birnie et al. (2005) emphasized that narrow technological, infrastructural approach without social contexts does not provide appropriate explanation for the problems of rural areas. Pigg and Crank (2005) claim that there are only a few evidence supporting that the intensive use of information and communication technology would lead to economic growth in underprivileged rural areas, even without major efforts for developing human resources and enhancing social cohesion. In Hungarian literature Nagy (2007) shares this opinion and searching access disparities draw attention to the importance of lifestyle and knowledge style factors.

Bruque and Moyano (2007) analyzing the penetration and usage of advanced information technology emphasize the effect of the socialization process, and that the usage of information technology heavily depends on the conditions of small enterprises' human resources. Studying Iranian small enterprises, Fathian et al. (2008) came to the same conclusion. Their results showed that accepting the tools of

the digital culture is primarily correlated with human resource and qualification/skills. Pleitner (1989) refers to the different cultural background of small enterprises, when analyses the usage of information revolution's technologies. Altobello Nasco et al. (2008), concerning the usage of information technology tools, emphasized the effect of social psychology factors.

Reviewing the literature has reassured that our approach is right. The effect of all the phenomena of the information revolution has not verified – nor locally, neither globally – the opinion of those who expected breakthrough on the front of social and economic inequalities purely by the spread of the tools. The digital divide, or in a newer and broader meaning, the digital inequalities (DiMaggio-Hargittai 2001) still exist, both in the dimensions of social and regional inequalities. We are aware of the fact these to kinds of inequalities are not independent from each other. The spread of new technologies in underprivileged rural areas is more difficult, and people living there have a more difficult access to these new technologies, but this is due to their cultural disadvantages as well. And these cannot be defeated with technical, infrastructural tools. A new methodology approach for measuring digital disparities proposed by James (2008) has a similar starting point. The fact that the usage of digital information is concentrated to socially favourable levels means that disparities are intensifying, even within one area, and this cannot be handled with traditional measuring methods. So statistics show a better picture than the reality.

Hereafter we discuss the location of telecottages in Hungary. Our initial hypothesis is that the telecottages do not live up to expectations in that sense that they do not manage to play an effective role in getting over the social and locational disadvantages. In our opinion this – among other things – is due to the fact the neediest settlements cannot get into or fall out of the network, because their human resources and social capital do not help them to reach and/or keep the establishment of telecottages. However we have to mention that this – as we have seen it – is in line with international experiences, and the biggest responsibility lies on the sector policy – but this could be the topic of another study.

2.3. Rural Hungary and Telecottages

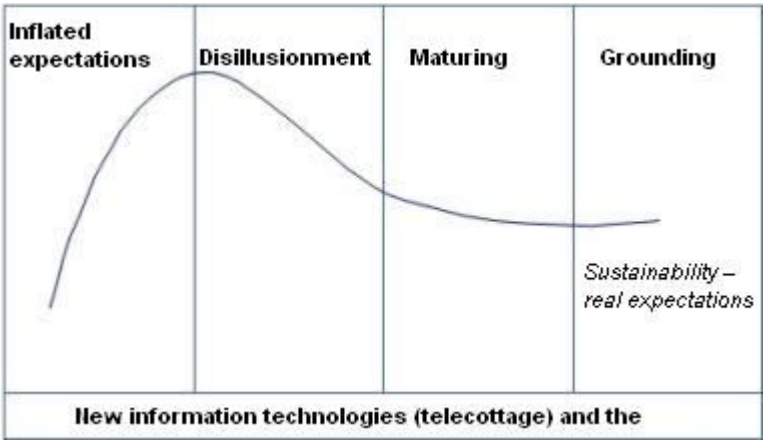
The purpose of our study is not to present the history of the telecottage movement or the evaluation of its operation. Relevant information can be found in the website of the Hungarian Telecottage Association (surveys) and in the studies of Erdősi (1992), Gáspár-Takáts (1997), Hohl (2006) and others.

At the start of telecottages and telecottage movement in Hungary, as we mentioned it in the introduction, there were serious expectations regarding the effects of information technologies (Gáspár 1999). In many cases these expectations were unfounded, and in our opinion, caused significant disadvantages to the movement. Later, both in Hungary and elsewhere in the world, the opinions and the approaches have changed. Jakobi (2007) regarding the spread of information and

communication technology tools draws attention to the enhancement of disparities in the existing social and economic processes, although believes that as a result of the activities of telecottages the information society can be expected to become more active. Grounding the point of views is significantly set back by the fact that the impact and efficiency assessment of telecottages has not been conducted yet in a scientific level (Hohl 2004).

Rao (2008) compares experiences gained in Indian telecottages – regarding social and social political acceptance – to the Gartner hype cycle that is well-known in the world in Jackie Fenn's explanation. The following figure shows the curve applied to Hungarian telecottages.

Figure 1. Gartner hype cycle relevant to telecottages



Source: own creation based on Rao (2008)

Experiences gained in Spanish telecottages confirm that from the point of view of telecottages' efficiency, namely getting over regional disadvantages, the social grounding into the local communities is vital (del Aguila Obra et al. 2002). Hungarian telecottages are more or less in the phase of maturing and grounding.

Unfortunately, as we have mentioned earlier, there are errors in both the methodology and the topics of the Hungarian impact assessments and surveys we know, so we can draw conclusion from them only in a restricted manner. There are also several problems in handling the database and in organizing the impact assessment. But this is mainly due to the indecisive appreciation and subsidy approach of the politics, not the lack of expertise. Grounding telecottages into the local society and its institutions – not surprisingly - bears characteristics of operation of the local society and politics. Sociological factors influencing the operation of telecottages get on in these circumstances.

Madon et al. (2009) – based on Brazilian examples – claims that the conditions of successful institutionalization of telecottages are not technological, but social factors, like positive appreciation of the local society, enhancement of social activity, etc. Similarly, an OECD report also originates from cultural and social factors the Scandinavian telecottages' success over Scottish ones in underprivileged areas (OECD 1996). Kleine (2009) observed that when entrepreneurs in underdeveloped regions of Chile employ the services of telecottages (telecentros), because of poor knowledge and skills they cannot work off their handicaps. The government of Chile assumes that they cannot enhance the social and regional cohesion. Lentz and Oden (2001) have conducted an analysis in the Mississippi Delta region, and drew the conclusion that regional policy makers have to understand that the most important issue is not simply increasing access to the technologies, but the development of the knowledge and skills required to use them. Apparently, human factors play a core role in the efficiency of telecottages, and this does not include only the leaders, managers and employees, but also the social and cultural characteristics of the region. The change and/or transformation of these factors in Hungary can only take place as a result of a process overspanning several generations. Everywhere, but especially in Eastern Europe politicians are interested in short term successes spanning from election to election, furthermore the economic elite can claim its interests in a more powerful way than underprivileged settlements, so long term investments into the development of human resources mostly are not compelling enough.

3. Methods and Results

Our core purpose, as we have mentioned, is to examine the location of telecottages and its correlation with the social and economic characteristics of the region and the settlement. With this in view we have combed together two databases and then chose the indicators of our analysis. The dependent variable was the presence of telecottages. That lead us to distinguish settlements and regions, which had a telecottage, but currently don't; which earlier had had a telecottage and currently also have one; which earlier had hadn't a telecottage, but currently have one; and which haven't had a telecottage at all. These databases did not provide us information about the operation of telecottages, but we did not intend to obtain such information.

Those settlements that had had or currently have a telecottage were considered to be “involved”, since once they have successfully applied for establishing a telecottage. This variable divided the settlements into two groups: involved and not involved. The initial rough analysis shows some interesting results. It turned out that 21.6% of settlements in the Central region are involved, and on the

other end of the scale we found Northern Hungary with 10.6% of settlements involved. Altogether there was or is a telecottage in 15.8% of the country. This result makes it probable that not neediness is the only factor that interplays in the successful establishment of telecottages.

Relationship between presence / current state of telecottages and size of the settlements is shown in the table below.

Table 1. Number of settlements in respect of telecottage involvement and size (2004-2007)

Size of the settlement	**No telecottage**	**Active telecottage 2004-2007**	**Inactive telecottage, 2007**	**New telecottage 2004-2007**	**TOTAL**
less than 1000 inhabitants	1453	131	93	38	1715
1001-3000 inhabitants	793	66	59	17	935
3001-10000 inhabitants	281	32	21	7	341
more than 10001 inhabitants	114	8	13	8	143
TOTAL	2641	237	186	70	3134

Source: own creation

We also conducted analyses with multiple variables to closely observe the effect of certain demographics, social and economic factors, filtering out their interdependence. One of the applied methods was discriminance analysis, where the dependent variable was "involvement" based on the two groups of the model. Out of 47 input variables the method found 5 whose effect remained significant even after filtering out the interdependence. These variables were the followings: aging index; rate of unemployment in active population; rate of children under 18; rate of permanently unemployed people among the unemployed and the rate of operating non-profit organizations. The characteristics of demographic structure, the presence of unemployment and the strength of civil society are the best indicators of evolution the involvement. The following table shows the role of these variables.

Table 2. Relations between standardized canonical discriminance function of involvementof settlements in establishing telecottages and the variables (2007)

Variables effecting involvement of settlements	**Coefficient of discriminance function**
aging index	,774
rate of unemployment	,475
rate of children under 18	,477
rate of permanently unemployed people	-,444
rate of operating non-profit organizations[6]	-,501
Significance level of the function	,000
Value of the function - involved settlements	-,296
Value of the function - not involved settlements	,061

Source: own calculation

The classification function also worked well, 84% of instances were classified correctly, belonging to the original group.

The results of discriminance analysis demonstrates that involvement proved to be significant in those settlements, where the rate of unemployment is lower, the aging index is more favourable and the presence of non-profit organizations is typical. In other words these settlements are in a more favourable situation, and those, where the rate of unemployment is high, the rate of non-profit organizations is low, the rate of children under 18 is also high, but the aging is typical as well, are less likely to become involved. Although the value of aging index is over 100% in both groups, but it is averagely 134% in settlements with telecottages, and 147% in settlements that are not involved. So we can claim with high probability that underprivileged settlements are small settlements even respecting telecottages. This is also confirmed by the evolution of the settlement size variable (population) in the logistic regression analysis, where correlation is quite obvious.

[6] Community Information Centres

Table 3. Beta results of logistic regression of telecottage involvement by settlement size
(2007)

Type of settlement	**Beta value**	**Significance level**
less than 1000 inhabitants	,-1.413	P=0,000
1001-3000 inhabitants	-,883	P=0,003
3001-10000 inhabitants	-,480	P=0,086
more than 10001 inhabitants	,107	P=0,694

Source: own calculation

As we can see, among settlements with less than 3000 inhabitants it is much less likely that we find such a settlement where there was or is a telecottage. Most socially and economically underprivileged small settlements lack the institution of telecottages, they have no direct experience with it, but they would be those that really need it.

Based on the available data we have also classified the "involvement" in another way. According to our database there are 2387 settlements that have never had a telecottage. There are 66 settlements that established telecottage between 2004 and 2007. There are 183 settlements where the telecottage was inactive between 2004 and 2007, and there are 240 settlements, where the telecottage was active between 2004 and 2007. We have also conducted the discriminance analysis for these groups with the same independent variables, like in the previous example. The 4 groups of settlements were significantly discriminated by density of population and the rate of active population. These 2 variables could estimate the adherence to certain groups quite well, since in 83.7% of cases settlements were classified into their original group. The most successful classification took place in the group where settlements haven't had a telecottage so far. In there the rate was 99.3%.

In this analysis we have identified 2 functions, and the following table shows that the new telecottages have the strongest positive correlation with the first dimension, and the second dimension has the same characteristic but with negative sign.

Table 4. Relations between standardized canonical discriminance functions of involvement of settlements in establishing telecottages by population categories and the variables (2007)

	Identifying canonical functions	
Involvement	**High density of population and lower rate of active population**	**Low density of population and moderate rate of active population**
no telecottage	,-0.03985	,-0.01823
active telecottage between 2004 and 2007	,0.08316	,0.105
inactive telecottage in 2007	,0.153	,0.164
new telecottage between 2004 and 2007	,0.714	,-0.179
The relation between the variables and the function (function coefficients) and the significance level of the relation		
rate of inhabitants between 18 and 59	,308 /0,000	,999 /0,000
density of population	,866 /0,000	-,586 /0,000

Source: own calculation

These results also confirm what we have claimed so far. The location of telecottages by type of settlements does not favour the underprivileged ones. If we want to explore the reasons, of course, further analyses would be needed, namely in quite sensitive fields. Anyhow we can see that the old sociological proverb is proven to be true: “Who he is poor is the poorest.”

4. Summary

The international literature – after a short period of enthusiasm – represents a more realistic way in observing the role and possibilities of information technology for diminishing the digital divide or digital disparities. There are several tendencies indicating that these disparities are still very significant, and there are data supporting the fact that this divide is deepening. Besides several other efforts the telecottage movement came to stay in many regions of the world, but as long as it is

treated as simple technical issue, it cannot develop its influence. From Canada to India the literature is united in the question that the success of telecottages depends on social, or sociological if you like and sociologically understood cultural factors.

This is the situation that prevails in Hungary as well, and data shows that the level of local human and social capital plays an important role in the presence of telecottage by type of settlements. And the settlements that would principally need those benefits are probably sidelined. This is due to, on the one hand, the fact that they have a lower level of human resource and social capital, and, on the other hand, the wrong approach of the relevant sector policy. Based on the analyses and results it is quite probable that geographical characteristics of digital disparity will sustain, moreover, the digital divide – along certain dimensions – could widen further in Hungary, too. Supporting telecottages and the digital technology is, of course, just one aspect of developing the underprivileged regions. We have to provide a much more space for long term ideas among the measures of rural development.

References

Aguila, O. A. – Bruque, C. S. – Padilla, M. A. 2002: The economic and organizational aspects of telecentres: the Spanish case. *Technovation,* 22, 785–798. p.

Albert, F. – Dávid, B. – Molnár, Sz. 2007: Internet használat és kapcsolati erőforrások időbeni alakulása Magyarországon. *Szociológiai Szemle,* 17, 3-4, 93-114. p.

Altobello, ., S. – Grandón, T. E. - P. Mykytyn, Jr. P. 2008: Predicting electronic commerce adoption in Chilean SMEs. *Journal of Business Research*, Vol. 61, Issue 6, 697-705. p.

Angelusz, R. – Tardos, R. 2004: Túl az egyötödön - A tudás - vagy élménytársadalom felé? *Jel – Kép,* 2, 3-33. p.

Berkeley, N. – Clark, D. – Ilbery, B. 1996: Regional Variations in Business Use of Information and Communication Technologies and their Implications for Policy: Case Study Evidence from Rural England. *Geoforum*, Vol. 27, Issue 1, 75-86. p.

Birnie, R. V. - Geddes, A. - Bayfield, N. - Midgley, J. L. - Shucksmith, D. M. - Elston, D. 2005: Improving the rural data infrastructure of Scotland: an overview. *Land Use Policy*, Vol. 22, Issue 2, 145-152. p.

Bruque, S. – Moyano, J. 2007: Organisational determinants of information technology adoption and implementation in SMEs: The case of family and cooperative firms. *Technovation*, Vol. 27, Issue 5, 241-253. p.

Cavanagh, A. 2009: From Culture to Connection: Internet Community Studies. *Sociology Compass*, 3/1 1–15. p.

Cavill, M. 1997: New Information Technologies for Australia. In: Droege, Peter (Ed): Intelligent Environments. Spatial Aspects of the Information Revolution. Elsevier B.V. 77-86. p.

Cleevely, D. – Walsham, G. 1980: Telecommunications models: Planning for regional development in LDCs Alert. *Telecommunications Policy* Vol. 4, Issue 2, 108-118. p.

Corea, S. 2007: Promoting development through information technology innovation: The IT artifact, artfulness, and articulation. *Information Technology for Development,* Vol 13, Issue 1, 49-69. p.

DiMaggio, P. – Hargittai E. 2001: From the 'Digital Divide' to 'Digital Inequality': Studying Internet Use as Penetration Increases.Working Paper #15, Summer.Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University.

Edwin. B. P. 1978: Communication satellites for rural development. *Telecommunications Policy*, Vol. 2, Issue 4, 309-315. p.

Erdősi, F. 1992: A falusi települések fejlesztése teleházak segítségével. In: *Telematika.* Távközlési Kiadó, Budapest.

Mohammad, F. – Akhavanb, P. – Hooralia, M. 2008: E-readiness assessment of non-profit ICT SMEs in a developing country: The case of Iran. *Technovation*, 28, 578-590. p.

Forgács, T. 2008: Távmunka-ház: Alternatíva a vidékfejlesztésben. *A falu, Autumn,* Agroinform, Budapest, 51-56. p.

Gáspár, M. 1999: Teleházat minden faluba? *A falu,* Winter, 63-67. p.

Gáspár, M. - Takáts M. 1997: Építsünk teleházat! Hungarian Telecottage Association, Erdei Iskola,

Gillespie, A. 1987: Telecommunications and the development of Europe's less-favored regions. *Geoforum*, 18, 3, 229-236. p.

Gillespie, A. E. - Robins, K. 1989: Geographical inequalities: the spatial bias of the new information technologies. *Journal of Communication*, 39, 7-18. p.

Grimes, S. 2000: Rural areas in the information society: Diminishing distance or increasing learning capacity? *Journal of Rural Studies*, 16, 13-21. p.

Hohl, F. 2004: A teleházak hatás- és hatékonyságvizsgálatának elméleti megalapozása. *A falu,* Spring, 85-96. p.

Hohl, F. 2006: A teleházak árképzésének és földrajzi elhelyezkedésének összefüggései. *Információs Társadalom,* 2, 98-108. p.

Homoki, M. 2002: Internet használat, Tudás - hátrányos társadalmi térben. *Jel – Kép,* 2, 79-101. p.

Jakobi, Á. 2007: Az információs társadalom térbelisége. ELTE Regionális Tudományi Tanszék, Budapest.

James, J. 2008: The Digital Divide Across All Citizens of the World: A New Concept. *Social Indicators Research*, 89, 275-282. p.

James, J. 2002: Low-cost information technology in developing countries: current

opportunities and emerging possibilities. *Habitat International*, Vol. 26, Issue 1. 21-31. p.

Kleine, D. 2009: The ideology behind the technology – Chilean microentrepreneurs and public ICT policies. *Geoforum,* 40, 2, 171-183. p.

Korsching, P. F. 2001: New technologies for rural America: boon or bane? *Technology in Society*, Vol. 23, Issue 1, 73-77. p.

Lengyel, Gy. – Lőrinczi, L. – Siklós, V. – Füleki, D. 2003: Hidak a digitális szakadék fölött. *Jel - Kép* 3, 25-43. p.

Lengyel, Gy. – Eranusz, E. – Füleki, D. – Lőrincz, L. – Siklós, V. 2004: A cserénfai kísérlet 2. 79-96. p.

Lentz,R.G - Oden, M. D. 2001: Digital divide or digital opportunity in the Mississippi Delta region of the US. Telecommunications Policy. 25, 291-313. p.

Madon, S. – Reinhard, N. – Roode, D. – Walsham, G. 2009: Digital inclusion projects in developing countries: Processes of institutionalization. *Information Technology for Development,* Vol. 15, Issue 2, 95 – 107. p.

Malecki, E. J. 2003: Digital development in rural areas: potentials and pitfalls. *Journal of Rural Studies*, 19, 201-214. p.

Mehta, S. – Kalra, M. 2006: Information and Communication Technologies: A bridge for social equity and sustainable development in India. *The International Information & Library Review*, Vol. 38, Issue 3, 147-160. p.

Molnár, Sz. - Z. Karvalics, L. – Pintér, R. 2007: Leszakadóban? Kormányzati reform és információs társadalom. *Információs társadalom,* 1, 8-16. p.

Mowlana, H. 1984: The myths and realities of the "information age": A conceptual framework for theory and policy. *Telematics and Informatics*, Vol. 1, Issue 4, 427-438. p.

Nagy, R. 2007: Új lencsék egy új társadalmi jelenség vizsgálatában: A digitális egyenlőtlenségek kutatásának átfogó szemléletéről. *Szociológiai Szemle,* 17, 1-2, 16-28p.

Ngwenyama, O. – Morawczynski, O. 2009: Factors affecting ICT expansion in emerging economies: An analysis of ICT infrastructure expansion in five Latin American countries. *Information Technology for Development,* Vol. 15, Issue 4, 237-258. p.

OECD 1996: *OECD Workshop on Information Infrastructure and Territorial Development*. Doc 34710, Paris.

Pándi, B. - Takács Gy. P. 2006: Kutatási jelentés: Mennyiség, eloszlás és minőség – helyzetkép az eMagyarország pontokról. *Information Society,* 4, 104-128. p.

Pintér, R. 1999: A hivatali Web-oldaltól az intelligens város átfogó programjáig. Intelligens Települések Országos Szövetsége, Budapest.

Pick, J. B. – Azari, R. 2008: Global digital divide: Influence of socioeconomic, governmental, and accessibility factors on information technology. *Information Technology for Development,* Vol. 14, Issue 2 , 91 - 177. p.

Pigg, K. E. - D. Crank, L. 2005: Do Information Communication Technologies Promote Rural Economic Development? Community Development. *The Journal of Community Development Society*, Vol. 36, Issue 1, 65-76. p.

Pleitner, H, J. 1989: Small firms and the information problem. *European Management Journal*, Vol. 7, Issue 4, 442-450. p.

Premkumar, G. – Roberts, M. 1999: Adoption of new information technologies in rural small businesses. *Omega*, Vol. 27, Issue 4, 467-484. p.

Rao, S. S. 2008: Social development in Indian rural communities: Adoption of telecentres. *International Journal of Information Management*, 28, 474-482. p.

Rodriguez, F. - Wilson, E. J. III. 2000: Are poor countries losing the information revolution? infoDev Working Paper. The World Bank, Washington, DC.

Rüdiger, K. – McVerry, A. 2007: Exploiting Europe's knowledge potential: 'Good work' or 'could do better' Knowledge Work and Knowledge Workers in Europe The Work Foundation. London.

Varga, Cs. 2000: Régiók az információs társadalomban In: A területfejlesztés feladatai az ezredfordulón és az információs társadalom. Hazai Térségfejlesztő Kft., Budapest.

Varga, Cs.: A tudástársadalom eszménye INCO 8. http://www.inco.hu/inco8/fooldal.htm

Railway Escape Velocity

Balázs Kotosz[1]

The borders of the agglomerations can be located by the administrative units, but these limitations do not reflect the borders of functional spaces. The aim of this paper is to show and apply a special locating method by the example of Budapest. The main idea is the motorway effect (the shortest way is generally not the quickest) applied to passenger railway traffic.

The travellers of the agglomeration towards cities far from the capital have a choice: going back to the centre and take a direct fast train or take a stop-train for getting to the first station out of the agglomeration where fast trains stop. We can terminate the settlements from where the return is quicker. This boundary can be perceived as the limit of the gravitational space of the main city.

Keywords: time-space map, railway traffic, gravitational law, agglomeration location

1. Introduction

The borders of the agglomerations can be located by the administrative units, but these limitations do not reflect the borders of functional spaces. The aim of this paper is to show and apply a special locating method by the example of Budapest.

The structure of the paper is the following. In the first part we present the motorway effect, the gravitational model and the gravitational law, as well as the theoretical background of time-space maps. Here we offer a new method to investigate our problem called "railway escape velocity". The second part is considered to the practical problems of the applied method, while in the third part we summarize the empirical evidence of the calculations. So, the problem is discussed in sociological and engineering perspectives (Wegener 2001).

[1] Balázs Kotosz, PhD, assistant professor, Corvinus University of Budapest, Faculty of Economics, Hungary.

2. Theoretical background

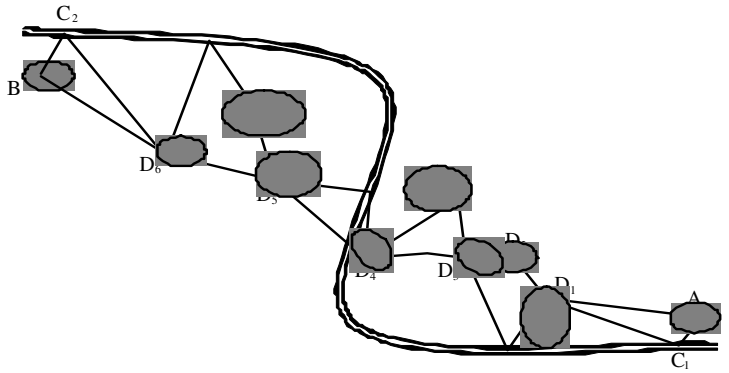
2.1. The motorway effect

The basic idea of the paper is the motorway effect. In transport analysis it is obvious that the shorter way is not always the faster, on the road it is quicker to find the nearest motorway, and to follow it to the nearest exit to the target. (See Figure 1: The shortest way from A to B is via D1 – D2 – D3 – D4 – D5 – D6, but it can be faster to avoid these settlements and take the motorway from C1 to C2.)

This model is also popular in economic policy theory: if the economy is not in an equilibrium growth path, the first step should be a stabilization program, and then the equilibrium growth path is easily sustainable.

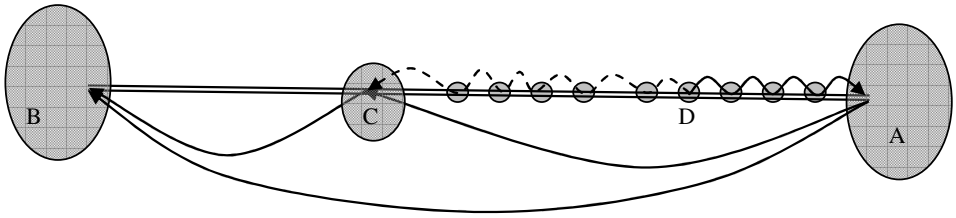
In the railway traffic the motorway effect is little bit different. In passenger transport, generally different types of trains run parallel between cities. Stop trains stop at (almost) every station, fast and mainly Intercity or high-speed trains serve only largest cities. Additionally, stop trains' itinerary is not longer than 60-100 km. Thereby, if travellers of the agglomerations (D) of large cities (A) would like to get to cities out of this 60-100 km circle, they have to change train somewhere (C) (See Figure 2). We have to mention that in some cases, long-run stop trains also circulate, but mostly it is faster to change. Practically, the only question is the place of the change. Is it faster to go back to the centre (D→A→B) or to crawl out of the agglomeration by a stop train (D→C→B)? It depends on the distance from the centre, the density and the speed of the trains and on the current time. We have to remark that these journeys may be atypical (Levinson-Kumar 1993, Ecostat 2006).

Figure 1. The motorway model in road transport



Source: own creation

Figure 2. The motorway model in railway passenger transport



Source: own creation

2.2. The physical model

The physical analogy is clear, if the kinetic energy of a body is small, the body cannot leave the gravitational zone of the other body. For the first sight, it is a case of the gravity model often applied in spatial analysis. The gravity model is really a wholesome method in the modelling of flows (see Fotheringham-Haynes 1988, Reilly 1929, Rodrigue et al. 2006), but the possible applications are limited (for this limits see Dusek 2003; Rodrigue et al. 2006). If our question was whether the traffic from D is larger to A than to B, the gravity model would be applicable. As Dusek (2003) enlightens it, the gravitational model and the gravitational law have the same roots, but not all phenomena of the gravitational law can be handled by the model. As in our case it is the law we apply, only some of Dusek's warnings should be minded.

In almost all spatial analysis we face the problem point of the territories: they have to be analysed as points. Where is the middle of a country, region, or even a village? This problem is not essential in the railway gravity case, because the size of small stops is negligible (regarding the used distances), and the journey time from great cities or capitals is counted from the relevant railway station.

Our model by the physical analogy is that of escape velocity. In the physics, the necessary speed of a body to leave the gravitational space of another body can be calculated by the following form:

$$v_2 = \sqrt{\frac{2 \cdot G \cdot M}{r}} \tag{1}$$

where G is the gravitational constant (6.6742*10-11 Nm^2kg^{-2}), M is the mass of the body being escaped from, and r is the distance between the centre of the body and the point at which escape velocity is being calculated. (Holics 1986) In socio-economic analysis, the escape velocity is a growing function of the socio-economic importance (mass) of the centre, and a decreasing function of the distance from the centre. In the adaptation to the railway traffic, the formula is as following:

$$v_{er} = c \cdot \frac{M^{\alpha}}{r^{\beta}} \tag{2}$$

where v_{er} is the "railway escape velocity", c is a constant, α and β are parameters. The v_{er} has a complex meaning of the frequency and average speed of stop trains, while M is definitively influenced by the importance of the centre in the national/international railway system. In this applied model, c, α and β have to be calibrated by the empirical evidence. For a starting point, we can assume that α=1/2 and β=1/2, as it is in the physical model. In the literature we could not find any precedent for the measurement and calibration.

The main problem is the definition of the mass, as the importance of a city can be represented by several measures. The illustration of railway mass of the city is more complicated. The number of train departures from the city or from the relevant station, in absolute or in relative terms can be the appropriate measures.

The distances are measured along the railway line (in km), but the optimization was made by time. Sometimes even the ranks of distance and time are not correlated, so the real space and the time space are distinct. (Janelle 1969, Janelle 1975, Knowles 2005)

2.3. Time-space maps

The time-space and the geographical space are quite often different. Geographical distances – mainly in the case of relevant distances of our paper – are more stable, than necessary time to get from one point to another. The distance in time terms depends on many factors, including the mean of transport, the state of the infrastructure, etc. Maps including time data are constructed to show the accessibility – in the form of isochronous maps or average accessibility time maps. (Dusek-Szalkai 2007) Another possibility is the strictly defined time-space map, where the original map is transformed as the distance of two points is proportional to the time-distance. (Dusek-Szalkai 2008)

2.4. Detailed description of possible zones

The space between points A and C can be separated to different zones: (1) the always attracting area (***R1***), (2) the typically attracting area (***R2***), (3) the typically not attracting area (***R3***), and (4) the never attracting area. So $\mathbf{R1} \subseteq \mathbf{R2} \subseteq \mathbf{R3}$. The exact definition of ***R2*** and ***R3*** sets can be found in chapter 3. Similar delimitation is not known in the literature (Kotosz 2010), other methods use journey time and costs. See (Levine et al. 2005), or (Iacono et al. 2008).

3. Methodological questions

We have several points to make them clear about the applied method.

In this paper, all calculations are made for journeys from the agglomeration to a far city. As the timetable is not symmetrical, the results of backward journeys can be slightly different. This is a typical problem of the time-space asymmetries.

All calculation of this paper is based on the official, first printed version timetable of the Hungarian State Railways (MÁV) for each timetable year (e.g. we call 2007 timetable that one valid from December 2006 to December 2007). As all timetable data is on minute exactness, we used this in all calculations. The timetable reading was manual, we did not use any journey planner software – in the random traveller case it would not be efficacious. The calculations were made in Excel.

3.1. The choice of the target city

If the target is close, some long-run stop trains can reach it or fast and Intercity trains do not halt there, and the results are misleading. Specially, in the case of the Western European high-speed trains, journey of two hours means 5-600 km, so applicable target is far from the starting-point. If the target is too far and/or has few direct connections from the main city, the results will better flash the contingency of the railway traffic, but not the gravitational zone. We suppose that a viable distance is that of a journey of three hours, in Eastern European context, this is 200-300 km.

3.2. Rational/random traveller

The starting time of the travel can be chosen in two different ways. First, we can suppose a rational traveller who arrives to the station just in the minute of the journey. In this case, we use a pure travelling time that excludes the density of trains. Generally, local stop trains run more frequently than Intercity or international trains, thereby in this model the rational traveller arrives to the local station when she gets the last stop train providing connection to the next fast train. Second, we can assume a random traveller who arrives to the station at random time (in this paper it means uniform distribution, e.g. she arrives with the same probability at 2:05 AM and at 8:05 AM). This version prefers frequent connections between the centre and the far city. For example, in the first case, if we have only one fast train per day, the average journey time is the same as we have hourly connection; but it is not true anymore in the second case.

3.3. Periodicity

At this point, another periodicity of the timetables has to be mentioned. Some stop trains do not circulate every day, so for the seven days of the week, separated calculations have been made. Finally, the average journey time was calculated by the uniform distribution principle, with the same probability for the days of the week.

3.4. Transfer

We can decide whether we consider any time for transfer. In double-track lines, it may happen that a stop train arrives and in the same minute the fast train starts to the opposite direction. We applied a five minutes period for the transfer – any connection shorter than five minutes are neglected. For an average traveller without huge baggage, this time could be enough. At this point, we have to mention that delays are also neglected; it is an important, but hardly measurable factor in this analysis.

3.5. Indicator of limitation

Finally we have to decide the critical indicator of the limitation. At a chosen moment, we can surely see which direction is better. But what means generally? First, we can calculate the average inward and outward journey time; and the lower will be better. Second, in the case of the random traveller, we can compare the time periods of the day when one or other direction is faster; then the direction having more than 12 hours a day is the winner. From a practical point of view, it may lead to favour the direction of the first train in the morning. In the case of the rational traveller this method is less complicated: we have to count the better choices.

4. Empirical evidence

Originally, the technical frame of the research was constructed for the Budapest-Győr-Wien line. First, we demonstrate the details of this line, after we synthesize the characteristics of two special lines, finally we show the maps made on the basis of the numerical analysis.

4.1. Budapest-Győr-Wien line

Table 1. Summary output for the Budapest-Wien line (2005)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	273	173	94.9	344	47.6	-
Budaörs out	253	245	62.1	407	37.3	0.0
Budaörs in	265	212	75.3	380	41.8	100.0
Budaörs opt	-	211	75.4	379	46.2	-
Törökbálint out	249	240	62.3	402	37.2	27.8
Törökbálint in	269	216	74.6	385	41.9	72.2
Törökbálint opt	-	215	73.1	383	44.9	-
Biatorbágy out	241	233	62.1	395	36.6	44.4
Biatorbágy in	277	223	74.5	392	42.4	55.6
Biatorbágy opt	-	215	72.0	382	43.7	-
Herceghalom out	234	226	62.2	388	36.2	64.8
Herceghalom in	284	231	74.0	399	42.7	35.2
Herceghalom opt	-	214	70.3	380	42.5	-

Source: own creation

We summarize here the particular results for the Budapest-Győr-Wien railway line, so we are curious about the Budapest agglomeration and the target city is Wien. The distance of the two cities is 273 km by rail. All times and distances are calculated from the terminus of the line (Budapest-Keleti). As all these international trains stop at Budapest-Kelenföld, journeys from the agglomeration via Budapest can be shorter.

First, about the 2005 year. The running time of the six direct trains was between 170 and 179 minutes. For rational and also random travellers, the average journey time from Biatorbágy with returning to Budapest is shorter than the direct way out. Additionally, in most cases the inward travel is quicker than outward. From Herceghalom (otherwise the last settlement in Pest county) all indicators change; the better choice "generally" is not returning to Budapest. Thereby the limit of the gravitational zone is clearly between Biatorbágy and Herceghalom.

Table 2: Summary output for the Budapest-Wien line (2006)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	273	174	94.0	345	47.4	-
Budaörs out	253	239	63.5	403	37.7	0.0
Budaörs in	265	189	84.0	369	43.1	100.0
Budaörs opt	-	189	85.0	369	49.5	-
Törökbálint out	249	234	63.8	398	37.6	18.4
Törökbálint in	269	194	83.1	374	43.1	81.6
Törökbálint opt	-	193	83.7	372	48.4	-
Biatorbágy out	241	227	63.6	391	37.0	20.4
Biatorbágy in	277	201	82.7	381	43.6	79.6
Biatorbágy opt	-	197	83.2	372	46.9	-
Herceghalom out	234	220	63.7	384	36.6	30.6
Herceghalom in	284	208	81.9	388	43.9	69.4
Herceghalom opt	-	202	80.7	373	44.9	-
Bicske out	224	191	70.3	357	37.6	83.7
Bicske in	294	209	84.5	387	45.5	16.3
Bicske opt	-	189	73.9	355	44.1	-

Source: own creation

In 2006, the supply of Budapest-Wien trains was broadened by an additional train. The running time from Budapest to Wien varied between 156 and 185 minutes. At the same time, in out-of-peak times, a rhythmical timetable of stop trains was introduced. By default, this rhythm was not in accord with other outgoing trains, so the limit of the agglomeration moved out, it was between Herceghalom and Bicske (where inland fast trains stop).

In 2007, the situation turned round. Over the addition of two international trains (with running time between 176 and 182 minutes), the outgoing timetable is rhythmical, with good connections towards Győr (change in Tatabánya). The stop trains from the agglomeration arrive to Budapest at the same minute when international trains depart, so travellers should wait to the connection almost one hour. As Table 3 suggests, if inward stop trains went 5 minutes earlier, the average inward travel time would be much shorter (the general time saving would be 55 minutes), so the end of the gravitational zone would be more far from Budapest. In that year, there is not clear solution for the limitation. The rational traveller's average journey time from Törökbálint is shorter in the case of outward travel, but in more than half of the cases (possible starting times) the inward journey is faster. The random travel-

ler – on the average – should start inward. Nevertheless, the limit should not be between two stops, but in Törökbálint.

Table 3. Summary output for the Budapest-Wien line (2007)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	273	179	91.3	325	50.4	-
Budaörs out	253	224	67.6	372	40.8	47.6
Budaörs in	265	217	73.3	357	44.6	52.4
Budaörs opt	-	208	74.2	349	49.9	-
Törökbálint out	249	219	68.1	367	40.7	47.6
Törökbálint in	269	222	72.7	362	44.6	52.4
Törökbálint opt	-	208	74.2	351	49.1	-
Biatorbágy out	241	212	68.1	360	40.2	71.4
Biatorbágy in	277	229	72.6	369	45.1	28.6
Biatorbágy opt	-	205	72.8	351	47.9	-

Source: own creation

In 2008 and 2009, only minor changes have been added to the timetable, not disturbing significantly our results, the situation seems to be conserved.

4.2. Budapest-Székesfehérvár-Szombathely/Nagykanizsa line

The choice of target city is open in this case, as the railway lines towards the middle Transdanubia are separated in Székesfehérvár (or over the city). According to the methodological background, the two potential target cities are Szombathely and Nagykanizsa. Nagykanizsa is a classical target city, as all trains use the same itinerary; while Szombathely is applicable for modelling high speed trains (the Intercity trains circulate via Győr). The latter fact and the difference of the average speed of the two lines a priori forecast pushing out of the boundary of the agglomeration. If we would like to estimate the boundary, this duplicity creates a new problem of how to synchronize the two results. For the estimate of the railway escape velocity this duality does not pose a problem.

We pursued the analysis for the period 2005-2008. Although we had the 2009 timetable, but during the most part of the year special – generally weekly changing – timetables were in function due to current reconstruction works.

Table 4. Summary output for the Budapest-Nagykanizsa line (2005)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	221	203	65.2	441	30.1	-
Bfok-Belváros out	212	255	50.8	482	26.5	51.0
Bfok-Belváros in	221	259	52.1	473	28.0	49.0
Bfok-Belváros opt	-	246	53.5	471	34.1	-
Nagytétény out	205	248	50.4	476	25.9	74.5
Nagytétény in	229	268	52.1	481	28.6	25.5
Nagytétény opt	-	244	52.5	470	33.0	-
Érd alsó/felső out	201	241	51.0	466	25.9	81.6
Érd alsó/felső in	233	267	53.3	476	29.4	18.4
Érd alsó/felső opt	-	239	52.8	461	33.6	-
Tárnok out	197	242	49.7	464	25.6	93.9
Tárnok in	237	274	52.7	482	29.5	6.1
Tárnok opt	-	240	51.8	461	32.6	-
Martonvásár out	188	235	48.0	457	24.7	93.9
Martonvásár in	246	282	52.3	490	30.1	6.1
Martonvásár opt	-	…	…	…	…	-
Baracska out	185	…	…	…	…	98.0
Baracska in	249	…	…	…	…	2.0
Baracska opt	-	…	…	…	…	-
Pettend out	180	…	…	…	…	100.0
Pettend in	254	…	…	…	…	0.0
Pettend opt	-	…	…	…	…	-

Source: own creation

Contrary to the Budapest-Győr-Wien line, the Budapest-Székesfehérvár line has a long (14 km) section inside the territory of Budapest. By the different logic of timetables, in the last five years, there were periods when the boundary of the agglomeration was inside of Budapest.

4.2.1. The Budapest-Nagykanizsa line

The limit of the agglomeration is determined by stop-trains between Budapest-Székesfehérvár and Budapest-Martonvásár. The latter influence only inward travel possibilities (they do not have any connections outwards) and they circulate only on workdays.

The Nagykanizsa line has a very specific train, connecting Budapest and Nagykanizsa on Friday afternoon. This train cannot be reached from any stations be-

tween Budapest and Székesfehérvár (with the last stop train one may get the previous fast train also). In some years, this fact was true into one direction, but in 2008 into both directions.

The number of direct trains was between 6 and 7 in the analysed period. The average speed was continuously decreasing, caused by the state of the tracks. In some years, a better timetable could cut back the journey time. The reconstruction works between Tárnok and Székesfehérvár in 2009-2010 will hopefully improve the connection.

Table 5. Summary output for the Budapest-Nagykanizsa line (2006)

Start	**Distance (km)**	**Rational traveller**		**Random traveller**		**Best choice rate (%)**
		average time (min)	**average speed (km/h)**	**average time (min)**	**average speed (km/h)**	
Budapest	221	210	63.2	479	27.7	-
Bfok-Belváros out	212	250	50.8	521	24.4	28.6
Bfok-Belváros in	221	251	52.8	494	26.8	71.4
Bfok-Belváros opt	-	238	54.2	493	32.2	-
Nagytétény out	205	244	50.4	514	23.9	71.4
Nagytétény in	229	258	53.3	502	27.4	28.6
Nagytétény opt	-	239	52.1	498	30.9	-
Érd alsó/felső out	201	229	52.7	492	24.5	83.7
Érd alsó/felső in	233	257	54.4	505	27.7	16.3
Érd alsó/felső opt	-	228	53.7	491	30.9	-
Tárnok out	197	224	52.9	487	24.3	100.0
Tárnok in	237	263	54.2	511	27.8	0.0
Tárnok opt	-	224	52.5	487	24.3	-

Source: own creation

Table 6. Summary output for the Budapest-Nagykanizsa line (2007)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	221	234	56.6	488	27.2	-
Bfok-Belváros out	212	268	47.4	519	24.5	16.1
Bfok-Belváros in	221	256	51.8	498	26.6	83.9
Bfok-Belváros opt	-	253	51.3	495	31.8	-
Nagytétény out	205	261	47.1	512	24.0	25.0
Nagytétény in	229	259	53.0	503	27.3	75.0
Nagytétény opt	-	253	51.7	498	30.7	-
Érd alsó/felső out	201	242	49.9	482	25.0	96.4
Érd alsó/felső in	233	255	54.8	497	28.1	3.6
Érd alsó/felső opt	-	242	49.6	482	25.4	-
Tárnok out	197	245	48.2	484	24.5	100.0
Tárnok in	237	267	53.2	508	28.0	0.0
Tárnok opt	-	245	48.2	484	24.5	-

Source: own creation

Table 7. Summary output for the Budapest-Nagykanizsa line (2008)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	221	234	56.6	475	27.9	-
Bfok-Belváros out	212	258	49.3	501	25.4	18.4
Bfok-Belváros in	221	247	53.7	481	27.6	81.6
Bfok-Belváros opt	-	244	53.2	480	32.1	-
Nagytétény out	205	252	48.9	494	24.9	18.4
Nagytétény in	229	247	55.7	485	28.3	81.6
Nagytétény opt	-	243	54.9	483	31.0	-
Érd alsó/felső out	201	237	50.9	482	25.0	95.9
Érd alsó/felső in	233	247	56.7	487	28.7	4.1
Érd alsó/felső opt	-	237	51.4	482	30.6	-
Tárnok out	197	239	49.4	481	24.5	89.8
Tárnok in	237	257	55.4	495	28.7	10.2
Tárnok opt	-	238	50.8	481	30.0	-
Martonvásár out	188	…	…	…	…	100.0
Martonvásár in	246	…	…	…	…	0.0

Source: own creation

We have to see that in 2006 (just like in 2005), the absolute attraction zone (R1 – from where the inward travel is always the best choice) is missing.

4.2.2. The Budapest-Szombathely line

This line has been a new challenge for the modeller. As the passengers of inward and outward travelling have different fast trains (while to Nagykanizsa they tried to reach the same fast train, now they are different – fast train via Veszprém, Intercity via Győr), the optimization process must be changed. In the row of Budapest we show the results of Intercity trains (rational travellers minimize the journey time). The optimal journey data may be peculiar because of this difference.

Table 8. Summary output for the Budapest-Szombathely line (2005)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	234	166	84.7	371	37.8	-
Bfok-Belváros out	224	248	54.1	473	28.4	20.0
Bfok-Belváros in	235	216	65.2	418	33.7	80.0
Bfok-Belváros opt	-	213	65.5	408	40.0	-
Nagytétény out	217	243	53.6	467	27.9	20.0
Nagytétény in	242	225	64.5	448	32.4	80.0
Nagytétény opt	-	216	65.9	432	38.6	-
Érd alsó/felső out	213	235	54.3	460	27.8	10.0
Érd alsó/felső in	246	210	70.1	407	36.3	90.0
Érd alsó/felső opt	-	210	68.3	396	41.0	-
Tárnok out	209	230	54.5	455	27.6	10.0
Tárnok in	250	220	68.2	422	35.5	90.0
Tárnok opt	-	218	66.3	411	39.3	-
Martonvásár out	200	223	53.7	448	26.8	45.7
Martonvásár in	259	227	68.3	430	36.1	54.3
Martonvásár opt	-	218	58.8	417	38.1	-
Baracska out	197	220	53.8	445	26.6	77.1
Baracska in	262	237	66.3	443	35.5	22.9
Baracska opt	-	218	55.2	426	36.8	-
Pettend out	192	215	53.6	440	26.2	100.0
Pettend in	267	242	66.2	448	35.8	0.0
Pettend opt	-	215	53.6	440	36.2	-

Source: own creation

In the 2005-2008 period, in both directions 5 trains were circulated that makes easy the comparison (there is no problem with the random traveller bias). The average speed between Budapest and Szombathely is much better than between Budapest and Nagykanizsa. This fact helps the railway to have a strong position against road transport.

In this case, the Martonvásár-Budapest trains add a strong support for inward travels for the stops of these trains, so the typical limit of the best choice based agglomeration is around Martonvásár.

Table 9. Summary output for the Budapest-Szombathely line (2006)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	234	175	80.1	379	37.1	-
Bfok-Belváros out	224	246	54.6	471	28.5	0.0
Bfok-Belváros in	235	208	67.7	414	34.1	100.0
Bfok-Belváros opt	-	208	67.7	402	40.8	-
Nagytétény out	217	240	54.2	465	28.0	0.0
Nagytétény in	242	214	67.7	419	34.6	100.0
Nagytétény opt	-	214	67.7	407	40.4	-
Érd alsó/felső out	213	231	55.3	457	28.0	6.1
Érd alsó/felső in	246	200	74.0	403	36.7	93.9
Érd alsó/felső opt	-	199	73.5	390	42.6	-
Tárnok out	209	226	55.5	452	27.8	12.2
Tárnok in	250	206	73.2	408	36.8	87.8
Tárnok opt	-	204	72.2	394	41.9	-
Martonvásár out	200	216	55.5	440	27.3	26.5
Martonvásár in	259	214	72.8	417	37.3	73.5
Martonvásár opt	-	209	68.1	399	40.8	-
Baracska out	197	212	55.6	436	27.1	100.0
Baracska in	262	239	65.8	443	35.5	0.0
Baracska opt	-	212	55.6	419	37.6	-

Source: own creation

The results of 2007 show how the agglomeration is getting out and out. For the random traveller, the average journey time is the same from Pettend, even if the distance by inward travel is 39% longer.

Table 10. Summary output for the Budapest-Szombathely line (2007)

Start	Dis-tance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	234	171	82.0	376	37.3	-
Bfok-Belváros out	224	247	54.5	473	28.4	0.0
Bfok-Belváros in	235	198	71.4	401	35.2	100.0
Bfok-Belváros opt	-	198	71.4	389	42.6	-
Nagytétény out	217	241	54.0	468	27.8	0.0
Nagytétény in	242	198	73.2	402	36.1	100.0
Nagytétény opt	-	198	73.2	390	43.2	-
Érd alsó/felső out	213	233	54.8	461	27.7	0.0
Érd alsó/felső in	246	186	79.4	401	36.8	100.0
Érd alsó/felső opt	-	186	79.4	390	44.2	-
Tárnok out	209	228	54.9	456	27.5	32.7
Tárnok in	250	206	72.7	414	36.3	67.3
Tárnok opt	-	205	67.6	397	42.3	-
Martonvásár out	200	218	55.0	443	27.1	60.0
Martonvásár in	259	215	72.2	423	36.8	40.0
Martonvásár opt	-	203	66.0	401	41.5	-
Baracska out	197	214	55.2	439	26.9	60.0
Baracska in	262	222	70.9	428	36.8	40.0
Baracska opt	-	203	66.0	404	40.8	-
Pettend out	192	208	55.4	433	26.6	100.0
Pettend in	267	227	70.6	433	37.0	0.0
Pettend opt	-	208	55.4	433	39.6	-

Source: own creation

The results of 2008 were shocking. The limit of the agglomeration (based on the average journey time) is around Baracska, but at some points of time, from the board of Lake Velencei is better to go back to the capital, instead of starting outwards. We have to remark that the average journey times are not typical ones in 2008, the dispersion of the data is relatively large.

Table 11. Summary output for the Budapest-Szombathely line (2008)

Start	Distance (km)	Rational traveller		Random traveller		Best choice rate (%)
		average time (min)	average speed (km/h)	average time (min)	average speed (km/h)	
Budapest	234	175	80.1	385	36.4	-
Bfok-Belváros out	224	246	54.5	468	28.7	5.7
Bfok-Belváros in	235	188	75.0	395	35.7	94.3
Bfok-Belváros opt	-	187	75.1	384	43.4	-
Nagytétény out	217	241	54.0	463	28.1	5.7
Nagytétény in	242	189	76.7	398	36.5	94.3
Nagytétény opt	-	188	76.8	386	43.9	-
Érd alsó/felső out	213	234	54.7	455	28.1	5.7
Érd alsó/felső in	246	196	75.3	404	36.5	94.3
Érd alsó/felső opt	-	194	75.5	391	43.2	-
Tárnok out	209	229	54.9	450	27.8	5.7
Tárnok in	250	201	74.6	409	36.7	94.3
Tárnok opt	-	198	74.9	394	42.9	-
Martonvásár out	200	219	54.7	441	27.2	40.0
Martonvásár in	259	209	74.3	417	37.2	60.0
Martonvásár opt	-	205	69.0	397	42.2	-
Baracska out	197	215	54.9	437	27.1	60.0
Baracska in	262	213	73.8	421	37.4	40.0
Baracska opt	-	205	65.2	399	41.8	-
Pettend out	192	209	55.1	431	26.7	60.0
Pettend in	267	218	73.5	426	37.6	40.0
Pettend opt	-	…	…	…	…	-
Kápolnásnyék out	189	205	55.3	427	26.6	60.0
Kápolnásnyék in	270	221	73.3	429	37.8	40.0
Kápolnásnyék opt	-	…	…	…	…	-
Velence out	186	202	55.2	424	26.3	80.0
Velence in	273	225	72.8	433	37.8	20.0
Velence opt	-	…	…	…	…	-
Velencefürdő out	185	199	55.8	421	26.4	100.0
Velencefürdő in	274	228	72.1	436	37.7	0.0
Velencefürdő opt	-	…	…	…	…	-

Source: own creation

4.2.3. Suggestion for zone limitations on Budapest-Székesfehérvár line

In Table 12 we summarise the limits of the different zones of the agglomeration. R1 stands for always better to go back to Budapest (the absolute attraction zone), R2 for the average or typical better zone (based either on average journey time, or on best choice), and R3 is the first station with always starting outwards label. In the case of ambiguity for R2, we computed the mean of the distances from Budapest.

According to the empirical evidence, the limit of the R1 zone should be Budapest-Kelenföld, so it can be linked to the planned inner zone of Budapest (congestion fee zone). The following limits can be Tárnok (the last morphologically homogenous settlement with Budapest), and Baracska (as the most far settlement with an important ratio of communing people to Budapest).

Table 12. Suggestion for zone limitations (with distances from Budapest-Kelenföld)

Zone	2005	2006	2007	2008	Suggestion
R1	—	—	—	—	Budapest-Kelenföld 0
	—	Nagytétény 12	Érd 16	—	
R2	—	—	Érd 16	Érd 16	Tárnok 20
	Martonvásár 29	Baracska 32	Baracska 32	Pettend 37	
R3	Pettend 37	Tárnok 20	Tárnok 20	Martonvásár 29	Baracska 32
	Pettend 37	Baracska 32	Pettend 37	Velencefürdő 44	

Source: own creation

4.3. Budapest-Vác-Szob- Štúrovo-Bratislava

The Budapest-Štúrovo-Bratislava line is very sensitive to the stops of international trains. In 2005, the stop-trains did not cross the state border, while the international trains did not stop in Hungary (except for Budapest); accordingly, for a journey of Szob-Štúrovo (neighbouring stations on the two sides of the border) one had to travel to Budapest and back. In that case, the boundary of the agglomeration is the state border. The physical analogy suggests Budapest to be a black hole attracting every object. In 2006 and 2007, three stop-trains crossed the state border, and an international fast train was *de facto* stop-train in Hungary, the gravitational zone decreased significantly. Since 2008, international trains have a stop in Vác, pulling in the boundary of the agglomeration.

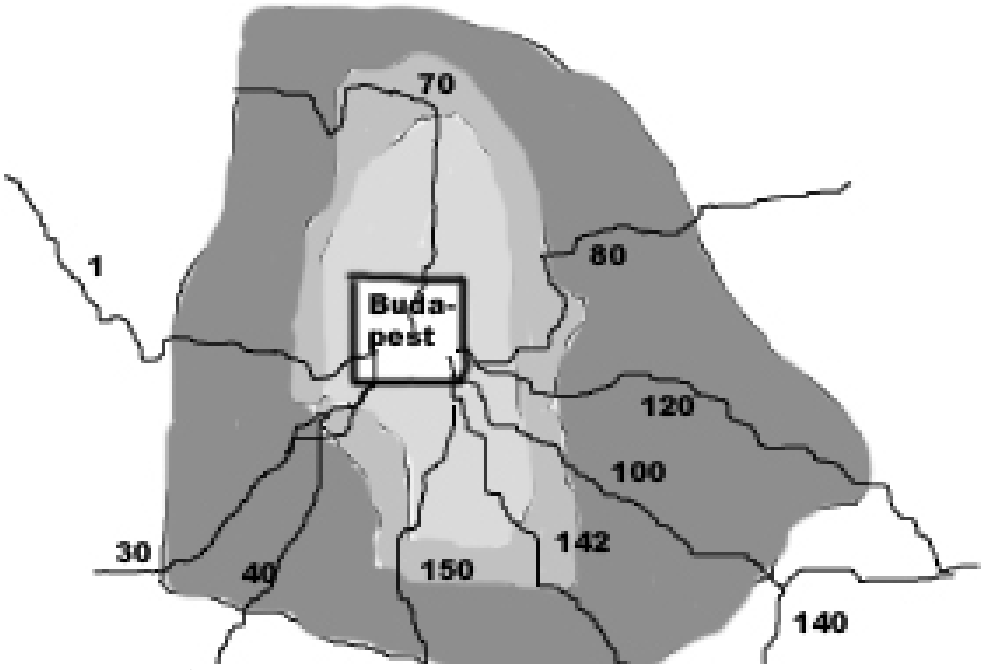
4.4. General empirical results

The empirical evidence highlights the sensitivity of the analysis. The limit of the gravitational zone depends on the frequency and the scheduling of the trains, not definitively on the speed difference of stop and fast trains. A good or bad timetable to one or other direction may halve or double the zone. The responsibility of (timetable) scribes is serious to the local population.

When the railway escape velocity is computed, the transfer time has to be included in the case of stop trains. Thereby the railway escape velocity is a relative measure (in the form of a ratio). The denominator includes the average speed of inwards travel; the numerator is for the average speed of outward travel. If we use the average speed for railway escape velocity, the same measure must be used for the delimitation of the agglomeration.

In Figure 3, we show a map of gravitational zones.

Figure 3. Example map of gravitational zones



Source: own creation

4.5. Limitations of the model

We have to acknowledge that the applied method have a series of deficiencies. For the analysis of the gravitational zone, we have the following problems:

- the location can only be made along the main railway lines (the map is not continuous),
- where the scanned city has special physical geographical nature (f.i. it is in a narrow valley, or on the coast), the ineluctable asymmetry of the railway system will distort the analysis (some obstacles are passed by other means of communication),
- the asymmetry of timetables results in an asymmetric delimitation,
- national borders have distortion effects.

While modelling of railway escape velocity, our problems are:

- the definition and measuring of mass is problematic, the quantity of necessary information is huge,
- the railway escape velocity is a strongly complex measure; one average speed is not satisfactory.

5. Summary

The first phase of the research on the railway escape velocity helps to measure and define the elements of the escape velocity. The achieved results are suitable to clearly see the possible problems of the method and to develop the adequate techniques to aim our goals. For the full calibration we need the analysis of all railway lines around Budapest, so the constant and the parameter of the distance can be calibrated. For the calibration of the mass parameter, at least several important railway nodes have to be scanned.

Our model is applicable to discover the gravitational space of any important city with extended railway network. Furthermore, we can draw maps of agglomeration and proposed fare zones. The latter may be constructed even in the lack of objective counting of travellers and without the use of gravity models (based on flows).

By the empirical evidence, well-constructed timetables with good transfer possibilities compensate the deteriorative infrastructure. The road transport is equal to the random traveller (departure is possible at any point of time by car), so the comparable average speed of 25-40 kmph of the railway transport is not competitive.

References

Dusek, T. 2003: A gravitációs modell és a gravitációs törvény összehasonlítása. *Tér és Társadalom*, n. 1, pp. 41-58.

Dusek, T.; Szalkai, G. 2007: Területi adatok ábrázolási lehetőségei speciális kartogramokkal. *Területi Statisztika*, n. 1, pp. 3-19.

Dusek, T.; Szalkai, G. 2008: *Differences between geographical space and time spaces: theoretical consequences and Hungarian examples*. 48th Congress of

the European Regional Science Association, Liverpool, UK, August 27-31 2008.

Ecostat 2006: *A gazdasági növekedés hosszú távú előrejelzése. A gazdasági szerkezet termelékenység, munkaerő-kereslet, globális előrejelzés*. Ecostat, Budapest; 2006.

Fotheringham, A. S.; Haynes, K. E. 1988: *Gravity and Spatial Interaction Models*. SAGE Publications, London.

Holics, L. (szerk) 1986: *Fizika*. Műszaki Könyvkiadó, Budapest.

Iacono, M.; Levinson, D.; El-Geneidy, A. 2008: Models of Transportation and Land Use Change: A Guide to the Territory. *Journal of Planning Literature* n. 4, pp. 323-340.

Janelle, D. G. 1969: Spatial reorganization: a model and concept. *Annals of the Association of American Geographers*, n. 2, pp. 348-364.

Janelle, D. G. 1975*: Time space convergence and urban transportation issues.* In Blong, C. K. (ed) *Systems Thinking and the Quality of Life*. The Society for General Systems Research, Washington DC, pp. 594-600.

Knowles, R. D. 2005: *Transport shaping space*. Paper presented at Annual Meeting of The Association of American Geographers, Denver, Colorado, USA.

Kotosz B. 2010: Gravitációs modellek a területi statisztikában. In: *Gondolatok közös javainkról*. Budapesti Corvinus Egyetem Közgazdaságtudományi Kar, Budapest, pp. 147-161.

Levine, J.; Inam, A.; Torng, G-W. 2005: A Chioce-Based Rationale for Land Use and Transportation Alternatives:Evidence from Boston and Atlanta. *Journal of Planning Education and Research*, n. 24, pp. 317-330.

Levinson, D. M.; Kumar, A. 1993: Multi-modal Trip Distribution: Stucture and Application. *Transport Research Record*, n. 1466, pp. 124-131.

Reilly, W. J. 1929: *Methods for the Study of Retail Relationship*. University of Texas, Texas.

Rodrigue, J-P et al. 2006: *The Geography of Transport Systems*, Hofstra University, Department of Economics & Geography, Hofstra, USA.

Wegener, M. 2001: New spatial planning models. *International Journal of Applied Earth Observation and Geoinformation,* n. 3, pp. 223-237.

Methodology and estimation of tourism-related road toll expenditures: experience from the Slovene Tourism Satellite Accounts

Matevz Raskovic[1]

The implementation of the complex Tourism Satellite Accounts (TSA) methodology for Slovenia has posed considerable challenges in different areas of tourism statistics. Particularly, due to lack of data in some tourism areas (i.e. transit visitors, same-day visitors, etc.), various estimation methods had to be developed to fill in the gap, as it is often the case with complex socioeconomic phenomena. Within the TSA methodology, transportation-related expenditures, such as gas and road toll expenditures represent an important product and service category, often ranking among top 5 tourism expenditure categories. This is especially true for small countries like Slovenia, with strategic geographical position, making ideal transit and on-the-way tourism destinations to other top tourist destinations (i.e. Croatia). The purpose of this paper is to provide experience, estimation tools and best practices in evaluating tourism-related road toll expenditures for different visitor categories in Slovenia for the year 2006. As such, the goal of the paper is to explain, how tourism-related road toll expenditures have been estimated within the TSA methodology for Slovenia for 2006 and to outline an estimation approach, which can be used either within the TSA or individually on its own.

Keywords: Tourism Satellite Accounts, road toll expenditures, estimation tools, best practices, Slovenia

1. Introduction

According to the United Nations World Tourism Organization (UNWTO) tourism is today in 4 out of 5 world countries among top 5 national industries. Globally, in 2008, international tourism generated over 642 billion EUR of revenues and represented over 30% of world exports of services (UNWTO 2009). Yet despite its importance, in terms of its impact on both the national and world economy, evaluating tourism, because of its demand driven phenomenon, is often hard and far from being methodologically exact. While the implementation of the Tourism Satellite Ac-

[1] Matevz Raskovic, PhD candidate, assistant and researcher, University of Ljubljana, Faculty of Economics, Slovenia.

counts (TSA) methodology is today believed to be the most in-depth and wholesome quantitative evaluation of tourism on the national economy, representing also an extension of the methodological approach of the System of National Accounts (SNA), the lack of appropriate data for some tourism-related expenditures and visitor categories (i.e. transit visitors, same-day visitors, etc.) means that evaluating these selected tourism expenditures today is still as much an art, as it is science. Nowhere is this more apparent than in the area of forecasting different tourist demand components (Chu, 2008). In this context, the development of comprehensive and consistent measurement approaches to compensate for the lack of reliable and available data in tourism, also for country comparison purposes, is one of the primary motivations for the use of TSA today (Frechtling 1999, Rivera 1999, Sharma-Olsen 2005). It is because of this, that sharing best practices and outlining transparent estimation approaches is of vital importance, particularly for newcomer countries, which are implementing the TSA methodology for the first time, and face considerable estimation challenges and questions, which the TSA Methodological Recommended Framework (Eurostat 2008) does not address in sufficient detail.

In Slovenia, tourism represents *"one of the leading economic activities, with a profound impact on national economy, regional development and employment"* (SURS 2009). Thus, tourism and its economic impact must be taken into consideration from the viewpoint of state, regional and economic development (Stynes 1999). However, despite its importance, the overall impact of tourism on Slovenia's national economy begun to be evaluated only with the implementation of the TSA methodology for the year 2000 only in 2004. Having said this, considerable progress has been made for 2003 and 2006 TSA estimations in the year 2007, where several new estimation approaches have been developed to compensate for the lack of available primary statistical data in some areas.

The structure of the TSA methodology is based on measuring the complexity of tourism in a national economy, and based also on its balance between demand for tourism-related products and services on one hand, with regards to their supply on the other hand (Rašković – Zagoršek 2008). It must also be noted that tourism measurement is particularly difficult and challenging, since tourism is implicit in many different industries, but does not comprise 100% of one single industry (Sacks 2004). Having said this, the general ideas of the TSA methodology are: (1) the analysis of all aspects of demand for goods and services, wholly or partly associated with tourism within the economy, (2) the observance of the operational interfacings with the supply of such goods and services within the reference economy, and (3) the description of supply and demand sides with other economic activities (Eurostat, TSA: Recommended Methodological Framework 2008).

Within the TSA methodology transportation-related expenditures such as gas and road toll expenditures represent an important product and service category, often ranking among top 5 expenditure categories. This is especially true for small coun-

tries, with strategic geographical positions, making ideal transit and on-the-way tourism destinations, where most tourist and visitor arrivals are done by road. Indeed, Slovenia's geographic position, lying between East and West, North and South, and neighboring to Croatia, one of Europe's top summer tourism destinations makes it a popular by-car transit tourism destination. Having said this, an estimated 12 million transit cars passed Slovenia in 2006 (Zagoršek et al. 2008). In addition, close to 87% of all tourist arrivals were by-road arrivals (Zagoršek et al. 2008). This is also consistent with other research, which has shown that *"the ability of visitors to travel around destination areas is crucial for tourism"* (Dickinson-Robbins 2008).

The purpose of the paper is to provide experience, estimation tools and best practices for evaluating tourism-related road toll expenditures for different visitor categories (i.e. domestic and foreign tourists, domestic and foreign same-day visitors and transit visitors). Let us here note that the term visitor is a broader term then tourist, addressing both tourists with an overnight stay, as well as transit and same-day visitors (without an overnight stay). As such, the goal of the paper is to explain, how tourism-related road toll expenditures have been estimated within the TSA methodology for Slovenia for 2006 and to outline an estimation approach, which can be used either within the TSA framework or individually. In the first part of the paper, a brief overview of the TSA methodology is outlined, since understanding the logic and methodological structure is crucial for subsequent estimations. This is followed by an overview of the various tourism-related visitor categories for Slovenia in the second part. The third part of the paper outlines a series of 'physical flows' (i.e. number of check-ins at road-toll stations) which provide a quantitative basis for our evaluations and estimations. The fourth part outlines and describes our estimation approach to the evaluation of tourism-related road toll expenditures, followed by limitations of the research and recommendations for implementation of the estimation approach in the fifth part, followed by a short conclusion in the sixth part.

2. A brief overview of the TSA methodology

France was the first country to start developing a methodological extension of the System of National Accounts (SNA), aimed first at evaluating the housing sector (Širše et al. 2004). By the 1970s France started to use the term *satellite accounts* for a methodological approach aimed at those accounting practices in specific horizontal areas, which were not correctly identified in the SNA, but could be approached methodologically as satellite sub-systems (Eurostat, TSA: Recommended Methodological Framework 2008).

By the beginning of the 1980s the UNWTO commissioned the preparation of Spain's tourism accounts, based on the SNA, which became one of the key documents at the UNWTO General Assembly in New Delhi in 1983 (Eurostat, TSA: Recommended Methodological Framework 2008). By 1985 Sweden was the first

country to partially implement the TSA methodology, however only for the supply side (Širše et al. 2004). By the beginning of the 1990s the Organization for Economic Cooperation and Development (OECD) began developing its Tourism Economic Accounts (TEA). In 1991 the Ottawa International Conference on Travel and Tourism Statistics held by the WTO provided a basis for the issue of Recommendation of Tourism Statistics in 1993, outlining the TSA methodology.

The first complete TSA were implemented by Canada in 1994, involving both the supply and demand sides. In 1999 the UNWTO conference at Nice presented a Conceptual framework for the TSA adopted both by the UNWTO, OECD and EUROSTAT. By the year 2000 EUROSTAT, OECD and UNWTO jointly issued the TSA: Recommended Methodological Framework, which is still in use today. In 2002 EUROSTAT issued in addition A Manual for use and implementation of TSA (Širše et al. 2004). Today the TSA methodology has started to be applied even for specific areas of tourism (i.e. TSA for the gambling industry) and for specific tourism regions, as well as providing important basis for economic policy and tourism decision making.

2.1. The basic idea of the TSA methodology

As already mentioned, the TSA methodology builds on the concept of SNA, which are the most complete and extensive representation of economic processes and activities within a national economy. The TSA comprehensively analyze myriad aspects of tourism-related demand for products and services, and compares it with the supply side. This approach is underpinned by the uniqueness of tourism being very much a demand driven socio-economic phenomenon, thus mostly following an expenditure approach. In addition, TSA further evaluate how tourism-related supply of products and services further directly and indirectly impacts other economic actors, industries and activities, and thus evaluates the total impact of tourism on a national economy.

Having said this, the TSA consist of 10 individual tables, which measure the demand and supply side of (1) tourism-related products and services, (2) tourism-connected products and services, and (3) non-tourism connected products and services, indirectly tied to some extent with tourism. The TSA further explore the impact of (a) inbound domestic and foreign tourism, (b) outbound domestic tourism, and the impact of tourism on (c) employment and evaluate tourism in terms of its (d) GDP percentage, (e) employment, (f) added value, (g) capital investments, (h) fiscal revenues (taxes, social expenditures, etc.) and its impact on the (i) balance of payments. The application of TSA enables a clearer understanding of the impact of tourism on a particular economy, provides a basis for more effective and efficient eco-

nomic and developmental policies and strategies, and enables a world-wide comparison of data on tourism and its impact on various national economies.

Tables 1, 2 and 3 of the TSA represent the most important component of tourism-related demand, outlined by various types of visitor categories – the final cash expenditures. In this context Table 1 evaluates inbound expenditures by different foreign visitor categories (i.e. same-day visitors, transit visitors and tourists, spending at least 1 night in the country of reference). Table 2 evaluates domestic tourism expenditures by different visitor categories, while Table 3 evaluates outbound domestic tourism expenditures by different visitor categories. Table 4 of the TSA represents a summary of all tourism-related cash expenditures, generated by inbound and outbound, domestic and foreign visitor categories. In addition, it also takes into account expenditures in kind and tourism-related social transfers. The final result is thus internal tourism-related demand in cash and kind. This represents the basis for evaluating tourism-related share in a country's GDP and tourism-related added-value.

Table 5 of the TSA represents the supply of tourism and non-tourism related industries in the national economy. Their output is divided into (1) tourism-specific product and service categories (i.e. hotel services, food and beverage expenditures etc.), (2) tourism-connected products and services (i.e. postage, insurance etc.) and (3) non-specific products and services (i.e. fuel expenditure, retail etc.). Table 6 of the TSA summarizes the data on tourism-connected demand and supply, and adds estimates on the level of tourism-related imports and net taxes from tourism-related production. Table 7 of the TSA evaluates tourism in terms of tourism-related employment and number of jobs, while Table 10 represents an overview of selected physical tourism-related flows, such as the numbers of different visitor categories, number of overnight stays and number of arrivals. Tables 8 and 9 and seen as auxiliary tables within the TSA, and look at capital investment flows, connected with tourism.

Based on the outlined structure of the TSA methodology and its complexity, the UNWTO (2000) points to 3 important principles in the process of TSA implementation. First, a wider stakeholder support must be obtained in order to facilitate cooperation and information sharing. Second, both conceptual and organizational flexibility are needed in the process of the actual development of the TSA. It is in this area particularly, that sharing of best practices in terms of estimating specific constructs is vital and that this paper partially addresses. Third, the implementation of the TSA has a strong learning curve and should be implemented on a gradual step-by-step basis. It is in this area that the OECD (Draft OECD Guidelines for a Tourism Satellite Account 1999) emphasizes the need for establishing priorities in estimation procedures and obtaining high quality information (Sharma - Olsen 2005).

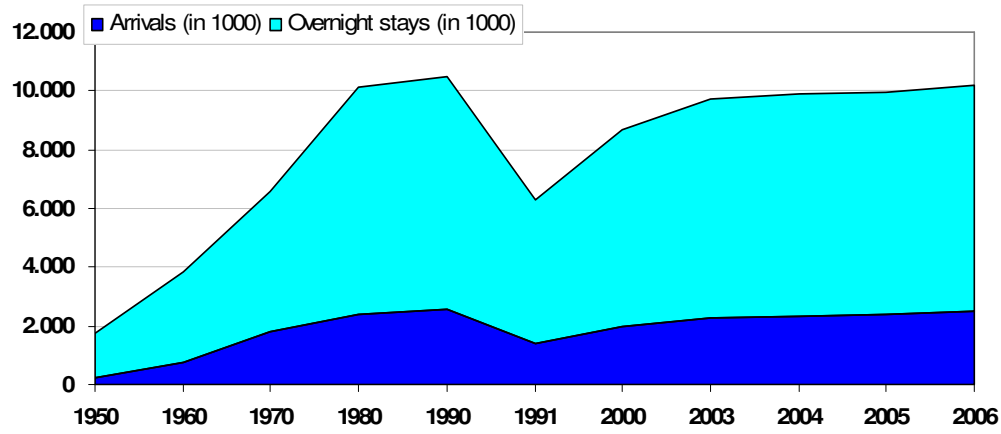
3. Relevant visitor categories for Slovenia for 2006

Since the paper address the estimation of tourism-related road toll expenditures for Slovenia for the year 2006, all relevant tourism information will be presented for 2006 as well.

3.1. A brief overview of Slovene tourism up to 2006

According to the data from the Slovene National Statistical Office (now on denoted as SURS) the most prosperous year for Slovene tourism was 1986, when Slovenia backed by a strong and 'patriotic' Yugoslav 22-milliom market recorded 2.8 million arrivals and over 9.2 million overnight stays (SURS 2009). Following Slovenia's independence in 1991 and conflicts in neighboring Croatia and Bosnia and Herzegovina, visitor numbers plummeted and began to pick up only by the mid 1990s. Since the mid 2000 the arrival of low cost airlines further increased visitor numbers, bringing visitors closer to the 1986 peak. Figure 1 displays a comparison of arrivals and number of overnight stays in the period 1950 up to 2006, while Figure 2 displays the ratio of domestic to foreign tourist arrivals and overnight stays from 1991 up to 2006.

Figure 1. Comparison of tourism arrivals and overnight stays in the period between 1950 and 2006 (in 1000)

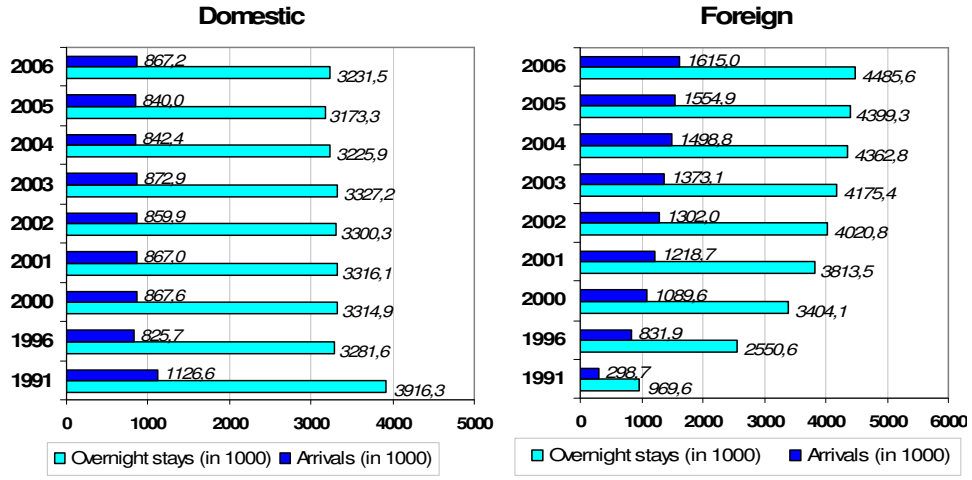


Source: SURS, Statistical Yearbook, 2007; own depiction

As the data from Figure 2 shows, the number of domestic tourist arrivals increased by more than 5% from 1996 to 2006, while the number of overnight stays fell by

around 2% respectively. In the same period the number of foreign tourist arrivals increased by 94%, while the number of overnight stays rose by some 76% (SURS, Statistical Yearbook 2007).

Figure 2. Comparison of domestic and foreign tourism arrivals and overnight stays in the period between 1991 and 2006 (in 1000)



Source: SURS, Statistical Yearbook, 2007; own depiction

With regards to the country of origin, German tourist overnight stays have decreased from 23.5% in 1995 to 13.9% in 2006, while the share of overnight stays by Italian tourists has increased from 15.9% in 1995 to 19.8% in 2006. These two groups of tourists also represent the most important tourists in terms of country of origin. In large part the increase of Italian tourists may be attributed to a ban on gambling in Italy and a thriving gambling industry in Slovenia. Hence, the gambling industry is also an important driver of same-day tourist visitors. The most prominent rise of the share of overnight stays is seen for UK tourists, whose share of overnight stays has increased almost 2.5-fold between 1995 and 2006, mostly because of low cost airlines. In addition, the share of overnight stays by US tourists has almost doubled from 1.3% in 1995 to 2.4% in 2006 respectively (SURS 2009).

3.2. Relevant tourism-related visitor flows for Slovenia for 2006

The TSA methodology distinguishes between the following tourism-related visitor categories, which are relevant to the evaluation of tourism-related road toll expenditures: (1) *domestic tourists*, (2) *foreign tourists*, (3) *domestic same-day visitors*, (4) *foreign same-day visitors* and (5) *foreign transit visitors*. Here, a tourist is referred

as a visitor with at least one overnight stay in the country of reference. While precise statistical data is usually available for domestic and foreign tourists incurring an overnight stay, less statistical data is available for same-day and transit visitors, which most often have to be estimated, as was the case also for Slovenia.

The number of (1) **domestic tourists** is based on data from SURS (Statistical Yearbook 2007), which registered 867,200 domestic tourist arrivals and 3.2 million domestic tourist overnight stays for 2006. In addition to inbound domestic tourist trips (made almost exclusively by car) 1.9 million car trips were also made by Slovenes as tourist abroad (outbound) and starting from Slovenia (SURS, Statistical Yearbook 2007).

Multiplying these 1.9 million outbound car trips by an average of 2.8 household members, we get some 5.3 million outbound domestic tourists travelling by car. Similarly, the data for (2) **foreign tourists** is also based on data from SURS (Statistical Yearbook, 2007), which recorded 1,615,000 foreign tourist arrivals and 4,485,600 overnight stays for 2006. Applying the 87% share of incoming tourist arrivals by car, 1.4 million foreign tourists are thought to have arrived by road in 2006.

The number of (3) **domestic same-day visitors** were estimated at about 9.8 million for 2006. This estimation is based on the number of domestic same-day car trips from the Survey of tourism trips by domestic visitors for 2003 (latest possible data) commissioned by the Slovene Tourism Organization. Estimates for 2006 were based on the extrapolation of 2003 structures. Based on the Survey on tourism travels of domestic population for 2003 conducted by SURS and extrapolated for 2006, an estimated 4.1 million same-day trips were made by Slovenes, out of which 3.5 million were made by car (86.2%). Multiplying this share by the average number of household members in Slovenia (2.8 based on the last 2002 census) we get some 9.8 million by-car domestic same-day visitors.

Among all visitor categories the estimation of the number of (4) **foreign same-day visitors** (excluding transit visitors, who were estimated separately) was hardest, due to lack of systematic statistical data collection. Because of Slovenia's size and geographical position, it is a popular one-day visitor destination for many tourists from Croatia, Austria and Italy. Because they do not generate an overnight stay and usually visit just one or two top tourist attractions in Slovenia, official statistics and has not been able to develop a systematic way of their tracking. In addition, the division between same-day and transit visitors, with the latter only crossing Slovenia to their final tourist destination, makes it even harder. Having said this, based on estimates from the main tourist attractions in Slovenia and more thorough data on foreign same-day gambling visitors, which constitute a sizable portion of all foreign same-day visitors (mainly from Italy, where gambling is prohibited by law), we have estimated some 2.1 million same-day foreign visitors for 2006 (Estimation of TSA for 2003 and extrapolation for 2006, 2007). Due to the nature of same-day

tourism visits, only a negligible share (< 5%) of these visitors is thought to have arrived by non-road transport (i.e. train).

While (5) **foreign transit visitors** are sometimes grouped with foreign same-day visitors, Slovenia's geographic position and incoming visitor structure call for a division between the 2 categories. According to the survey data from SURS (Statistical Yearbook, 2007) almost 87% of all incoming tourists arrive to Slovenia by road. In addition, according to data from border crossings, an estimated 38.2 million transit visitors (tourism and non-tourism related) passed Slovenia in 2006 by car, which makes for some 12 million transit cars in 2006. Most of them in the late spring, summer and early autumn seasons, going to Croatia; one of Europe's top summer tourist destinations. Thus an estimated 65% of all individual person border crossing are thought to be transit-related (Estimation of TSA for 2003 and extrapolation for 2006, 2007).Of these transit-related visitors about 60% are thought to be tourism-related which brings us to some 22.9 million tourism-related foreign transit visitors in 2006. Table 1 in the end summarizes the presented data or estimations for all the 5 relevant visitor categories for 2006.

Table 1. Summary of relevant data and estimations for all 5 visitor categories for Slovenia for 2006

Visitor category	Number of by-car visitors for 2006
Domestic same-day visitors	~ 9.8 million
Domestic tourists	0.867 million (+ 5.3 million outbound)
Foreign transit visitors	22.9 million
Foreign same-day visitors	~ 2.1 million
Foreign tourists	1.4 million

Source: SURS, Statistical Yearbook, 2007; Zagoršek et al., 2008; own estimations

The provided data and estimations of the 5 relevant visitor categories in Table 1 was later on used to divide the total estimated amount of all road toll expenditures, among the 5 relevant visitor categories. This is described in greater detail in the next section of the paper.

4. The methodological approach

In the next section, the paper outlines the general methodological approach adopted in the estimation of tourism-related road toll expenditures for Slovenia in 2006. In this context, a top-to-bottom approach was used in the sense that, first the cumulative share of total tourism-related road toll expenditures was estimated using a differential approach (explained in greater detail further on) and second, this expendi-

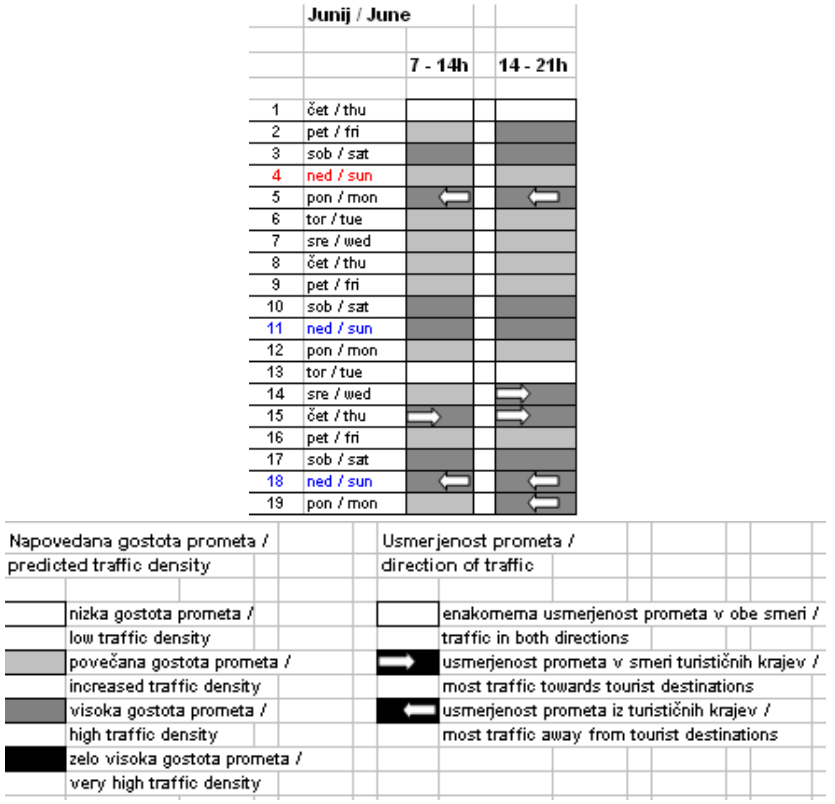
ture amount was later divided between particular visitor categories (i.e. domestic tourists, foreign same-day visitors, transit visitors, etc.).

4.1. Relevant input sources

The main methodological approach to the evaluation of cumulative tourism-related road toll expenditures is based on comparing (1) physical flows from road toll stations in **tourist and non-tourist times (seasons) of year** and (2) comparing flows between **regular working days, weekends and holidays**. In all cases the difference between tourist vs. non-tourist, as well as regular work days vs. other days was corrected by a factor and attributed to tourism. Hence, the difference obtained from this comparison was then used as an input for the estimation of cumulative tourism-related road toll, as a percentage of total (tourism and non-tourism) road toll expenditures in 2006. Having said this, 3 key information sources provided the basis for our estimates.

The first information sources represented (1) **actual physical data on the number of crossings at all road toll stations** across Slovenia gathered systematically by DARS, a state-run Motorway Company of the Republic of Slovenia. All together, data was gathered from 26 road-toll stations in Slovenia, which registered almost 60 million crossings by all non-carriage vehicles (i.e. commercial trucks etc.). In this part, data was compared for first 2 classes of road toll, thus including (a) personal vehicles (toll class R1), as well as other vehicles up to 3.5 tons (toll class R2). The second information source was the so called (2) **Traffic calendar**, which displays a day-by-day classification of physical flows, with direction of flows, and classifies each day into one of the following categories: (1) low traffic density, (2) increased traffic density, (3) high traffic density and (4) very high traffic density. Figure 3 displays a short excerpt from this calendar for illustration.

Figure 3. Excerpt of the Traffic calendar for 2006, showing direction of traffic and traffic density



Source: DARS, Traffic calendar, 2007a

Based on this information source tourism-related physical flows were calculated as the difference (hence the differential approach) between traffic flows in different density periods, week days and times of year. This is explained in greater detail also in Table 2.

Table 2. An excerpt from actual physical data on the number of crossings at road toll stations collected by DARS for 2006 (road toll class R1)

Date / Toll station	**KOZINA**	**TOROVO**	**VRANSKO**	**Day**	**Season**
16. 1. 2006	912	13.632	12.087	*Monday*	NT
17. 1. 2006	935	13.585	12.535	*Tuesday*	NT
18. 1. 2006	999	13.896	12.856	*Wednesday*	NT
19. 1. 2006	1.130	15.376	13.578	*Thursday*	NT
20. 1. 2006	1.227	17.281	16.042	*Friday*	NT
21 1. 2006	1.021	17.679	12.516	*Saturday*	NT
22. 1. 2006	1.007	14.466	10.176	*Sunday*	NT
...	**...**	**...**	**...**	**...**	**...**
24. 7. 2006	1.903	19.436	14.956	*Monday*	T
25. 7. 2006	1.622	18.297	13.587	*Tuesday*	T
26. 7. 2006	1.723	19.490	14.268	*Wednesday*	T
27. 7. 2006	1.732	22.286	14.959	*Thursday*	T
28. 7. 2006	2.458	29.085	17.815	*Friday*	T
29. 7. 2006	3.603	33.684	20.966	*Saturday*	T
30. 7. 2006	2.591	26.346	17.467	*Sunday*	T

Source: DARS, Traffic calendar, 2007a.*NT: non-tourist season; T: tourist season

Pertaining to the excerpt data shown in Table 2, let us just point, that the classification of tourist seasons pertained to all school holidays (i.e. Christmas, Easter, etc.), as well as the annual summer holiday period between July and August. This division was also employed by DARS in their Traffic calendar. In terms of type of day (i.e. regular working vs. non-working) Tuesdays through Thursdays were classified as typical regular working days. Mondays and Fridays were not classified as regular working days, since they often hold tourists and visitors on extended weekend holidays, as well as student commuters.

The third source of data was linked to the (3) **total amount of collected road toll** by DARS in Slovenia for 2006. According to their annual report for 2006, DARS collected some 36.3 billion SIT or approximately 151.4 million EUR (DARS, Annual report 2007b). Thus, the cumulative of all tourism-related road expenditures for 2006 is shown as the percentage of all road toll collected by DARS, which is then further divided between different visitor categories based on their physical flows. The next section of the paper gives a more detailed description of the methodological approach.

4.2. Estimating the cumulative tourism-related road toll expenditures for 2006

Based on the data on road toll crossings the average number of daily crossings was calculated separately for tourist and non-tourist seasons, as well as for regular working days and non-working days. The general assumption was thus that traffic on Tuesdays, Wednesdays and Thursdays in non-tourist season is the best indicator of non-tourism related traffic. Table 3 represents a brief excerpt of the outlined comparison for regular working days in tourist and non-tourist season for illustrative purposes, and for 3 examples of different road toal stations in Slovenia.

Table 3. An excerpt from a comparison of the total number of crossings and average number of daily crossings in tourist and non-tourist seasons for regular working days (road toll class R1)

Season	Variable	KOZINA	TOROVO	VRANSKO
Non-tourist season	*Cumulative*	134,640	1,864,053	1,724,860
	Average daily	1,213	16,739	15,539
Tourist season	*Cumulative*	60,499	732,236	576,295
	Average daily	1,681	20,340	16,008

Source: DARS, Traffic calendar, 2007a; own calculations

Based on the provided average daily number of crossings in non-tourist and tourist seasons for work days depicted in Table 3, we can see sizable differences between the average number of daily crossings in tourist and non-tourist working days for the three displayed road toll stations. However, we believe that even in non-tourist work days a portion of road traffic can still be tourism-related. Having said this, Table 4 displays the calculated difference between average daily number of crossings in non-tourist and tourist seasons, where the basis for non-tourist comparisons was estimated with the help of factors. A similar approach was also used for road toll group R2.

Table 4. A comparison of average daily crossings in tourist and non-tourist seasons for working days (road toll class R1) and corresponding tourism-related factors

	Variable	**KOZINA**	**TOROVO**	**VRANSKO**
Non-tourist season	*Average daily*	1,213	16,739	15,539
	Tourism factor	0.2	0.25	0.25
	*Factored average daily**	1,213*(1-0.2)= = 971	12,595	11,654
Tourist season	*Average daily*	1,681	20,340	16,008
Average daily difference		710	7,745	4,354
Number of working days		250	250	250
Total difference		710*250= = 177,500	1,936,250	1,088,500

Source: DARS, Traffic calendar, 200a7; own calculations

A slightly different approach was however used for non-working days[2]. Here the difference between non-tourist and tourist non-working days was estimated as the minimum number of crossings in either non-tourist or tourist seasons. This was further corrected by a selected factor (for the degree of estimated tourism) and further multiplied by the number of non-working days. The obtained estimation was then compared to the actual number of crossings on non-working days, both in tourist or non-tourist seasons. The obtained difference was used as an input in further calculations. Table 5 displays an excerpt from the estimation of tourism-related road toll for road toll class R1 on non-working days.

[2] For this purpose 115 Saturdays, Sundays and holidays were taken into account.

Table 5. A comparison of average daily crossings in tourist and non-tourist seasons for
non-working days (road toll class R1) and corresponding factors

	KOZINA	TOROVO	VRANSKO
Minimum daily crossings in non-tourist season	829	11,101	9,339
Minimum daily crossings in tourist season	655	8,184	6,302
Tourism factor	0.15	0.20	0.20
Factored minimum daily crossings	655*(1-0.15)= = 557	6,547	5,042
Number of non-working days	115	115	115
Total factored crossing on all non-working days	115*557[3]= = 64,026	752,928	579,784
Sum of all crossings on non-working days (in non-tourist and tourist season)	185,685	2,057,272	1,703,755
Difference between sum of all crossings and total factored crossings on all non-working days	121,659	1,304,344	1,123,971

Source: DARS, Traffic calendar, 2007a; own calculations

In the last part, the calculated cumulative differences between tourist and non-tourist seasons for working days (Table 4) were multiplied by the average price of the road toll crossing. These prices were provided by DARS. Since the data in Table 4 is provided only for the road toll class R1 Table 6 in the next section illustrates the final calculation for working days for road toll class R1. A similar approach was also used for class R2 as well.

[3] In the multiplication process non-rounded numbers with two decimals were used.

Table 6. An excerpt of cumulative expenditures for 3 illustrated road toll stations for working days (road toll class R1) in Slovene Tolars (SIT)

	KOZINA	**TOROVO**	**VRANSKO**
Total difference	177,500	1,936,250	1,088,500
Average price of crossing	690 SIT	380 SIT	310 SIT
Cumulative expenditures	122,500, 771 SIT	735,795,327 SIT	337,414,461 SIT

Source: DARS, Traffic calendar, 2007a; own calculations

A similar approach was also employed for the non-working days, where an additional factor was used, based on the importance of the road toll station in the commuting flows on non-working days (Mondays, Fridays and weekends). Table 7 illustrates the final calculation for non-working days for road toll class R1. A similar approach was also used for class R2.

Table 7. An excerpt of cumulative expenditures for 3 illustrated road toll stations for non-working days (road toll class R1) in Slovene Tolars (SIT)

	KOZINA	**TOROVO**	**VRANSKO**
Difference	121,659	1,304,344	1,123,971
Average price of crossing	690 SIT	380 SIT	310 SIT
Additional importance factor	0.75	0.85	0.80
Cumulative expenditures	62,958,532 SIT	421,303,112 SIT	278,744,808 SIT

Source: DARS, Traffic calendar, 2007a; own calculations

For the end, Table 8 displays the estimated cumulative tourism-related road toll expenditures based on the illustrated differential approach for all 26 road toll stations and both road toll classes (R1 and R2). Let us also point out at the end that smaller factors were used to estimate the share of tourism-related road toll in the road toll class R2, which inlcudes vehicles up to 3.5 tons, corresponding to lower shares of tourism in this vehicle category.

Table 8. Joint total tourism-related road toll expenditures for Slovenia in 2006 in Slovene Tolars (SIT)

	Class R1	**Class R2**
Working days	4,154,685,467 SIT	850,913,216 SIT
Non-working days	4,628,597,464 SIT	196,359,187 SIT
JOINT TOTAL		**9,830,555,334 SIT**

Source: DARS, Traffic calendar, 2007a; own calculations

Based on the outlined methodological approach, a joint total of about 9.8 billion SIT was estimated for all tourism-related road toll expenditures in Slovenia in 2006. This represents 27.09% of all road toll expenditures (36.3 billion SIT) collected by DARS in 2006. Such a high share of tourism-related road toll expenditures vis-a-vis all road toll expenditures may be attributed to Slovenia's central geographical position, as well as a high number of transit visitors.

4.3. Division of estimated tourism-related road toll expenditures among visitor types

In the last section, Table 9 displays a summarized disaggregation of the total estimated tourism-related road toll expenditures of 9.8 billion SIT or about 27% of all road toll expenditures for 2006 (DARS, Annual report, 2007b). The disaggregation is based on the physical flows of the 5 visitor categories. For domestic (inbound) and foreign by-car tourists, a factor of 2.5 was used to estimate their road toll expenditures, since the duration of their tourist stay lasted for several days.

Table 9. Summarized disaggregation of total tourism-related road toll expenditures according to selected visitor categories in Slovene Tolars (SIT)

	Table 1 (foreign)	**Table 2a (domestic; inbound)**	**Table 2b (domestic; outbound)**
Same-day	335,732,822 SIT	658,777,444 SIT	327,528,707 SIT
Transit	6,665,507,034 SIT	N/A	N/A
Tourists	484,020,471 SIT	818,821,768 SIT	540,140,632 SIT
JOINT TOTAL			**9,830,555,334 SIT**

**Due to rounding up the joint total may not be the exact amount of all the disaggregated estimates according to visitor category.*

As can be seen from Table 9, the biggest share (about 68%) of all tourism-related road toll expenditures is estimated to be paid by foreign transit visitors. The second biggest category is the domestic inbound tourist category (8.3%).

5. Limitations of the research and recommendation for implementation

In the absence of concrete statistical data on the level of individual visitor categories the only possible approach was the top-to-bottom approach, meaning also a higher degree of data aggregation and higher implicit error. Here, the widely advocated differential approach has received much attention in these types of tourism-related expenditure estimations. In this context, we have tried to use additional factors of considerations to take into account the relative importance of a particular road toll station (either en route to neighboring Croatia and a specific Slovene tourist destination or not). These factors were however assigned based on in-depth interviews and discretion of the estimators and should be given more consideration in the future, as well as tailored to the particular visitor structure of a given country.

Next, the distinction between pure working days (Tuesday, Wednesday and Thursday), as opposed to other working days (Monday and Friday) was an intuitive one and based on strong student and work commuting on the weekends and extended weekends (particularly relevant with students). Also, several professions work regularly on weekends, meaning there is also a share of working commuters on the weekends as well. Next, the distinction between tourist and non-tourist seasons is based solely on the Traffic calendar of DARS and thus, does not take into account i.e. older (retired) tourists and visitors that plan their trips outside peak tourist seasons.

In terms of assigning the cumulative tourism-related road toll expenditures to specific visitor categories, the assumption of tourists spending 2.5 times more on road toll, as opposed to same-day visitors and transit visitors is again in many ways a discretionary decision and should be re-examined. Furthermore, one might also argue that domestic tourists and domestic same-day visitors, having local knowledge tend to use local roads, thus avoiding strong traffic on motorways in peak tourist seasons and paying less road toll. These are just some of the key areas for future improvements of the outlined methodology.

5.1. Recommendations for implementation

With regards to specific recommendations for the implementation of the described estimation approach to tourism-related road toll expenditure estimation, the key starting point are concrete data on the number of crossings at specific road toll sta-

tions. In the absence of these physical data, which may not be available in your country, data on traffic density or traffic counting may also be used as a substitute. However, without a concrete day-by-day and all-year-around information, our estimation approach cannot be applied. We also recommend, in the first stage of the process, a thorough overview of all relevant information sources and the identification of missing data.

In the next step, applying the so called differential approach, a clear distinction should be made between tourist and non-tourist times of year (seasons). Within this context, we strongly advise also making the distinction between working and non-working days, since the issue of commuting must be dealt with. In-line with the differential approach, the difference between tourist and non-tourist days may to large part be attributed to tourism. However, we recommend the use of factors to take into account also shares of business trips and commuting with seasonal components in so called tourist seasons. Here, the applied factors should be based on assessments by experts and in-depth interviews and should have good argumentation. Also, the importance of individual road toll stations should be further considered and appropriate factors applied, since some road toll stations due to their geographical position have higher degrees of tourism-related expenditures, also on working and non-tourist season days.

Once the cumulative tourism-related road toll expenditure has been estimated, particular attention should be given to assigning appropriate shares among individual visitor categories. Here, the structure of inbound and outbound visitors is very important. Particular attention must also be given to outbound domestic tourists and visitors going by car to other tourist destinations, since some of the road toll is also spend in their home country. It must also be noted that the division between various visitor categories should be made based only on data from by-car visitors and not visitors arriving also by other modes of transport. While Slovenia may have the majority of visitors coming by road, other countries may not. In countries where car rental may also be an important aspect of tourist travel, this data should also be taken into account. In Slovenia, this was however not the case.

With regards to the distinction made between same-day visitors and transit visitors let us clearly note that the TSA methodology does not specifically call for this distinction and groups them as one category. This division has been made for Slovenia mainly because transit visitor represent such an important visitor and tourism expenditure category. For your country and making this distinction, please first consider the importance of transit visitors within all tourism-related visitors.

Last but not least, when assessing the available data on tourists and visitors understanding the structure of inbound and outbound visitor categories is essential an may have a direct impact on selecting appropriate factors in the estimation approach. Furthermore, a good understanding of how various visitor categories travel (i.e. travel modes) is also an essential component of our estimation approach. In the end, appropriate feedback should also be made to official statistics offices and col-

lectors of tourism-related statistical data, to ensure in the long run also better quality of data gathering, which would be more tailored to the needs of the TSA methodology. Here, the broader two-way stakeholder approach outlined by the UNWTO is an essential part of the TSA implementation processes, as well as the emphasized learning curve effect.

6. Concluding remarks

As the results of tourism-related road toll estimations within the TSA methodology have shown, tourism represents or represented an important part of the Slovene road toll system, where tourism-related road toll expenditures were estimated at 27% of total road toll expenditures in Slovenia for 2006. In this context, the foreign transit visitors are the single biggest driver of tourism-related road toll expenditures. This is a direct consequence of the structure of visitors and Slovenia's central geographical position, which lies en route to Croatia, one of Europe's top summer tourist destinations.

Following a strong learning curve, the Slovene task team, commissioned by the Ministry for Economy of the Republic of Slovenia focused on developing new estimation approaches in the area of transportation-related tourism expenditures (namely gas and road toll), since they represent a significant expenditure category for Slovene tourism. Using a differential approach, which has been advocated in the estimation of other tourism expenditure areas by other countries, more insight was hopefully reached. While there is undoubtedly plenty room for fine tuning of the estimation approach, the presented approach is (based on a through overview of all possible TSA-related literature) the first transparent attempt of tourism-related road toll expenditure estimation, taking into account the complex visitor structure of a particular country.

For the end, let us just also briefly note that it comes as no surprise that the Slovene government, being very much aware of this structure and results from the 2006 TSA, decided to introduce the vignette system in the second part of 2007. First, it introduced only 2 types of vignettes, namely the 6 months vignette (35 EUR) and the 12 months vignette (55 EUR). Slovenia was severely scolded by the European Commission for introducing such a discriminating vignette system, mainly for the foreign transit visitors, same-day visitors and tourists. As a result of this, it had to reintroduce in 2008 a more diversified portfolio of vignettes, which were less discriminating against foreign visitors and tourists. While most foreign tourists and visitors today still argue that the vignette system is unfair, the price of a 10-day vignette is at par with a return trip from north to the coastline or to neighboring Croatia under the old road toll system. Still, foreign transit visitors remain the biggest payback source for the vast loans taken on by DARS to finance the completion of

the Slovene motorway system, while at the same time causing sever congestion at holiday peaks, not to mention the negative externalities of the environment.

References

Chu, F-L. 2008: Analyzing and forecasting tourist demand with ARAR algorithm. *Tourism Management*, 12, 1185-1196.

DARS. 2007a: *Traffic calendar.* [http://www.dars.si/Dokumenti/Medijsko_sredisce/Publikacije_DARS/Na_poti_in_prometni_koledar_53.aspx]. Accessed on April 14th 2007.

DARS. 2007b: *Annual report 2006*.

[http://www.dars.si/Dokumenti/O_nas/Financno_sredisce/Letna_porocila_DARS_13.aspx]. Accessed on June 16th 2007.

Dickinson, J. E. – Robbins, D. 2008: Representations of tourism transport problems in rural destination. *Tourism Management*. 29, 1110-1121.

Eurostat. 2008: TSA: Recommended Methodological Framework (TSA: RMF 2008). Eurostat: Luxembourg.

Frechtling, D. 1999: The Tourism Satellite Account: Foundations, Progress, and Issues. *Tourism Management*. 20, 163-170.

OECD. 1999: *Draft OECD Guidelines for a Tourism Satellite Account*. OECD Publications: Paris, France.

Rašković, M. – Zagoršek, H. 2008: *Evaluating fuel expenditures by same-day and transit visitors.* Presented paper at the conference: An Enterprise Odyssey: tourism, governance and entrepreneurship. Zagreb: Faculty of Economics and Business.

Rivera, G. 1999: Mexico's Experience of Setting up its Tourism Satellite Account. *Tourism Economics*. 5, 315-330.

Sacks, A. 2004: *The Alaska Tourism Satellite Account: A comprehensive analysis of the economic contribution of travel & tourism*. Global Insight: Alaska, USA.

Sharma, A. – Olsen, M. D. 2005: Tourism Satellite Accounts: Implementation for Tanzania. *Annals of Tourism Research*. 32, 2, 367-385.

Stynes, D. J. 1999: Approaches to Estimating the Economic Impact of Tourism: Some examples. *Economic impact approaches*. 1, 1-18. Michigan State University reports.

Širše, J. – Vidjen, T. – Kalin, J. – Černič, I. 2004: *Implementation of Tourist Satellite Accounts in Slovenia: Estimations of economic impact of tourism in Slovenia with Tourism Satellite Accounts methodology (TSA)*. Ministry for Economy of the Republic of Slovenia: Ljubljana.

SURS. 2007: *Statistical yearbook 2006*. SURS: Ljubljana.

SURS. 2009: [http://www.stat.si/publikacije/pub_statinf1.asp?podrocje=21]. Accessed on December 12th 2009.

UNWTO. 2000: *Report of the secretary general*. [http://www.world-tourism.org/statistics/tsa_project/Basic%20concepts%20of%20the%TSA.pdf]. Accessed on December 13th 2009.

UNWTO. October 2009: *World Tourism Barometer*. UNWTO: Washington, D.C.

Zagoršek, H. – Jaklič, M. – Bregar, L. – Hribernik, A. – Rašković, M. 2008: *Economic importance of tourism in Slovenia in 2003 and extrapolation of trends for 2006: using the Tourism Satellite Account methodology*. Ministry for Economy of the Republic of Slovenia: Ljubljana.

Analysis of the Presence of Traditional Agricultural and Food Industry Products and Regional Dishes in the Choice of Dishes Offered in the Southern Great Plain Touristic Region

Dénes Sándor[1]

The background of the study is based on previous studies dealing with the areas of Transdanubia and Central Hungary. Those works have proved the hypothesis of author that local ingredients and dishes represent minimum role on menus. The question is how similar or different the choice of menu of Southern Great Plain is from the previously examined areas of Hungary. Additional questions may also arise in terms of the usage of local ingredients available near catering units.

This piece of work has studied several menus of Southern Great Plain, Hungary. The material was collected from different types of restaurants, 'csárda'-s and inns. The methodology was two-tier: data collecting on the one hand, and the detailed examination of differences between the catering units on the other hand: the number of dishes, the right order of dishes, the appearance of local ingredients and dishes, the consideration of seasonality, the usage of different cooking methods and group of dishes.

The study reflects the main problems of countryside restaurants in Hungary. Only a minimum number of dishes could be connected to local dishes out of more thousands. However, there would be a chance to make a comparative analysis between the areas of Transdanubia and Central Hungary.

All in all, dishes are not as precise and perfect as the 'old ones' used to be; that is why it is important for every expert to keep traditions alive in daily routines and by way of strategic means too.

Keywords: Southern Great Plain touristic region, catering, local ingredients, traditions

1. Introduction

The relevance of the topic is given by the global advance in tourism highlighting the choice of dishes offered by catering units. The length of stay for domestic and foreign guests may range from one day to multiple weeks, and consequently the

[1] Dénes Sándor, Principal lecturer, Budapest Business School, Faculty of Commerce, Catering and Tourism (Budapest)

number of occasions they have meals also varies. At the same time, a guest staying one night takes a single hotel service but usually has meals three times. Restaurants, small restaurants (vendéglő in Hungarian) and roadside restaurants (csárda in Hungarian) should preferably offer *regional dishes* – prepared at least partly from *local raw material* – in their menu.

The analysis of the studied topic is a current issue since, in accordance with a present hypothesis, one of the problems is that procuring raw materials is much easier and cheaper in hypermarkets. This is the reason why catering companies partly do their shopping in these commercial units and do not spend time on finding local products. Another problem is that local producers cannot always guarantee a constant level of quality and quantity of requested products. *Lack of regional dishes* appearing *in the menu* of catering units poses an additional problem.

Bases of the literature background 'Appearance of Regional Dishes in the Food Choice of the Southern Transdanubian Touristic Region' (Sándor D. 2008) and 'Analysis of Transdanubian Supply of Dishes – local raw materials, dishes' (Sándor 2009) provide proper ground for the present study focusing on the similar characteristics of the Southern Great Plain Touristic Region. Earlier works covering the Southern Transdanubian region, the entire Transdanubia and the Budapest-Middle-Danube region verified the hypothesis of the Author that *local raw materials* and *regional dishes* play an *insignificant role* in the restaurant *menus*.

A professional paper featuring an important collection of catering traditions says: 'With the process of river regulation, draining marshlands and also with the development of agrarian culture, growing, collecting and using certain foods simply 'went out of fashion'; these plants, however, used to be important raw materials of cooking and eating, including: buckwheat, turnip, millet, mangel, trapa natans, pumpkin seed oil, pickled carrot, spined loach fish, old herbs etc.' (Szigeti 1999). Additional traditional and regional agricultural and food industry products are presented in the professional study [Farnadi 2002].

Present study tries to given an answer to the question how the choice of dishes in the hot cuisine catering units of Southern Great Plain region differs from that in the previously studied ones. It is also a question whether catering people of the region take advantage of procuring and using locally grown fruits and vegetables. *Hungary has gastronomic traditions* we should take care of by including authentic local foods and dishes of the region in catering units' menus.

2. Goals

The study aims to present the menu supply of the Southern Great Plain Touristic Region in the years 2005-2008 by a field research of a non-representative sample. The main goal is to survey the choice of dishes and then to evaluate the supply of diverse hot cuisine units – restaurants, small restaurants, and roadside restaurants –

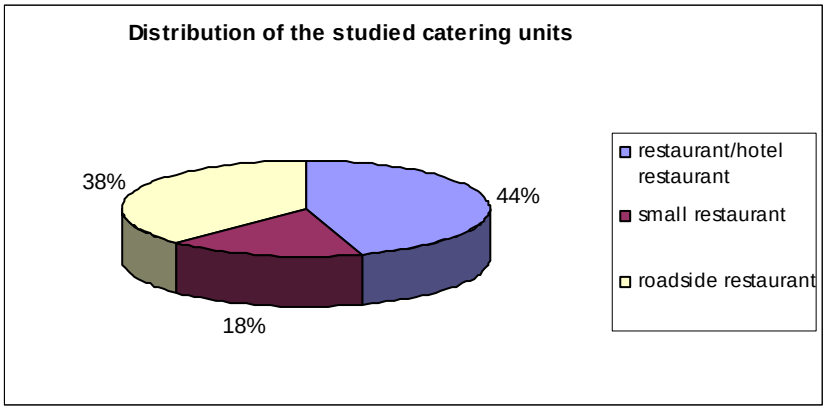
by professional aspects. Accordingly we may have a view of the quantitative and qualitative parameters of the dishes listed on the menus.

Among others, the *analysis covers* the logical dish grouping and listing system, the number of dishes, the diverse use of raw materials and preparation methods, and the supply of *local specialities*, traditional foods, creative menu compilations and explanation of ambiguous denominations used in menus.

3. Research Methods

The *analysis* of the choice of dishes is *based on the menus of 18 restaurants, 7 small restaurants and 15 road restaurants* used in the years 2005-2008. Catering units were randomly selected by a non-representative sample of the Southern Great Plain Touristic Region. Figure 1 shows the distribution of studied catering units.

Figure 1. Distribution of studied catering units



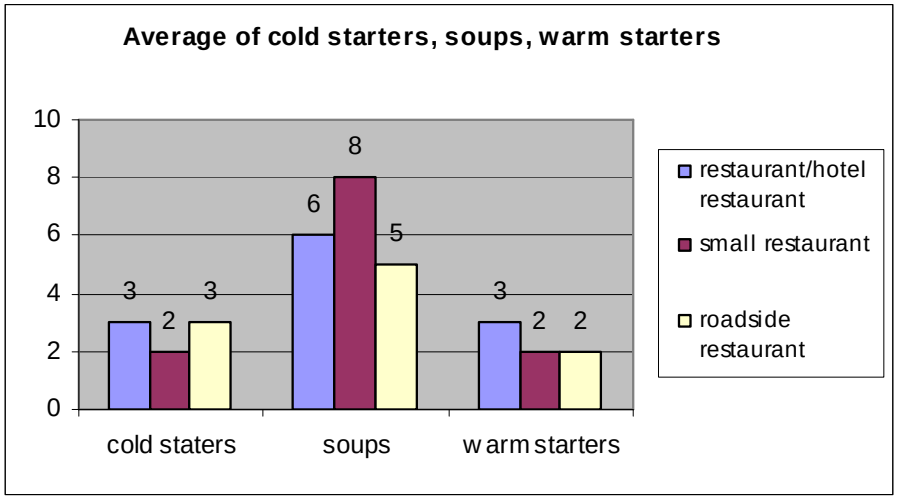
Source: own research

After electronically recording the names and quantity of dishes listed on the menus, meals of certain dish groups were aggregated. Next the average number of meals in each dish group was calculated with handling different types of catering units (restaurants, small restaurants, roadside restaurants) separately. To provide a more efficient visualization, several diagrams are included to present different distributions and average values. Dish groupings and dishes of the menus can be examined by their total and average numbers.

The total figures of soups, cold and hot starters (18 restaurants, 7 small restaurants and 15 roadside restaurants) show that the average number of cold

starters by types of catering units is as follows: 3 in restaurants and roadside restaurants and 2 in small restaurants (see Figure 2). *The average number of soups is 6 in restaurants; 8 in small restaurants; and 5 in roadside restaurants*. The range of soups, i.e. the average number of soups on menus shows that preparing and offering 5-8 different soups might be too much. In case of hot starters, their average number is 3 in restaurants and 2 both in small restaurants and roadside restaurants. The number of cold and hot starters is sufficient in all the three types of studied units; the choice of only 2 different cold and hot starters may be too narrow for small restaurants and roadside restaurants.

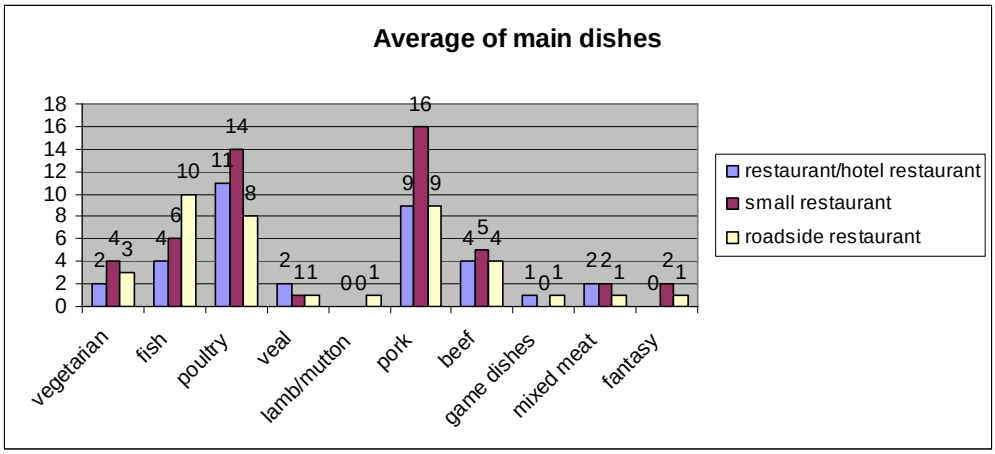
Figure 2. Average number of soups, cold and hot starters in the different hot cuisine units



Source: own research

Figure 3 shows the average number of main courses in restaurants, small restaurants and roadside restaurants in the Southern Great Plain. Accordingly *pork and poultry dishes are the most popular with an average supply of 11* different dishes. Note that a typical small restaurant enlists 16 different dishes of pork on its menu, while the same figure for restaurants and roadside restaurants is 9. Pork dishes are followed by dishes made of fish, beef, veal and game with an average range of 7, 4, 1, 1 dish, respectively.

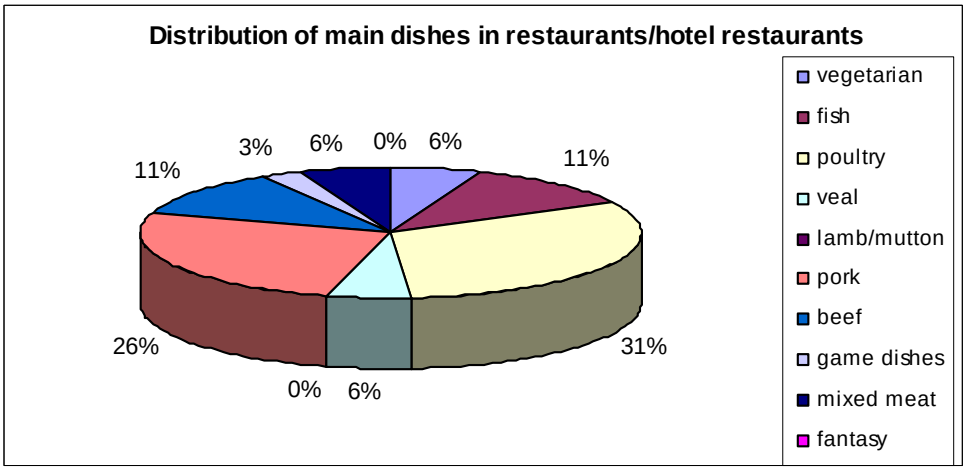
Figure 3. Average number of main dishes in restaurants, small restaurants and roadside restaurants



Source: own research

Figure 4 present the distribution of dishes offered as main courses in restaurants. Dishes prepared of poultry amount to 31%, while fish, pork and beef dishes give altogether 48% of the main courses. Dishes made of the above four types of meat provides accordingly 79% of the total main course supply. These meals are dominated by poultry and only 21% of main courses include vegetarian, veal, lamb/mutton, game dishes etc.

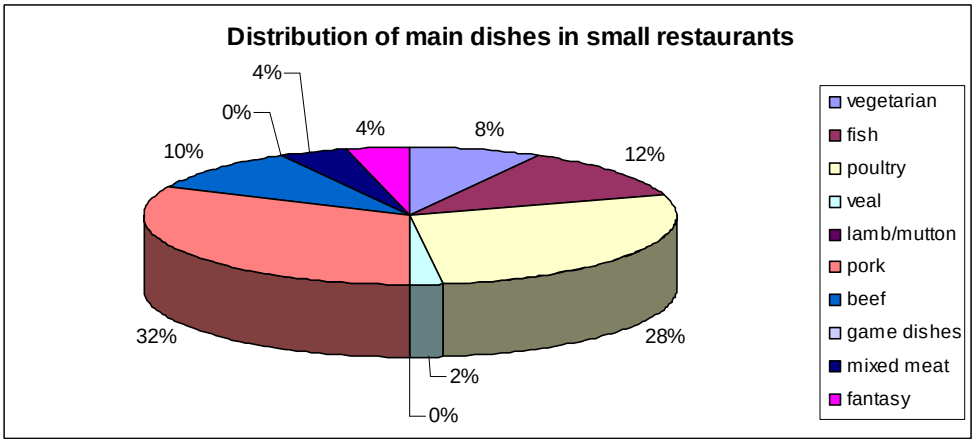
Figure 4. Distribution of main dishes in restaurants



Source: own research

Figure 5 shows the distribution of main dishes in small restaurants. Pork dishes give 32% of the total range, poultry dishes amount to 28%, dishes made of fish contribute with 12% to total main dishes, beef dishes are responsible for 10%, while vegetarian dishes account for 8% of total supply which latter can be deemed a relatively high share. However, only 10% of space remains for other categories of main courses, which allows for a quite limited range of other meals.

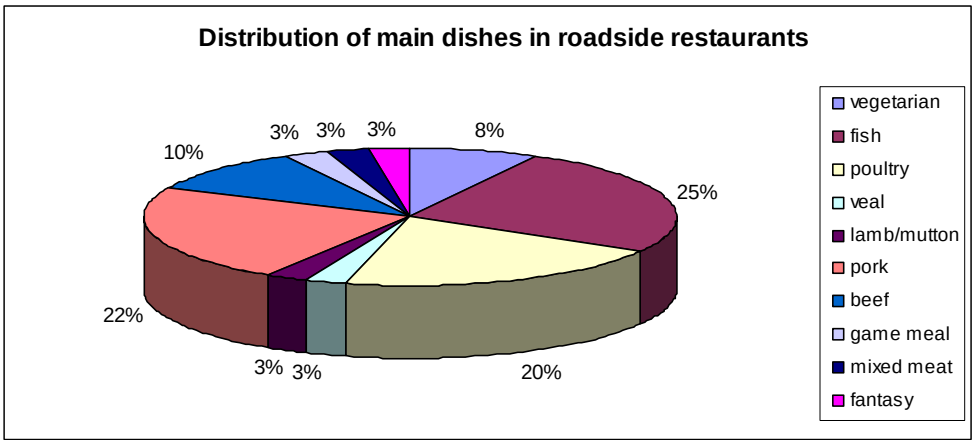
Figure 5. Distribution of main dishes in small restaurants



Source: own research

Figure 6 shows the distribution of main course meals in roadside restaurants. Fish, pork and poultry dishes amount to two thirds of total main course supply. The high number of fish dishes, amounting to a quarter of all main courses, comes from roadside fish restaurants. The high number of pork and poultry dishes appropriately reflects the balance of demand and supply. However, only a narrow choice remains for other dish categories.

Figure 6. Distribution of main dishes in roadside restaurants

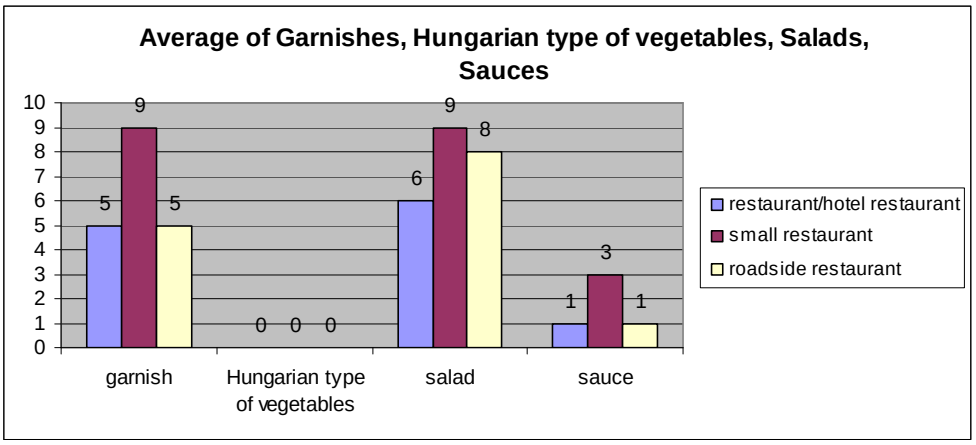


Source: own research

The total amounts of garnishes, Hungarian type of vegetables, salads and sauces (18 restaurants, 7 small restaurants and 15 roadside restaurants) indicate that the average number of different garnishes is 5 both in restaurants and roadside restaurants and 9 in small restaurants (see Figure 7). *The average range of garnish is 7*, which can be deemed quite wide, but garnish offered as a separate group of meals often poses problems. Of course it is possible to include the choice of garnish in the menu, but when offering a complete dish creation as meat-garnish-sauce, a separate category for garnish makes no sense. The average number of Hungarian type of vegetables is 0 in restaurants, small restaurants and roadside restaurants. The choice of this type of dishes is extremely poor. The lack of demand might also affect the scarcity of supply. At the same time, it would be vital the selection to include Hungarian type of vegetables prepared of high quality seasonal raw materials in a modern preparation way.

Average numbers in case of salads: 6 in restaurants, 9 in small restaurants, and 8 in roadside restaurants. In accordance with *Hungarian traditions, offering 6-9 different salads*, which includes also pickles, is typical. In this case, pickles (including pickled cucumber, pickled paprika, etc.) can be listed in one group; still the selection might be unique. Should this be the case, of course these pickles must be offered at the same price. By the similar principle, salads (tomato, cucumber etc.) could be listed in a different line. Average number of sauces is 1 in case of both restaurants and roadside restaurants, while 3 in small restaurants. Here comes the question: why it is necessary to include a separate group of dishes of sauces.

Figure 7. Average number of garnish, Hungarian type of vegetables, salads and sauces in different hot cuisine catering units

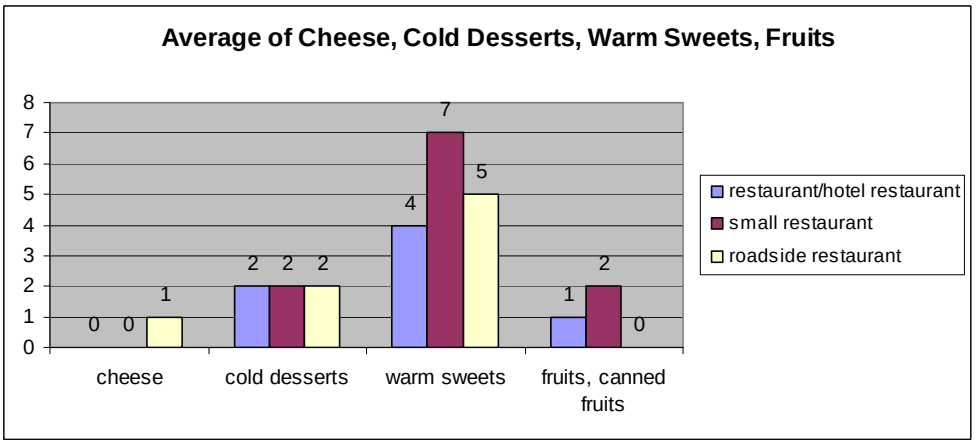


Source: own research

The total values of cheese, cold desserts, warm sweets and fruits (based on the data of 18 restaurants, 7 small restaurants and 15 roadside restaurants) indicate that the average number of cheeses is 0 both in restaurants and small restaurants and 1 in roadside restaurants (see figure 8). The number of cheeses appearing in menus is quite low, but catering units occasionally offer a selection of cheeses including more types of cheese. *The average number of cold desserts is 2* in all the three types of catering facilities. This figure is rather low. *The average number of warm sweets is 4 in restaurants, 7 in small restaurants and 5 in roadside restaurants*. The width of range of warm sweets is 7 in restaurants, a rather high figure.

The average range of fruits and canned fruits is 1 in restaurants, 2 in small restaurants and 0 in roadside restaurants. These figures are quite meaningful. Concerning the already studied regions, offering 2 kinds of fruits in a restaurant has represented the widest choice so far. At least 1 kind of fruit should be offered in roadside restaurants as well.

Figure 8. Average number of cheese, cold desserts, warm sweets, fruits and canned fruits in the different hot cuisine units



Source: own research

4. Results

After processing the data from 40 menus, the primary goal of surveying the choice of dishes in the region can be summarized as follows: the quality of menus in terms of their content reflects a quite diverse picture. *There are restaurants where quests are awaited with nearly 300 different dishes*. In-depth analyses suggest that producing such a wide range of meals at a consistently high level is a great challenge for any kind of hot cuisine unit regardless of venue and time.

Among the studied units, however, Kisködmön Restaurant in Gyula is one of those operating with the highest quality offering altogether 20 dishes excluding desserts, but its menu is a unique, professional curiosity. The supply includes historical dishes and preparation methods involving high quality raw material. According to the findings of the six studied regions, the smallest number of dishes is offered in the Restaurant of Szeged Novotel Tisza. However, the 15 dishes included in their menu can provide a suitable selection for almost all different guest demands including also vegetarians. The choice of dishes could have been expanded with one or two more hot starters and some warm sweets. Managing raw materials is simpler this way and this makes consistently preparing dishes of high quality easier.

In some cases, dish groups are listed in menus in an order fashionable some decades ago. As a consequence of this, guests may encounter desserts earlier than main courses or salads.

It is also rather old fashioned to use the categories of readymade dishes and freshly made dishes. Foreign speaking guest may think in this case that the translated denomination of readymade dishes might refer to "meals not cooked at that time". At the same time, several hot cuisine units apply logical listing order of the meals – based on the timeline of consuming.

Fortunately, nearly a hundred dishes prepared of Hungarian raw materials including goose liver, goose cracklings, ewe cheese, ewe curd, salami, sausage, bacon etc. can be found. Regional dishes appear on the menus only at an insignificant frequency. *Out of the 3000 dishes surveyed only 40 dishes can be somehow connected to a given region.* Often only the denomination of the dish refers to its regional character e.g.: kecskeméti kakas rétes-béles, Bean Soup Szénás Style, Betyár Soup of Gyula, Hungarian Great Plain Soup with Dumplings, variations of chicken breast: Szatymaz Style, Maros Style, Félegyháza Style, Great Plain Style, Balástya Style etc.

In accordance with Hungarian traditions, dishes starting the meal (cold/warm starters and soups) are *dominated* first of all *by Soups*. With regards to main courses in restaurants, the number of poultry dishes is the highest, while in small restaurants and roadside restaurants the same stands true for pork dishes. Basic supply of vegetarian dish is still one-sided and quite limited (including only breaded cheese and mushroom).

The range of garnish covers 7 different types on the average, however, the question may arise why catering units don't offer directed main course-garnish pairs? As of international practise, main course is a complete dish including meat (or without meat), garnish and sauce. Fortunately, some of the surveyed menus included no separate garnish category. Strange and surprising, the number of *Hungarian type of vegetables is less than minimal, that is zero* in all the three types of units (restaurants, small restaurants, roadside restaurants). On the other hand, the average number of salads offered to main courses is 6 in restaurants, 9 in small restaurants and 8 in roadside restaurants. The range of sauces is not too extensive as well, with an average of 2: 1 in restaurants; 3 in small restaurants and 1 in roadside restaurants. In addition, as it has already been mentioned in the analysis of the South Transdanubian Region, the composition of the most often appearing Tartare sauce is not exactly identical with the sauce offered in international cuisines.

The average choice of cheese is 0 in restaurants and small restaurants, and 1 type of cheese is usually offered in roadside restaurants. *The standard favourite of cold desserts are ice cream cups, sponge cake Somlo style and chestnut puree*. Warm sweets are dominated by pancakes (60%) largely contributing to the base of this category. In roadside restaurants, cottage cheese pasta appears in the menu as the most frequent pasta. The choice of fruits, including also canned fruits, is rather narrow, the average figure being 1 in restaurants, 2 in small restaurants and 0 in roadside restaurants. Why are tasty Hungarian fruits missing from the menu?

An improvement can be detected in the fact that more hundreds of different dish versions are offered in restaurants, small restaurants and roadside restaurants. There are some places where interesting fantasy names are given to these dishes, e.g.: Favourite of Rózsa Sándor, Favourite of the Roadside Restaurant's Host, Jackals' Favourite.

In many cases there are formal mistakes in the menus, e.g.: the name of the place, the manager name or the name of the chef are missing, the word classification is used instead of categorization which was the terminology on force in the surveyed period, etc. Also spelling mistakes and misuse of terminology often appear in the menus, e.g.: "baconszalonna", caviar, tuna fish, Gordon blue, beafsteak, etc. The proper use of these terms is as follows: bacon (or "húsos szalonna"). Caviar is the roe of a certain species of fish, namely sturgeon; otherwise it should be named salted roe or salmon caviar in case of salmon. Frozen fillet of sea fish is often referred to as tuna fish but these are generally hake fillets. The proper name of the dish is Cordon Bleu, while it should be spelt as beefsteak in Hungarian menus.

5. Conclusion, Suggestions

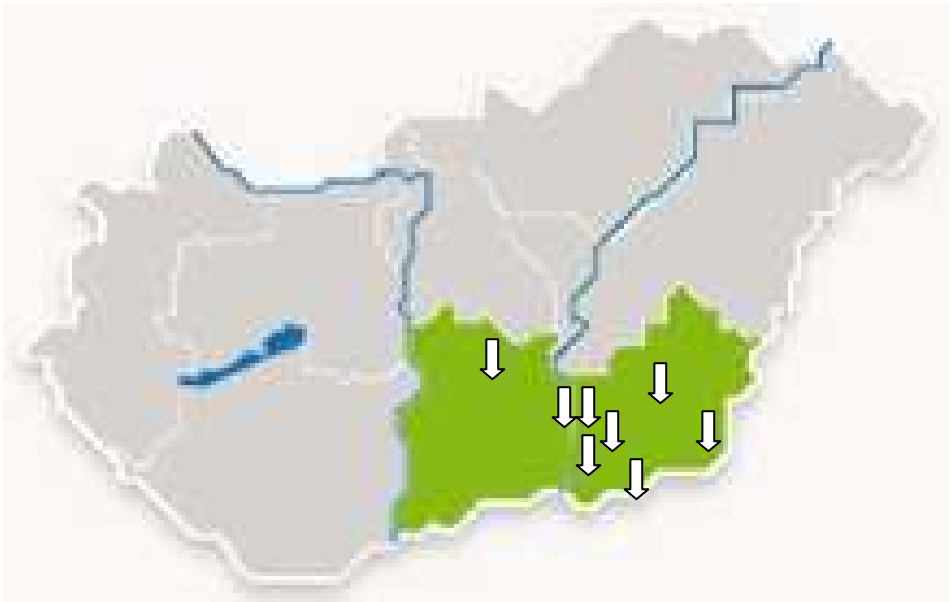
The following conclusions may be made from the analysis of choice of dishes offered in the menus of 18 restaurants, 7 small restaurants and 15 roadside restaurants in the years 2005-2008:

- A too extensive range of dishes listed in the menu may have an adverse effect on preparing dishes in consistently high quality
- It would be practical *to keep selection within reasonable limits at the same time providing outstanding quality dishes* for the guests
- Applying international menu formatting experience would be useful
- The terms "Readymade dishes" and "Freshly made dishes" should not be used any more, instead dishes should be included among main courses based on their main raw material
- Range of vegetarian dishes should be extended and made more diverse, and tasty vegetables dishes should be prepared with less or no 'roux' (thickening agent of flour and oil)
- The category of garnishes and sauces is not necessary when garnish and sauce are added to all main courses
- *Selection of cheese and fruits should be reasonably widened*
- Range of cold desserts should be made more flexible, there should be other desserts in addition to the magic trio of ice cream cup, chestnut puree and sponge cake Somlo Style

- Pancake dominancy of restaurant warm sweets should also be weakened by expanding the choice of this category
- Fantasy dishes cause no problem when an explanation of the dish is added
- *Range of regional dishes should be widened*, since domestic and foreign tourists are mostly interested in local foods, accordingly it is a shame that out of 3000 dishes only 40 can be attached to the region in any way
- Adding the name of a geographic location to that of the dish does not make the food an authentic local speciality in itself
- It is recommended to include traditional foods in the menus by applying professional literature and the related knowledge of local people
- Using local traditional and regional raw materials is an outstanding possibility for expanding the range of regional dishes. Some of the local traditional and regional raw materials can be seen in Figure 9.

Traditional and regional raw materials, products of the South Great Plain Region: goose liver of Orosháza, sausage of Csaba, sausage of Gyula, grey catfish of River Tisza, apricot of Kecskemét, peach of Szatymaz, garlic and onion of Makó, tomato sweet pepper of Szentes

Figure 9. Traditional and regional raw materials and products in the Southern Great Plain Region



Source: own creation based on http://www.nvteredmeny.hu/del-alfoldi-regio/index.html and Farnadi 2002

- A further possibility is to use local wines when preparing the dishes, however, it is not enough to indicate only 'wine' in the menu, it should be detailed e.g.: Venison Stew with paprika flavoured with 2007 Cabernet Franc Barrique of Birkás Cellar
- Areas to improve the most include the usage of an extremely limited range of raw materials and preparation methods, menu compositions often lacking creativity
- Therefore it would be practical to dismiss the most regular choice of dishes available in almost each and every restaurant and offer regional dishes prepared of local raw materials in the menus
- More attention should be paid to meeting the formal and spelling requirements of the menus

6. Summary

All in all, current choice of dishes should be updated and a wider range of regional dishes should be offered; the above goals require the cooperation of all professionals in their everyday practise to observe our traditions in the future as well.

References

Alsósoproni, N. F.-né 1926: *Dunántúli szakácskönyv*. Pannónia Sajtóvállalat részvénytársaság, Szombathely, 313. p.

Farnadi, É. 2002: (felelős szerk.) Cserhalmi, ZS. – Csoma, ZS. – Némethyné, U. H. – Sipos, B. – Szabó, E. (szerk.) *Hagyományok – Ízek – Régiók program II*. FVM AMC Kht. Budapest. 365 p.

Sándor, D. 2002: *Ételismeret.* BGF. KVIFK., Budapest, 323. p.

Sándor, D. 2008: *Tájjellegű ételek megjelenésének elemzése a dél-dunántúli turisztikai régió ételválasztékában.* Tudományos évkönyv BGF., Budapest, 194-208. p.

Sándor, D. 2009: Dunántúli ételkínálat – helyi alapanyagok, ételek – elemzése. *Economica,* Tudományos közlemények, Szolnok, 50-57. p.

Sándor, D. 2009: *Tájjellegű ételek megjelenésének elemzése a dél-dunántúli turisztikai régió ételválasztékában*. Turizmus Bulletin, 14-18. p.

Sándor, D. 2009: Helyi alapanyagok és tájjellegű ételek (2), A közép-dunántúli turisztikai régió ételválasztékának elemzése. *Kulinárium Gasztronómiai és Szállodaipari Szaklap,* 28-29. p.

Szigeti, A. 2003: *Népi konyha Dunántúli tájak ételei*. Mezőgazda Kiadó, Budapest, 384 p.43/1998. (VI. 24.) IKIM rendelet a vendéglátó üzletek kategóriába sorolásáról, valamint ártájékoztatásáról

http://www.nvteredmeny.hu/del-alfoldi-regio/index.htmldate of access: October 30th, 2009.

Development of the statistical process control methods

Boris Radovanov[1] *– Aleksandra Marcikić*[2]

Statistical process control represents a statistical procedure using control charts to see if any part of production process is not functioning properly and could cause poor quality. Process control is achieved by taking periodic sample from the production process and plotting these sample points on a chart, to see if the process is within statistical control limits. Statistical process control prevents quality problems by correcting the process before it starts producing defects. This paper encompasses an application of statistical process control methods with numerous modifications in order to make possible appropriate process quality improvement of the soft drink production line via detecting out-of-control process or unusual patterns in a sample. Used methods corresponding with the total quality management require a never-ending process of continuous improvement that covers people, equipment, suppliers, materials and procedure. The end goal is perfection, which is never achieved but always sought.

Keywords: Statistical process control, total quality management, control charts

1. Introduction

Quality can be described as the most important factor in the long-term profitability and success of the firms. Therefore, it can not be underestimated or overlooked by any firm, regardless of its size or assets.

Globalization and foreign competition changed the business environment and created higher expectations for products and services. Quality not only allows for product discrimination, it also has become a marketing weapon focusing on the consumer as the most important part of the production line (W. Edwards Deming) (Russel-Taylor 1998). This results in a commonly used definition of quality as a service's or product's fitness for intended use. At the other side, we assume another perspective known as producer's quality perspective which means a quality during

[1] Radovanov Boris, Msc, teaching assistant, University of Novi Sad, Faculty of Economics Subotica, Serbia.
[2] Marcikić Aleksandra, teaching assistant, University of Novi Sad, Faculty of Economics Subotica, Serbia.

production focuses on making sure that the product meets the specification required by the design. Perhaps, the most important role in achieving and improving required quality level play employees throughout the direct involvement in the management process referred to as participative problem solving. A quality circle is one of the most commonly used employee-involvement program and presents a small voluntary group of employees and their supervisors. Using team approach, it represents the brainstorming to generate ideas as a group technique for identifying and solving problems.

2. Total quality management

To make sure that products and services have the quality they have been designed for, a commitment to quality throughout the organization is required. Total quality management represents a set of management principles that focus on quality improvement as the driving force in all functional areas and at all levels in a company. However, the commitment to quality must begin at the top and spread down through the organization. Although it is popular to say that quality is everyone's responsibility in a company, total quality management generally requires a total commitment from the top management to monitor and maintain quality throughout the organization.

3. Statistical process control

A major topic in statistical quality control is statistical process control. It represents a statistical procedure using control charts to see if any part of production process is not functioning properly and could cause poor quality.

Walter Shewhart developed the technical tools that formed the beginning of statistical quality control. He and his colleagues introduced the term quality assurance as a commitment to quality throughout the organization using statistical quality control methods first time applied at Bell Telephone Laboratories. Later on W. Edwards Deming changed the focus of quality assurance from the technical aspects to more of a managerial philosophy. Today US companies that have been successful in adopting Total quality management concept train all their employees in statistical process control methods and make extensive use of statistical process control for continuous process improvement. Through the use of statistical process control, employees are made responsible for quality in their area, to identify problem and either correct them or seek help in correcting them.

Process control is achieved by taking periodic sample from the production process and plotting these sample points on a chart, to see if the process is within

statistical control limits. Statistical process control prevents quality problems by correcting the process before it starts producing defects.

All processes contain a certain amount of variability that makes some variation between units inventible. There are two reasons why a production process might vary. The first is the inherent random variability of the process, which depends on the equipment, engineering, the operator and the system used for measurement. This kind of variability is a result of natural occurrences. The other reason for variability is unique causes that are identifiable and can be corrected. These causes tend to be nonrandom and, if left unattended, will cause poor quality.

4. Production quality improvement

This chapter encompasses an application of statistical process control methods in order to make possible appropriate process quality improvement via detecting out-of-control process or unusual patterns in a sample.

Total quality management requires a never-ending process of continuous improvement that covers people, equipment, suppliers, materials and procedure. The end goal is perfection, which is never achieved but always sought.

4.1. Quality measure and sample size determination

First concern of quality management after they noticed a potential quality problem is choosing a right quality measure. The quality of a product or a service can be evaluated using either an attribute or a variable measure. An attribute is a product or a service characteristic that can be evaluated with a discrete response. Attributes can be evaluated quickly and represents a qualitative classification. Even if quality specifications are complex and extensive, a simple attribute test might be used to determine if a product is or is not defective. A variable measure is a product or a service characteristic that is measured on a continuous scale. Because it is a measurement, a variable classification typically provides more information about the product.

Difference between attribute and variable measure requires different sample size determination. In general, larger sample sizes are needed for attribute charts because more observations are required to develop usable quality measure. Variable control charts require smaller sample sizes because each sample observation provides usable information. After only a few sample observations, it is possible to compute a range or a sample average that reflect the sample characteristics.

It may also be important that the samples come from a homogeneous source so that if the process is out of control, the cause can be accurately determined. If production takes place on either one of two machines, mixing the sample observations between them makes it difficult to ascertain which operator or machine

caused the problem. If the production process encompasses more then one shift, mixing the sample observation between shifts may make it difficult to discover caused a quality problem in the process.

Example presented in this paper contains weight as a variable measure in monitoring quality of soft drink production process. This variable contains not only control limits from product design specifications, but also natural variations that can not be exceeded by the actual products. Natural variations known as design tolerances are design or engineering specifications reflecting customer requirements for a product. In Serbia the Regulation in quality and other requirements for soft drink products is standardized by the Serbian customer society. The mentioned Regulation provides natural tolerances for different weights of product; for example product in 1 liter package has 2% deviation and a product in 2 liter package has 1,5% deviation allowed. Considering all above mentioned facts we choose one production line and one work shift in this analysis to create numerous samples where each contain 5 actual products. Every production process has a different probability that shift in the process will be detected known as $\alpha = (1 - \beta)$, or probability of not detecting process quality shift β. β depends on the sample size, population variance σ and difference between actual mean μ_t and hypothesized mean μ_0. This coefficient presents the input value in calculation of ARL (average run length) as the expected number of samples to be taken before control chart indicates a shift in a process level.

Probability of not detecting process quality shift can be calculated using normal distribution probability formulae as follows:

$$\beta = P\left\{z < (\mu_0 - \mu_t)/(\sigma/\sqrt{n}) + z_\alpha\right\}$$

To calculate the number of samples necessary to detect some shifts it can be used the formulation below:

$$ARL = \frac{1}{1-\beta}$$

The presented production line has input values $\mu_0 = 1$, $\mu_t = 0{,}998$, $\sigma = 0{,}01$, $n = 5$. So this part of process has its own probability of not detecting shift and average run length according to the stated input values.

$\beta = P(z < 2{,}0922) = 0{,}9818$

ARL = 54,95

There is a very high percentage of not detecting shift level probability in samples. Hence average run length needs to be high and exceed the number of 54,95 samples. According to the level of ARL, quality examination of production line will use 60 equal samples taken in the same period interval from the same production line.

4.2. Control charts

Walter Shewhart developed a simple but powerful tool to separate two or more items known as the control chart. Control charts represent a basic statistical process control tool to measure performance of a process. Control charts are used to investigate the variability of the process and this is essential when assessing the capability of a process. Data are often plotted on a control chart in the hope that this may help to find causes of problem. These charts visually show if a sample is within statistical control limits. Control limits are the upper and lower bands of a control chart.

There are four basic types of control charts, two for attribute measure (p-chart and c-chart) and two for variable measure (mean-chart and R-chart). This chapter, as it is noticed before, contains the charts for variable measure. Beside the two basic types, this paper contains development of different modification in order to spot the smaller shifts in a process. However, looking at the figures alone will not give the reader any clear picture of the safety performance of the business.

4.2.1. R-chart

In this chart, the range is the difference between the smallest and largest value in a sample. This range reflects the process variability instead of the tendency toward a mean value. Control limits in this type of control chart are presented via formulas as fallows:

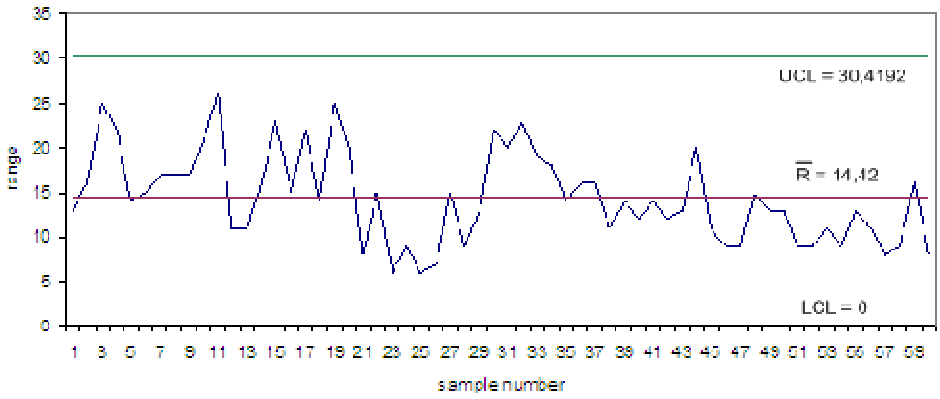
$$UCL = D_4 \cdot \overline{R} \quad LCL = D_3 \cdot \overline{R}$$

$\overline{R}$ is the average range, and also a central line of the chart, for the samples:

$$\overline{R} = \frac{\sum_{i=1}^{k} R_i}{k}$$

where k is the number of samples and D_3 and D_4 are factors for determining control limits that have been developed based on range values rather then standard deviations.

Figure 1. Range control chart



Source: own creation

Figure 1 indicates that the production process is in the control. There are no significant differences between ranges in each sample. Therefore, the variability observed is a result of natural random occurrences. Although individual values are all different, as a group they form a pattern that can be described as a normal distribution.

4.2.2. Mean chart

The mean of each sample is computed and plotted on the chart, so the points are the sample means. The central line of the chart is the overall process average, the sum of the averages of k samples:

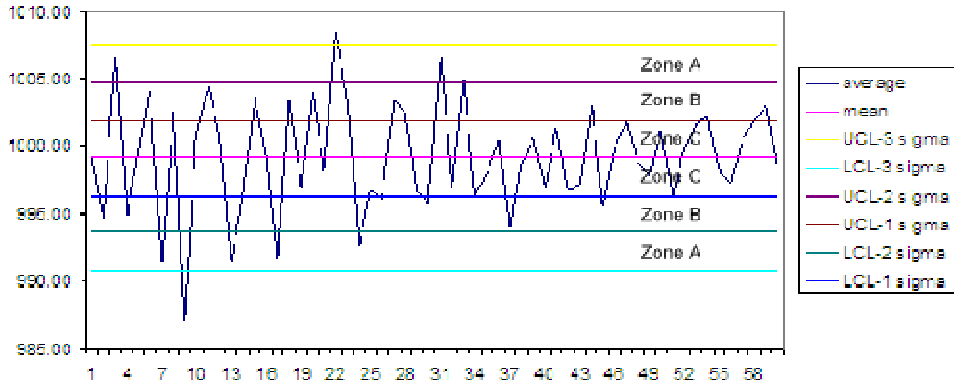
$$\overline{\overline{x}} = \frac{\overline{x}_1 + \overline{x}_2 + \cdots + \overline{x}_k}{k}$$

When the mean-chart is used in conjunction with an R-chart, the allowing formulas for control limits are used:

$$UCL = \overline{\overline{x}} + A_2 \cdot \overline{R} \quad LCL = \overline{\overline{x}} - A_2 \cdot \overline{R}$$

Where A_2 is a tabular value used to establish the control limits. Values of A_2 were developed specifically for determining the control limits for mean-chart and are comparable to 3-standard deviation limits.

Figure 2. Mean chart and zones for pattern test



Source: own creation

Obviously, there are two sample averages that are outside of specified control limits. Here we have a situation that the process averages are not in control like variability presented in Figure 1. This example illustrates the need to employ the R-chart and mean-chart together. In the R-chart none of the ranges for the samples were close to the control limits. However, ranges for some samples were relatively narrow, whereas means of those samples were relatively high. Hence the use of both charts together provided a more complete picture of the overall process variability.

If it is important to take a closer look on a control chart and to decide if pattern is nonrandom or random the pattern test is in wide use. One type of pattern test divides the control chart into three zones on each side of a center line, where each zone is one standard deviation wide. Figure 2. shows the numerous points (samples) in zones A and B (specially in the first part of samples) imply that nonrandom pattern exists and the cause should be investigated.

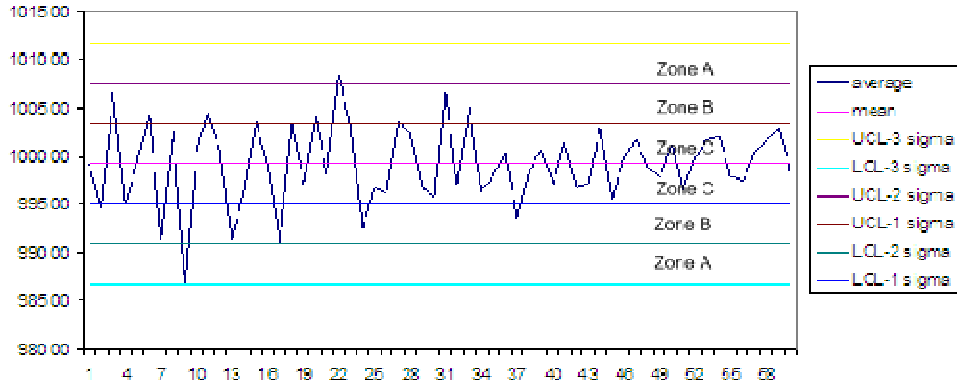
This kind of process pattern needs to be investigated using just the mean-chart with a standard deviation and the following formulas for establishing control limits are:

$$UCL = \overline{\overline{x}} + z \cdot \sigma_{\overline{x}} \ , \ LCL = \overline{\overline{x}} - z \cdot \sigma_{\overline{x}}$$

The sample standard deviation is computed as:

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

Figure 3. Mean chart using standard deviation



Source: own creation

Figure presented above shows corrected control limits according to rule of 3σ. After the correction there are no samples beyond control limits, but there still exists a nonrandom pattern of used samples because of numerous sample averages in zones A and B. The points plotted in zone C, also known as green zone, show stability in a process. The sample points plotted in zone B (yellow zone) suggests there may have been an assignable change and that another sample must be taken in order to check. Any points plotted in zone A (red zone) indicate that the process should be investigated and that, if action is taken, the latest estimate of the mean and its difference from the original process mean or target value should be used to assess the size of any correction.[3]

Interpretation of the nonrandom pattern presented through the shift in process level may be caused by the introduction of new production process or new workers, changes in methods, row materials or machines, a change in the inspection methods etc. The presented time series are stationary and uncorrelated according to the augmented Dickey-Fuller test of unit root (ADF = 7,1379 > χ^2(1%) = 3,5457, the null hypothesis is rejected and production process sample series is a stationary process) and autocorrelation and partitial autocorrelation functions are not showing any specified model formulation. They are showing that the process is in control and stabile.

The control process suspected to posses a nonrandom pattern needs to be investigated by modified control charts. These control charts are useful when it is important to detect small shift in a process.

4.2.3. Modified control charts

[3] Oakland, John (2003) Statistical Process Control – fifth edition, Butterworth – Heinemann, Oxford, p 120

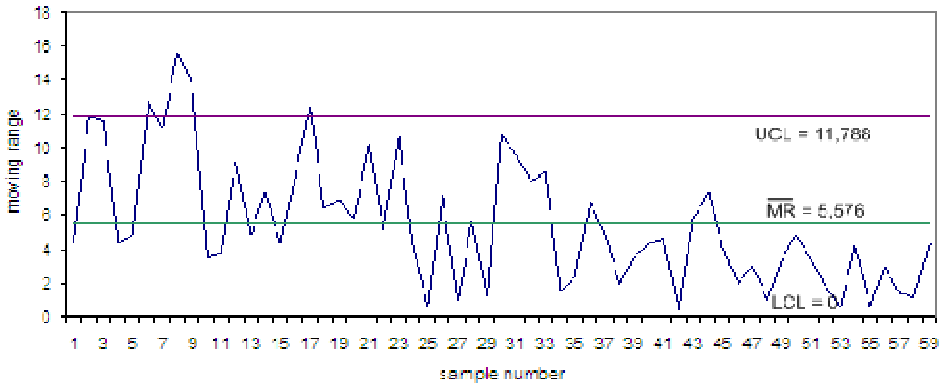
1. Here we give a presentation of the Shewhart control chart for individual measurements known as moving average control chart. Mean charts are not able to detect the difference between routine or natural variations and exceptional variations. However, moving range charts serve to reinforce the mean charts by detecting changes in process variation even if the process average does not change, and providing the ability to discuss the amount of routine variation in a process making the difference between upper and lower natural limits.

Fallowed formulas are used in calculation:

$$\text{Center line } \overline{MR} = \frac{1}{k-1}\sum_{i=1}^{k} MR_i$$

$$UCL_{MR} = \overline{MR} + 3 \cdot d_3 \cdot \sigma_x \quad LCL_{MR} = \max(0, \overline{MR} - 3 \cdot d_3 \cdot \sigma_x)$$

Figure 4. Moving range control chart



Source: own creation

The moving range chart divides variations and presents only the random (routine) variations. Figure 4. shows that the first part of the total observed samples has several points beyond control limits and confirms the above mentioned explanation about the type of pattern samples have. The second part of the time series performs a stationary series contained inside the permitted limits.

2. Exponentially weighted moving average (EWMA) control chart is a statistic for monitoring the process that averages the data that gives less and less weight to data as they are further removed in time. By the choice of weighted factor λ, the EWMA control procedure can be made to a small or gradual shift in a process, whereas Shewhart control procedure can only react when the last data point is outside control limits.

$0 < \lambda \leq 1$ is a constant that determines the depth of memory of the EWMA procedure, or parameter that determines the rate at which older enter into the calculation of the EWMA statistic. Thus, the larger value of λ gives more weight to recent data and less weight to older data, a smaller value of λ gives more weight to older data. Values of λ in the interval $0{,}05 \leq \lambda \leq 0{,}25$ work well in practice. The use of smaller values of λ is appropriate to detect smaller shifts.

The calculated statistic is:

$$EWMA_i \rightarrow z_i = \lambda \cdot x_i + (1-\lambda) \cdot z_{i-1} \quad \text{where} \quad z_0 = \mu_0$$

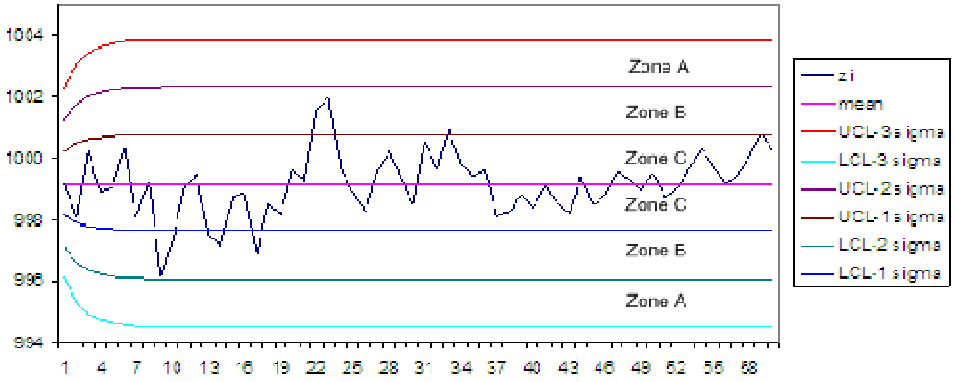
$$UCL = \mu + L \cdot \sigma \cdot \sqrt{\frac{\lambda}{(2-\lambda)} \cdot (1-(1-\lambda)^{2i})}$$

$$CL = \mu$$

$$LCL = \mu - L \cdot \sigma \cdot \sqrt{\frac{\lambda}{(2-\lambda)} \cdot (1-(1-\lambda)^{2i})}$$

where L is a number of σ in control limits.

Figure 6. Comparison of raw and exponentially weighted data

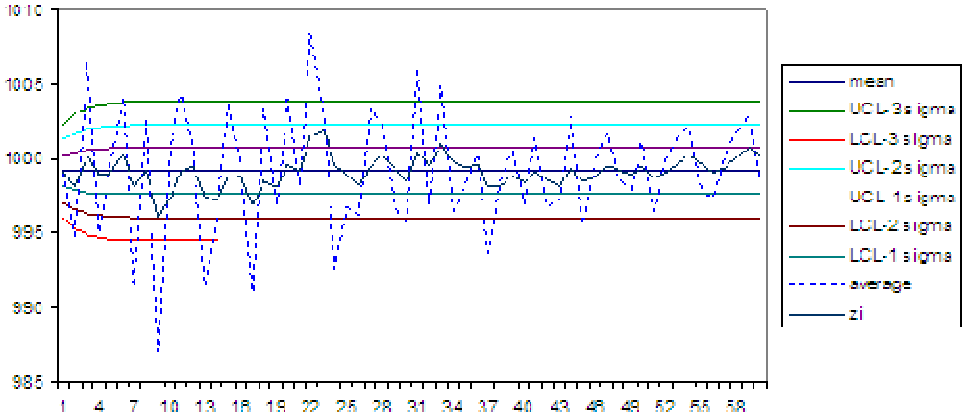


Source: own creation

Figure 5. with λ = 0,25 indicates that the process is in control because all EWMA points lie in ± 2σ control limits. Also there is no consecutively big number of sample points in Zone B which indicates the existence of only random or natural variations especially in the first part of the sample. Advanced explanations could use

the comparison between original sample averages and exponentially weighted moving average data presented in Figure 6.

Figure 6. Comparison of raw and exponentially weighted data



Source: own creation

Last figure shows the ability of exponentially weighted data to smooth the affects of known, uncontrollable noise in the data. Many short-term fluctuations may be large, but they are purely indicative of process instability. Appropriate choice of lambda could determine a control chart less sensitive to those short-term fluctuations.

3. Moving average control charts are used to monitor processes over time. MA chart is efficient in detecting small shifts and to evaluate stability of process.

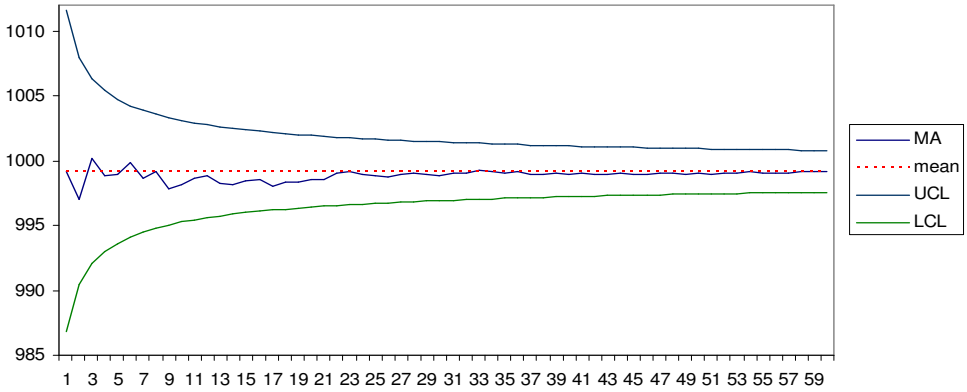
Moving average of span w is $M_i = \frac{x_i + x_{i-1} + \cdots + x_{i-w+1}}{w}$

Where the calculations of the control limits are:

$$UCL = \mu_0 + \frac{3\sigma}{\sqrt{w}} \quad LCL = \mu_0 - \frac{3\sigma}{\sqrt{w}}$$

According to the above formulas control limits are asymptotically approaching to center line simultaneously decreasing the level of permitted variations.

Figure 7. Regression control chart



Source: own creation

This picture clearly presents the lack of variations in the second part of sample group. Moving average line is constantly approaching to the center line and it is showing no sample points out of the retrenched control lines. This is another evidence of improvement in production line process and it could be said that the introductory period for machines or workers is past now and presently the company possesses a completely equipped production line for continuous process improvement.

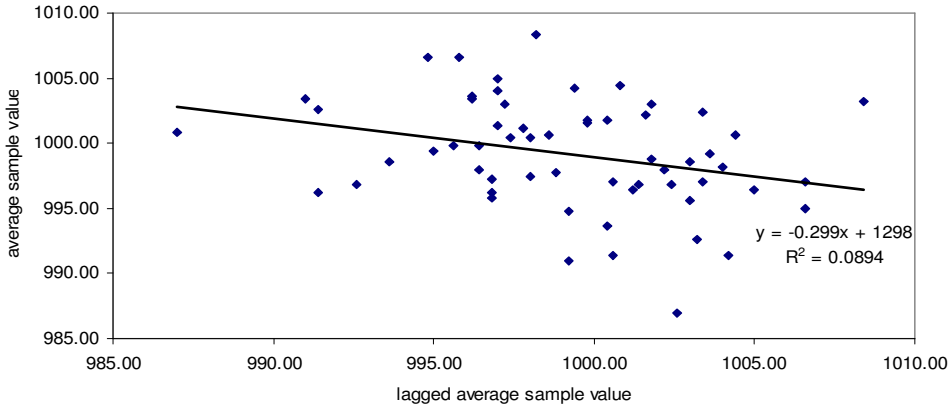
4. The regression control chart is useful when monitoring between two aspects of a production process is important. Using scatter diagram with a regression line could be of great help in identification of outlets in specified regression.

In this paper the following model specification form is used:

$$x_t = \xi + \varphi \cdot x_{t-1} + \varepsilon_t$$

As we can see, the basic autoregressive model is presented here, where ξ and φ are the regression coefficients and ε is residual or model error.

Figure 8. Regression control chart



Source: own creation

Based on simple autoregressive process this figure connects the actual and lagged variable of sample averages. R^2 shows only small percentage of explaining variable variations and this redirect us on the significance of random variations in the explanation of sample values.

5. The modified control charts mentioned above take into account part of the previous data, but the technique which uses all information available is the Cumulative Sum or CUSUM method. This type of charts is one of the most powerful management tools available for the detection of trends and slight changes in data. This chart is useful for detecting a short- and long-term changes and trends. Their interpretation requires care because it is not the actual CUSUM score which signifies the change, but the overall slope of the graph. For this reason the method is often more suitable as a management technique then for use on the operation level.

The method of cumulative differences and plotting them has great application in many fields of management and they provide powerful monitors in such areas as:

- Forecasting
- Absenteeism production level – detection of slight changes
- Maintenance performance
- and many other in which data must be used to signify changes (Oakland 2003)).

Calculation of the CUSUM score may be represented by the formula:

$$Sr = \sum_{i=1}^{n} (x_i - t)$$

Where

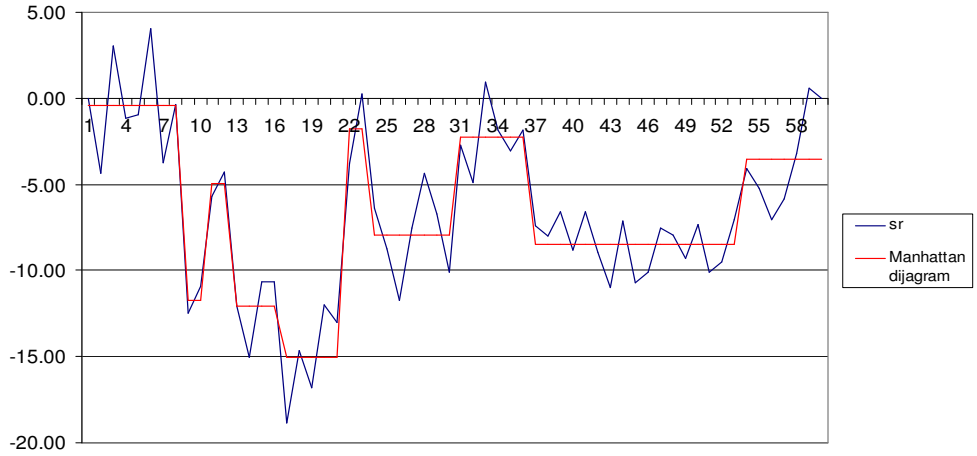
Sr is a the CUSUM score of the *n*th sample

x_i is the result from the individual sample *i* (may be sample mean)

t is called the target value.

The choice of the value of *t* is dependent upon the application of the technique and it is clear that the choice of the *t* value is crucial to the resulting CUSUM graph.

Figure 9. Cumulative sum chart and Manhattan diagram



Source: own creation

This figure present cumulative sum chart with average value as the target value and the Manhattan diagram as the average process mean with time. Because of the constant changes in CUSUM slope, the observations are changing level with many samples below target value. However, second part of the presented observations shows horizontal slope with no significant changes.

Also, CUSUM chart can be used in categorizing process output. This may be for the purposes of selection for different processes or assembly operations, or for dispatch to different customers with slightly varying requirements. To perform the screening or selection, the CUSUM chart is divided into different sections of average process mean by virtue of changes in the slope of the cumulative sum plot. This information may be represented on a Manhattan diagram, named after its appearance. It shows clearly the variations in average process mean over the time scale of the chart.

4.3. Process capability

Another use of control charts not mentioned till now is to determine process capability. Process capability is the range of natural variation in a process known as C_p. It is sometimes also referred to as the natural tolerances of a process. It is used by product designers and engineers to determine how well a process will fall within design specification. In other words, this analysis refers to the infirmity of a process and can be helpful in including development activities prior to manufacturing in analyzing variability relative to product requirements or specifications. Process capability measures potential capability in the process, whereas C_{pk} measures actual capability.

Commonly accepted process capability indices include:

$$C_p = \frac{USL - LSL}{6 \cdot \sigma}$$

$$C_{pu} = \frac{USL - \mu}{3 \cdot \sigma} \text{ upper specification limit only}$$

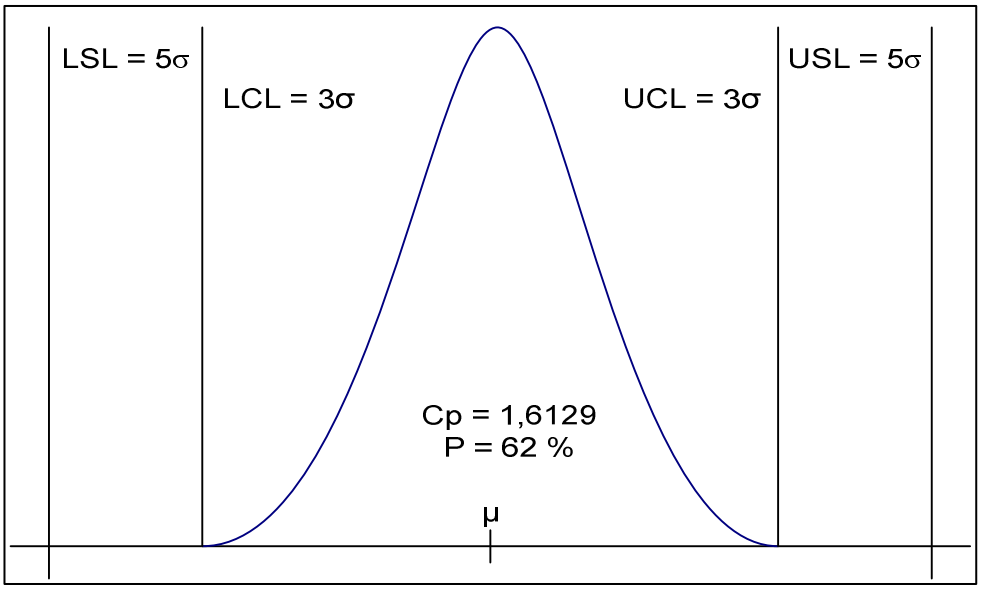
$$C_{pl} = \frac{\mu - LSL}{3 \cdot \sigma} \text{ lower specification limit only}$$

$$C_{pk} = \min(C_{pu}, C_{pl}) \text{ actual process capability}$$

Most capability indices estimates are valid only if the sample number is large enough and assume that the population of data values is normally distributed. Large enough is generally more then 50 independent sample values used in data analyzing.

Specification limits are design or engineering specifications reflecting customer requirements for a product and are set on the level of ± 2 % of the specified product cubage. Process capability is C_p = 1,6129, where upper and lower specification limits are C_{pu} = 1,6782 and C_{pl} = 1,5476. According to this values C_{pk} = 1,5476. Corresponding to previous measures it is possible to estimate percentage of the specification band that the process uses up, $P = \left(\frac{1}{C_p} \right) \cdot 100$. P = 62% of the total allowed natural tolerances.

Figure 10. Process capability and natural tolerances



Source: own creation

Watching this figure we are able to translate process capability into rejects or product failures. Comparison of normal distribution and natural tolerances gives us a probability that any point exceeds the control limits. First, $p_{3\sigma} = 0{,}0027$ or 0,27% of failures according to specified control limits, and second, two-sided process capability specification of 5σ is $p_{5\sigma} = 0{,}6$ parts per million which is the failure level satisfied the customer specifications.

5. Conclusion remarks

Undoubtedly, efforts undertaken for production process quality improvement will continue. Production line presented in this paper shows that after the methods, machines or workers introduction period had finished, the production has become stabile with constant needs for tracking sudden process changes. Due to this, control charts will require periodic revisions of the control limits and central line in order to make an accurate decision in quality improvement.

As we demonstrated before, the use of the numerous control charts is not only desirable but also necessary due to variety of advantages and disadvantages amongst them.

References

Dale, G. B.-van der Wiele, T.-van Iwaarden, J. 2007: Managing Quality, Blackwell Publishing, Carlton.

Douglas, C. M. 2005: Introduction to Statistical Quality Control, John Wiley and Sons, Inc. New York.

Engineering Statistics Handbook
http://www.itl.nist.gov/div898/handbook/pmc/section1/pmc16.htm, 23.06.2009.

Heizer, J.-Render, B. 2008: Principles of Operations Management, Pearson Education, Inc. Upper Saddle River, New Jersey.

Oakland, J. 2003: Statistical Process Control – fifth edition, Butterworth – Heinemann, Oxford, 230 p.

Russell, S. R.-Taylor, W. B. 1998: Operations Management – Focusing on Quality and Competitiveness, Prentice Hall, Inc. Upper Saddle River, New Jersey.

Zu X.-Fredendall D. L.-Douglas J. T. 2008: The Evolving Theory of Quality Management – the Role of Six Sigma, Journal of Operations Management, Volume 26, number 1, 630-650. p.

Critical evaluation of sustainability indices

Hajnalka Ván[1] – Judit Gébert[2]– György Málovics[3]

Questions of welfare and development have been an interesting aspect for economic researchers for a long time. Though most economists agree that the most often used indices of the SNA (GDP, GNP) struggle with a serious deficit in welfare measuring, these indices are still "doing well" as welfare indicators within the current economy and economic policy decision-making. The reason for this is partially that, according to some opinions, maximization of these economic indices is a necessary but not sufficient condition of welfare growth. On the other hand convincing welfare measuring alternatives are missing.

In our study we will critically analyse three welfare indicators (GDP, Human Development Index (HDI), Sustainable Society Index (SSI)) based on the dilemmas connected to welfare theories appearing in welfare economics. Our conclusion is that (1) valid single-dimensional indicators can not necessarily be created for measuring welfare, and (2) over a certain level the GDP per capita indicator does not go hand in hand with the extent of welfare increase indicated by other welfare indicators based on broader information basis.

Keywords: Sustainability, Gross Domestic Product (GDP), Human Development, Index (HDI), Ecological Footprint (EF), Genuine Savings (GS), Sustainable, Society Index (SSI)

1. Introduction

The information base of the SNA indicators does not (or only to very limited extent) cover the natural environment, the natural processes evoked by economy, nor human welfare. Knowing this, in the past decades - and most particularly since the 1990's - more and more organisations and researchers have been trying to work out indicators and indicator systems reflecting sustainability and sustainable

[1] Hajnalka Ván, University of Szeged, Faculty of Economics and Business Administration, Institute of Finances and International Economic Relations Division of Accountancy (Szeged)

[2] Judit Gébert, University of Szeged, Faculty of Economics and Business Administration (Szeged)

[3] György Málovics, University of Szeged, Faculty of Economics and Business Administration, Institute of Economics and Economic Development Division of Regional Economic Development (Szeged)

development. Nothing can prove it more than the fact that the 2003 researches mention more than 500 attempts of elaborating welfare sustainability indicators (indicator sets) (Böhringer-Jochem 2007). Nowadays every important geopolitical (EU, UNO, OECD) has its own sustainable development – and welfare as a part of it – indicator set.

In contempt of all these attempts, traditional indicators of the SNA are those determining mainstream economics analyses on one hand and economics decision-making on the other. There are more reasons for this. On one hand, according to some opinions increasing these economic indicators is a necessary, though not sufficient condition of welfare growth, on the other hand there are no convincing alternative indicators for welfare measurement and supporting economic policy decision-making. At the same time, statements above can not be considered generalized, they are rather questionable instead. The structure of our study in order to critically investigate these statements is as follows: First we will deal with a couple of welfare theories with radically differring messages, then, based on these messages we will determinate a system of viewpoints which will be later used for analysing the indicators investigated. Following this we will critically analyse the indicators of GDP, HDI and SSI based on the roughed theories, dilemmas and system of viewpoints in order to see to what extent can these (based on the level of our current knowledge) be considered valid in measuring welfare. In the third part of our study we will analyse the relationship of these indicators on national level using statistic methods. Finally, we will draw consequences based on this analysis.

2. Theoretical basis of welfare - approaches to welfare in economics

2.1. Economics theories of welfare and the resulting dilemmas with regard to measuring welfare

Defining the term welfare and its theories reaches back to a long history. In our study we will focus on those three theories that can be bound to indicators analysed by us.

The most widespread theory in welfare economics is the *preference utilitarianism* which - as indicated by its name - takes satisfying the individual preferences into consideration. This is a formal theory, ie. it provides only methodology for reaching welfare but does not determine it explicitly. The core idea of this theory is: „If the individuals are entirely following their own interest, they will prefer *x* to *y* in case and only if they think *x* is better for them than *y*. In case they are well informed, their supposition will become true and *x* will be better for them than *y* in case and only if they prefer *x* to *y*." (Hausmann-McPherson 1997,). Preference satisfaction is thus intentionally quite far from undeterminable areas of

emotions and it presumes that people know what is good for them and it is appropriately mirrored by their preference system. This theory belongs to neoclassic mainstream economics and it creates a base for most of economic policy nowadays.

Preference utilitarianism basically operates with ordinal concept of usefulness, ie. it strives to find out whether one state is better or more useful than the other. To measure and compare welfare, however, we might need also a cardinal scale, it means to assign a value to particular preferences (Dasgupta 2001). Money is obviously the simpliest solution for economists. It means we should take a look at how much people would pay for the realization of a particular preference. This point of view is called *materialistic utilitarianism*.

Another theory – also having a significant literature background - is the theory of *elementary goods*. The founder of this stream is John Rawls. According to him welfare must be measured in social goods that are instruments for the individual to live a complete life in the society. Rawls says that elementary goods are the following: freedom, law, power, opportunity, income, property and social basics of self-esteem (Rawls 1971). The command of freedom rights is the primarily important among these.

Since then many economists have tried to set a list of elementary human goods starting-up from Rawls' elementary goods. Using results of different sciences resulted in similar or absolutely different results. The lists offer quite a wide range of goods: they contain material goods such as income, drinking water, food, health; psychological factors such as complacence, happiness, self-realization; more complex terms such as education, safety, human rights as well as more abstract factors such as opportunity or freedom (Alkire 2002). In setting up the elementary goods there can be seen a tendency that along with material goods, even more emphasis is given to environmental protection, human rights, participation in public life or equality and freedom.

Another substantive – it means explicitly defining welfare - theory can be associated to the name of Amartya Sen, a Nobel-prize economist: it is the *capabilities approach*. The term "capability" is explained by Sen as real freedom enjoyed by people. The most significant difference between his and Rawls' elementary goods is that the concept of capabilities does not contain only instruments necessary for welfare, but also relevant human features that determine to what extent can the individual use his elementary goods in order to achieve his goals. Among capabilities the freedom enjoyed by individuals, which is not only an instrument but also a goal of welfare for Amartya Sen, is of special importance. According to Sen's theory the main advantage of the capabilities approach in comparison to utilitarian approaches is that it leans on a wider information base (Sen 2003), meaning that it takes such information and elements of welfare into consideration that were left out from materialistic utilitarianism due to strong reduction. Among these we can find health or education for instance.

The common factor of substantive theories is that they keep the resources of welfare in mind in one form or another, they are specifically interdisciplinary, pluralistic, it means they approach the term of human welfare from more aspects, thus they catch such aspects that can not - or can only be with difficulties - measured with money. They receive most of criticism because they want to take too much into consideration and it goes at the expense of welfare measurability.

Based on the theories sketched above building our criteria system is as follows: as a starting point we will use Sen's term of information base, it means we will criticise the indicators based on what information they include from the most widely understood term of welfare. In line with this we will investigate which information is left out by indicators, because the information left out is as important in index evaluation as the information included (Sen 2003).

Based on the width or scarcity of the information base we can state the following about the welfare theories: the most frequent accusation of utilitarian approaches is that they are over-reduced, they cover a narrow information base, ie. there is valuable information that is not taken into consideration during welfare measurement. In comparison to this, the theory of elementary goods, or the capabilities approach leans on a much wider base, it takes many kinds of welfare elements into consideration, however, this multi-dimensionality often causes difficulties with measurability.

Based on the width or scarcity of information base we have made up the following criteria system:

- **Which measuring devices are favoured by indicators.** The first visible set of problems when bringing welfare theories to practice is the question of welfare measurability and its unit of measurement. Materialistic utilitarianism uses money as a unit of measurement and valuation of different goods is done by the market. This obviously results in significant information base narrowing as most welfare elements - health for instance - are not easily measured with money. In comparison to this, substantive theories recommend using more units of measurement, thus trying to avoid information loss.
- **Mental or physical welfare are taken into consideration.** In the long list of elementary goods, but also among the Sen-like capabilities we can find factors belonging to physical welfare, such as healthy lifestyle or drinking-water and food supply, as well as categories determining mental categories such as complacence, happiness, the feeling of holding one's life in one’s own hands or optimism with regard to future events. Clean environment - which definitely contributes to human health - can also be considered as belonging to physical welfare. Though measuring mental welfare is one of the most problematic questions. We can state that their consideration increases the width of the information base.

- **Do they take such important aspects of welfare like social equality, applying human rights or clean and healthy environment into consideration?** Aggregated indicators are often criticised that during aggregation the question of equality is being overlooked. This is true for all kinds of equality: equal distribution of material goods as well as equal opportunities between sexes. We could see with substantive theories that individual values, human rights, positive freedom rights are even more stressed, therefore these make up an important element of the information base.

3. Critical evaluation of welfare indices

Based on the theoretical dilemmas roughed above we endeavour to show the particular "paradigmatic" elements in particular indicators (chosen by them), as well as evaluate those from the point of view of their information base - it means to what extent they can be considered valid in welfare measurement. As most indices or indicator systems do not unambiguously fit a welfare paradigm, we do no have the opportunity to unambiguously assign these to the theories roughed above. Next we will take the indices chosen for analysis, we will briefly introduce them and at the end we will deal with the criticism specifically touching the indicators.

3.1. GDP

With regard to the information content of the traditional accountancy systems such as the System of National Accounts (SNA) or European System of National Accounts (ESA) it can be said that those do not take action outside the market (for example housework, work in the voluntary sector and services provided by the environment) into consideration. Both of them are based on the assumption that natural resources form an inexhaustible property, and that nature can entirely accept all waste created during mining and other works, it means it is built on the theory that economic growth can reach no natural limits (Giovanelli 2004). We can meet with even more criticism related to information content of review system indicators (GDP/GNP) (Szlávik 2006, van de Bergh 2007, Dabóczy 1998a, 1998b). With regard to the aspects of our study we can state the following about these indicators: based on our first aspect here comes the main criticism of the GDP: it *evaluates in money* and deals only with measuring economic performance. On one hand evaluating in money faces serious theoretical dilemmas itself, on the other hand it excludes (or includes only for impacts on small processes) those elements of welfare about which we can not unambiguously state that they are closely connected with income growth. We can include health or healthy and good quality natural

environment here. This leads us further to our next aspect: elements describing welfare as good mental or physical state. GDP does obviously *not take mental states into consideration* and from among the physical elements of welfare only income is being dealt with. We can use the same criticism based on the third aspect: GDP can in no case be *connected with social equality* as it does not deal with distribution of income nor with applying human rights. To sum it all up, we can say about GDP that it leans on the theory of utilitarianism and it excludes much information from the evaluation, therefore it offers only an extremely narrow information base. Therefore, it can be considered a reductionist welfare indicator with very narrow information base.

3.2. HDI – Human Development Index,

One of the most famous and influencing indicator belonging to substantive theories is the Human Development Index (HDI). This indicator focuses on the human being instead of material goods as it „considers the human development a process of broadening an individual's possibilities of choice" (quotes Garami 2008). Creating the HDI is usually connected with Amartya Sen[4] whose capabilities approach provides a theoretical frame for the indicator. The main advantage of HDI is that it tries to catch the human development in more dimensions thus giving a more complex picture about the change of standard of living than the one-dimensional approaches. At the same time it has a smaller data demand in comparison with complex indicator systems; therefore, its value can be calculated for many countries and can be easily applied also in political decision-making.

HDI uses three dimensions for measuring development that represent the possibilities of choices of people:

1. long and healthy life,
2. education,
3. economic performance (HDR 2007/2008).

The first component is being approached using life expectancy at birth. This indicator gained the least criticism; therefore, it remained unchanged (Husz 2001). They had first tried to determine the second component as the rate of literate population first, but this approach got much criticism (Husz 2001) so in 1995 the combined gross education rate was added which covers the rate of all visitors of elementary, secondary and higher education institutions to the entire population (6-23 years of age) in %. This way the education dimension is obtained as the weighted mean of these two values, the number of literate population being two-thirds and the education rate being one-third of the total value. The index attempts to measure the material goods abundance using the purchasing power parity GDP per capita where

[4] De Benicourt (2002) for example denies that any theory could be in relation with creating the HDI.

a logaritmic formula has been established for HDI. It was probably this component that received most criticism later on.

The problem arising from different units of measurement of the three components was resolved in the way that instead of absolute indicator values the relative values of particular countries were compared side by side. According to this conception, performance growth of the countries was, however, bound to performance of other countries. Because of this problem following the criticism since 1994 extremes of all indicators have been recorded and data of particular countries were correlated to these extreme values. Thus the annual values of these indices have become comparable (Husz 2001).

From the values of dimensions obtained this way the HDI has been calculated as follows:

$$HDI = 1 - \frac{I_1 + I_2 + I_3}{3}$$

Analysing the HDI based on our system of viewpoints using the first point of view we can declare that this indicator does not expresses the welfare change in money but in a *point value* between zero and one. Its preferred measurement device thus is a scoring system where the education and long and healthy life factors were not reduced by financial evaluation. The selected units of measurement obviously can - and they definitely do - face methodology criticism, nevertheless, we can state for sure that using more dimensions is a step towards information base increasing (Husz 2001). With regard to taking mental and physical states into consideration we can say that though HDI does not handle mental states, it takes *three different elements of physical welfare* into consideration: health, education and income. This obviously increases the mass of information included, nevertheless, it always excludes such factors as clean environment or drinking-water and food supplement. One of the most important criticisms of the indicator belongs to our third point of view: *social differences do not appear explicitly* in the indicator. It implicitly includes the question of distribution with the logaritmic formula used for GDP calculation, though. The same way we get information about the fulfilment of human rights only indirectly: about right for education via the education dimension and about healthcare via the factor of long and healthy life. However, types of political freedom are missing as such from the indicator information base.

To summarize what we already know: the HDI per the Sen-like capabilities approach determines the development by growth of possibilities, the human possibilities and such welfare are being approached by the three emphasized factors. By involving health and education the HDI provides much broader information base for the index than the emphasized indicators of the SNA. At the same time its information base can be still considered quite narrow from the point of view of welfare, considering that certain information important of welfare are not or only

very indirectly included (we can mention inequality, political freedom rights, mental states, clean and healthy environent here). Next, the averaging used during conversion of partial values of individual components to a final HDI value can easily result in the fact that countries in entirely different situations can appear as similarly developed.

3.3. SSI – Sustainable Society Index

The SSI summarizes the most important elements of a nation's sustainability and quality of life in a simple and transparent way. *A society is considered sustainable when it is able to fulfil the needs of the current generation in a way that does not endanger future generations' ability to fulfill their needs, and where all the people have the possibility to develop their capabilities in a balanced society freely and in harmony with their environment.*

The indicator takes twenty-two factors into consideration (grouped into five categories), calculated by using data of scientific institutions and international organisations. The twenty-two factors are converted to a scale ranging from zero to ten in the following way: in case the factor fully corresponds with the sustainability criteria set, it has the value of ten points, in case it does not correspond at all, it is worth zero points. Determining the criteria was simplier for some factors (e.g. the portion of undernourished people to the whole population should be zero), while it was more complicated for others (mostly with environmental factors). From the five groups the first three were taken 1/7 of their weight while it was 2/7 of weight for the last two groups (indicators within the groups appear with equal weight) (van de Kerk-Manuel 2008a).

In the first group we can find factors of personal development, healthy life, sufficient food supply, sufficient amount of drinking-water, the appropriate health conditions, the possibility of taking part in public education as well as the equality between sexes. The second group includes factors of clean environment, air quality, quality of surface water and the quality of soil. The third group measures factors important from the point of view of a settled society: good government, unemployment, population growth, distribution of income and government debt are considered. The fourth group contains factors of sustainable use of resources, recycling of waste, utilization of renewable hydropower and renewable energy consumption. The fifth group contains factors important from the point of view of Earth sustainability, area of land covered by forests, preservation of biodiversity, emmission of greenhouse gases, ecological footprint and international cooperation.

With respect to the nature of SSI welfare sustainability indicator and the fact that our study concentrates merely on measuring welfare, during our analyses the starting point was not the entire SSI indicator, only its version narrowed to measuring welfare („well-being SSI"). From the five SSI indicator groups roughed out above the first and the third contains social indicators, while the second, the

fourth and the fifth contains environmental indicators (excluding the international cooperation) (for indicator details see van de Kerk–Manuel 2008b). The „well-being SSI" indicator created by us contains the social indicator groups as well as the one from the environmental indicator groups in connection with clean environment. The latter because this indicator is not primarily connected with sustainability (ie. capability of ensuring welfare in the future) but with the current welfare as the environmental indicators it contains influence the welfare in present directly at a particular place (see the Sen-Stiglitz material).

Based on our evaluation viewpoints it can be said that the well-being SSI uses a chiselled *point system* as the unit of measurement, thus financial evaluation is supressed. Measuring the consumption preferred by the materialistic utilitarianism is consciously denied as the GDP per capita is intentionally excluded from among the welfare indicators. Based on our first viewpoint this is definitely a step towards increasing the information included. Based on our second viewpoint we can say that SSI takes the wide offer of *physical welfare elements* explicitly into consideration: food, drinking-water, health, clean environment, etc. However, SSI also *handles the mental states as excluded information*. From some factors - for example unemployment or good government - we can deduce general satisfaction though. In accordance with our third viewpoint the strength of the indicator is that in measuring welfare it *emphasizes the particular types of social differences* (income distribution, equalitiy between sexes) and thus also justice as well as certain human rights.

To sum it all up, we can say about the index that from the three indicators analysed by us the SSI offers the broadest information base. Only mental states belong to excluded information, therefore, criticism can be first of all given to factor weighting or point system methodology. At the same time the fact that three welfare components compose one well-being SSI value brings up the problem discussed with regard to HDI, according to which countries in entirely different situations appear to be similarly developed (this problem permanently appears when calculating the SSI value as the information loss caused by averaging is even bigger here).

Finally, we would like to draw attention to one more viewpoint common for all three indicators - or rather generally with regard to welfare indicators and indicator systems based on statistical data collection. **It is the too general nature of welfare and sustainability indicators and indicator systems based on statistical data** (Simon 2003, Pataki without publication data.). As Pataki draws with regard to these indicators: „Though these indicators and associates make a big step towards a more complex and real measuring of the society, everything has certain limits. First of all it is that they are rather general, a social group, the society or its well delimitable part can not identify itself in it as a whole. This process results in customized, broad indicators - or more precisely its well selected aggregation - in which a region or a town makes common steps to sustainable development in accordance with the society." (Pataki, without publication data).

4. Analysis

The present analysis, related to well-being indexes, has been done on the basis of a 150-element data series. The main analytical objective was to study the existence of any reported correlations between well-being indexes and the interpretation method serving this correlation. The relative GDP and HDI data refer to 2005, while the SSI data relates to 2006 (as the latter indicator was not published in years prior to 2006). HDI values are based on data provided by the 2007/2008 Human Development Report and documents dealing with HDI trends and indicators (1980-2007), including the latest data (http://hdr.undp.org/en/statistics/data/). The GDP data source is the 2007/2008 annual Human Development Report[5]. The basic data for the well-being SSI values have been provided by the official homepage for this indicator (http://www.sustainablesocietyindex.com/ssi-data.htm). Original data is intentionally used in the present work; in a case of data shortage, calculations were based on the average value of a specific region (as indicated several times in this analysis).

The basic sample data proves that the average HDI value is 0.7, which indicates a moderate development level. As shown with a regional division, North-America and Europe (EU and non-EU) are highly developed; Latin-America and the Caribbean, Middle East, Central Asia and Asia-Pacific are moderately developed, while Africa - with its 0.52 HFI value - is very underdeveloped. With regard to relative GDP data, North-America and Europe (EU and Non-EU) are ranked in the first three places; intermediate locations, i.e. the medium-developed category based on HDI, is made up of the 3 aforesaid regions; and Africa is the most underdeveloped of all. The difference for this indicator compared with HDI is due to the comparative diversity of the regional development rankings (Table 1). The mainstream is unchanged in respect of SSI values, as the separate categories include the same groups of countries as the other two indicators. However, the internal category ranking changes here: Europe is first, North-America is second, and these are followed by non-EU member states. The Middle East and Central-Asia, Latin-America and the Caribbean and Asia-Pacific come ahead of Africa, which is ranked last as regards region. In short, the reported indicators uniformly separate developed, moderately- and a less-developed regional groups – though such indicators do also lead to various well-being-rankings within regional groups (Table 1).

[5] As GDP is one of the HDI components, relative GDP data of HDI has been used to avoid errors related to data diversity published by various organizations. The HDI index of several countries - Norway, Ireland, the United States, Singapore, Kuwait, and the United Arab Emirates - has been calculated with a $40 000 GDP per capital ratio. This data has been replaced in the analysis by the corresponding data of the Human Development Report.

Table 1. Regional data and range based on the analyzed well-being indicators

Region		HDI	Well-being rank by HDI (mean)	GDP (in PPP)	Well-being rank byGDP (mean)	Well-being SSI	Well-being rank by well-being SSI (mean)
Africa	Mean	0,5163		2983,6764		4,4776	
	Std. Deviation	0,1281		3666,0129		0,8528	
	Minimum	0,3302		283,2481		2,6988	
	Maximum	0,8375	7	14635,5937	7	6,6121	7
Middle East and Central Asia	Mean	0,7558		9857,1152		6,3124	
	Std. Deviation	0,1246		8773,9084		0,6528	
	Minimum	0,3473		1036,3554		5,0676	
	Maximum	0,9293	5	26321,0000	4	7,7402	4
Asia-Pacific	Mean	0,7100		9162,8517		5,5194	
	Std. Deviation	0,1414		11188,5660		1,1077	
	Minimum	0,5272		904,4509		3,6583	
	Maximum	0,9666	6	33820,6802	5	7,8549	6
Latin America and the Caribbean	Mean	0,7917		7773,0051		5,7218	
	Std. Deviation	0,0536		4287,2773		0,6727	
	Minimum	0,6910		1130,7481		4,1295	
	Maximum	0,8720	4	19998,8528	6	6,8778	5
North America	Mean	0,9588		38261,0165		7,4895	
	Std. Deviation	0,0055		5132,1577		0,2710	
	Minimum	0,9549		34632,0330		7,2979	
	Maximum	0,9627	1	41890,0000	1	7,6812	2
Europe (EU)	Mean	0,9133		25508,1959		7,5202	
	Std. Deviation	0,0455		9225,7965		0,4427	
	Minimum	0,8240		9837,4869		6,6670	
	Maximum	0,9614	2	38505,0000	2	8,4633	1
Europe (Non-EU)	Mean	0,8347		15293,6566		7,3163	
	Std. Deviation	0,0735		12981,7493		0,6017	
	Minimum	0,7124		2319,5085		6,5497	
	Maximum	0,9683	3	41420,0000	3	8,4156	3
All	Mean	0,7127		10623,8201		5,8089	
	Std. Deviation	0,1815		11319,5672		1,3683	
	Minimum	0,3302		283,2481		2,6988	
	Maximum	0,9683		41890,0000		8,4633	

Source: own creation

After basic data analysis, attention will focus on the main research issue i.e. the existence of any correlations between well-being indicators. As proved by the correlation matrix (Table 2), the well-being indicators are significantly, strongly and positively related to each other.

Table 2. Correlation matrix for the analysed well-being indicators

Correlations

		HDI value	GDP per capita, PPP (constant 2007 international $	Well-being SSI
HDI value	Pearson Correlation	1	,787**	,862**
	Sig. (2-tailed)		,000	,000
	N	150	149	150
GDP per capita, PPP (constant 2007 international $	Pearson Correlation	,787**	1	,765**
	Sig. (2-tailed)	,000		,000
	N	149	149	149
Well-being SSI	Pearson Correlation	,862**	,765**	1
	Sig. (2-tailed)	,000	,000	
	N	150	149	150

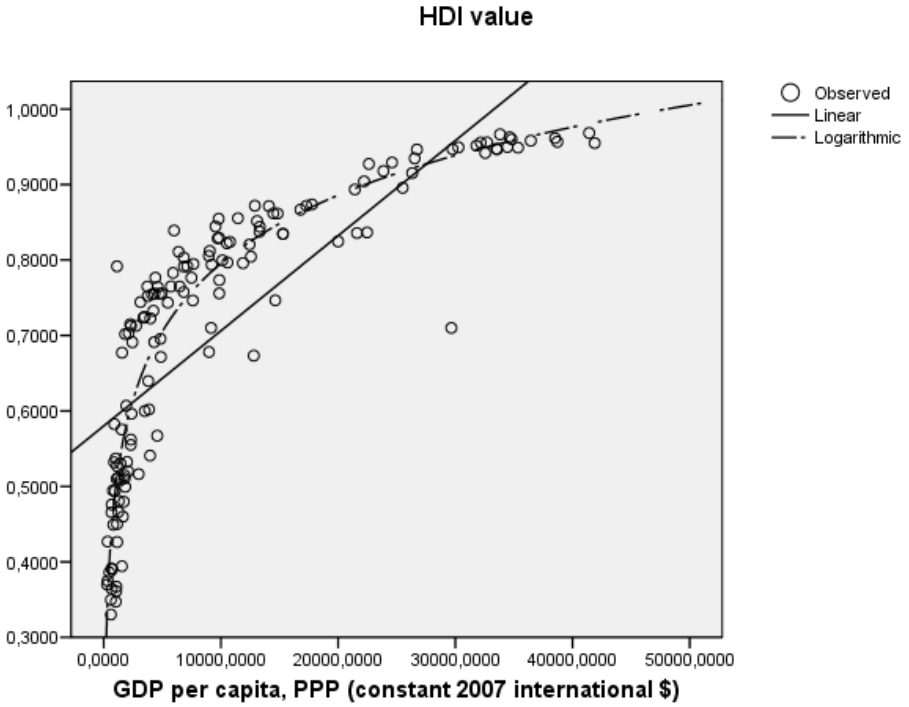
**. Correlation is significant at the 0.01 level (2-tailed).

Source: own creation

A regressive model estimate has been applied to seek out the correlations between specific indicators. Linear, logarithmical and exponential models will be later inserted into this analysis to demonstrate the direction of the correlations between indicators.

With an **analysis of GDP-HDI correlations**, we can say that these two indicators are logarithmically correlated (explanatory strength rate: R2=0,845) (Table 3), so a GDP increase proportionally follows human development for a period if the latter is measured in HDI (Figure 1). With countries having a low GDP per capita, this rise has a growing tendency, yet it slows down after the inflection point of the model; and after a point, the GDP increase hardly contributes at all to any HDI increase. It is interesting to note that happiness research results indicate a similar correlation between economic performance per capita and well-being - if this goes to equal happiness (Layard 2005) the possession of capital assets increases personal and social well-being for a period, but above a certain level it comes to a virtual halt.

Figure 1. Correlation between the development of GDP and HDI indexes (cross-section analysis relating to the countries of the Earth based on 2005 data)



Source: own creation

Table 3. Principle features of the regressive model inserted in cases of correlation between GDP and HDI indexes (cross-section analysis relating to the countries of the Earth based on 2005 data)

Model Summary and Parameter Estimates

Dependent Variable: HDI value

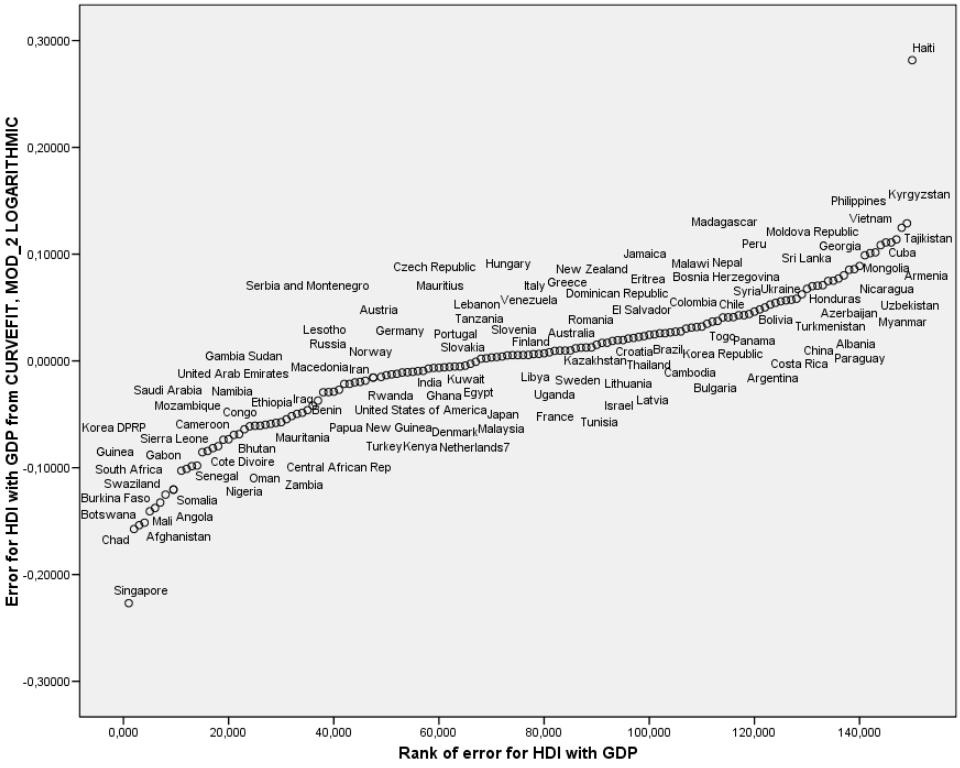
	Model Summary					Parameter Estimates	
Equation	R Square	F	df1	df2	Sig.	Constant	b1
Linear	,619	238,837	1	147	,000	,580	1,26E-005
Logarithmic	,864	930,460	1	147	,000	-,405	,130
Exponential	,521	159,586	1	147	,000	,566	1,83E-005
Logistic	,521	159,586	1	147	,000	1,766	1,000

The independent variable is GDP per capita, PPP (constant 2007 international $.

Source: own creation

Additionally to general relations, there are exceptions - nations that most differ from the trends – that may provide us with important information relating to analysis of the correlation of well-being indexes. Exceptions - i.e. countries less adaptive regarding the logarithmic trend - can be studied by describing the rank places of the residual tags occurring during model insertion (Figure 2).

Figure 2. Deviation from the logarithmic regressive function



Source: own creation

In addition to the significant deviations, the figure also shows that the number of decreasing deviations exceeds the number of increasing ones - and this typically relates to developing countries. Therefore, the relatively high GDP of many developing countries can be associated with a comparatively low HDI level. Consequently, relative GDP well-being is an over-estimate compared with HDI, even in cases when the inflection point of the model has not been reached - though without consideration of the aforesaid result, which supposes that GDP does not bring about the well-being level seen in HDI after a specific period. So a higher

GDP does not automatically equal a higher well-being level according to HDI (Also see Sen 2003).

As information for the upper or lower position of the countries on the model is also important, residual tags are not to be regarded as absolute values. As shown in the analysis, Haiti and Singapore are marked as less adaptive countries. Thus, further analysis of these countries is, unfortunately, rejected, as the relevant HDI values have been replaced by the average value of the specific region due to data shortage. It is possible to analyse the following exceptional countries: Chad, Botswana, Afghanistan, Tajikistan and Kyrgyzstan. As seen in the analysis (Table 4), Chad, Botswana and Afghanistan are situated in the lower part of the model i.e. their relative GDP is associated with having a low HDI value. Detailed data reveals that this situation comes about due to the low educational level in Chad, whilst it is caused by low life expectancy in Botswana and Afghanistan. In contrast, Tajikistan and Kyrgyzstan are to be found above the line i.e. the well-being level calculated in HDI is relatively high and is accompanied by a relatively low GDP in these countries owing to a good education level; we must also make a comparison with economic opportunities (indicated in relative GDP) in both countries, and a relevantly high life expectancy should also be noted. A comparison of countries concerning different groupings gives further interesting results: (1) well-being diversity in two countries that have similar relative GDP values (Chad, Tajikistan); (2) a difference in HDI component values in the case of countries with very correlative HDI values; finally, (3) economic development disparities as indicated in GDP in countries with similar HDI values (Botswana, Tajikistan).

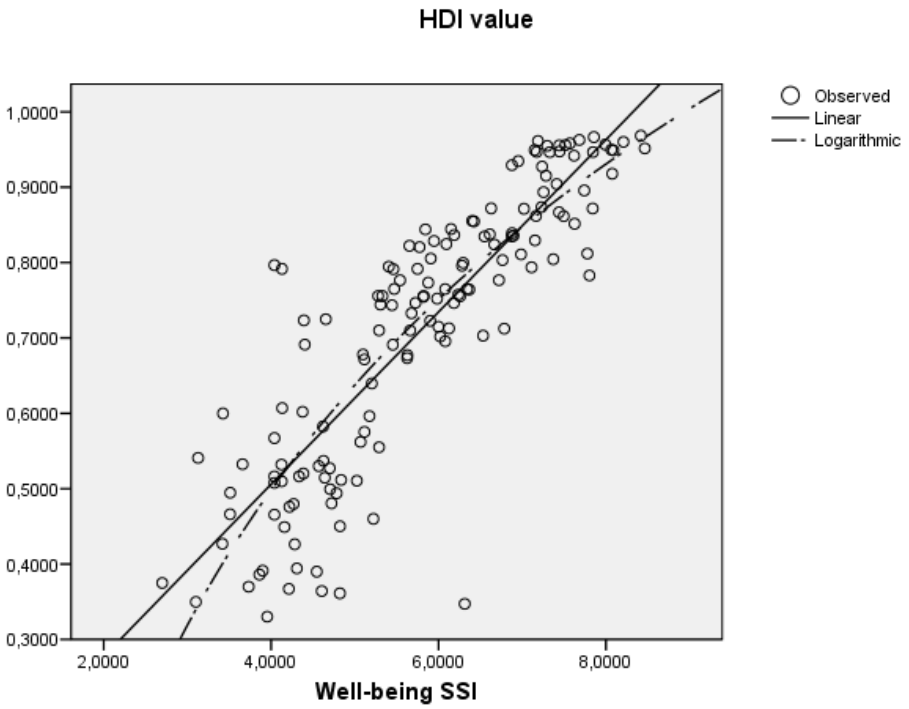
Table 4. Detailed GDP and HDI values of less adaptive countries when it comes to the logarithmic trend

		HDI	Life expectancy at birth (years)	Adult literacy rate (% of population aged 15 years and over)	Combined gross enrolment ratio (% of the population of the theoretical age group for education)	GDP	Life expectancy index	Education index	GDP index
The regression modell overestimate	Chad	0,3943	48,476	31,7648	36,53400392	1554,542	0,391267	0,333545	0,457946
	Botswana	0,6733	50,921	81,3997	70,62144985	12799,77	0,432017	0,778069	0,809821
	Afghanistan	0,3473	42,88	28,00347654	50,0510204	1036,355	0,298	0,353527	0,390271
The regression modell underestimate	Tajikistan	0,6772	65,607	99,5	70,78379431	1563,96	0,676783	0,895946	0,458954
	Kyrgyzstan	0,7018	67,12	98,7	77,49740849	1829,527	0,702	0,918325	0,485131

Source: own creation

The linear model (R2=0,743) has the most notable explanatory strength i**n a correlation of HDI with the well-being SSI**. However, the explanation strength of the logarithmic model is almost identical (R2=0,734) (Table 5, Figure 3).

Figure 3. Correlation between the development of HDI and well-being SSI indexes (cross-section analysis with reference to the countries of the Earth based on 2005 HDI data and 2006 SSI data)



Source: own creation

Table 5. Principle features of the inserted regressive model in the case of HDI and well-being SSI indexes (cross-section analysis with reference to the countries of the Earth based on 2005 HDI data and 2006 SSI data)

Model Summary and Parameter Estimates

Dependent Variable: HDI value

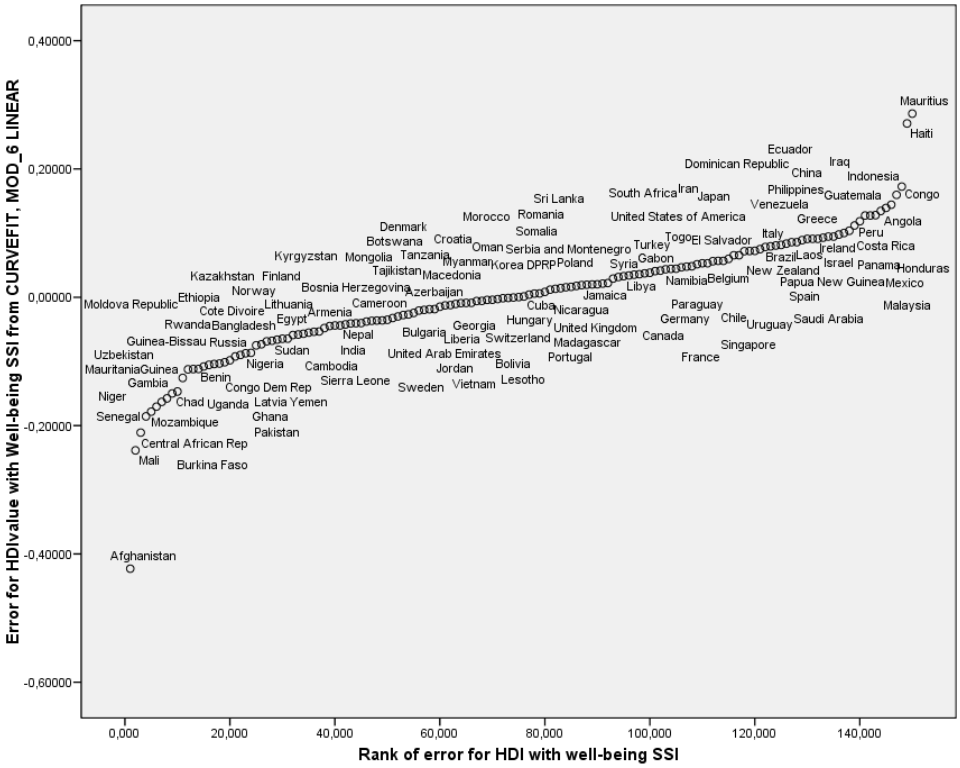
	Model Summary					Parameter Estimates	
Equation	R Square	F	df1	df2	Sig.	Constant	b1
Linear	,743	427,108	1	148	,000	,049	,114
Logarithmic	,734	408,069	1	148	,000	-,366	,624
Exponential	,683	319,573	1	148	,000	,251	,173
Logistic	,683	319,573	1	148	,000	3,989	,841

The independent variable is well-being SSI.

Source: own creation

Should the analysis be completed with residual tags as described in the GDP-HDI correlation analysis, there will be countries that are to be regarded as relatively less adaptive in relation to the linear trend line (Figure 4).

Figure 4. Deviations from the linear regressive model



Source: own creation

As shown by the figure, Afghanistan, Mali and the Central African Republic are located on the lower side - i.e. they have high well-being SSI points and a relatively low HDI. In contrast, Haiti and Mauritius exceed the line, i.e. low well-being SSI is associated with their high HDI. The well-being SSI values gained by Mauritius and the HDI value of Haiti have been replaced by the regional average value due to data shortage, so it is impossible to carry out a more profound analysis with regard to deviations.

However, the reason for diversities can be seen in the case of Mali and the Central African Republic (Table 6). These countries have a relatively low HDI and

comparatively high well-being SSI points[6]. It would be quite logical to presume that this deviation can be associated with the environmental component, which is completely ignored by HDI. However, the situation is precisely the opposite: both countries fail to meet the requirements of clean environment as related to the well-being SSI index. Moreover, this deviation is primarily caused by the dimension of the well-being SSI index, which should correspond most of all to HDI i.e. human development, due to its major concern i.e. human development. Meanwhile, its SSI index defines the concept of human development ever more differently from HDI (See the comprehensive, theorical part of this study). The third component of SSI (a well-balanced society), whose information content - equalling that of clean environment - is not explicitly indicated by HDI, is a little lower than the final well-being SSI value.

Table 6. The HDI and well-being SSI data of countries less adaptive to the linear trend

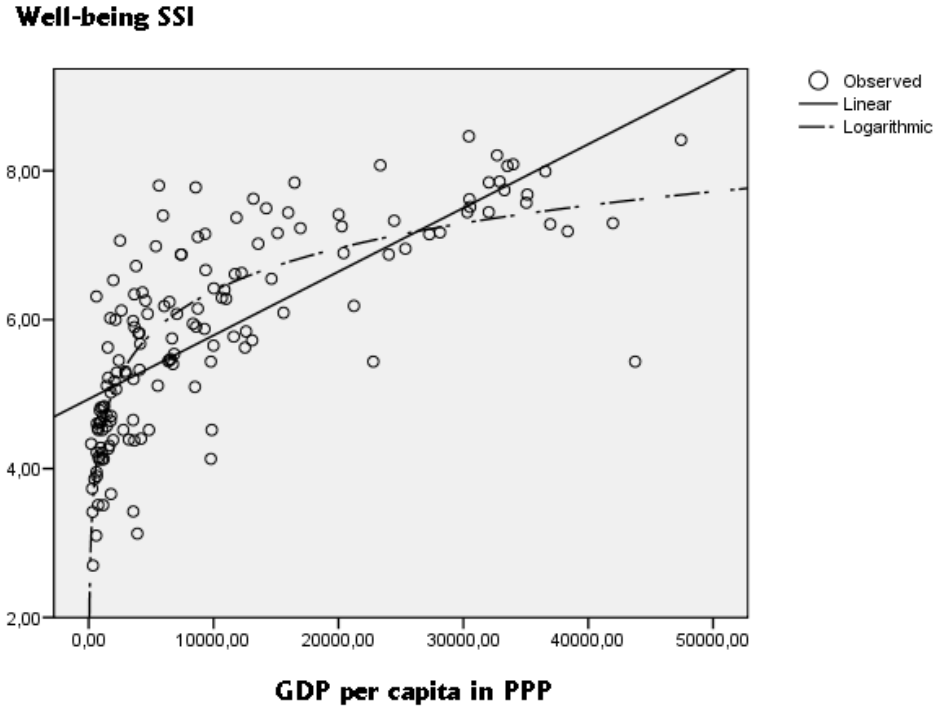
		HDI components								Well-being SSI components		
	HDI	Life expectancy at birth (years)	Adult literacy rate (% of population aged 15 years and over)	Combined gross enrolment ratio (% of the population of the theoretical age group for education)	GDP	Life expectancy index	Education index	GDP index	Well-being SSI	Well-balanced society SSI komp.	Personal development, SSI komp. 2006	Clean envrionment SSI komp. 2006
Mali	0,3612	47,354	26,17655464	42,62671424	1062,9635	0,372567	0,316599	0,394502	4,8213	4,3022	6,5940	3,5676
Central African Rep	0,3642	46,16	48,57356259	28,71530617	682,1324	0,352667	0,419541	0,320465	4,6051	4,3750	5,9112	3,5291

Source: own creation

The correlation between the well-being SSI and GDP, best described via a logarithmic model (R2=0,698) (Table 7), is analysed last in the present study. GDP growth runs in parallel with a well-being increase as indicated by the well-being SSI for a specific time period. The well-being marked by the well-being SSI does not follow GDP growth if a certain GDP level is exceeded (Figure 5).

[6] Even the similarity of the HDI point values for the two countries is contentious in view of the quite significant differences seen in the HDI value as regards the disparity between some of the HDI components. Take into account the previous critics on the reductionist feature of HDI.

Figure 5. Correlation between the development of GDP and well-being SSI indexes (cross-section analysis with reference to the countries of the Earth based on 2005 GDP data and 2006 SSI data)



Source: own creation

Table 7. Principle features of the inserted regressive model in the case of GDP and well-being SSI indicators (cross-section analysis with reference to the countries of the Earth based on 2005 HDI data and 2006 SSI data)

Model Summary and Parameter Estimates

Dependent Variable: well-being SSI

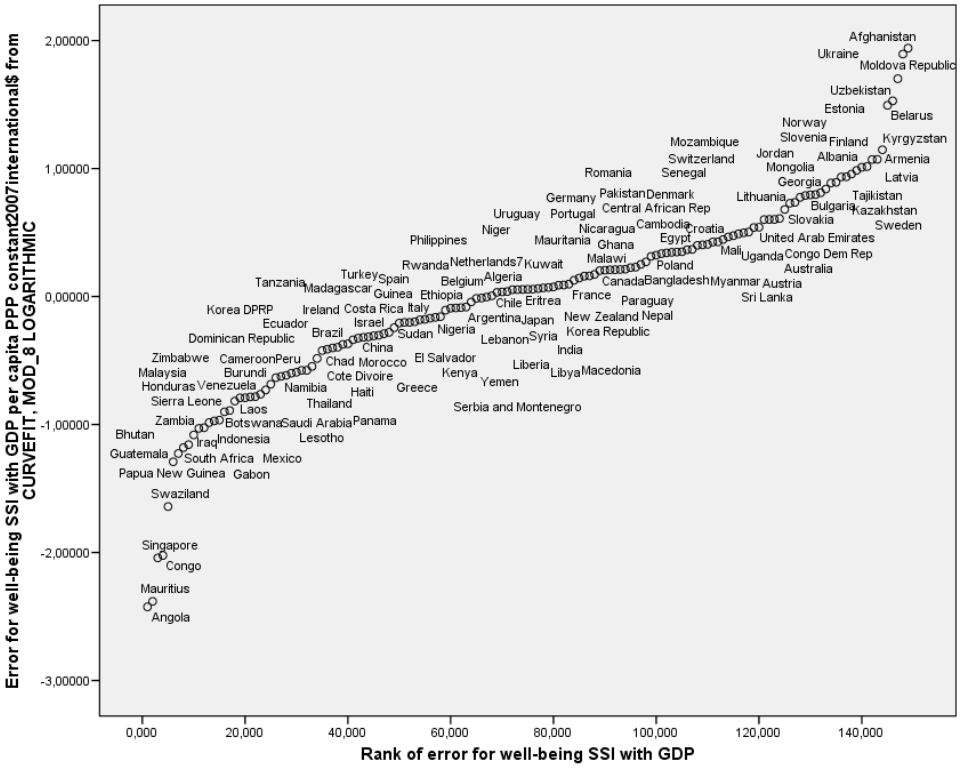
	Model Summary					Parameter Estimates	
Equation	R Square	F	df1	df2	Sig.	Constant	b1
Linear	,585	206,965	1	147	,000	4,841	9,22E-005
Logarithmic	,698	340,425	1	147	,000	-1,755	,882
Exponential	,519	158,547	1	147	,000	4,778	1,58E-005
Logistic	,519	158,547	1	147	,000	,209	1,000

The independent variable is GDP per capita, PPP (constant 2007 international $.

Source: own creation

Should the analysis be completed with residual tags as described in the GDP-HDI correlation analysis, there will be countries that can be regarded as being relatively less adaptive to the linear trend line (Figure 6).

Figure 6. Deviations from the logarithmic regressive model



Source: own creation

As shown in the figure, Angola, Mauritius, Singapore, Swaziland and the Congo are situated on the lower side of the model i.e. they can be characterized by possessing a relatively high GDP and low well-being SSI - unlike Afghanistan, the Ukraine, Uzbekistan, the Moldavian Republic and Belarus, situated above the line, where we have a comparatively low GDP associated with a higher well-being SSI. As the well-being values of Mauritius, Singapore and Swaziland have been replaced by the regional average due to data shortage, these countries will not be part of this analysis. Detailed data for other countries are included in Table 8.

Table 8. Precise well-being SSI and GDP values of countries less adaptive to the logarithmic trend

		GDP per capita, PPP (constant 2007 international $)	Personal developm ent, SSI komp. 2006	Clean envrionme nt SSI komp. 2006	Well-balanced society SSI komp.	Well-being SSI
The regression modell overestimate	Angola	3948,3903	3,9356	2,4734	2,9734	3,1274
	Congo	3504,3508	4,3256	3,3667	2,5794	3,4239
The regression modell underestimate	Ukraine	5912,3584	8,5172	7,649	7,1661	7,7774
	Uzbekistan	2118,8273	8,6939	7,43	7,2835	7,8025
	Belarus	9044,0186	7,1811	5,9368	6,4767	6,5315
	Moldova Rep.	2319,5085	7,4022	7,5889	5,3621	6,7844

Source: own creation

Two conclusions may be drawn from the table. On one hand, SSI values may refer to countries possessing relatively different well-being states - for instance, the similar final well-being SSI values of Angola and Congo cover notable deviations between the different SSI dimensions, as in the case of Belarus and the Moldavian Republic, too. On the other hand, the well-being states suggested by the well-being SSI - the same SSI index values for different countries - may also appear with very diverse relative GDP values (as clearly proven given a comparison of both Uzbekistan-Ukraine and Belarus-the Moldavian Republic).

5. Summary

This study has focused on three indexes associated with each other in economics and with the concept of well-being in the case of the well-being SSI. As stated on the basis of an analysis of the theoretical background of such indicators, the little information basis of GDP in terms of well-being is broadened by the HDI, while the well-being SSI (i.e. the information mass required to outline national well-being levels) is additionally given special attention. In view of the amount of information that has been ignored, both indicators can be regarded as over-reductionist owing to their top-down feature and information that is lost in the process of component aggregation related to indicator creation.

The cross-section statistic analysis outlines a strong statistic correlation between the indicators. Whilst the well-being SSI and HDI are linearly correlated,

these indicators have a logarithmic correlation with GDP. Consequently, the relative GDP - i.e. economic well-being increases – will bring about the well-being growth as indicated by the well-being SSI and HDI (a major information basis in relation to aspects of well-being). However, this does not refer to a higher relative GDP. In addition, a study of exceptions - the nations which most differ from the trend that takes on board the correlation between indicators - reveals that, on one hand, the higher relative GDP situated on the side before the inflection point of the model is often associated with a lower HDI or well-being SSI ratio; while, on the other hand, identical and relative GDP values may refer to a different HDI or SSI well-being level. So the present study supports the theory of moving from an economy-focused interpretation of well-being towards a more complex approach which integrates actual well-being-related scientific knowledge both in economics and eco-political areas.

References

Alkire, S. 2002: Dimensions of Human Development, *World Development*. 30/2, 181-255. p.

Böhringer, C. – Jochem, P. E. P. 2007: Survey: Measuring the immeasurable — A survey of sustainability indices. *Ecological Economics*, 63, 1-8. p.

Dabóczi, K. 1998a: A mérhető balgaság, avagy miért nincs olaj a közgazdaságtan lámpásában? *Kovász*, 2, 2, 32-57. p.

Dabóczi, K. 1998b: Fogalmak által megcsalatva - Kísérlet céljaink és eszközeink embert szolgáló meghatározására. *Kovász*, 2, 3, 16-31. p.

Dasgupta, P. 2001: *Human Well-Being and the Natural Environment*. Oxford University Press, Oxford.

Garami E. 2008: *A humán erőforrás területi különbségei*. 281-298. p.

Giovanelli, F. 2004: *Environmental accounting as a sustainable development tool*. Committee on the Environment, Agruculture and Local and Regional Affairs, Italy.

Hausmann, D. M. – McPherson, M. S. 1997: *Economic analysis and moral philosophy*. Cambridge University Press, Cambridge.

Human Development Report 2007/2008 2008, UNDP New York – Oxford, website: http://hdr.undp.org/en/reports/global/hdr2007-2008/, downloded: 09. 12. 2009.

Husz, I. 2001,:Az emberi fejlődés indexe. *Szociológiai Szemle*, 2, 72-83. p.

Layard, R. 2005: *Happiness: Lessons from a New Science*. Penguin Press, New York.

Pataki, Gy. (without publication data): Ha már mérjük magunkat. *Merítő*. Website: http://www.tabulas.hu/cedrus/2001/08/merito.html downloded: 02. 07. 2009.

Rawls, J. 1971: *A Theory of Justice.* Massachusetts: Belknap Press of Harvard University Press, Cambridge.

Sen, A. K. 2003: *A fejlődés, mint szabadság*. Európa, Budapest.

Simon, S. 2003: *Sustainability Indicators*. ISEE Encyclopedia.

Szlávik, J. 2006: A nem fenntartható növekedés és a fenntartható fejlődés jellemzői. in Bulla, M –Tamás, P. (szerk.): *Fenntartható fejlődés Magyarországon – Jövőképek és forgatókönyvek.* ÚMK, Budapest, 196-211. p.

van de Kerk, G. – Manuel, A. 2008a: Sustainable Society Index. SSI-2008. *Sustainable Society Foundation.*

van de Kerk, G. – Manuel, A. R. 2008b: A comprehensive index for a sustainable society: The SSI — the Sustainable Society Index. *Ecological Economics*, 66, 228-242. p.

Van den Bergh, J. C. J. M. 2007: *Abolishing GDP*. Tinbergen Institute Discussion Paper. Available at SSRN: http://ssrn.com/abstract=962343.

Methodology of Innovation Surveys

Zsuzsanna Szunyogh[1]

The needs for data on innovation increased together with the recognition of the importance of innovation activities. The article gives an overview of the history of innovation statistics focused on methodological developments in the EU countries and in Hungary. The methodological guidelines for innovation surveys, the Oslo Manual was published by the OECD the third time in 2005. Internationally comparable data on innovation have been available for the past ten years, and non technological – marketing and organizational – innovation have only been measured in most countries since 2006. EU legislation on innovation statistics came into force in 2004 and harmonised the surveys – named CIS – in this field.
Lots of results have been achieved but some problems are to be solved. More information on innovation is required, for example on co-operation and skills needed for innovation. Micro data and connected data sources should be used for deeper analyses. The creation of more composite indicators could give a wider picture of different connections in the field of innovation. Knowledge of innovation at enterprise level has to be improved and closer cooperation between data users and data producers would help the development of innovation statistics.

Keywords: innovation statistics, Oslo Manual, CIS

1. Introduction

Lately, the importance of innovation statistics has significantly increased. Today, innovation[2] is undoubtedly a major driver of the economy. Consequently, it is of high priority to acquire all related information. To understand the data on innovation, it is very important to know what kind of methodology is behind them. To meet the growing information needs, it is essential to develop a detailed methodological background. Significant improvements have been made in this area, but still there are many problems to be solved. This article describes the progress, outlines the shortcomings and formulates proposals for their elimination.

[1] Zsuzsanna Szunyogh, Professional Advisor, Hungarian Central Statistical Office (Budapest)
[2] The present definition of innovation by the Oslo Manual: An innovation is the implementation of a new or significantly improved product (good or service), or process, a new marketing method or a new organizational method in business practices, workplace organization or external relations. (Oslo Manual, 3rd Edition page 46)
Author expresses her thanks to Annamária Inzelt for the useful professional advice.

2. The history of innovation statistics

The main drivers of development of innovation statistics are the OECD and EUROSTAT. The work in this field is led and partly financed by these two organisations. Data collections specifically focusing on innovation, providing internationally comparable data do not have a long history.

Until the end of the 1970's, quantitative information on innovation were derived from R&D statistics and were only referred to technological development. The reason for this approach was that the technology push model (Schumpeter 1980) of innovation was widely accepted. It means that all innovation activities were based on R&D. Nowadays, it is obvious that most innovations have other, sometimes more important sources than R&D, for example acquisition of machinery, equipment and software, or external knowledge, training.

In 1980, the OECD organized a conference on innovation measurement, and in 1986 established a working committee with the task of identifying metrics for innovation. Seven countries (the United Kingdom, Canada, Germany, the United States, France, the Netherlands and Italy) developed the method based on their experience for a survey to be carried out in the Nordic countries. Based on the results of this survey a proposal for NESTI (National Experts of Science and Technology Indicators) was developed in 1989. The working group of NESTI – with the active participation of a Hungarian delegate, Annamária Inzelt, in collaboration with the OECD Secretariat, prepared the methodological guidelines for innovation surveys known as the Oslo Manual. Its first edition was published in 1992. The Oslo Manual provides a harmonized framework for measuring innovation and it is used in all the OECD countries and is also a model for other countries.

The first edition of the Oslo Manual placed emphasis on technological innovation and mainly dealt with manufacturing. It was not suitable for measuring innovation in the service sector and non-technological innovation, although their importance had increased. It was the main reason for the revision of the Oslo Manual. The second version of the Manual came out in 1997, but its weaknesses revealed in a short time. Not only product[3] and process innovation[4] must be measured but the other forms of innovation, like marketing[5] and organizational[6] innovation too.

[3] A product innovation is the introduction of a good or service that is new or significantly improved with respect to its characteristics or intended uses. This includes significant improvements in technical specifications, components and materials, incorporated software, user friendliness or other functional characteristics. (Oslo Manual, 3rd Edition, page 48)

[4] A process innovation is the implementation of a new or significantly improved production or delivery method. This includes significant changes is techniques, equipment and/or software. (Oslo Manual, 3rd Edition page 49)

[5] A marketing innovation is the implementation of a new marketing method involving significant changes is product design or packaging, product placement, product promotion or pricing. (Oslo Manual, 3rd Edition page 49)

The results of surveys and the needs for detailed information on innovation drew attention to other contexts of the issue. The revised Oslo Manual came out in 2005. The main changes are that this publication includes the measurement methods for non-technological innovation forms, – like marketing and organizational innovation – and not only technological improvements have been accepted as product or process innovation (OECD 2005, Katona 2006). If the methodology is changed, it is very important to take the aspects of comparability into account.

The first CIS (Community Innovation Survey) type survey (CIS1) was carried out – based on the Oslo Manual – in 1993. The original purpose of the survey was to obtain data on innovation outputs, in addition to a range of innovation inputs and activities that were not based in formal R&D. It was a pilot survey aiming to test the methodology. Internationally comparable data were only provided by CIS2 in 1997-98. After CIS2, CIS was organised every second year, namely CIS3 in 2000, which was the first fully harmonized at EU level.

The year 2004 was extremely significant, because the Commission Regulation (EC) No 1450/2004 on innovation statistics came into force (Európai Közösségek Bizottsága 2007). This regulation orders not only what kind of data should be compiled and what kind of variables need to be reported to the EUROSTAT, but it determines the methodology for data collection too. It means that the main characteristics of innovation surveys are the same in all the EU countries:

- Target population: all enterprises with more than 10 employees in selected NACE categories (agriculture, construction and some service industry are not covered
- A harmonized questionnaire is used (the core questionnaire is exactly the same)
- The survey is repeated regularly every other year.

The "Blu Sky II" conference entitled "What Indicators for Science, Technology and Innovation Policies in the 21st Century" was organized by the OECD in 2006. This forum focused on the measurement of innovation and it outlined how to construct policy-relevant indicators to respond to the changing needs.

In 2008 CIS2006 was carried out. The questionnaire of this survey – according to the 3rd Oslo Manual – included two new modules: one for measuring organisational innovation and the other for marketing innovation. First of all the policy-makers were interested in environmental innovation. To get information on this topic, CIS2008 has a separate part for it.

[6] An organisational innovation is the implementation of a new organisational method in the firm's business practices, workplace organisation or external relations. (Oslo Manual, 3rd Edition page 51)

3. Innovation statistics in Hungary

The first innovation survey was carried out in Hungary out in 1994 in the frame of Innovation Research Institute led by Annamária Inzelt. It was a pilot survey aiming to test the methodology recommended by the Oslo Manual (Inzelt 1995). In 1999, another pilot survey was organized by the Innovation Research Institute focusing on knowledge intensive service sectors (Inzelt 2001). The goal of this survey was also to improve methodology.

The first innovation survey at national level was carried out by the Hungarian Central Statistical Office (HCSO) in 2000, in order to measure innovation activities in manufacturing (KSH 2001). The primary aim was to test the questionnaire and to adapt the methodology. The method used was based on the Oslo Manual. The definitions were new and sometimes unknown for the data providers. The new questionnaire was long and it was to respond on a voluntary basis, which resulted in a very low response rate (under 20%).

The second innovation survey at national level, CIS3 was fully harmonized with the OECD's and EUROSTAT's recommendations and its data were more reliable than from the previous one (KSH 2003). "CIS light" with fewer questions was implemented in 2003.

In accordance with EU legislation Hungary has been carrying out CIS every other year since 2004. CIS4 was organized by HCSO in 2005. CIS4 was the first obligatory one to respond, and its response rate was higher than 70%. Hungarian surveys cover more than 19,000 enterprises, their sample size is app. 6,000 and include all enterprises over 100 employees and the smaller ones by the sample selection (KSH 2006).

Besides the HCSO's innovation surveys, there are others organized by different research institutes which are in most cases focused on special areas like industrial parks or microenterprises.

4. The strengths of present innovation statistics

There is a separate survey especially for the collection of data concerning innovation.

According to theoretically supported methodological guidelines this CIS type survey measures technological and non technological innovation.

The standardized methodology and the harmonized core questionnaire are used in every EU countries. The core questionnaire includes the following parts:

- General information about the enterprise
- Product (good or service) innovation and its developers, share of turnover of new products
- Process innovation and its developers

- Ongoing or abandoned innovation activities for product and process innovations
- Innovation activities and expenditures for product and process innovation
- Sources of information and co-operation for innovation activities
- Innovation objectives
- Organisational innovation
- Marketing innovation
- Basic economic information on enterprise
- EU countries are to report the set of indicators to EUROSTAT on a mandatory basis. (Until 2004, the implementation of surveys was based on "gentlemen's agreement".)

The quality of data on innovation is controlled, based on a quality report system prescribed by EUROSTAT and it is obligatorily used by the EU countries. The data from CIS are mostly internationally comparable. The regional innovation data are available in most countries, first of all to help policymakers to define innovation strategy.

A Safe Centre was established by EUROSTAT, where the microdata sets from innovation surveys are available for researchers for deeper analyses.

The method of measurement is continuously developing; from time to time the experts of this area discuss the problems and necessary improvements which are based on the increasing needs of data users. There is a difficult task to supply more and more data without increasing the burden of data providers and data producers. One of the solutions of this problem is that the structure of the questionnaire is modular. In addition to the core questions, every survey has a module, dealing with different topics. (For example the module for environmental innovation in CIS2008, carried out in 2009-2010.)

The description of strengths may seem short, but the content behind them is momentous and very important. It has been achieved by means of enormous efforts.

5. The weaknesses of innovation statistics

One of the basic problems is the definition of innovation itself. The definition is not very strictly defined, a bit "light". It is difficult to identify what is new or significantly improved. In general, it is rather the smaller companies that are faced with the problem. They often think that innovation could only be used by bigger enterprises. The other problem is that the definition is not as widely known as it should be. Sometimes the term of R&D and innovation are used synonymously.

There are some problems connected with the usefulness of indicators. The most widely used indicator in innovation statistics is the average ratio of innovative enterprises. It is not enough correctly useable for two reasons. One of them is that this indicator does not show how innovative the firms were, e.g. when the enterprise

introduced one single product with a share of cca.1% of the turnover, it is counted equal with another firm introducing more products with higher share and in addition having process innovation. The average ratio is not fully appropriate for comparisons because the structure of economy (by NACE categories and by size of enterprises) affects its size. The countries could be very different in this respect. The use of standardization method would be preferable.

The other group of problems is related to the lack of information. Only little information is collected on the impacts of innovation. It would be very important to know more exactly how innovation affects productivity and the profit of companies. Other data sources than CIS should also be used for analysing this. It is a fact that the possibility of connecting different databases is very limited. The rules of data protection are strict in most countries. Generally National Statistical Offices are only allowed to connect data and analyze them at enterprise level. OECD had a micro data project. Its results were published and give a good example of use of CIS data for econometric analysis (OECD 2009).

EUROSTAT provides access to anonymised micro data for scientific purposes but this is not a proper solution for the above mentioned problem because the anonymisation method could cause the loss of important information and the EU countries micro data are sent to EUROSTAT only on a voluntarily basis, it means that some of the national datasets are not available at the Safe Centre.

It is known that co-operation plays a very important role in innovation but the related information is insufficient. CIS collects information only on type of partners and not on type, frequency and content of co-operation.

Human resources are key elements of innovation. There are no questions in CIS about this topic. Innovation scoreboard only includes data on the number of people graduated at ISCED 5 and 6 levels and the number of participants in lifelong learning. These data are not detailed and are not in direct connection with innovation activities of enterprises. CVTS[7] could provide valuable information. CVTS are carried out every five year in the EU counties, based on harmonized methodology providing internationally comparable data. There is a possibility to use data from this survey and CIS together. The reference periods of the two surveys are not the same, but the relatively slow changes in these fields would not cause significant disturbing effects in the analyses. EUROSTAT plans to introduce a new module in CIS questionnaire on skills needed for innovation but it has not yet been worked out in details.

It is evident that different types of innovation have effects on each other but only a little information is available on that. Most product innovations go hand in hand with process innovations, and technological innovations very often need non-

[7] CVTS is Continuing Vocational Training Survey carried out every five years in EU countries with harmonized methodology

technological innovations. The connection between different types is bidirectional. Deeper analyses are needed to discover links between several innovation activities.

It is known that innovation could have lot of sources. It is to be regretted that there is no information on the importance of different sources. The CIS questionnaire only includes a list of the sources without their order. Data on innovation expenditure are collected and have some usefulness in this respect but these are not reliable enough because enterprises in general do not have separate files on their investments, costs by different purposes and cannot show appropriately the weights of separate sources. R&D expenditures are the most solid data but it is not too simple to divide them by correlation to innovation.

An overall problem in innovation statistics is the weak co-operation between data producers and data users, namely the determining group, the policymakers. The data needs are sometimes not defined properly or come out very late to take into consideration for surveys. The period between the emergence of needs and their satisfying is long, because of the necessary methodological preparation work.

For satisfying the needs the length of questionnaire and the burden of respondents must be taken into account too.

These days the efficiency of functioning of organizations in the government and higher educational sector is a key question. There is no information on innovation activities in the public sector. A special approach is needed for measuring innovation in this area. The OECD's and EUROSTAT's expert groups have begun the methodological preparation work but only pilot surveys have been carried out in a few countries until now.

Last but not least, the problem of used indicators has to be mentioned. Most of them are simple and not composite (OECD 2008). (A composite indicator means the combination of answers to several questions in creation of an indicator, like dual – product and process – innovators.) Composite indicators can provide a deeper picture of innovation.

6. Conclusions

The needed improvements follow from the weaknesses. Some of them were mentioned above. The main message of these problems is that more information should be used. It does not mean that CIS needs to be enlarged. The use of different connected data sources could be one of the solutions. Separated surveys on smaller samples could be the answer to the question concerning co-operation.

The knowledge of innovation also needs to be diffused at enterprise level. It is not a statistical task, it needs a contribution of governor institutions.

It is clear that innovation activities are more important than their measurement itself, but without measurement it is impossible to evaluate results and define the strategy

which helps to become more innovative with economically and socially beneficial results.

References

Európai Közösségek Bizottsága 2007: A Bizottság jelentése a Tanácsnak és az Európai Parlamentnek az 1608/2003/EK európai parlamenti és tanácsi határozat végrehajtásáról

Inzelt, A. 1995: *Az Osló kézikönyv magyarországi alkalmazhatóságának értékelése az 1994. évi innovációs felvétel alapján*. Zárótanulmány, OMFB, Budapest.

Inzelt, A. 2001: Kísérlet az innovációk mérésére a szolgáltatási ágazatokban *Külgazdaság*. 45, 4. 35-51. p.

Katona, J. 2006: *Az Oslo kézikönyv harmadik kiadásának kiértékelése*. NKTH, Budapest.

KSH 2001: *A feldolgozóipar innovációs tevékenysége*. Központi Statisztikai Hivatal, Budapest.

KSH 2003*: Innováció 1999–2001*. Központi Statisztikai Hivatal, Budapest.

KSH 2006*: Innováció 2004*. Központi Statisztikai Hivatal, Budapest.

OECD 2005: *Oslo Manual 3rd Edition*, OECD Publications, Paris.

OECD 2008: Innovation indicators, working paper, presented in NESTI meeting, 16-19 June 2008, Paris.

OECD 2009: *Innovation in Firms OECD* www.oecd.org/publishing/corrigenda

Schumpeter, J. A. 1980: *A gazdasági fejlődés elmélete*. Közgazdasági és jogi könyvkiadó, Budapest.

Possible approaches of the Quarterly GDP Estimation

Klára Anwar[1]

Nowadays, quarterly national accounts - in particular quarterly GDP - represent an essential instrument and source of information for both economic and monetary policy as continuously underlined by the major users such as the European Central Bank (ECB), government authorities, economic actors and forecasters.
At the end of 2006, Hungarian Central Statistical Office (HCSO) has introduced full set of current price quarterly calculations based on statistical and administrative data sources. According to this approach, on production side quarterly GDP is not obtained directly but as a difference of output and intermediate consumption plus taxes less subsidies on production. For this compilation, intermediate consumption has been estimated quarterly though there is no direct data source available for this figure. Under continuous methodological work, I would like to explore and analyse the possible approaches to estimate intermediate consumption quarterly.

Keywords: quarterly national accounts, GDP, intermediate consumption, methodology

1. Introduction

The HCSO has started to establish the methodological work on quarterly national accounts in 1993. Its actuality and importance is well defined in the European System of Accounts as "they are the only coherent set of indicators, available with a short time-lag, able to provide a short term overall picture of both non-financial and financial economic activity." (Eurostat 1996)

At the beginning, quarterly GDP was estimated by volume projection method on production side, i.e. available proxy indicators were applied to extrapolate the value added series at two-digit level of NACE[2], the available nomenclature for economic activities at that time. Volume indices were aggregated by using the shares of industries in the total value added of the base year. Thus GDP from the production side was not available at current prices. (KSH 2002) While from the expenditure side the estimation was based on current and constant price estimation from the beginning. This fact evolved a methodological gap between the two

[1] Klára Anwar, statistical counsellor, Hungarian Central Statistical Office, Hungary
[2] NACE is the European classification system of the economic activities by industries (abreviation of Nomenclature générale des Activités économiques dans les Communautés Européennes).

approaches. (Anwar 2008) However, some European recommendations also expected the use of current price approach on both sides. The first results of this approach were published by HCSO in December 2006. (HCSO 2009)

Nevertheless, there are continuously growing demands against the quarterly GDP, like it should be more and more timely and also more and more accurate, comprehensive, and also reasonably detailed, generates continuous research work and improvements.

2. Data sources and methods

Due to obtain value added at current and constant prices both output and intermediate consumption have to be estimated. Its reason is provided by one of the well known National Accounts principle called double deflation. According to it, value added at constant prices can be obtained by deducting the separately deflated intermediate consumption from output. The ESA 1995 therefore defines value added at constant prices "as the difference between output and intermediate consumption at constant prices" (Eurostat 1996). This can be expressed in the following equation:

$$VA = \sum P_O \cdot Q_O - \sum P_{IC} \cdot Q_{IC},$$

where P_O and Q_O are prices and quantities for output and P_{IC} and Q_{IC} are prices and quantities for intermediate consumption.

Similarly to most of the EU countries, no short term data are available for the calculation of intermediate consumption, thus an alternative estimation method should be applied (OECD 1996). According to international methodologies, the application of the intermediate consumption to output ratio from the last available annual accounts is a generally applied method. My purpose is to improve or confirm the quarterly intermediate consumption estimation.

Therefore I analyzed the relationship between the annual intermediate consumption and the annual gross output for non-financial corporations sector at the A6 ESA classification level (Eurostat 1996) at current prices for the period of 1995-2008. I prepared my analysis on current price data due to the chain-linking methodological concept, i.e. the current price of the previous year provide the comparable price for the actual year using annual weighting (Anwar–Szőkéné Boros 2008). As a first step, I chose A6 aggregation level and if a good relationship can be proved at this level than a detailed analysis may be followed.

The A6 classification means 6 groups of branches in the following grouping: agriculture and fishing; industry (mining; quarrying; manufacturing and electricity); construction; trade, repair, hotels, restaurants; transport, storage and communication; financial intermediation and real estate activities; public administration, education,

health, social work and other community, social and personal service activities. The first group of A6 was deducted from the observation due to the fact that the intermediate consumption of agriculture practically remains stable within the year in spite of any changed in its output. While I added the branches, total to the observed series.

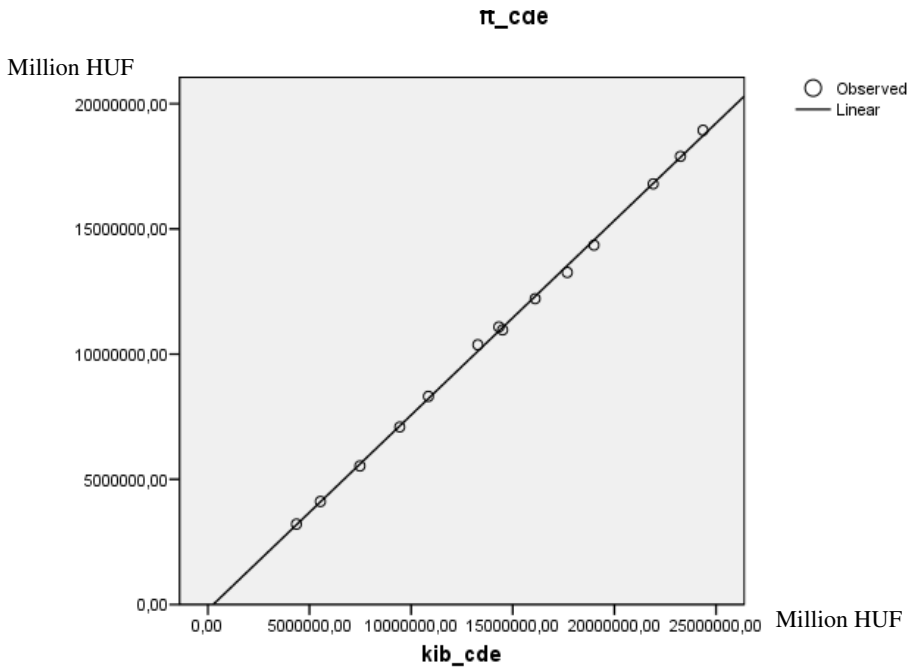
For this study I applied SPSS 15.0 software. I used linear and log linear regressions as well to describe the relationship but similar results were reached. Than, I have performed some time series analysis.

3. Results

Each observed series have 14 elements as HCSO compiled consistent national accounts time series back to 1995. Though more elements would be suitable for any examination, the consistent time series are only available for this period. Each series were examined separately to find out the relationship between the two types of the figures.

In case of the **industry – mining, manufacturing and electricity, gas and water supply** (CDE) R^2 value showed a very strong relationship i.e. variance of the intermediate consumption (IC) can be explained 99.9% by the output, and the rest 0.1% is random error. (See figure 1)

Figure 1. The relationship between output *(kib_cde)* and intermediate consumption *(ft_cde)* for industry



Source: SPPS software results.

The *F* statistic was highly significant, therefore the null hypothesis was rejected according to the coefficients' test to zero value. The results of *t* statistics showed also the same, it was highly significant as well. It was so nice to be true. I checked the Durbin-Watson test. Its value (see table 1) explained everything.

Table 1. **Model Summary[b]**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	1,000[a]	,999	,999	155308,470	,815

a. Predictors: (Constant), kib_cde

b. Dependent Variable: ft_cde

Source: SPPS software results.

Though there is a thumb rule that claims if the result of Durbin-Watson test is under the R^2 value means that there is false regression between the examined

figures, if the number of the observed population is low, the results should be tested by other methods (Darvas 2001).

A possible method to avoid false regression is to set up regression for the differences. This resulted appropriate statistics. Beside the high R^2 value (99.6%) and the significant *F* and *t* statistics, the Durbin-Watson test showed also an acceptable (Hunyadi 2001) result (1.523) beside 1.0097 and 1.3404 critical values, i.e. residuals are linearly independent. The significance level of Kolmogorov - Smirnov test (0.990) showed that the residual distribution is normal.

In case of **construction** (F) the deterministic coefficient was similar to that of the industry, and the *F* and *t* statistics also explained high significance between the 2 figures. The result of the Durbin-Watson test was 1.192 here that fell in the indifferent zone, between the critical values of 1.04495 and 1.35027, i.e. according to this test; no decision could be taken. So I tested the regression for the differences and an acceptable Durbin-Watson figure (1.928) was reached. Furthermore the significance level of Kolmogorov - Smirnov test (0.640) described normal distribution of residuals.

The third observed group of industries was **trade, repair, hotels, restaurants, transport, storage and communication** (GHI). In this case correlation statistics showed that output may explain the intermediate consumption variance by 99.7%. This was justified by significant *F* and *t* statistics. Even the result of Durbin-Watson test (1.812) and the significance of Kolmogorov -Smirnov test (0,832) confirmed this result. Thus the relationship between the two figures of these industries might be expressed by regression.

In case of **financial intermediation and real estate activities** (JK), the deterministic coefficient was significant as well by its extremely high, 99.9%, and it was confirmed by 19 621.69 *F* statistic and by significant *t* statistic value.

The Durbin-Watson test resulted 2.536, i.e. the null hypothesis that residuals do not correlate to each other could be accepted. The significance of Kolmogorov - Smirnov test (0.532) showed that the normal distribution of residuals is granted. Therefore, the relationship between the two figures might be expressed by regression in this case.

The forth case was the group of **public administration, education, health, social work and other community, social and personal service activities** (LMNO) branches. This showed similar results to those of the construction. Though the high deterministic coefficient (99.4%) was justified by the results of *F* test (2 091), that of *t* statistic (see Table 2) and by the significance of Kolmogorov - Smirnov test (0.957). The Durbin-Watson test result fell in the indifferent zone. Thus, the regression for the differences was tested and the new Durbin-Watson test explained (1.967) linearly independent residuals.

The last case was **the industries, total,** i.e. total output and total intermediate consumption. This showed similar facts. R^2 explained significant relationship between intermediate consumption and output that was supported by significant *F*

and *t* statistics values, and also by the significance of Kolmogorov - Smirnov test (0.434), while the Durbin-Watson test value (1.183) fell in the indifferent zone. Therefore the regression for the differences has been tested and acceptable Durbin-Watson figure (2.42) was obtained as according to the critical values this figure should be lower that 2.64973 and higher than 1.35027 i.e. there is no autocorrelation in the model. (Anwar-Ugrósdy 2009)

The results that 3 of the 6 studies were finished by undecisionable Durbin-Watson test induced me for further examination; however regression set up for differences provided acceptable results in these cases.

It can be also stated that the number of the observed population was relatively low; however it was the highest number of population that could be observed at that time as the quarterly GDP series start at 1995.

Even if the regression set up for differences gave acceptable results, this weakens the success of the model. This means that new indicators should be introduced into the analysis, or further examination is needed.

Therefore I applied time series analysis approach to investigate the relationship between the series. (See table 2. for the case of industry.)

Table 2. **Parameter Estimates**

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	,711	,280	2,545	,026
Regression Coefficients	kib_cde	,771	,007	117,296	,000

Melard's algorithm was used for estimation.

Source: SPPS software results.

Due to the fact that the level of the autocorrelation was very high in each case, differences should be calculated for each series to reach stationarity. These transformed series had acceptable autocorrelation level i.e. the application of one difference was enough, and even the significance of Kolmogorov - Smirnov test showed normal distribution of residuals. In spite of the expectation, the parameter of the autoregressive model (at lag one) was not significant in any of the observed series; however other statistics confirmed the significance of these models.

3. Conclusion

According to the above mentioned results, in case if there is not enough direct information available for intermediate consumption, it could be advisable to proceed GDP estimation on the following way: to obtain the intermediate consumption to output average ratio of the last period for a given branch and apply it in the estimation of the current period data.

Though according to the above mentioned results, further examination is also needed from another aspect. It seems a possible approach to filter out the trend factor from both output and intermediate consumption series. Then the relationship between the two types of the obtained series should be analysed again for each group of branches. This may lead to better coverage between the two variables and smaller relative error.

References

Anwar, K. 2008: Current Price Approach of Quarterly GDP Estimations from P roduction Side in Hungary, *Hungarian Statistical Review*, Vol. 86. Special No. 12., 45-57 pp.

Anwar, K. – Szőkéné Boros, Zs. 2008: A láncindexek alkalmazása a nemzeti számlákban, *Statisztikai Szemle*, Budapest, 86. Évf. 7-8., 713-731 pp.

Anwar, K. – Ugrósdy, Gy. 2009: New Aspect of the Methodological Improvement of Quarterly GDP Estimation, *Bulletin of the Szent István University*, Gödöllő (under publication)

Darvas, Zs. 2001: *Beezetés az idősorelemzés fogalmaiba*, jegyzet, Corvinus Egyetem, Budapest, 59-70 pp.

EUROSTAT 1996: *European System of Accounts*, ESA 1995, Luxembourg, 337-342, 426 pp.

EUROSTAT 2002: *Quarterly chaining of National Accounts*, Eurostat, B1/CN 516 e number manuscript, Luxemburg

Hunyadi, L. 2001: *Statisztikai következtetéselmélet közgazdászoknak*, KSH, Budapest, 407-419 pp.

HCSO 2009: *First release, Gross Domestic Product 2009 2nd quarter*, Budapest

KSH 2002: *A negyedéves bruttó hazai termék (GDP) számítási módszere* Magyarországon, Budapest

OECD 1996: *Quarterly National Accounts, Sources and Methods* Used by OECD Member Countries, Paris

The role of accounting in analyzing the company's performance

Đerđi Petkovič[1] *– Livia Rac*[2]

At a company it is unimaginable to have a successful management without available, understandable and up-to-date information. An organization's performance must be measured and analyzed, based on adequate information.

Accounting, as one of the company's information subsystem, has the task to provide the required information to different internal or external stakeholders by financial statements. Analyzing the used methods, six groups of valuation methods can be defined, which try to measure the company's value: balance sheet oriented, income statement oriented, mixed oriented, cash flow oriented, value creation oriented and options oriented methods.

The aim of this research is to find out which performance indicators can be designed using financial statements' items and how we can organize the accounting department in order to meet the managers' information needs. In connection with this, the quality of the information included in financial statements is also an interesting question.

Keywords: performance measure, indicators, financial statement, accounting

1. Introduction

Performance is present in every part of people or company's life. Although past years' and centuries' companies also had their strategic goals and in some way measured the achieved performance, nowadays performance is more emphasized if we talk about the success of management. For successful strategy implementation, we need the evaluation of the company's performance. Strategies vary from company to company depending on the company's size, financial power, possibilities, market conditions and other circumstances. Required performance measures can also differ from company to company. It's all right, but if these measures differ from each other, the question rises, how can we compare different companies' achieve-

[1] Prof. Đerđi Petkovič, PhD, professor, University of Novi Sad, Faculty of Economics Subotica, Serbia, pegy@ef.uns.ac.rs.
[2] Livia Rac, research assistant, University of Novi Sad, Faculty of Economics Subotica, Serbia, liviar@ef.uns.ac.rs.

ment? Applying international accounting standards (IFRS/IAS, GAAP) the differences between financial statements are minimized, but they still exist, as these regulations define only a framework for accounting.

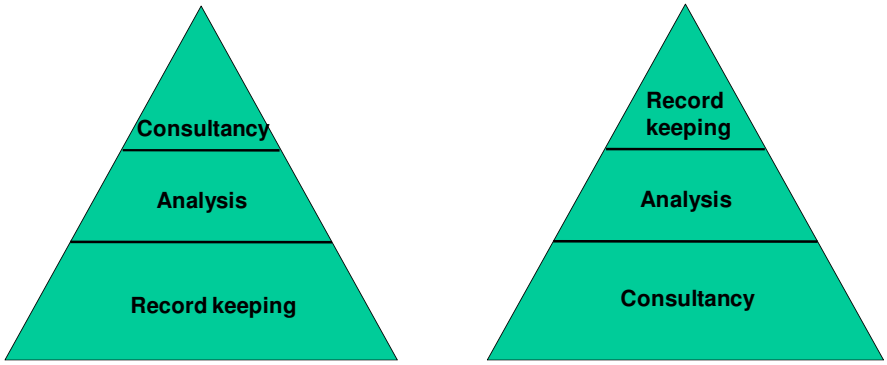
The aim of this research is to find out which performance indicators are suggested to be used by accounting literature. In order to get a practical review about this topic a case study was made at a small and a medium sized company. The aim, defined in abstract, that is to find out which performance indicators can be designed using financial statements' items, is not achieved. This research can be a starting point for other research activities with the aim to create a unique model for evaluating different companies' performance.

2. The role of accounting in analyzing company's performance

The basic part of company's information system is the accounting information subsystem. (Rac 2008) It collects, records, processes, analyzes and keeps data of business transactions, and as a final output, presents information in the form of financial statements or in other reporting forms. Based on accounting documents, which prove that business transactions did happen, every company must keep records, according to the law and international accounting standards.

Nowadays accounting must have a „modern" or "advanced" role, to provide consulting activities to the managers. (Figure 1.) At multinational and large companies this role of accounting is provided by controlling system. At small and medium sized companies this kind of controlling activities are not established. Of course the basic must be a record keeping activity, and then the analyzing activities based on the collected data, but accountants must accept this new role to be open for changes and new techniques in their accounting profession. Record keeping could be automated thanks to information technology's achievement, so accountants could have more time to analyze the information and to think about new possibilities and improvements. The first step of this improving process toward "modern" accounting is to estimate the current situation of accounting processes in the company. At the same time we must investigate the managers and the external stakeholders information needs, in order to find out what information must accounting function prepare, at which level of specification, in which form, how often and for which period of time. According to these requirements, we must create an optimal process of accounting activities by eliminating wastes and focusing only on those activities, which create value for the end-user (following the five lean-thinking principles: value, value stream, flow and pull, empowered people, perfection (Rac et al. 2009)). A good relationship between accounting and other business functions must be created, as we must consider the company, as a whole, where business functions are in active interactions.

Figure 1. Traditional and Advanced accounting



Traditional accounting Advanced accounting

Source: own creation

It's true, that the output of accounting is the financial statement, which is based on financial data, but we cannot forget the importance of non-financial data as well. This non-financial data is significant, because it speaks about the other aspects of production or business processes, not just the financial dimension is considered. If a company wants every employee and maybe partners (customer, supplier) to take an active part in the organization's life, to think about the possible improvement opportunities, the information about the achieved and planned performances must be understandable and accessible for everyone inside and outside the company. Accounting must collect and organize the records in the way to be available at all time, in any form (financial, non-financial) and for every internal or external stakeholder, as it is required, of course according to professional and legal regulations.

Due to globalization there is a big interest to provide comparable financial reports, in order to get new investors and to be transparent in the international level. In order to meet this need, the International Accounting Standards Board defined and is defining International Accounting Standards (IAS 1 – IAS 41) and International Financial Reporting Standards (IFRS 1 – IFRS 8) which are professional accounting regulations with the aim to assure the harmonization of the financial reports at the international level. As don't exist two very same companies, financial standards define only main evaluation techniques, the list of possible methods to calculate the amortization, depreciation, and other accounting problems. Legal regulations also give some directions, how to keep the records in the national level, according to professional regulations. As companies have the freedom to choose between some possible methods, their decisions must be written down, and be publicized together with the financial reports, in the form of Notes to the financial statements.

We must define in our accounting policies, whether certain expenses should be capitalized or expensed. Costs expensed appear on the financial statement as a cost that was incurred that period. On the other hand, costs capitalized are amortized over multiple years. Most expenses are clearly either expensable or capitalisable, but some could be treated either way according to the preference of the organization. Depends on what option is used, the financial result will differ.

In order to provide all the necessary information, we must organize our accounting and controlling department in the way to collect the data of business transactions by cost-pools or by activity-pools, by products and services, by departments, in relation with the target, comparing planned and achieved quantities. We must collect and after that provide financial and non-financial data as well. Our bookkeeping is as much accurate, as more detailed data is recorded. Parallel to this, the data specification costs, so we must always make a cost/benefit analyze. If it is possible, we must try to automate the routine tasks, to achieve the advanced accounting, which role is to give a support and assistance to other organization's department.

3. Valuation methods

It is stated in the previous chapter that companies are different. Due to this fact the organization of accounting function can be also different. As a result the companies' financial statements present the information about the organization's financial position in similar way, but we must be careful by comparing this information, as we have to consider the relevant circumstances and accounting policies.

The authors of the paper have analyzed the accounting literature on performance measures and indicators of successful companies. The results are colorful.

The main literature on valuation methods was presented in 1986, by Rappaport A., in his book Creating shareholder value: The new standard for business performance. (Rappaport 1986). His book was revised and printed in 1998 with the title Creating shareholder value – A guide for managers and investors. (Rappaport 2002). He deals with shareholder values, with the different objectives of shareholders as well as with the measures. It's very important not to focus on the short time objectives, but on the value creation processes. He stated that accounting measures have drawbacks, and analyzers must be aware of these malfunctions. These are: different accounting methods could be used; the change in capital engaged is not taken into account, and the time's financial value is not considered by most of the accounting measures. (Rappaport 2002). In his book different performance measures are described and analyzed. By Rappaport the financial result included in income statement is the poorest, as investments and risks are not considered, includes accounting distortion, doesn't notify the value maximization and doesn't project the value changes. ROI and ROE takes into account the investments, but the risk is not considered, includes distortion and doesn't have the ability to project the value creation.

The residual income and EVA (economic value added) methods focus on investment and risk, but other aspects are also not included. The change of residual income and change of EVA deals with investments and risks, excludes accounting distortion and can show the value maximization. The SVA (shareholder value added) method is the most perfect, as gives positive answer all the required aspects. (Rappaport 2002)

Johnson L. and Soenen L. (Johnson et al. 2003) in their study „Indicators of Successful Companies" identify the factors which are important for companies to achieve the defined performance. By them, financial performance is measured by three methods: Sharpe's ratio, Jensen's alpha and Economic Value Added (hereinafter EVA). The Sharpe's ratio is defined as the rate of return on a particular stock in excess of the risk free rate divided by the standard deviation of the returns on that stock during a certain time period. Jensen's alpha is defined as the realized rate of return on a security, and it should be a linear function of the risk-free rate of return plus a risk premium that is a function of the security's systematic risk, plus a random term. The EVA is calculated as net operating profit after taxes minus weighted average cost of capital, multiplied with the capital employed. The EVA is actually the amount of money, which remains after all providers of capital have been compensated, it is the residual income. In connection to these methods, they defined 10 different indicators which have bigger or smaller influence on the performance. These indicators are: (1) book-to-market ratio, (2) total assets, (3) sustainable growth rate, (4) capital structure, (5) liquidity, (6) cash conversion cycle, (7) earnings volatility, (8) profitability, (9) research and development expenditure and (10) advertising expenditure.

Cho and Pucik (Cho et al. 2005) investigated the relationship between innovativeness, quality, growth, profitability and market value in their study. Their hypothesis is that the higher the quality and the innovativeness are, the higher the performance will be. By them performance is measured in three different ways: growth performance (growth rates of total assets, total revenues and market capitalization), profitability performance (return on assets, return on equity and return on investment) and market value performance (market-to-book ratio and Tobin's q ratio).

Höbarth L. (Höbarth 2006) examined a three dimensional approach to performance measures. He combined different field of a company's success. He takes into account three measures: market performance, profitability and cash flow performance. The market performance is measured by the current market value which refers to the increase or decrease of the price of a listed stock during one period of time. This measure can provide some information about the company's future perspective. The profitability performance is measured by the return on investment ratio and shows how efficiently the invested money of the shareholders was employed. The cash flow performance is calculated as shareholder cash flow, which is a proportion of the sum of all dividends paid out to the shareholders in relation to the total market value. As Johnson and Soenen did, Höbarth also defined potential indicators which influence the performance. His list of indicators are similar to Johnson's list,

the new indicators that are not included on Johnson's list of indicators are: capital expenditure, auditor's opinion, current ratio and quick ratio as forms of liquidity ratios, sales percentage change, EBIT margin, domestic long term issuer credit rating, domestic short term issuer credit rating and common stock rankings. These additional indicators take into account the qualitative aspect of a company's performance.

In the 1960s Edward Altman dealt with the problem of bankruptcy. Investigating different indicators, finally he defined the so called ZETA Model (Z = 1.2A+1.4B+3.3C+0.6D+1.0E), where A refers to the ratio of working capital and total assets, B is equal to retained earnings divided by total assets, C means the ratio earnings before interest and tax divided with total assets, D is market value of equity divided by total liabilities and E refers to the ratio of sales and total assets. The lower „Z", the more likely the company is going bankrupt within two years. (Horn 1997)

In Serbian literature Rodić J. (Rodić et al. 2007) also defines three aspects of measuring the company's performance, as follows: profitability position, property position and financial position. A company's profitability position could be defined by analyzing its financial result, the risk of achieving the result, the profit margin, the financial power and the return on capital invested. All these information can be got from the income statement. Analyzing the structure of assets, the structure of fix assets, the structure of current assets as well as the assets' efficiency (turnover ratio) a company's property situation can be presented. The financial position refers to the ratio of current and quick liquidity, the solvency, indebtedness, the competence of reproduction and the ability to save the real value of the capital. This information could be gotten from the balance sheet.

Another author from Serbia, Stevanović N. (Stevanović et al. 2006) investigating the measurement issues, defined three aspects of measuring performance: ratio of liquidity, ratio of solvency and ratio of profitability. All these ratios have different sub-ratio indicators.

Examining the different studies' results, the authors create six general groups of valuation methods as follows:

1. balance sheet oriented methods,
2. income statement oriented methods,
3. mixed oriented methods,
4. cash flow oriented methods,
5. value creation oriented methods and
6. options oriented methods.

By Balance sheet oriented methods a company's value is estimated based on the value of its assets (assets can be evaluated by book-value, adjusted book value, liquidation value, re-purchase value and substantial value). The basic statement which provide information for this kind of methods is of course, the balance sheet, specialized in order to meet the stakeholders' informational needs. The advantage of

these methods is their simplicity as they do not need special calculations or additional costs to get the needed data. The main drawback is that these methods provide static and historic data; the focus is on the past. For making future decisions balance sheet oriented methods are not significant, but their data can be used as a starting point for further analyzing processes.

Using the information that is included in the income statement dynamic data is gotten, so one of the drawbacks of balance sheet oriented methods is eliminated. The company's value is defined through its earnings, sales, costs, revenues and incomes. This information is not static as refers to a certain period of time, but like balance sheet oriented methods, these methods are also historical, thus the future is not analyzed. Based on the information of income statement positions, some trends could be recognized, but it does not mean the real estimation of future transactions. This would be only a forecast, based on past period's information, not considering the new potential circumstances.

As income statement oriented methods also have drawback like balance sheet oriented methods, the mixed oriented methods try to eliminate these drawbacks by combining the previous two methods. Simple measures exist here such as: return on capital employed, return on investment and earnings per share. These methods are often used by investment analysts. As these methods use a mix of balance sheet's and income statement's positions a complex result could be calculated, but the basic drawback, that is, the information historical characteristic is still presented, so using these methods we could get a real picture about the current situation of the company. As it was mentioned in the previous part of the paper, if different accounting policies are used, we must be careful in comparing the given information between different companies. Different evaluation methods could easily lead to distorted interpretation of the regularly collected and prepared information.

The fourth group of evaluation methods doesn't focus on the past, but emphasizes the future. This group can be named as cash flow oriented methods. They estimate the future cash flow and discount them at a discount rate, equivalent to the specific flows' risk. These methods show the net present value of all the future cash flows. The focus is on the future, the company's value is evaluated based on future assumptions. The main concept is that the value of a company is defined on how much cash flow is generated in the future and how much the present value of these cash flows is. These methods' drawback lies in the task to define the appropriate discount rate. Determining the discount rate forecasts must be used. Our aim must be to provide as objective information as it is possible, so we must be very careful considering all the relevant future circumstances by determining the discount rate.

Value creation oriented methods are the fifth group of valuation methods, defined by the authors of the paper. The main method which is placed to this group is the EVA, which is a new and little bit complicated method. The economic value added is the difference between capital employed and net operating profit multiplied with the weighted average cost of capital. It refers to the amount that remains after

the company's shareholders and all other providers of capital have been compensated. This new concept says, it's not enough to create a profit, but profit has to exceed a certain level to create value. Other value creation oriented methods are: market value added and total shareholder return.

The newest and most complicated methods are the option pricing methods. They are very complex, thus rarely used in practice. Because of this reason authors put them into a separate group. Option pricing methods try to estimate the value of a company by valuing future events, which will occur with a certain probability. These methods are similar to the discounted cash flow methods, but more complex, as not use only the discount rate.

Following this categorization, the former described methods from accounting literature on performance measures can be also put in these groups. Methods described by Johnson R and Soenen L. (Johnson et al. 2003) are value creation oriented methods. Cho and Pucik (Cho et al. 2005), Edward Altman (Horn 1997), Rodić J. (Rodić et al. 2007) as well as Stevanović N. (Stevanović et.al. 2006) define mixed oriented methods to measure a company's performance from different aspects. They combine balance sheet's and income statement's data. Höbarth L. (Höbarth 2006) combines three aspects of company's success, the market performance, the profitability and the cash flow performance. His measures could be categorized as mixed oriented and one measure is cash flow oriented.

4. The importance to evaluate a company

As performance is present at everyday life, the need to evaluate a company is also presented at every transaction. In earlier parts of the paper different performance evaluation methods were described and categorized based on accounting literature. To successful strategy implementation, different objectives must be achieved. A manager or shareholder could control the strategy implementation process by using different financial and non-financial performance measures.

The evaluation process is not important at strategy level only, but at operation level too. Beside for strategy creation, control and correction, the company's evaluation process could be important in other fields of activities also. One of the basic tasks is to define the goods' or services' price. We buy and/or sell goods and services at certain price. This price mostly depends on the demand and offer of the goods, but the country's economic-political situation is also influencing the level of price. Beside these factors, the price of goods depends on the company's value too, which actually refers to its goodwill, brand or market image. In everyday transactions customer always defines the highest price to pay, and the vendor the lowest price at which he is prepared to sell. Defining these lowest and highest prices they take into account all the available information in connection with the prices and

goods or services. To define the best price, top or middle managers must know the value of the company.

If we speak about a listed company, the importance of a company's valuation is more evident. A company's value must be known in order to compare the share's price on the stock market with the aim to make the best decision whether to sell, buy or hold the shares. The information for the public about these listed companies' operations is given by share's price. Potential investors would like to know the company's potential value creation power in the future, while for banks it is very important to know the company's value while accord a credit.

Different stakeholders require different information for their decision making. As financial statements for external stakeholders have fix, by international accounting standards and legal regulations defined form, external stakeholders have to combine the data in different ways according to their needs, but keeping in mind the potential differences in valuation the assets, obligations, capital, revenues and expenditures.

For internal stakeholders the form of financial reports is not defined, managers can create unique reports. Managers must know the organization's value to be able to create a real and useful strategic plan for a longer period of time. Based on strategic planning, a short-period decisions are must be made in order to define exactly what has to be done. It is essential to define clearly the strategic goals, as all other activities are derived from them. According to strategy, different reports could be prepared. Managers must evaluate the past performance regularly to analyze if all the strategic aims are realized. Of course if there is a difference between the planned and realized values, further analyzes are required. Information about the past and potential performances is important for the managers because of subjective reasons also. At almost every organization a managers' bonus depends on value that is created by them. This system of bonus can create problems in evaluation, as often managers at the end of the year distort some data in order to have nicer result. As company operates on "going concern" principle, first or last these corrections come to light (like well-known accounting scandals).

For different objectives, derived from strategy, different performance valuation methods, measures and indicators are useful. Every company must find a unique set of indicators which the best way meet its needs. But if we would like to compare and analyze different companies' performance, we must establish a unique evaluation model, which applied to different companies, would give a comparable result. In order to be able to set this kind of model, beside the accounting literature, companies performance measures have to be analyzed, taking into account all the circumstances which influence the company's activity.

5. Practical issues

As it was described before, accounting literature has many different issues related to the evaluation of the company's performance. The question is, whether these methods are used in Serbia by small and medium sized companies. A case study is made on this topic, involving a small and a medium company at Vojvodina.
The examined small company is a trading company; its profile is technical accessories for houses. By its parameters, according to the Law of accounting and auditing in Serbia, it is a small company (Zakon o računovodstvu i reviziji 2006). The average number of employees is 32, yearly revenues are 1.750.000 Eur, and total assets are 1.050.000 Eur. It was established in 1992, and nowadays has 5 stores in four cities in North Vojvodina. At the company the book-keeping as well as the analyzing function is being done by the same accountant. The top manager has no much knowledge about accounting, so accountant must prepare simple reports about the company's past performance. They use income statement oriented method, as the report includes only revenues and expenditures. These categories are grouped only by main groups of revenues and expenditures and shown by months. Beside this report monthly by profit centers are shown purchased and sold stock and in connection with this calculated and realized price margin. There are not any indicators as for liquidity, profitability, cash flow, some future trends, or others.

The medium sized company was established in 1993 as company for production, trade, transport, export and import. Its basic profile is production of two series of bicycles, low cost and high quality bicycles. As this is a seasonal product, company trades with fitness gyms and skies during winter time. According to Law of accounting and auditing, company is medium sized, as the average number of employees is 125, yearly revenues are 5.700.000 Eur, and total assets are 4.600.000 Eur. At the examined medium sized company the analyzing process is more serious then at the small sized company is. There exists an accounting department with employee to make analyzes, plans as well as to correct the plans. They also emphasize the positions of income statement, but use some data from the balance sheet as well. We can conclude that the studied medium sized company uses mixed oriented method. Their reports are divided into two parts: performance measures at company level and performance measures at level of profit centers. Of course all the information is considered monthly. At company level they analyze the planned and realized revenues, all by goods or products and by the criteria if domestic or foreign, the planned and realized expenditures by types of costs, the operation result after costs of material and goods sold, the operation result after all expenditures and the financial result before interest and after interest. At company level the information focuses on the expenditures and revenues as well as on the financial result.

At the level of profit centers the average monthly costs of the store are calculated, as well as the cost structure in percentages, the revenues by goods or products,

the realized price margin as a percent of total revenues, the net financial result of the center, the turning point of the business and the stock turnover ratio.

The difference between the two studied companies is in the information depth, which influences the power of analysis. As they are not listed companies with shares, they couldn't use performance measures like earnings per share and similar measures. Considering the past accounting activities of these companies, we can say, that they are improving its accounting function by focusing on analyzing process. This must be a never-ending process with continuous improvement. Based on theoretical issues and case study the further step of this research is to create a questionnaire for small and medium sized companies.

6. Conclusion

Every company is interested in creating value for customers and at the same time achieving its strategic goals. To control, if the defined aims are achieved or not, managers must measure the company's past performance, not forgetting to project the future circumstances. In globalization managers must be prepared to challenges and is desirable to react proactively. All this is possible if they use accurate, up-to-date and useful information.

Accounting could play main role in preparing such required information, if this function is well organized and the accountants are well educated. A continuous improvement as well as a continuous education for accountants is needed, always trying to find out a more efficient and effective methods. Book keeping must collect and organize the information by cost pools, activity cost pools, profit centers, cost centers, in order to be able to give important and useful information for making successful decisions. In accounting literature many studies were made regarding the issue of performance measures and the indicators of success. We have to analyze these thesis, the context and assumptions of these studies, as we have to implement the ideas in our specific circumstances. For different purposes of the company, different performance measures are appropriate. The most widely used measures are: profitability measures, revenues and expenditures structure, ROA, ROE, market-to-book ratio, and financial and liquidity ratios. As factors of success are almost at every company different, if we would like to compare the companies, we must create a relative measure or some relative measures, combining performance indicators with the company's size, age, industry and other key factors.

The evaluation must be done in order to find out which value drivers are creating value for customer and which presents only wastes. Of course our attention must focus on eliminating the wastes.

There doesn't exist a perfect performance measure without disadvantages. We must be aware of these disadvantages and try to combine these performance categories in order to finally get the most suitable performance measure model.

References

Cho, H.J. – Pucik, V. 2005: Relationship between Innovativeness, Quality, Growth, Profitability and Market Value, *Strategic Management Journal,* 26, 6, 555 – 575 p.

Horn, J. 1997: *Finansijsko upravljanje i politika*, MATE, Zagreb.

Höbarth, L. 2006: Modeling the Relationship between Financial Indicators and Company Performance – An Empirical Study for US-listed Companies, *PhD Dissertation*, Wirtschafts Universitat, Wien.

Johnson, L. – Soenen, L. 2003: Indicators of Successful Companies, *European Management Journal*, 21, 3, 364 – 369. p.

Rac, L. 2008: Accounting Information Subsystem for Logistics Managing in the "DDDK" Company*, Strategic management*, 13, 4, 48 – 51. p.

Rac, L. – Petkovič, Đ. – Leković, B. 2009: Performance measurement of lean systems – challenge for accounting profession**,** *Strategic management,*14, in print.

Rappaport, A. 1986: *Creating shareholder value: The new standard for business performance*, Free Press, New York.

Rappaport, A. 1998*: Creating shareholder value – A guide for managers and investors*, Free Press, New York.

Rappaport, A. 2002: *A Tulajdonosi érték – Útmutató vállalatvezetőknek és befektetőknek*, Alinea Kiadó, Budapest.

Rodić, J. – Vukelić, G. – Andrić, M. 2007.: *Teorija, politika i analiza bilansa*, Poljoprivredni fakultet, Beograd.

Stevanović, N. – Malinić, D. – Milićević, V. 2006: *Upravljačko računovodstv*o, Ekonomski fakultet, Beograd.

Zakon o računovodstvu i reviziji, Službeni glasnik RS, 46/2006., Belgrade.

A methodology of assessment based on the interaction between practice-based paths and standardized enquiring instruments

Giuseppe Scaratti[1] - Silvio Ripamonti[2] - Alessandro De Carlo[3]

The object of this paper is to describe an intervention conducted using specific assessment tools created after an analysis of work practices and standardized instruments.

The experience was carried out at a company managing the yard of an important commercial port in Italy, and in this paper the authors also present the reasons underlying the intervention request and a brief explanation of the specific context in which this company operates.

After a description of the steps taken to enquire the work practices on the field, the article continues explaining how, starting from the results of the study about the practices, the authors created specific assessment techniques tailored for this particular situations. Such techniques consist of focus groups and role playing, expressly adapted to the context. A tool for evaluation, namely a grid containing different indicators was also created on the basis of the company needs. At the same time, the authors chose a set of standardized instruments with the aim of using them for integrating the qualitative results. These instruments were the M.P.P. test (Multidimensional Personality Profile) and some Organizational Citizenship scales. The set of the instruments was integrated to provide the company with a broader and more exhaustive set of information, in which the two types of tools gave a different contribution.

The complexity of the problems at stake makes the choice of a qualitative approach necessary. As it is known, anyway, this is unique and non repeatable. For this reason it is important to make it sounder by adding objective data, in order to gain new and strong elements for sustaining the argumentations and making the conclusions usable and controllable by the company and the scientific community.

Keywords: Work Practices, Qualitative Methodology, Quantitative Methodology, Methodology Interaction, Assessment

[1] Giuseppe Scaratti, *Catholic University of Milan, Italy*
[2] Silvio Ripamonti, *Catholic University of Milan, Italy*
[3] Alessandro De Carlo, *Catholic University of Milan, Italy*

1. Objectives of the paper

The objective of this paper is to show how instruments derived from a qualitative analysis were used together with statistically based instruments for creating an integrated system of assessment and feedback for organizations (Thornton & Rupp 2005). This operation adds the advantages of the usage of instruments that can give a standardized feedback (Eurich et al. 2009) both for the researchers during the process of assessment and for the company, which is also allowed to know where it stands in comparison with a larger sample of data, to the depth of qualitative instruments specifically crafted for the committing company (St-Onge et al. 2009). In this way no aspect of the process is neglected, creating value for the research, and there is also the possibility of a greater control of the processes by the researchers.

For doing this, an experience carried out in a real working environment is presented. During that, which was an assessment activity for a big company, these techniques were used.

2. The context

This article has its focus on the assessment experience carried out by a research group in the year 2008. Hence, considering the specific tools used for this intervention, it is important to specify the context in which it has been done.

The intervention which we shall discuss here was carried out at a multinational company which manages the yard of an important commercial port in southern Italy. The company's operations are focused on transshipment, that is, to deal with the movement of containers from long-distance ships onto smaller ships which have the possibility of docking in the smaller ports located in Italy and the Southern Mediterranean. The port serves the entire Mediterranean area and is considered a high quality facility not only for the local area, but for all of Italy. In fact, due to the number of containers handled, it is among the leading ports in the world in terms of goods handled.

To understand the functioning of the port, it is important to refer to the organizational structure and how the organizational chart is set up. A total of approximately 1,000 people work at the port. 800 of those individuals are yard employees with the task of dealing with the loading and unloading of arriving ships. The workers are organized into teams of about 10 people each, coordinated by a Team Supervisor. There are 7/8 ships present at the port's wharf at any given time, on which an equal number of teams work continuously, 24 hours a day. The work day is organized into three shifts, each of which has a Shift Foreman (SF) who coordinates the allocation of the teams working on the ships based on urgency and unforeseen events. The teams are then grouped into three large units, each of which is coordinated by an Operational Unit Supervisor (OUS). The Shift Foreman and the

Operational Unit Supervisor are in radio contact with the Port coordination center, which controls the operations through the monitors, keeping watch over the operations in the yard and the movement of the containers. From an organizational standpoint, this is a critical area of interaction. In fact, the operator at the monitor (Controller) has an important coordinating function because he has to decide how the teams work, on which ships, and with what priority. It should be stressed that all of the figures described above are constantly in contact with each other by two-way radio in order to exchange information and data on the operations which are being performed.
This method of work requires an efficient coordinating function which must be capable of allocating the human resources based on urgent needs, unforeseen events, and the importance of different clients.

A critical aspect is represented by the fact that, despite being monitored as regards their loading and unloading operations, the yard workers have a certain amount of autonomy which effectively allows them to influence the flow of containers based on how they manage routine operations and how they deal with the unforeseen situations which periodically arise.

2.1. The request for an intervention

The port is located in a very particular socio-cultural context, where an industrial culture, which could help the personnel to identify with the company's mission, is lacking.

The owners of the company decided to adopt a dual, challenging objective: that of monitoring and limiting the presence of organized crime, and guaranteeing that the majority of the port workers are from the local area, in order to make a contribution to the economic and social recovery of the area itself. However, these choices make it essential to pay particular attention to the care and development of the personnel, who in many cases, have not internalized an industrial-type working culture. The training of the personnel is thus an area which requires special attention at the company.

Following particularly critical events, which indicate an ambivalent involvement and merely utilitarian attitude on the part of the yard personnel (the laborer level) towards the company, the managers asked themselves how they could sustain positive engagement towards the work on the part of the company's employees. Behavior denoting a lack of participation in the company is often noted, indicated by a high level of absenteeism, lack of coordination between the work teams, and dependence on the orders coming from the hierarchy, without taking responsibility for dealing with critical situations which arise in the yard.

The company management has thus promoted an assessment process with the goal of launching a promotion program for middle management in order to enhance the roles of individuals with high potential. The idea is to then rely upon those individuals to sustain a positive organization of the work teams and a good level of engagement on the part of the laborers.

The request for intervention from the Port Management was driven by the intention to succeed in determining which skills are critical for the Team Supervisors who manage the work teams in the yard. The managers' initial intention was to assess those skills based on criteria of the workers' compliance with general skill levels which are normally measured by assessment programs, for which many different protocols and measurement methods have been developed by now. Once that initial request was made, the group at the Università Cattolica began a process of working together with the managers in order to reformulate the request. A shift was made, from the request to use only standardized tools to assess skills or personality traits, to the proposal to develop an assessment process based on the measurement of skills actually deployed by the workers in their daily operations in addition to the more traditional quantitative methodologies.

The advantages regarding depth, precision and explanatory capabilities of an assessment based on both qualitative and quantitative methods was discussed and accepted by the management despite the problem of a greater expense of time due to the necessity of creating the qualitative instruments after an extensive work of ethnolmetodological observation (Brummel et al. 2009; Krause et al. 2006).

3. The qualitative instruments

The first step taken for creating the qualitative instruments was the observation of work practices. These, as defined by Gherardi, are: "*a mode, relatively stable in time and socially recognized, of ordering heterogeneous items into a coherent set"* (Gherardi 2006). These practices influence and give rise to a specific mode of experiencing and constructing the organization through a process of continuous sense making and interpretation of what happens (Weick 1969, 1995). The working hypothesis adopted is that in real operational situations, it is possible to intercept the (more or less functional) concrete modes through which people develop responses to problems they must face, exchange and circulate knowledge and learning, and share and develop practices and recurrent systems of action with respect to the work needs with which they are faced.

The survey of the work practices was made possible by ethnographic observation work in the field, which allowed for taking note of the work processes, critical points, and real difficulties encountered by the men in the yard. The researchers negotiated with the port managers to devise an ethnographic observation plan which allowed for accompanying the yard teams in all phases of their work, on

each shift. The observations lasted for 14 days, divided into two sessions lasting one week each. For each day, each researcher was assigned to a team, and accompanied that team as it performed its work. The researchers also negotiated permission to take notes in the presence of the workers.

The observation process was guided by certain specific questions which, after having been discussed with the management, represented the viewpoints which the researchers adopted to direct their observations when they were involved in the work situation in the yard. The questions which guided the observation were aiming to assess how and through what processes does the passage from the inputs to the implementation of the company plans take place, what is the proximity/distance between the representation of the anticipated profiles/skills and their translation into practice in the system of action, what methods of promotion and management of knowledge in use are adopted and supported, what points of strength and weakness seem to characterize the organizational situation observed.

The review of the ethnographic forms was a process which initially involved the researchers, who attempted to identify macro-areas in order to understand the material that had been gathered. Recording the information on ethnographic forms allowed for an initial identification of certain critical areas with respect to which the personnel in the yard appeared to develop differentiated work practices, showing different levels of skills in managing unforeseen situations in daily work. This process consisted in the detailed analysis of the data gathered, with the purpose of registering the episodes which, in order to be dealt with, required the ability to activate a process of sense making by the groups of workers. The researchers then identified different types of critical episodes dealt with, which in order to be overcome required the activation of different areas of skills. The outcome of the first re-elaboration of the ethnographic forms was the preparation of a report for the Company Management which was discussed with the managers in order to decide which skills and abilities underlie the modes of work detected. With the managers' contribution, we reconstructed the data which emerged and identified which skills and practices were to be considered relevant and appropriate for monitoring during assessment, based on the company's goals and the management of the work.

Based on the results of this process of identifying skills, which was performed in close collaboration with the Management, specific assessment tools were devised which presented the workers with recurrent work situations and critical situations resulting from the real organization of the work at the port (Table 1).

Table 1: critical situations at the port

Areas and Skills	
Human resources management area	Management of work team groups Promotion of personnel motivation Management of informal working networks which arise based on immediate urgent needs
Knowledge Communication and Management Area	Competent exchange and knowledge circulation Communication styles
Commitment Area	Promotion of commitment to the company Support of commitment
Working Style and Ability Area	Management of priorities and urgent needs Technical knowledge Problem setting/problem solving Safety management

Source: own creation

The design of the setting and assessment tools took place starting with the analysis of the work practices made visible by observation in the field. The work on the practices detected saw active cooperation between the group of researchers and the clients, who participating in identifying a grid in which the dimensions to be assessed were listed by order of importance. The specific nature of this step is that the identification of each skill took place following an analysis of the practices used in the port yard, and that "context of use" for each of them was made explicit. The forms were analyzed, and the study which was performed jointly by the researchers and managers allowed for identifying the recurrent work practices associated with successful work in the yard. Subsequently, the skills necessary for effectively managing the work practices that are considered strategic were identified.

The process of monitoring the skills took place by presenting the individuals under assessment with some of the work situations and practices in use, detected during the phase of observation in the field.

The tools prepared for the assessment were of three types, as indicated below.

3.1. First type of assessment tool: Group discussion starting with forms which depict recurrent problems to be managed.

This type of assessment situation allowed for highlighting the mental schemes activated with respect to problematic situations, which the individuals under assessment were asked to comment on and discuss as a group. For example, in the situation reported in sequence 1. …, the communication process is confused because the Manager calls the Team Supervisor directly due to a slowdown of the lifts. The latter has not been notified by the Controller. We are describing a process, which occurs very often, where the establishment of the temporary decision-making unit cannot take place because a clear exchange of information between the actors is lacking; the manager intervenes, but without activating the Controller. Thus, what prevails is interaction between individuals who remain isolated, and the decision-making unit described above is not established. A communication process is created which does not activate the network consisting of the Shift Foreman and Controller, who each have a viewpoint on the organizational dynamic which is partial, but essential in order to develop an overview of the problem. The middle manager and Team Supervisor alone do not possess the wealth of information necessary to manage the episode in the best manner possible.

Furthermore, in the episode chosen in the box, there is a reference to a problem of managing the men in the yard. It often happens that slowdowns of a technical nature are taken as an opportunity to further slow down production and not work up to one's own abilities. The referenced event can serve to attempt to manage this type of recurrent problem in the simulation. Thus, in this situation, the Team Supervisor must indicate a type of thinking which allows for defining the situation as stated above, while having certain hypotheses in mind as to how to manage his own work group.

Sequence 1

Loading operations are underway involving a strategic ship for a potential new client. The Operational Unit Supervisors have directed that close attention be paid to the process in order to avoid any type of problems or delays. It is known, however, that the yard is full of containers, and this could slow the process down considerably. Thus, it is necessary for everyone to cooperate in order for the shift to perform well.

Shortly after the beginning of the shift, while doing a check of the cranes on the terminal monitor, the Shift Foreman calls the Team Supervisor to check on why certain lift trucks have long work times.

Shift Foreman: “Hey Giacomo, check on lift truck SC 17. The Dispatcher told me that its times are too long. We're at an average of 4.5, which is a bit too low. What's going on?

Team Supervisor: “OK, I'll check on it now.

Team Supervisor (calls the lift operator): “SC 17, what's going on? Your times are too long; is there a problem?

Lift operator: “Yeah, I'm picking up from position xxx, which is far away, you know. The problem was known. The yard is full and we have to do pick-ups in the rear sections.

Team Supervisor: “OK, got it, I'll inform the Dispatcher and Shift Foreman

Team Supervisor to the Dispatcher : “Hi, how come you didn't inform me of the delay with SC 17? The Shift Foreman called to tell me. Anyway, it's because the route is slow.”

The Shift Foreman again contacts the Team Supervisor. Shift Foreman: “can't we speed this up a bit?”

Team Supervisor: “we'll check now, we'll try, but you know how the situation is,

don't you?" Shift Foreman: "OK, OK." The environment is quickly "heating up" during the afternoon. Due to the request to increase the loading speed, which is slower than expected, the operators began to get nervous due to the tension. The Team Supervisor realizes that the far-away position of the containers is slowing down the loading times, but that this is being used as an excuse to justify a certain laxity by the lift operators, who are taking it "easy."
Delivery for discussion. At this point, how can the CO manage this situation?

3.2. Second type of assessment tool: Role Playing

The second type of tool proposed for the assessment is the use of role playing, where people are asked to identify themselves with a professional role and dialogue with their colleagues based on a specific situation which is presented. The role playing scenario was designed based on the analysis of the ethnographic forms, and points out critical situations for which the individuals under assessment are asked to explain, in a group discussion, how they would react if they were involved in such a situation.

In this case, based on a situation outlined starting with the results which emerged from the observations, they are asked to participate in a discussion in which they interpret the role of the Team Supervisor with given tasks.

In the case presented here, the problem is represented by the lack of motivation of the men in the yard, and the use of this role playing scenario clearly

identifies the hypotheses regarding the methods the Team Supervisors can use to motivate the workers. During the assessment, a meeting between the Shift Foremen is simulated, who comment on the situation described.

The Team Supervisor realizes that he is asking a lot of his team, and he also knows that the problems which are occurring have various different causes. There are slowdowns due to the full yard, technical malfunctions, and things aren't working the way they should. He also knows that his group is following procedures properly, but without doing its best. If only the lift operators were to put a bit more grit into it and the crane operators worked harder, the limitations and technical problems could be overcome or at least held in check, allowing the group to maintain a good level of productivity. If the team's mood is good, then it is possible to lessen the impact of the delays due to technical problems.

He decides to intervene with his workers at the end of the shift, to try to motivate them a bit….

Delivery for role playing.

You are a group of COs. Try to imagine what you would do in this situation. Decide on a common strategy to adopt, as COs, identifying and describing four potential actions, in order of priority.

You have 45 minutes.

3.3. Third type of assessment tool: Questionnaire for monitoring recurrent episodes

A final assessment tool, devised using the analysis of the practices, was the questionnaire for monitoring recurrent episodes. This tool was devised starting with the construction of boxes which depict the most common recurrent situations monitored by the research group. These situations were summarized and presented to the workers in the form of a multiple choice questionnaire. The possible answers were developed based on direct observation of behavior in the field. Each alternative

represents an actual form of behavior detected during observation. This tool thus requires the worker to put himself in a certain situation and identify the behavior considered most proper based on one's own knowledge.

One example of this tool is presented below.

There are slowdowns and delays in the loading and unloading operations in the yard (lines of lifts at the cranes, different needs in terms of loading speeds, critical situations,) which are affecting productivity. The CO implements certain solutions and:

Uses his visual control of the yard and suggests moving some lifts from one crane to another, because he knows that they will not be used to capacity.

Leaves it to the Operational Unit Supervisor to intervene and manage the problem: it's his responsibility.

Knows that the Controller is overloaded and therefore does not communicate the problem to him.

Having to manage multiple ships at once, he doesn't have time to deal with these things!!

He concentrates on providing support to the lift and crane operators who have to accelerate their work for certain ships, knowing that a special effort is required of them in these situations.

You can't take care of everything! With the yard full there are many moves and there just isn't time for communications which slow things down…

4. The Quantitative Methods

As it has been said through this article, in addition to qualitative methodologies and instruments, a group of quantitative instruments has been also used for making the assessment more broad and descriptive.

The tools that have been implemented for this assessment have been chosen for their capability of observing variables relevant for the working activity. Another characteristic of the questionnaires was the fact that they would measure different traits, attitudes and behaviors then the qualitative instruments crafted for this intervention, giving a real edge to the assessors in observing a much wider set of variables.

Specifically, the instrument used for this assessment have been the Multidimensional Personality Profile (Caprara et al. 2006), a scale for measuring commitment and one for measuring citizenship behaviors.

The Multidimensional Personality Profile (MPP) is a personality inventory specifically created for assessing traits that are relevant for working. This makes it a useful tool for any work related assessment. Furthermore, the dimensions it measures are different than the ones assessed with the observation-based instruments, or, in rare cases, just partially overlapping. This has been an important factor that led to the choice of it as the main quantitative assessment tool.
The test measures five fundamental areas of personality:

- Agenticity – the attitude of a person to be proactive during in different situations, to find solutions for problems and to have the willingness to change.
- Socio-emotive intelligence – the ability of a person to understand others, to feel empathy and to be able to show emotions
- Self-regulation – the ability of a person to be precise and reliable in accomplishing tasks, of being punctual and of managing his/her own emotions, feelings and desires.
- Ability to face critical situations – the ability of a person to think in an efficient way when faced with a difficult situation.
- Innovation – the ability of a person to find new solutions and of thinking in a different way, also the willingness to do it.

This test provided some insights of the deeper personality traits of the workers that would have been difficult to assess with the work related tasks discussed above. Particularly, the MPP test has been chosen also because its dimensions are substantially different than the ones measured with the qualitative instruments. In this perspective, this assessment shows a methodology in which the two types of instruments should not have correlation, as they are meant to observe a vast range of characteristics instead of proving with quantitative instruments what was found with qualitative ones.

Other two scales have been used for this intervention: a commitment scale (Meyer & Allen 1991, 1997) and a citizenship behavior (Organ 1988, Organ et al. 2005) one. The first scale measures the duty of the workers toward the employee's psychological attachment to the organization. All the three dimensions of organizational commitment were included in the scale:

- Affective Commitment: the employee's positive emotional attachment to the organization. An employee who is affectively committed strongly identifies with the goals of the organization and desires to remain a part of the organization.
- Continuance Commitment: The commitment due to the perception of the high costs of losing organizational membership, including economic costs and social costs
- Normative Commitment: The commitment due to of feelings of obligation.

The organizational citizenship behavior scale measures a special type of work behavior that are defined as individual behaviors that are beneficial to the organization and are discretionary, not directly or explicitly recognized by the formal reward system (Smith et al. 1983).

The quantitative instruments have been administered to the workers during the assessment in small groups. In fact everyday they were split in three groups doing different activities, one of those being completing the MPP test and the scales. The amount of time required for doing that was about 1 hour and 30 minutes.

5. Putting the instruments together

Having data obtained with these two types of instrument, the problem was to give a useful feedback to the organization.

For doing that a grid was created. This instrument allows a broad view of the results of all tasks completed by the assessed workers.

The scores have been given on a scale from 1 to 5 (the scores of the MPP have been transposed within these values too) and weights have been used for giving a greater importance to the qualitative part. Specifically the single dimensions measured with qualitative instruments were weighted following the points of greater interest of the company and the general score of the qualitative part was added to the quantitative for giving the total score.

An example of the grid presented to the management is presented above, weights are recognizable, as well as the single and total scores (Figure 1).

Figure 1. Assessment grid

Gioia Tauro - March 2008
ASSESSMENT CENTRE - UNIVERSITA' CATTOLICA DEL SACRO CUORE DI MILANO

Name	Id	Assessors	Role

Minimum score = 10
Maximum score = 50

SCORES								
Dimensioni	W.	Case 1		Case 2		Role playing		Pondered average
		Score	Pond.	Score	Pond.	Score	Pond.	
Group management	0,5		0,0		0,0		0,0	0,0
Leadership	0,5		0,0		0,0		0,0	0,0
Motivation promotion	1,5		0,0		0,0		0,0	0,0
Conflict managementi	1,5		0,0		0,0		0,0	0,0
Communication	0,5		0,0		0,0		0,0	0,0
Affiliation	1,5		0,0		0,0		0,0	0,0
Promotion of positive organizational cultures	0,5		0,0		0,0		0,0	0,0
Priorities management	1,0		0,0		0,0		0,0	0,0
Problem setting and problem solving	1,0		0,0		0,0		0,0	0,0
Safety	1,5	Focus	Score		Pondered			0,0
								0,0

TEST SCORES		
Dimensioni	Test	Score
Agenticity	MPP	
Socio-emotional intelligence		
Self-regulation		
Coping ability		
Innovation		
Committment	COM	
Citizenship	CIT	
		0,0

TOTAL SCORES		BALANCED
ASSESSMENT SCORES	0,0	0,00
TEST SCORES	0,0	0,00
	0,0	0,00

WEAKNESSES	STRENGHTS

Source: own creation

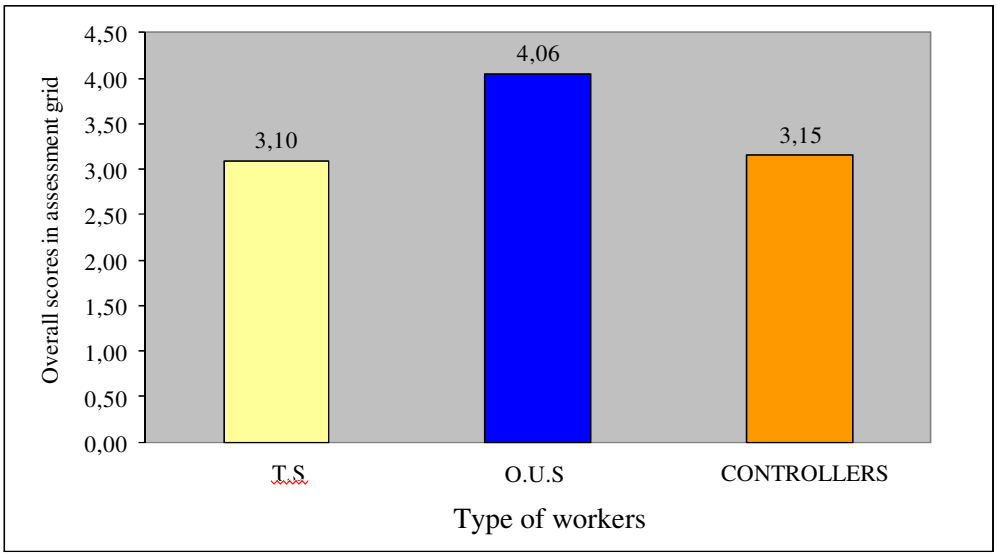
It is important to notice how an area for individual judgments by the assessors has been maintained. This allows the experience of the professionals to add more value to the results. In fact, all the tasks, both qualitative and quantitative, have been conducted with the supervision of at least one assessor, two for the dialogical tasks. This granted the reliability of the assessment and it allowed the professionals to formulate specific notes on every participant.

6. Results

Some results, being in line with the expectation of the company management are used as a tool of validation of the assessment, specifically, the company appreciated these correspondences.

The first fact is that the OUS personnel, already selected by the management from the port crew, scored higher points overall (Figure 2)

Figure 2. Overall Assessment Scores



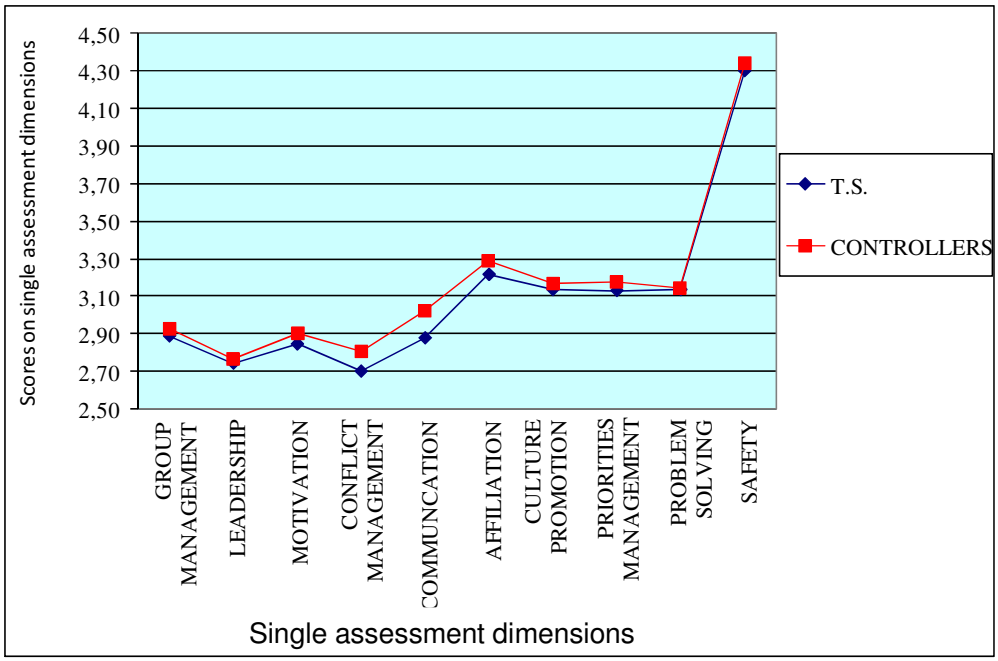
Source: own creation

This is in line with the expectations. The assessment allowed understanding which specific OUSs had outstanding qualities and might be ready for career advancement.

Another result that overall follows the expected path is the difference between dock workers and controllers in the abilities in which controllers are supposed to be stronger (Figure 3)

As it is possible to note, the controllers show higher abilities in communication and conflicts management, and this was expected. Instead, motivation and commitment have been discovered to be higher in the controllers groups with this assessment. Anyway the focus of this intervention was to rate the single workers, and that what has been done. Aggregate charts like these have been taken into account mostly for checking the consistency with expectation and data that the company already had.

Figure 3. Scores on single assessment dimensions



Source: own creation

The feedback from the organization has been very positive. They used the results of the assessment for adjusting their policies in human resources management, and this is still ongoing. They implemented a protocol for continuous assessment based on the findings of the research conducted for creating the instruments discussed in this paper. The organization is now working with the Università Cattolica for implementing a follow up of the intervention, which aims to be even more targeted on the specific environment and practices observed during the first contact.

7. Conclusions

This experience shows how it is possible to put qualitative and quantitative techniques together for adding value to an assessment process and how it is possible to use the results together for creating general and specific rankings that can be useful for the companies for acknowledging their strengths and weaknesses and for choosing courses of action.

The quantitative part, being corroborated by a vast database, gives also more strength to the company in its eventual actions toward the single workers and the groups that might follow the assessment. This is a fact that must been taken into consideration, especially in realities in which trade unions are strong.

On the other hand, the qualitative part was not only important for its assessment value, it was useful for the managers in order to become more familiar with the dynamics of the yard. The reflection on the ethnographic forms produced an outcome in terms of very in-depth knowledge of the real work processes and led to considerations with respect to new methods of organizing the work of the teams.

Furthermore there is the advantage for the researchers to have objective data that are useful for the scientific community. In fact the context in which this assessment has been carried on is very particular, as it has been said, and the data obtained can be helping for future scientists and professionals operating in similar contexts.

A fact is that there has been a high level of participation. Despite being an evaluation/development session, most of the participants participated actively, considering the work setting to be useful as training. The tests and scales created curiosity about the scores and about how they work, the parts derived from the observations have been perceived as a possibility for close exchange between professional figures who cooperate with each other daily, but do not have the opportunity to understand the different viewpoints used to interpret work situations.

References

Brummel, B. – Rupp, D. – Spain, S. (2009): Constructing parallel simulation exercices for assessment center and other forms of behavioral assessment. *Personnel Psychology*.62, 1, 137-170. p.

Caprara, G.V. – Barbaranelli, C. - De Carlo, N.A. – Robusto, E. (2006): *Il Multidimensional Personality Profile (MPP) – Un questionario di nuova generazione per la misura della personalità.* Franco Angeli, Milano.

Eurich, T.L. – Krause, D.E. – Cigularov, K. – Thornton, G.C. (2009): Assessment Centers: Current Practices in the United States. *Journal of Business & Psychology*; 24, 4, 387-407. p.

Gherardi, S. (2006): *Organizational Knowledge. The texture of Workplace Learning*. Blackwell, London.

Krause, D.E. – Kersting, M. – Heggestad, E.D. – Thornton, G.C. (2006): Incremental Validity of Assessment Center Ratings over Cognitive Ability Tests: a Study at the Executive Management Level. *International Journal of Selection and Assessment*. Vol. 14, Issue 4, 360-371. p.

Meyer, J. P. - Allen, N. J. (1991): A three-component conceptualization of organizational commitment. *Human Resource Management Review*, *1*, 61-89. p.

Meyer, J. P. - Allen, N. J. (1997): *Commitment in the workplace: Theory, research, and application*.

Sage, Thousand Oaks, CA.

Organ, D. W. (1988): *Organizational citizenship behavior: The good soldier syndrome. Issues in organization and management series.* Lexington, Manchester.

Organ, D.W. - Posdakoff, P.M. - MacKenzie, S.B. (2005): *Organizational Citizenship Behavior. Its Nature, Antecedents, and Consequences.* Sage, Thousand Oaks, CA.

Smith, C.A. - Organ, D.W. - Near, J.P. (1983): Organizational citizenship behavior: Its nature, and antecedents. *Journal of Applied Psychology*, 68, 653-663. p.

St-Onge, S. – Morin, D. – Bellehumeur, M. – Dupuis, F. (2009): Managers' motivation to evaluate subordinate performance. *Qualitative Research in Organizations and Management: An International Journal*. 4, 3, 273-293. p.

Thornton, G.C. – Rupp, D.E. (2005): *Assessment Centers in Human Resource Management*. Lawrence Erlbaum Assoc. Inc., Hardcover.

Weick, K. (1969): *The Social Psychology of Organizing*. Addison-Wesley, Reading, MA.

Weick, K. E. (1995): *Sensemaking in organizations*. Sage, Thousand Oaks, CA.

Vendor Selection by Application of Revised Weighting Method and Fuzzy Multicriteria Linear Programming

Tunjo Perić[1] - Zoran Babić[2]

Vendor selection in supply chain is a multicriteria problem that involves a number of quantitative and qualitative factors. This work deals with a concrete problem of flour purchase by a company that manufactures bakery products. The criteria for vendor selection and quantities supplied by individual vendors are: purchase costs, product quality and reliability of vendor. The problem of vendor selection is solved by a model that combines revised weighting method and fuzzy linear programming. The study points to the advantages of using the combination of the two methods in comparison to the separate use of one of them only.

Keywords: vendor selection, fuzzy linear programming, revised weighting method

1. Introduction

The problem of vendor selection and determination of material quantities supplied is the key element in the purchasing process in manufacturing which is one of the most important activities in supply chain. If all the selected vendors are able to meet the buyer's requirements completely, then the selection process becomes easier and is based only on the selection of the most suitable vendor in terms of purchasing costs, product quality, and vendor reliability. Nevertheless, practice shows that it is not good to rely on one vendor only. Therefore the management of the purchasing company generally enters into contracts with several vendors. Their number usually ranges from two to five for each sort of material. Also, there are cases when no vendor can meet the buyer's demand, or will not do it in order to protect his own business interests.

In principle there are two kinds of supplier (vendor) selection problem: The first is supplier selection when there is no constraint or in other words all suppliers can satisfy the buyer's requirements of demand, quality, delivery etc. In this kind of supplier selection the management needs to make only one decision - which supplier

[1] Tunjo Perić, PhD student, general manager, Bakeries Sunce, Rakitje, Rakitska cesta 98, 10437 Bestovje (Croatia)
[2] Zoran Babić, PhD student, full professor, Faculty of Economics, University of Split, Matice Hrvatske 31 (Croatia)

is the best one. The other type of supplier selection problem is when there are some limitations on suppliers' capacity, quality and so on. In other words, no supplier can satisfy the buyer's total requirements and the buyer needs to purchase some of the needed material from one supplier and some from another to compensate for the shortage of capacity or low quality of the first supplier. Consequently, the firm must decide which vendors it should contract with and it must determine the appropriate order quantity for each vendor selected.

In this paper we will discuss the second kind of supplier selection problem, but we will also provide a model which allows making both decisions simultaneously. The model combines two methods used in operational researches. The first of them, revised weighting method is used to determine the coefficient weights of complex criteria functions (cost, quality and reliability). Coefficients determined in this way present the coefficients of the objective functions in the fuzzy multi-criteria programming model providing the final selection and the quantity supplied from a particular vendor. The constraints in the multiple objective programming model are the total demand and the limitations of supplier capacities.

High competition forces companies to produce faster, at less cost, and better than their competitors, which cannot be done unless they select the best vendors. The increasing importance of vendors makes companies consider a number of criteria in vendor selection. A list of criteria for vendor selection can be seen in the classic study by Weber et al.(1991), or for example in the paper from Lin and Chen (2004) who generate a Generic Configuration Hierarchy (GCH) of supplier attributes that could be used as the basis for supplier selection criteria in any industry. They list no fewer than 138 such attributes.

The literature dealing with vendor selection uses various methods. Among the numerous studies dealing with this topic we will mention some more important ones. The AHP method was used for vendor selection in the following works: Narasimhan 1983, Nydick-Hill 1992 and Barbarosoglu-Yazac 1997. For vendor evaluation Weber-Ellram 1992, Weber-Desai 1996, and Weber et al. 1998 use the DEA method. The fuzzy AHP method for vendor evaluation is used in the studies by Haq-Kannan (2006) and Chan-Kumar (2007). For vendor evaluation and determination of supply quotas the AHP is used in combination of some methods of mathematical programming. Thus for instance Ghodsypour and O'Brien (1998) use the AHP method in combination with linear programming. Ge Wang et al. (2004) use the AHP and goal programming. Kumar et al. (2008) use the AHP method and fuzzy linear programming, while Kumar et al. (2004, 2005) use only fuzzy goal programming for that purpose.

Obviously, vendor selection is an important issue dealt with by numerous researchers. Great efforts are made to define appropriate models for vendor selection and determination of supply quotas from the selected vendors and to apply the adequate methods to solve such models.

The aims of this work are the following: (1) to point on the concrete example that vendor selection is a multicriteria problem, (2) to propose criteria for vendor selection, (3) to propose the model for vendor selection and determination of supply quotas by using the revised weighting method and fuzzy linear programming, and (4) to point to the advantages of the proposed model in comparison to the usual methods of vendor selection. The concrete example will be the problem of flour vendor selection by a bakery.

The rest of the paper is organized as follows: We will first present the methodology of vendor selection and determination of supply quotas by use of revised weighting method and fuzzy linear programming. Then the proposed methodology will be tested on the concrete example of vendor selection by a bakery. Finally, we will carry out sensitivity analysis of the obtained solutions. In the conclusion we will point to the advantages of using the proposed methodology in comparison to the use of revised weighting method, or fuzzy linear programming method only.

2. Methodology of vendor selection and determination of supplied quantity

For vendor selection and determination of supplied quantity we will use the revised weighting method and fuzzy linear programming (FLP). The revised weighting method is used to determine the coefficient weights of complex criteria functions. The coefficients obtained in this way are used as criteria functions coefficients in the LP fuzzy model. The main steps in the proposed model are:

1. Determining criteria for vendor selection,
2. Applying revised weighting method to determine the variable's coefficients in criteria functions,
3. Building and solving the FLP model to determine supply quotas from selected vendors,
4. Sensitivity analysis of the obtained solution.

2.1. Determining criteria for vendor selection

The first step in the proposed methodology is selection of criteria for vendor selection. Numerous criteria are stated in literature and their selection depends on the concrete problem (Weber et al. 1991). The most important criteria may certainly be: the total purchasing costs in a particular period, product quality offered by particular vendors, and vendor reliability. Each of these criteria is expressed through a number of sub-criteria, which can further be expressed through a number of sub-sub-criteria, etc. This reveals the hierarchical structure of criteria for vendor

selection, which directs us to apply the revised weighting method to solve this problem (Koski-Silvennoinen 1987).

2.2. The revised weighting method

We will give a brief outline of the basic propositions of this multicriteria method used in a large number of factual cases.

The main idea of the weight coefficient method as presented by Gass and Satty (1955) and Zadeh (1963) is to relate each criteria function with the weight coefficient and to maximize/minimize the weighted sum of the objectives. In that way the model containing several criteria functions is transformed into the model with one criteria function. It is assumed that the weight coefficients w_j are real numbers so that $w_j \geq 0$ for all $j = 1,\ldots,k.$ It is also assumed that the weights are normalized, so that $\sum_{j=1}^{k} w_j = 1.$ Analytically presented, the multicriteria model is modified into a monocriterion model and is called the weight model:

$$\max/\min \sum_{j=1}^{k} w_j f_j(\underline{x}) = \sum_{j=1}^{k} \sum_{i=1}^{n} w_j c_{ij} x_i \tag{1}$$

s.t

$$\underline{x} \in X, \tag{2}$$

where $w_j \geq 0$ for all $j = 1,\ldots,k,$ $\sum_{j=1}^{k} w_j = 1.$ To make the weight coefficients w_j express the relative importance of criteria functions f_j we propose linear transformation of criteria functions coefficients. To allow addition of weighted criteria functions we have to transform all of them either into functions that have to be maximized or into functions to be minimized. Linear transformation of criteria functions coefficients that have to be maximized is performed in the following way:

$$c_{ij}^{'} = c_{ij} / c_j^{*}, \tag{3}$$

where $c_j^{*} = \max_i c_{ij}.$ Obviously $0 \leq c_{ij}^{'} \leq 1.$

The criteria functions that have to be minimized will be transformed into functions to be maximized by taking reciprocal values of coefficients $1 / c_{ij}.$ Then

$$c_{ij}^{'} = \frac{1 / c_{ij}}{\max_i (1 / c_{ij})} = \frac{\min_i}{c_{ij}} = \frac{c_j^{\min}}{c_{ij}}. \tag{4}$$

Now we will normalize the coefficients $c_{ij}^{'}$ into dimensionless space $[0,1]$:

$$c_{ij}'' = \frac{c_{ij}'}{\sum_{i=1}^{n} c_{ij}'}, \quad j = 1,\ldots,k. \tag{5}$$

The above transformations allow us to obtain the weighted sum of criteria functions in which the weights reflect the relative importance of criteria functions.

It is to be noted that in the revised method of weight coefficients all theoretical results valid in the basic weight coefficient method remain valid. We will here present the three basic theorems in the light of the revised weight coefficient method.

Theorem 1: Solution of the weight model (1-2) is weakly Pareto optimal.

Proof: The proof will be shown for the case of maximization. Let $\underline{x}^* \in X$ be the solution of the weight model. Let us assume that the solution is not weakly Pareto optimal. In such a case there is solution $\underline{x} \in X$ so that $f_j(\underline{x}) > f(\underline{x}^*)$ for all $j = 1,\ldots k,$ because we have $w_j > 0$ for at least one j. Consequently, $\sum_{j=1}^{k} w_j f_j(\underline{x}) > \sum_{j=1}^{k} w_j f_j(\underline{x}^*)$. This contradicts the assumption that $\underline{x}^*$ is the solution of the weight model. Therefore, $\underline{x}^*$ is the weakly Pareto optimal solution.

Theorem 2: The solution of the weight model (1-2) is Pareto optimal if all the weight coefficients are positive, i.e. $w_j > 0$ for all $j = 1,\ldots k.$

Proof: Let $\underline{x}^* \in X$ be the solution of the weight model with positive weight coefficients. Let us assume that this solution is not Pareto optimal. This means that there is another solution $\underline{x} \in X$ so that $f_j(\underline{x}) \geq f(\underline{x}^*)$ for all $j = 1,\ldots k,$ and $f_j(\underline{x}) > f(\underline{x}^*)$ for at least one j. As $w_j > 0$ for all $j = 1,\ldots k,$ we get $\sum_{j=1}^{k} w_j f_j(\underline{x}) > \sum_{j=1}^{k} w_j f_j(\underline{x}^*)$. This contradicts the assumption that $\underline{x}^*$ is the solution of the weight model and therefore has to be Pareto optimal (Miettinen 1999).

Theorem 3: The unique solution of the model (1-2) is Pareto optimal.

Proof: Let $\underline{x}^* \in X$ be the unique solution of the weight model. Let us assume that it is not Pareto optimal. In that case there is solution $\hat{\underline{x}} \in X$ so that $f_j(\hat{\underline{x}}) \geq f(\underline{x}^*)$ for all $j = 1,\ldots k,$ and $f_j(\hat{\underline{x}}) > f(\underline{x}^*)$ for at least one j. Because all the coefficients w_j

are non-negative, we have $\sum_{j=1}^{k} w_j f_j(\hat{\underline{x}}) \geq \sum_{j=1}^{k} w_j f_j(\underline{x}^*)$. On the other hand, the uniqueness of $\underline{x}^*$ means that $\sum_{j=1}^{k} w_j f_j(\underline{x}^*) > \sum_{j=1}^{k} w_j f_j(\hat{\underline{x}})$ for all $\hat{\underline{x}} \in X$. These two inequations are contradictory, therefore $\underline{x}^*$ has to be Pareto optimal.

In this paper we use the revised weight coefficients method to reduce the number of complex criteria functions. This idea originates from Koski and Silvennoinen (1987). According to it, the normalized original criteria functions are divided into groups so that the linear combination of criteria functions in each group forms a new criteria function while the linear combination of new criteria functions form a further criteria function, etc. In this way we obtain a model with a reduced number of criteria functions. According to this each Pareto optimal solution of the new model is also Pareto optimal solution of the original model, but the reverse result is not generally true.

The weight coefficients method has some shortcomings so that it is not the most appropriate one to create a set of Pareto optimal solutions. The shortcomings are: (1) varying weight coefficients do not guarantee that we will determine all Pareto optimal solutions, and (2) the determined Pareto optimal solutions are those that are situated in the extreme points of the convex polyhedron but not those that connect the two extreme points. To determine the set of compromise solutions and the preferred solution, we will here use the fuzzy linear programming method.

2.3. Fuzzy linear programming (FLP)

The general multi-criteria programming model to solve the problem of determining the supply quotas by selected vendors can be presented as:

Find the vector $\underline{x}$ which minimizes criteria functions f_r and maximizes criteria functions f_s with

$$f_r = \sum_{i=1}^{n} c_{ri} x_i, \qquad r = 1, 2, \ldots, p \tag{6}$$

$$f_s = \sum_{i=1}^{n} c_{si} x_i, \qquad s = p+1, p+2, \ldots, q \tag{7}$$

s.t.

$$\underline{x} \in X_d, \tag{8}$$

where $X_d = \left\{ \underline{x} \mid g_l(\underline{x}) = \sum_{i=1}^{n} a_{il} x_i \geq b_l,\ g_p(\underline{x}) = \sum_{i=1}^{n} x_i = D,\ 0 \leq x_i \leq u_i,\ l = 1,\ldots,m, \atop i = 1,2,\ldots,n \right\}$, and x_i is the quantity ordered from the vendor *i*, *D* is the total demand in the given period, u_i is the upper limit of order from the vendor *i*, c_{ri} are the coefficients with variables in criteria functions that are to be maximized, such as: Total Value of Purchasing (TVP), product quality, vendor reliability, etc., while c_{si} are coefficients with variables in criteria functions that have to be minimized, such as total purchasing costs, etc., and a_{il} are coefficients in constraints, which can for instance be vendor flexibility in terms of delivery quotas, subjective rating of the vendor, etc., while b_l represents the lower limit of constraint $g_l(\underline{x})$.

Zimmermann 1978 solved the problem (6 – 8) by fuzzy linear programming approach. He formulated the fuzzy linear program determining for each criteria function f_j its maximal value f_j^+ and its minimal value f_j^-, solving:

$$f_r^+ = \max f_r, \quad \underline{x} \in X_a, \quad f_r^- = \min f_r, \quad \underline{x} \in X_d \tag{9}$$

$$f_s^+ = \max f_s, \quad \underline{x} \in X_d, \quad f_s^- = \min f_s, \quad \underline{x} \in X_a \tag{10}$$

f_r^-, f_s^+ will be obtained by solving the multicriteria model as linear programming model separately minimizing or maximizing single criteria functions. $\underline{x} \in X_d$ means that solutions belongs to feasible set X_d, while X_a is a set of all optimal solutions obtained by solving single criteria functions.

As for each criteria function f_j its value is changed linearly from f_j^- to f_j^+, that value can be observed as a fuzzy number with linear membership function $\mu_{f_j}(\underline{x})$.

Consequently, the MCLP model (6 – 8) with fuzzy goals and fuzzy constraints can be presented as:

$$\tilde{f}_r = \sum_{i=1}^{n} c_{ri} x_i \lesssim f_r^0, \qquad r = 1,2,\ldots,p \tag{11}$$

$$\tilde{f}_s = \sum_{i=1}^{n} c_{si} x_i \gtrsim f_s^0, \qquad s = p+1, p+2,\ldots,q \tag{12}$$

s.t.

$$\tilde{g}_l(\underline{x}) = \sum_{i=1}^{n} a_{il} x_i \gtrsim b_l, \quad l = 1,\ldots,m \tag{13}$$

$$\tilde{g}_p(\underline{x}) = \sum_{i=1}^{n} x_i =\approx D,\ 0 \le x_i \le u_i,\quad i = 1,\ldots,n. \tag{14}$$

In this model the sign $\approx$ indicates fuzzy environment. The symbol $\lesssim$ denotes the fuzzy version $\le$, and is interpreted as "essentially smaller than or equal to", the symbol $\gtrsim$ is interpreted as "essentially greater than or equal to", while the symbol $=\approx$ is interpreted as "essentially equal to". f_r^0 and f_s^0 represent the aspiration levels of criteria functions that will be achieved by the decision maker.

Assuming that the membership functions based on preference or satisfaction are linear, we can present the linear membership functions for criteria functions and constraints as follows:

$$\mu_{f_r}(\underline{x}) = \begin{cases} 1 & \text{for} \quad f_r \le f_r^- \\ (f_r^+ - f_r(\underline{x}))/(f_r^+ - f_r^-) & \text{for} \quad f_r^- \le f_r(\underline{x}) \le f_r^+,\ r = 1,2,\ldots,p \\ 0 & \text{for} \quad f_r \ge f_r^+ \end{cases} \tag{15}$$

$$\mu_{f_s}(\underline{x}) = \begin{cases} 1 & \text{for} \quad f_s \ge f_s^+ \\ (f_s(\underline{x}) - f_s^-)/(f_s^+ - f_s^-) & \text{for} \quad f_s^- \le f_s(\underline{x}) \le f_s^+,\ s = p+1, p+2,\ldots,k \\ 0 & \text{for} \quad f_s \le f_s^- \end{cases} \tag{16}$$

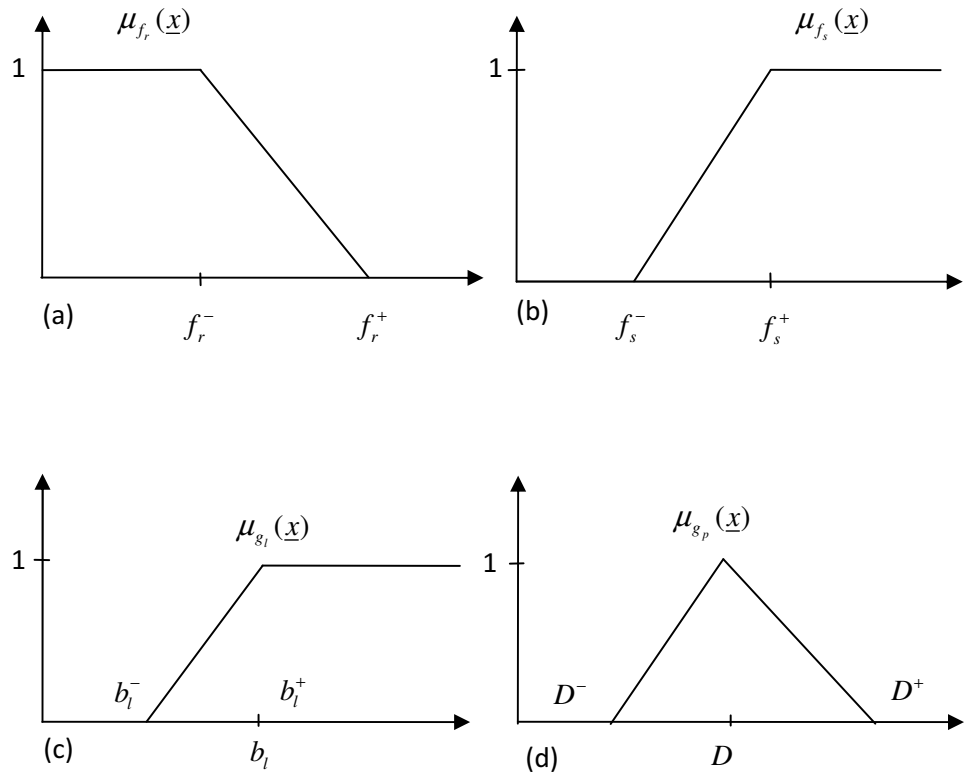
$$\mu_{g_l}(\underline{x}) = \begin{cases} 1 & \text{for} \quad g_l(\underline{x}) \ge b_l \\ (g_l(\underline{x}) - b_l^-)/(b_l - b_l^-) & \text{for} \quad b_l^- \le g_l(\underline{x}) \le b,\ l = 1,2,\ldots,m \\ 0 & \text{for} \quad g_l(\underline{x}) \le b_l^- \end{cases} \tag{17}$$

$$\mu_{g_p}(\underline{x}) = \begin{cases} 1 & \text{for} \quad g_p(\underline{x}) = D \\ (g_p(\underline{x}) - D^-)/(D - D^-) & \text{for} \quad D^- \le g_p(\underline{x}) \le D \\ (D^+ - g_p(\underline{x}))/(D^+ - D) & \text{for} \quad D \le g_p(\underline{x}) \le D^+ \\ 0 & \text{in other cases,} \end{cases} \tag{18}$$

where $b_l^- = b_l - d_l$, and $b_l^+ = b_l + d_l$, and $D^- = D - p_1$, $D^+ = D + p_2$. d_l are subjectively determined constants expressing the limits of allowed deviations of l inequation (tolerance interval) and p_1, p_2 are subjectively determined constants expressing the limits of allowed deviations of equation $g_p(\underline{x})$.

The graphic presentation of membership functions looks like this:

Figure 1. Fuzzy linear membership functions for target functions and constraints: (a) minimization of criteria functions, (b) maximization of criteria functions, (c) constraints of the type $\geq$, (d) constraint of the type $=$.



Source: own creation

In the fuzzy programming model, according to Zimmermann's approach, the fuzzy approach represents the average intersection of all the fuzzy sets that represent fuzzy criteria functions and fuzzy constraints. The fuzzy solution for all the fuzzy goals and fuzzy constraints is given as follows:

$$\mu_D(\underline{x}) = \left\{ \left\{ \bigcap_{j=1}^{k} \mu_{f_j}(\underline{x}) \right\} \cap \left\{ \bigcap_{i=1}^{m} g_l(\underline{x}) \right\} \cap \mu_{g_p}(\underline{x}) \right\}. \tag{19}$$

The optimal solution ($\underline{x}^*$) is:

$$\mu_D(\underline{x}^*) = \max_{\underline{x}\in X_D} \mu_D(\underline{x}) = \max_{\underline{x}\in X_D} \min\left[\min_{j=1,\ldots,k} \mu_{f_j}(\underline{x}), \min_{l=1,\ldots,m} \mu_{g_l}(\underline{x}), \min \mu_{g_p}(\underline{x}) \right]. \tag{20}$$

The optimal solution ($\underline{x}^*$) of the above model can be obtained by solving the following linear programming model (Zimmermann 1978):

$$(\max)\lambda \tag{21}$$

s.t.

$$\lambda \le \mu_{f_j}(\underline{x}), \quad j = 1,2,\ldots,k \tag{22}$$

$$\lambda \le \mu_{g_l}(\underline{x}), \quad l = 1,2,\ldots,m \tag{23}$$

$$\lambda \le \mu_{g_p}(\underline{x}) \tag{24}$$

$$0 \le x_i \le u_i, \quad i = 1,\ldots,n; \ \ \lambda \in [0,1], \tag{25}$$

where $\mu_D(\underline{x})$ is the membership function for the optimal solution, $\mu_{f_j}(\underline{x})$ represents membership functions for criteria functions, $\mu_{g_l}(\underline{x})$ represents membership functions for constraints of type $\ge$, and $\mu_{g_p}(\underline{x})$ represents a membership function for constraint of type $=$. In this model the relation between constraints and criteria functions is totally symmetrical (Zimmermann 1978), and here the decision maker cannot express the relative importance of criteria functions and constraints.

In order to express the relative importance of criteria functions and constraints we have to solve the so called weight additive model in which weights present utility functions of criteria functions and constraints (Bellman-Zadeh 1970, Sakawa 1993, Tiwari et al. 1987 and Amid et al. 2006).

The convex fuzzy model proposed by Bellman and Zadeh 1970 and Sakawa 1993 and the weight additive model, by Zimmermann (1978) is

$$\mu_D(\underline{x}) = \sum_{j=1}^{k} w_j \mu_{f_j}(\underline{x}) + \sum_{l=1}^{m} \beta_l \mu_{g_l}(\underline{x}) + \gamma \mu_{g_p}(\underline{x}), \tag{26}$$

$$\sum_{j=1}^{k} w_j + \sum_{l=1}^{m} \beta_l + \gamma = 1, \quad w_j, \beta_l, \gamma \ge 0, \tag{27}$$

where w_j, β_l and γ are weight coefficients representing the relative importance between the fuzzy criteria functions and fuzzy constraints.

To solve the above fuzzy model we will use the following linear programming model:

$$(\max) f = \sum_{j=1}^{k} w_j \lambda_{1j} + \sum_{l=1}^{m} \beta_l \lambda_{2l} + \lambda_3 \tag{28}$$

s.t.

$$\lambda_{1j} \le \mu_{f_j}(\underline{x}), \quad j = 1,2,\ldots,k, \tag{29}$$

$$\lambda_{2l} \le \mu_{g_l}(\underline{x}), \quad l = 1,2,\ldots,m, \tag{30}$$

$$\lambda_3 \le \mu_{g_p}(\underline{x}) \tag{31}$$

$$0 \le x_i \le u_i, \quad i = 1,\ldots,n; \tag{32}$$

$$\lambda_{1j}, \lambda_{2l}, \lambda_3 \in [0,1], \quad j = 1,2,\ldots,k; l = 1,2,\ldots,m, \tag{33}$$

2.4. Sensitivity analysis

Sensitivity analysis has to indicate robustness of the obtained solutions in vendor selection and in determination of the quantities supplied from them. After the application of FLP it is necessary to test the sensitivity of the obtained solutions caused by changes in criteria weights.

3. Case study

3.1. Criteria for vendor selection

Vendor selection and determination of quantities supplied by the selected vendors is a multicriteria problem. The most important issue in vendor selection is the choice of criteria for their evaluation. A large number of criteria that can be used in vendor selection are offered in literature. Which criteria will be chosen by the decision maker depends on the kind of problem to be solved. In this study we will consider criteria that can be used by producers of bakery products when selecting flour vendors.

Criteria used for evaluation of flour vendors can be:

- flour purchasing costs,
- flour quality, and
- vendor reliability.

Flour purchasing costs involve unit cost and transportation costs expressed in monetary units per ton.

Flour quality criteria important for bread production are expressed by the following parameters:

- General characteristics of flour (moisture in %, ash in %, acidity level ml/100 grams and wet gluten in %),
- Farinograph (water absorption %, dough development in minutes and mellowness in FU),
- Extensigraph (energy in square centimeters after 60 minutes of dough resting, elasticity in mm and resistance in extensigraph units (EU), and

- Amylograph (peak viscosity in AJ). Indicators of swelling time, temperature maximum and gluten formation time are not significant for bread production technology, therefore are not taken into account here.

It is very important to use appropriate methods for flour analysis consistently.

When contracting flour supply, it is important to find reliable vendors, i.e. those that are assumed with a high degree of certainty that will not get into financial difficulties which could result in supply discontinuation. To evaluate vendor reliability we can use indicators of their solvency, financial stability, indebtedness, liquidity, and financial performance.

Solvency indicators may be: total cash inflow in the last 30 days, average balance in the last 30 days, the amount of credit allowed, data on the number of continuous days of frozen account, number of frozen account days in the last 180 days. These data can be obtained from the SOL 2 form issued by the bank in which the vendor's main account is opened. In our opinion, it is risky to do business with suppliers with a frozen account, or with those that have had a frozen account in the last 180 days, and such vendors should be eliminated before the selection process.
Indicators of financial solvency, indebtedness, and liquidity can be: coverage of fixed assets and stocks by capital and long term resources, share of capital in source of funds in %, annual indebtedness factor, total assets turnover coefficient, general liquidity coefficient, short term receivables collection period, average sale period in days.

Indicators of financial performance are: coefficient of total revenue and expenditure ratio, share of profit in total income in %, share of profit in assets in %, and profit per employee in monetary units.

Decision maker's subjective evaluation can also be one of the indicators of vendor reliability. This indicator becomes very important in cases when company has a long standing business relationship with the vendor.

It is to be noted that it is not advisable to do business with unreliable vendors. In most cases practice shows that vendor reliability and their product quality are correlated, so that the vendors ranked high in terms of quality are usually also ranked high in terms of reliability. Indicators of solvency, indebtedness, and liquidity, as well as indicators of financial performance can be obtained from the form BON-1 issued by the central financial agency that keeps records of all legal entities based on their financial statements.

Vendor reliability criteria should include those indicators that in the period covered by the contract of purchase can have a negative effect on delivery of goods. A large number of vendor reliability indicators will make the decision making difficult. It would be hard to adequately evaluate vendor reliability without support of experts and application of quantitative methods. Therefore we will here propose quality and reliability criteria for whose application collecting data will not be a problem.

3.2.Data required for vendor selection and determination of supply quotas

We will show the example of vendor selection for a bakery. It is to be noted that in production of bread and bakery products the purchase of flour is contracted for the period of one year, from harvest to harvest (which usually does not correspond to the calendar year). After the harvest flour producers have the information on the available wheat quantity, price and quality which allows them to define the price, quality and quantity of flour they can supply in the subsequent one-year period.

In the one-year period the bakery plans to consume 4000 tons of flour Type 550. The company contacts 4 potential flour suppliers and defines the upper limit of flour supplied by a single vendor in the amount of 1500 tons. The proposed prices of flour and transportation costs (Criterion C1) are shown in the Table 1. The potential vendors supply the data on flour quality that they have to maintain throughout the contract period (Criterion C2). It is to be noted that the quality of flour depends on the wheat sort and quality and on technology used in flour production. The vendors also should supply data on their reliability by the forms SOL-2 and BON-1 (Criterion C3). The Tables 2 and 3 show the flour quality indicators and vendor reliability. The weights expressing the relative importance of criteria and sub-criteria are given in brackets, and they are determined by the decision maker where in every group of sub-criteria the sum of weights is 1.

Table 1. Purchasing costs for flour Type 550

Vendor	Purchasing price in MU/ton (B1)	Transportation cost in MU/ton (B2)	Total purchasing costs per ton in MU
1	2300	100	2400
2	2180	120	2300
3	2090	110	2200
4	2120	130	2250

Source: own creation

Table 2. Quality indicators fo flour Type 550

Quality indicators	Criteria weights	Vendor 1	2	3	4
General characteristics of flour (A1)	**(0.20)**				
Moisture in % (B3)	min (0.30)	13.53	13.27	13.49	13.33
Ash in % (B4)	min (0.20)	0.57	0.549	0.53	0.486
Acidity level in ml/100 grams (B5)	min (0.10)	1.5	1.5	1.6	1.8
Wet gluten in % (B6)	max (0.40)	26.7	25.8	25.1	24.0
Farinograph (A2)	**(0.30)**				
Water absorption in % (B7)	max (0.40)	60.8	59.8	58.5	61.1
Degree of mellowness in FJ (B8)	min (0.60)	70	65	85	60
Extensigraph (A3)	**(0.30)**				
Energy u cm^2 (B9)	max (0.40)	81	104	87.2	107.3
Elasticity in mm (B10)	max<190 (0.30)	137	162	180	165
Resistance (B11)	max (0.30)	395	280	235	350
Amylograph (A4)	**(0.20)**				
Peak viscosity in BU (B12)	max (1.00)	1054	860	1275	1325

Source: own creation

Table 3. Vendor reliability indicators

Reliability indicators	Criterion	Vendor			
		1	2	3	4
Financial stability, indebtedness and liquidity (A5)	**(0.60)**				
Coverage of fixed assets and stocks by capital and long term resources, (B13)	max (0.20)	1.12	0.88	0.87	0.92
Share of capital in source of funds in %, (B14)	max (0.10)	49.36	23.6	48.92	49.69
Indebtedness factor, number of years (B15)	min (0.10)	7	19	13	19
Total assets turnover coefficient (B16)	max (0.10)	0.65	0.49	0.52	0.35
General liquidity coefficient (B17)	max (0.30)	7.17	1.19	1.07	0.75
Short term receivables collection period, in days (B18)	min (0.20)	86	101	102	58
Performance indicators (A6)	**(0.40)**				
Coefficient of total revenue and expenditure ratio (B19)	max (0.20)	1.06	1.03	1.03	1.02
Share of profit in total income in % (B20)	max (0.30)	4.81	1.85	2.66	1.02
Share of profit in assets in % (B21)	max (0.20)	3.14	0.91	1.39	1.01
Profit per employee in mu (B22)	max (0.30)	60538	21189	12370	15446

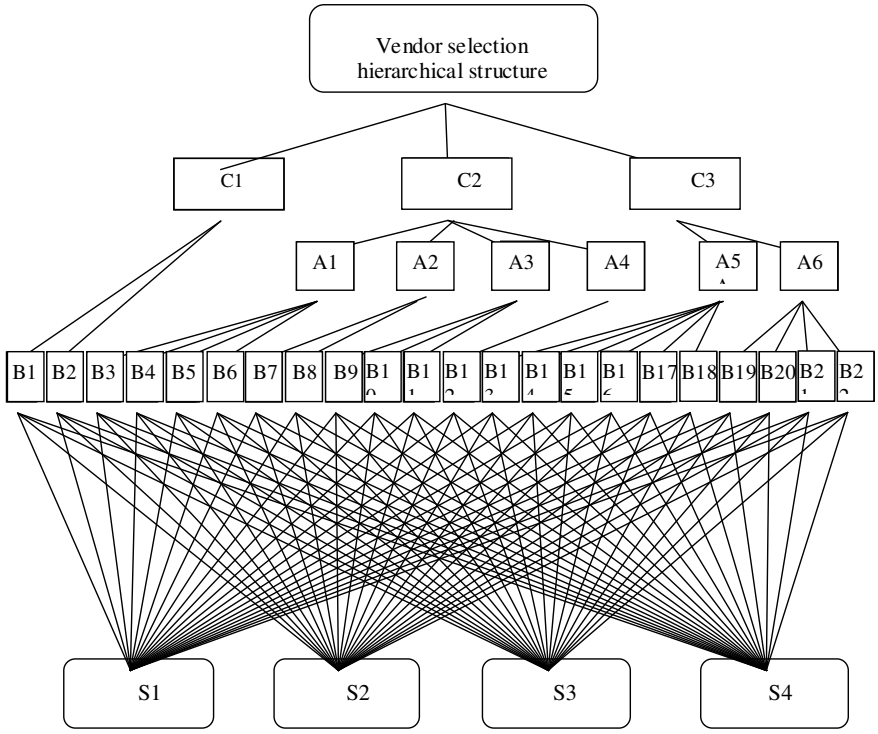
*Source:*own creation

According to the indicators from the form SOL-2 potential vendors have an active current account that has not been frozen in the last 180 days.

3.3. Application of revised weighting method

Considering the data from the Tables 1, 2 and 3 we form a hierarchical structure of goals and criteria for vendor selection. The hierarchical structure is shown in the Figure 2.

Figure 2. Hierarchical Structure of Suppliers Selection



Source: own creation

The hierarchical structure in our example consists of five levels as shown in the Figure 2. Level 1 represents the vendor general efficiency (or total value of purchasing - TVP), Level 2 represents criteria for vendor selection, Level 3 represents criterion sub-criteria, Level 4 represents sub-criterion sub-criteria, and Level 5 represents the available alternatives (vendors).

After decomposition of the problem and formation of the hierarchical structure of goals and criteria, we have applied a revised weight coefficients method to calculate the coefficients of cost, quality and reliability functions. By application of the relation (3) and (5) the cost function coefficients are normalized. The following weights are obtained:

Table 4. Normalized coefficients of cost function

Variable	Coeff. c'_{i1}	Coeff. c''_{i1}
x_1	1	0.262295
x_2	0.958333	0.251366
x_3	0.916667	0.240437
x_4	0.9375	0.245902

Source: own creation

The quality function has a hierarchical structure and has to be maximized. Sub-criteria B3 to B12 are grouped into 4 sub-criteria sets. According to the data on coefficients weights, their linear transformation and normalization into the interval $[0,1]$ is carried out. The normalized coefficient values are shown in the following table:

Table 5. Normalized coefficient values with variables for sub-criteria B3-B12

Var.	c''_{iB3}	c''_{iB4}	c''_{iB5}	c''_{iB6}	c''_{iB7}	c''_{iB8}	c''_{iB9}	c''_{iB10}	c''_{iB11}	c''_{iB12}
x_1	0.247674	0.233287	0.265193	0.262795	0.253122	0.245874	0.213439	0.212733	0.313492	0.233496
x_2	0.252527	0.242211	0.265193	0.253937	0.248959	0.264788	0.274045	0.251553	0.222222	0.190518
x_3	0.248409	0.250894	0.248619	0.247047	0.243547	0.202485	0.229776	0.279503	0.186508	0.282455
x_4	0.25139	0.273608	0.220994	0.23622	0.254371	0.286853	0.28274	0.256211	0.277778	0.293531

Source: own creation

Using the data on coefficient weights with variables of grouped sub-criteria and weight coefficients with sub-criteria A1, A2, A3 and A4, and by applying the relation (1) we calculate the coefficients with criterion C2 variables:

Table 6. Normalized coefficient weights with quality criterion variables

Variable	Coeff. c''_{i2}
x_1	0.244824
x_2	0.241625
x_3	0.241354
x_4	0.272198

Source: own creation

Reliability criterion coefficients are calculated in a similar way:

Table 7. Normalized coefficient weights with reliability criterion variables

Variable	Coeff. c''_{i3}
x_1	0.397097
x_2	0.191739
x_3	0.208131
x_4	0.203032

Source: own creation

3.4. FLP model building and solving

As there are constraints in terms of capacity or limited quantity supplied by a single vendor, we have to form a MLP model to determine the quantities to be supplied by selected vendors. Considering the data on normalized coefficient weights with variables of cost, quality, and reliability functions, the total demand for flour in the given period and limited quantities supplied from single vendors, we form the following MLP model:

Minimization of purchasing cost

$$(\min) f_1 = 0.262295x_1 + 0.251366x_2 + 0.240437x_3 + 0.245902x_4 \tag{30}$$

Maximization of flower quality

$$(\max) f_2 = 0.244824x_1 + 0.241625x_2 + 0.241354x_3 + 0.272198x_4 \tag{31}$$

Maximization of vendor reliability:

$$(\max) f_3 = 0.397097x_1 + 0.191739x_2 + 0.208131x_3 + 0.203032x_4 \tag{32}$$

s.t.

Total needed flour quantity, limited quantities supplied, and non-negativity of variables:

$$g_1 = x_1 + x_2 + x_3 + x_4 = 4000 \tag{33}$$

$$g_2 = x_1 \leq 1500 \tag{34}$$

$$g_3 = x_2 \leq 1500 \tag{35}$$

$$g_4 = x_3 \leq 1500 \tag{36}$$

$$g_5 = x_4 \leq 1500 \tag{37}$$

$$x_1, x_2, x_3, x_4 \geq 0 \tag{38}$$

Model (29-38) is a multi-criteria linear programming model where the coefficients of the goal functions are obtained in the first stage of problem solving by application of the revised weighting method.

Model (29-38) is first solved by linear programming method optimizing separately each of the three criteria function on the given set of constraints. The results are given in the Payoff table:

Table 8. Payoff values

Solution	(min) $f_1(\underline{x})$	(max) $f_2(\underline{x})$	(max) $f_3(\underline{x})$
x_1^*	980.8745	1011.953	808.4835
x_2^*	1013.662	1017.158	1091.933
x_3^*	1000.00	1001.465	1110.874

Source: own creation

It can be seen that the obtained solutions differ and that we have to choose a compromise solution. This work proposes methodology for vendor selection and determination of supply quotas by application of fuzzy linear programming on the model (29-38) in which the functions f_1, f_2 and f_3 are optimized on the given set of constraints. The application of FLP requires determination of the highest and lowest value for each criteria function. These values represent the aspiration levels in FLP. The lowest and highest values for criteria functions are shown in the following table

Table 9. Fuzzy goals

Criteria	Value-I	Value-II
f_1	980.8745*	1013.662
f_2	1001.465	1017.158*
f_3	808.4835	1110.874*

Source: own creation

Based on the above data we calculate the linear membership functions:

$$\mu_{f_1}(\underline{x}) = \begin{cases} 0 & \text{if} \quad f_1(\underline{x}) \geq 1013.662 \\ 1 - \dfrac{(f_1(\underline{x}) - 980.8745)}{(1013.662 - 980.8745)} & \text{if} \quad 980.8745 \leq f_1(\underline{x}) \leq 1013.662 \\ 1 & \text{if} \quad f_1(\underline{x}) \leq 980.8745, \end{cases} \tag{39}$$

$$\mu_{f_2}(\underline{x}) = \begin{cases} 0 & \text{if} \quad f_2(\underline{x}) \leq 1001.465 \\ 1 - \dfrac{(1017.158 - f_2(\underline{x}))}{(1017.158 - 1001.465)} & \text{if} \quad 1001.465 \leq f_2(\underline{x}) \leq 1017.158 \\ 1 & \text{if} \quad f_2(\underline{x}) \geq 1017.158, \end{cases} \tag{40}$$

$$\mu_{f_3}(\underline{x}) = \begin{cases} 0 & \text{if} \quad f_3(\underline{x}) \leq 808.4835 \\ 1 - \dfrac{(1110.874 - f_3(\underline{x}))}{(1110.874 - 808.4835)} & \text{if} \quad 808.4835 \leq f_3(\underline{x}) \leq 1110.874 \\ 1 & \text{if} \quad f_3(\underline{x}) \geq 1110.874. \end{cases} \tag{41}$$

Based on the calculated membership functions the model (29-38) can be transformed into the following linear programming model:

$$(\max)\lambda \tag{42}$$

s. t.

$$\lambda \leq \mu_{f_1}(\underline{x}) \tag{43}$$

$$\lambda \leq \mu_{f_2}(\underline{x}) \tag{44}$$

$$\lambda \leq \mu_{f_3}(\underline{x}) \tag{45}$$

$$x_1 + x_2 + x_3 + x_4 = 4000 \tag{46}$$

$$0 \leq x_1, x_2, x_3, x_4 \leq 1500 \tag{47}$$

$$0 \leq \lambda \leq 1. \tag{48}$$

The model (42-48) is a linear programming model. By solving it we obtain the following optimal solution:

$$\lambda_{\max} = 0.6708, \quad x_1 = 987.7088, \quad x_2 = 12.2912, \quad x_3 = 1500, \quad x_4 = 1500,$$

$$f_1 = 991.6692, \quad f_2 = 1015.113, \quad f_3 = 1011.317.$$

The fuzzy technique applied in the model (42-48) solving does not take into account the subjective importance of criteria functions. In order to include the subjective importance of criteria functions for the decision maker we solve the model (28-33), where we determine the criteria weights: $w_1 = 0.40,\ w_2 = 0.40$ and $w_3 = 0.20.$ We obtain the following solution:

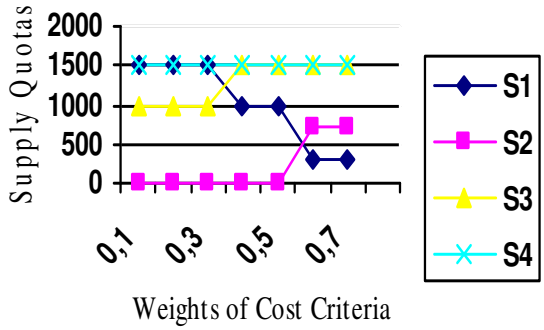
$\lambda_1 = 0.6667,\quad \lambda_2 = 0.8722,\quad \lambda_3 = 0.6791,\quad x_1 = 1000,\quad x_2 = 0,\quad x_3 = 1500,\quad x_4 = 1500,$

$f_1 = 991.8035,\ f_2 = 1015.152,\ f_3 = 1013.842.$

3.5. Sensitivity analysis

We will show sensitivity analysis of the quantities supplied by selected vendors according to changes in weights given to individual criteria by the decision maker. The selected vendors supply quotas as the consequence of increased weights in purchasing cost criterion with reduced weight in product quality criterion and keeping the weight in vendor reliability criterion at the level of 0.20 (case I) are shown in the following figure:

Figure 3. Vendors' supply quotas (case I)



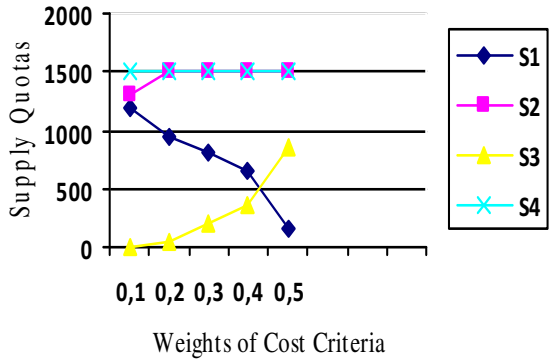
Source: own creation

It is obvious that the supply quota from the vendor S4 does not change no matter what the increase of cost criterion importance is which means that the vendor S4 is not sensitive to changes in cost criterion with simultaneous decrease in the quality criterion. The supply quota from this vendor remains 1500 t, which shows that the vendor S4 is the best in terms of both cost and quality criteria. The increase

of importance of the cost criterion from 0.10 to 0.50, with simultaneous decrease of importance quality criterion from 0.70 to 0.30 does not change the supply quota from the vendor S2 (it remains zero). However, the increase in importance of the cost criterion from 0.50 to 0.60 with simultaneous decrease of the quality criterion importance from 0.30 to 0.20 causes a significant increase of supply quota from this vendor, from 0 to 704.5 tons. The vendor S1 is negatively sensitive to the increase of cost criterion importance in the interval of 0.30 to 0.70, while the vendor S3 is positively sensitive to the increase of cost criterion importance from 0.30 to 0.40, whereby further increase of cost criterion importance does not affect this vendor's supply quota. The vendors S1 and S3 are not sensitive to the increase of cost criterion importance from 0.10 to 0.30, while the vendor S2 is not sensitive to the increased importance of the cost criterion from 0.10 to 0.50.

It is interesting to observe the changes in supply quotas caused by the changes in weights of cost and reliability functions with the constant weight of 0.40 for quality function (case II). The graph presenting these relations is shown in the Figure 4.

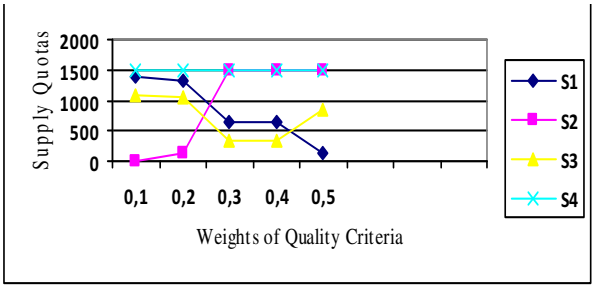
Figure 4. Vendors' supply quotas (case II)



Source: own creation

Changes in supply quotas caused by changes of weights in quality and reliability functions with the constant cost function weight 0.40 (case III), is shown in the following figure:

Figure 5. Vendors' supply quotas (case III)



Source: own creation

Observing the above figures we can conclude that the vendor S4 is not sensitive to changes in criteria functions importance coefficients. As this vendor is the best in terms of all the criteria we should consider increasing the supply quotas from this vendor. The supply quota from S2 increases from zero to 852.75 tons and in the case when the weight of cost function rises from 0.40 to 0.50 with the weight of quality function remaining at 0.40 and the weight of reliability function dropping from 0.20 to 0.10. In all the other cases this vendor's supply quota is equal to zero, which makes us conclude that we should avoid purchase from this vendor. The supply quotas from vendors S1 and S3 are sensitive to changes in criteria functions importance. However, the vendor S3 is positively sensitive to changes in criteria functions weights, which suggests that we should consider the possibility of increasing the supply quota from this vendor.

4. Conclusion

Solving the concrete example by application of the proposed methodology we can make a number of conclusions presenting the advantages of using the revised weighting method and FLP in comparison to the application of only one of them.

The revised weighting method allows efficient reducing of complex criteria functions into simple criteria functions. For DM, it is easier to determine weighting coefficients if he/she deals with few criteria functions than if he/she deals with a large number of them. If there are a large number of criteria and sub-criteria, there is a high probability of error in determining of weighting coefficients.

The weight coefficient's method applied alone has some shortcomings so that it is not the most appropriate one to create a set of Pareto optimal solutions. The shortcomings are: (1) varying weight coefficients do not guarantee that we will determine all Pareto optimal solutions, and (2) the determined Pareto optimal

solutions are those that are situated in the extreme points of the convex polyhedron but not those that connect the two extreme points. To determine the set of compromise solutions and the preferred solution it is better to use the fuzzy linear programming model.

When solving the MLP model the use of fuzzy technique proves to be very efficient. The efficiency of the fuzzy technique in solving the model can be seen in the possibility to define weights for criteria functions that express the decision maker's preferences. However, if you deal with complex criteria functions it is complicated to use the FLP method alone because of arising problems by determination of weighting coefficients.

Application of revised weighting method and FLP to solve the problem of vendor selection and determination of supply quotas allows a simple sensitivity analysis of the obtained solutions. The proposed methodology can be used in solving similar business problems.

References

Amid, A. - Ghodsypour, S. H. - O'Brien, C. 2006: Fuzzy multiobjective linear model for supplier selection in supply chain. *Int. J. Production Economics*, No. 104, 394-407. p.

Barbarosoglu, G. - Yazac, T. 1997: An application of analytic hierarchy process to the supplier selection problems. *Production and Inventory Management Journal*, *1st quarter*, 14-21. p.

Bellman, R. G. - Zadeh, L. A. 1970: Decision making in fuzzy environment, *Management Sciences* 17, 141-164. p.

Chan, F.T.S. - Kumar, N. 2007: Global supplier development considering risk factors using fuzzy extended AHP-based approach. *Omega: Int. J. Management Science*, Vol. 35, 417-431. p.

Gass, S. - Satty, T. 1955: The Computational Algorithm for the parametric Objective Function. *Naval Research Logistics Quarterly 2*, 39-45. p.

Ge Wang - Samuel, H.H. - Dismekes, J. P. 2004: Product driven supply chain selection using integrated multicriteria decision-making methodology. *Int. J. Production Economics*, Vol. 91, 1-15. p.

Ghodsypour, S. H. - O'Brien, C. 1998: A decision support system for supplier selection using an integrated analytic hierarchy process and linear programming. *Int. J. Production Economics*, Vol. 56, 199-212. p.

Haq, A. N. - Kannan, G. 2006: Fuzzy analytical hierarchy process for evaluating and selecting a vendor in supply chain model. *Int. J. Advanced Manufacturing Technology*, Vol. 29, 826-835. p.

Koski, J. - Silvennoinen, R. 1987: Norm Methods and Partial Weighting in Multicriterion Optimization of Structures. *International Journal for Numerical Methods in Engineering* 24, No. 6, 1101-1121. p.

Kumar, P. - Shankar, R. - Yadav, S. S. 2008: An integrated approach of Analytic Hierarchy Process and Fuzzy Linear Programming for supplier selection. *Int. J. Operational Research*, Vol. 3, No. 6, 614-631. p.

Kumar, M. - Vrat, P. - Shankar, R. 2004: A fuzzy goal programming approach for vendor selection problem in a supply chain. *Computers & Industrial Engineering*, Vol. 46, Issue 1, 69-85. p.

Kumar, M. - Vrat, P. - Shankar, R. 2005: A fuzzy goal programming approach for vendor selection problem in a supply chain. *Int. J. Production Economics*, Vol. 101, 273-285. p.

Lin, C. R. - Chen, H. S. 2004: A fuzzy strategic alliance selection framework for supply chain partnering under limited evaluation resources. *Computers in Industry 55, 2*, 159-179. p.

Miettinen, K.M. 1999: *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Dordrecht.

Narasimhan, R. 1983: An analytic hierarchy approach to supplier selection. *Journal of Purchasing and Materials Management*, Vol. 19, 27-32. p.

Nydick, R. L. - Hill, R. P. 1992: Using the analytical hierarchy process to structure the supplier selection procedure. *Int. J. Purchasing and Materials Management*, Vol. 28, 31-36. p.

Perić, T. 2008: *Multi-criteria Programming - Methods and Applications (in Croatian)*. Alka script, Zagreb.

Perić, T. - Babić, Z. 2009: Determining Optimal Production Program with a Fuzzy Multiple Criteria Programming Method. *Proceedings of International MultiConference of Eneireers and Computer Scientists IMECS 2009*, 2006-2013. p.

Sakawa, M. 1993: *Fuzzy Sets and Interactive Multiobjective Optimization*. Plenum Press, New York.

Tiwari, R.N. - Dharmahr, S. - Rao, J. R. 1987: Fuzzy goal programming – an additive model. *Fuzzy Sets and Systems 24*, 27-34. p.

Weber, C. A. - Current, J. R. - Benton, W. C. 1991: Vendor selection criteria and methods. *European Journal of Operational Research*, Vol. 50, 2-18. p.

Weber, C.A. - Ellram, L. M., 1992: Supplier selection using multi objective programming: a decision support system approach. *Int. J. Physical Distribution and Logistics Management*, Vol. 23, 3-14. p.

Weber, C. A. - Desai, A. 1996: Determination of paths to vendor market efficiency using parallel co-coordinators representation: a negotiation tools for buyers. *European Journal of Operational Research*, Vol. 90, 142-155. p.

Weber, C. A. - Current, J. R. - Desai, A. 1998: Non co-operative negotiation strategies for vendor selection. *European Journal of Operational Research*, Vol. 108, 208-233. p.

Zadeh, L. 1963: Optimality and Non-Scalar-valued Performance Criteria. *IEEE Transactions on Automatic Control 8*, 59-60. p.

Zimmermann, H. J. 1978: Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and System 1*, 45-55. p.

Zimmermann, H. J. 1987: *Fuzzy Sets, Decision Making and Expert Systems*. Kluwer Academic publishers, Boston.

Zimmermann, H. J. 1993: *Fuzzy Sets Theory and its Applications*, Kluwer Academic Publishers, Boston.

Using Massively Parallel Processing in the Testing of the Robustness of Statistical Tests with Monte Carlo Simulation

Tamás Ferenci[1] *– Balázs Kotosz*[2]

In this paper, we will examine the application of the Monte Carlo method in the testing of the robustness of statistical tests. The very computation-intensive Monte Carlo testing was implemented using the so-called GP-GPU method, which utilises the video cards' GPU (Graphical Processing Unit) to perform the necessary calculations.

The robustness of a statistical test is defined as its property to remain valid even if its assumptions are not met. (We call a test valid if its significance level is equal to its Type I Error Rate.) One way to investigate the robustness of a statistical test (especially useful if the test's complicated structure makes analytic handling infeasible or impossible) is to employ the so-called Monte Carlo method. Here, we generate many random samples (meeting or violating the assumptions, which we can arbitrarily set), perform the statistical test many times on them, and then check whether the empirically found Type I Error Rate converges to the specified significance level or not. This way, we give up exactness for the complete insensitivity to the complexity of the examined statistical test.

This method (by requiring enormous amount of random number generations and statistical testings for reliable results) is very computation-intensive; traditionally only supercomputers could be used effectively, limiting the availability of this method. However a new approach, called GP-GPU, makes it possible to harness the extremely high computing performance of – even ordinary, widely available – video cards found in every modern PC.

We implemented a framework that performs the abovementioned MC-testing of the robustness of statistical tests. We call our program "framework", because it can be easily expanded with virtually any statistical test (to which no GP-GPU knowledge is needed), and be tested with very high performance – even with widely available tools.

As an example, we performed the analysis of the well known Student's t-test; and – using this as a starting point – we demonstrated the main advantages of our framework.

Keywords: robustness of statistical tests, Monte Carlo method, GP-GPU, simulation

[1] Tamas Ferenci, MSc, Corvinus University of Budapest, Faculty of Economics, Department of Statistics (Budapest)
[2] Balazs Kotosz, PhD, assistant professor, Corvinus University of Budapest, Faculty of Economics (Budapest)

1. On the robustness of statistical tests and its empirical investigation with Monte Carlo simulation

1.1. Introduction

Statistical hypothesis testing, a method of inductive statistics, is of central importance in both theoretical and applied statistics (Hunyadi–Vita 2006). Since its introduction by Ronald Fisher, Jerzy Neyman and the two Pearsons in the early 20th century (Hald 1998) it became a cornerstone of modern statistics, widely used in nearly every area of nature-, life- and social sciences and technological practice. (Despite receiving challengers, mainly the Bayesian approach – for an introduction see (Lee 2009) – in the late 20th century.)

Statistical hypothesis testing is performed using so-called *statistical tests*, of which hundreds (Hunyadi 2001) have been developed throughout the decades for a variety of purposes. These tests are usually based on *assumptions*, typically regarding the populations from which the samples are coming on which the test operates. The test behaves the way it was specified only if these assumptions are met.

To illustrate the abovementioned, Table 1 shows some statistical tests for testing the equality of the means (expected values) of two populations, from which independent samples are available. Table 1 also gives the assumptions of these statistical tests.

Table 1. Selected statistical tests for testing the equality of population means (expected values)

Name of the test	Assumptions
Student's *t*-test	Populations normally distributed, having equal variances
Welch-test	Populations normally distributed
Mann-Whitney *U*-test	Populations arbitrarily distributed, but PDFs having the same shape
Brunner-Munzel test	None

Source: own creation

As it can be well seen, tests being lower in Table 1 have fewer assumptions. This is not a universal advantage however: these tests also have less statistical power. (So, when selecting a statistical test for a purpose[3], it has to be carefully considered, which are the assumptions that can be either a priori accepted or tested (preferably on an independent sample), and select a test that is based *only* on those – no less (loss of power), no more (wrong behaviour).)

[3] See (Ferenci 2009b) for a brief discussion of these questions in connection with a (biostatistical) application.

1.2. Validity and robustness of statistical tests

In the previous subsection, we said "The test behaves the way it was specified only if these assumptions are met". This can be made more precise as follows.

A test is guaranteed to be *valid* only when its assumptions are met. The validity of a statistical test is defined as its property of having Type I Error Rate equal to its significance level.

When a statistical test's assumptions are not met, it is no longer mathematically guaranteed that the test is valid. However, statistical tests can show quite different behaviours when being operated on samples violating the assumptions: some tests still remain valid to some extent, even if their assumptions are not met. This property of a statistical test is called *robustness*; we call a statistical test *robust*, if it has this property.

Validity can be obviously checked not only when the assumptions are met, but also when they are violated (logically it needs to be – necessarily – confirmed only in the former case). Hence the checking of robustness can be viewed as an extension of the checking of validity: we can conclude on robustness by performing many checkings of validity (with assumptions violated to different extent).

1.3. Using Monte Carlo (MC) method to examine the robustness of a statistical test

One way to check the robustness of a statistical test is to analyse its mathematical structure: suppose some assumption-violating property of the samples (which can be formalised mathematically), expose the test to such samples, and – still purely mathematically – derive how the test would operate on those samples. (E.g. what Type I Error Rate would it produce with a given significance level.)

The advantage of this method is that it results in *exact analysis*. However, there is one serious limitation: it becomes increasingly *infeasible* (if not impossible) when used on tests that have more and more complex algebraic structure – typical for today's statistical tests. This might even make exact, mathematical analysis impossible.

An alternative approach is to use the so-called *Monte Carlo* (MC) method for checking robustness: it gives up the exactness, but in exchange it provides a testing method which is completely insensitive to the complexity of the examined statistical test (Rubinstein–Krose 2008).

MC method (which is a general approach used in many other areas of scientific research) is based on the idea of *empirically exploring* a system: instead of describing the system with rigorous mathematical formalism, it is investigated stochastically, i.e. the system is faced with many randomly selected inputs, and its "operation" is reconstructed by the outputs it generates as an answer to the random inputs. Consequently, this is a stochastical approach, and it can be effective if a very high number of random inputs are given, so that the system can be mapped appropriately.

The MC method is employed in applications where deterministic examination is infeasible or impossible due to the complexity of the problem. (A classical example for the application of MC is high dimensional numerical integration: while it is almost impossible to perform a numerical integration in 100 dimensions with classical mathematical tools (10^{100} points would be needed just to have 10 integration point on each axis!), it is well possible with MC, in which points are randomly selected in the 100 dimensional space and then checked whether they are "above" or "below" of the function – this avoids the exponential increase in the number of required points.)

In our context (checking the validity/robustness of statistical tests), this means the following. We generate a random sample and then apply the statistical test. Then, we record the result (accept or reject) and repeat this procedure ("test-cycle") many times. Thus, we obtain a Type I Error Rate, which we can compare to the significance level: as stated, the test is valid if these two are equal. An important side note is however, that "equality" has to be understood in its statistical sense: as Type I Error Rate was determined empirically (i.e. by testing samples), this rate will be exposed to statistical fluctuation. (Caused by the fact that the rate depends on which samples we actually test.) When comparing this with the fixed significance level, we can never say "surely" that they are equal; rather we can only draw a statistical conclusion. (Using another statistical test, as a matter of fact…) However, we can very well aim to draw a conclusion that they are "extremely highly likely" to be equal. This can be achieved by using very high number of random samples. Thus, we actually examine whether the empirically found Type I Error Rate *converges to* the significance level as the number of test-cycles increases.

This operation makes it understandable why it is also called MC-simulation.

A critical aspect of this approach is that we have full control on the properties of the samples. (As we generate them, we can arbitrarily parameterise the random number generator algorithm.) Hence for any statistical test, we can generate samples that meet the test's assumptions, and samples that do not. From this point, the completion of the MC method for checking validity/robustness is quite straightforward: if we generate samples that *meet* the investigated test's assumptions, we can test validity, if we generate samples that *do not meet* these assumptions, we can test robustness.

Moreover, if the extent, to which the assumptions are not met, can be quantified, we can also *iterate through* different levels of departure from the assumptions, hence "mapping" the robustness of the test. (A prime example for this is the violation of normality, which can be well quantified with skewness and kurtosis. We will return to this question later in this paper.)

1.4. Our demonstrational example: checking the robustness of the t-test for non-normality

We chose Student's two independent samples *t*-test as a *demonstrational example* for MC-simulational testing. It is solely a demonstration: the results obtained will not be of scientific value, as this test has been completely analysed long ago – the method, the MC-simulational testing itself is interesting now. And for demonstration, it seemed logical to choose a simple, well-known statistical test (Hunyadi–Vita 2006). This is in fact the most well-known statistical test, and perhaps also the most widely used ever developed. It can be used to check the equality of the means (expected values) of two populations, from which independently drawn samples are available.

The test assumes (among others) that the populations are *normally distributed*. (See also Table 1.) Our aim will be to check robustness in respect to this assumption. (I.e. we will let every other assumption be met, but violate this one.) Furthermore, we will quantify the violation of this normality assumption (by the skewness and kurtosis of the population's distributions), so we will also be able to "map" the robustness: to check how validity alters when the test's assumptions are more and more violated. Thus we will be able to answer the question: how valid is the *t*-test when the populations are non-normal, i.e. how robust it is to the violation of population normality?

1.5. Problem statement

As already stated, we will use MC-simulation to answer this question. In this context, it means that we will randomly generate many samples coming from distributions having specified skewness and kurtosis, and test them. By recording empirical Type I Error Rate, we can judge the validity of the test. And by iterating through different levels of skewness and kurtosis, we can test robustness.

The problem is that this is very *time-consuming*, even on modern personal computers. This can be traced back to the fact that even modern PCs have inadequate computing performance when using very high number of test cycles. (Which is needed however to achieve low statistical fluctuation in empirical Type I Error Rate, a question which we already discussed.) If multiple parameters govern the level of violation of the assumptions (two in our demonstrational example: skewness and kurtosis), and we want to check *every possible combination* of them, it will also exponentially raise the number of needed tests.

In this paper we use a novel approach to handle this situation. Traditionally, the only resolution would have been to employ a supercomputer, grid computing etc., available only to a few researchers. Now, we will show a solution which is *widely available* (both technically and financially), but still produces *impressive performance* (sometimes magnitudes higher computing performance than the CPU!),

and can be used to MC-simulational testing of the robustness of statistical tests. We will discuss this in Section 2.

2. On GP-GPU computing and its usage in the MC-simulational examination of statistical tests

2.1. Introduction

Since the introduction of the category in the early 1980's, *video cards* (Cserny 1997) constitute a regular part of Personal Computers (PCs). These devices (typically built as a standalone expansion card, communicating with the host computer through some standardised bus) are responsible for generating visual output from the computer (typically displayed on monitors).

Since the mid-1990's, an increasing demand appeared for video cards that can perform image-generation related calculations themselves, instead of relying on the computer's central processing unit (CPU), thus relieving that. This was especially true when 3D image generation became a commonplace (due to computer games, above all). These required tremendous amount of calculations, but only from a very few types, so it seemed logical to design a hardware *specifically* for that purpose (instead of loading the general purpose CPU with this task), which is usually named *hardware acceleration*. Hence hardware for this purpose, so-called Graphical Processing Units (GPUs) began to widely spread – typically VLSI (Very Large Scale Integration) chips, designed specifically to accelerate image generation-, especially 3D image generation-related calculations.

The early-2000s saw an enormous development in the field of GPUs. They became more and more complex (as an illustration: the number of transistors (Hagerdoorn 2009) grew from 1 million (Voodoo, 1996) to 1.5 billion (NVidia GT200, 2009)!), they were able to perform more and more sophisticated tasks in the field of hardware acceleration. Because these were highly specialized designs (as we already mentioned, only a few types of calculations have to be done by a GPU, compared to the general purpose CPU), their *processing power* grew in a much faster rate than that of CPUs. (As only a few types of calculations have to be done… but these have to be done very fast.)

This, along with the fact, that the GPU's more and more complex structure also widened the range of possible calculations, rose up the idea of using the GPU for applications that are *non-graphical*, but require *high computing capacity*. This led to a technique called General Purpose GPU, or GP-GPU (Owens et al. 2007). Note however that the term "general" is a bit misleading: it is not "general" in the sense we named CPU so. While theoretically GP-GPU in fact aims to run any program on GPU, practically this is only useful when the program fits to the structure of the GPU.

The key point in that structure is the *very highly parallel* nature: a GPU consists of processing units which are very simple themselves, but a very high number of them are used. Thus logically those programs can be effectively recoded to run on GPU (with the GP-GPU approach) which can be well parallelized… just like Monte Carlo simulation! (In addition to that, a range of other tasks might be solved effectively with GP-GPU, from signal processing to cryptography.) For these tasks, GPUs can be used essentially as "supercomputers".

This approach can be related to the Massively Parallel Processing (MPP) concept in the theory of computer architectures.

2.2. Usage of GP-GPU in the MC-simulational examination of statistical tests

It has to be strongly emphasized, that even cheap, widely available video cards (more precisely: their GPUs) can provide *extremely high* computing performance for applications that can utilise it. As we just pointed it out, Monte Carlo simulation is a prime example for this, so we decided to undertake this (i.e. GP-GPU) approach to resolve the problem of obtaining high computing performance (which is needed in MC-simulation), but still creating a solution available to virtually any interested user.

The video card we selected was based on NVidia's GeForce 9600 GT GPU (from the GeForce 9 series), based on the G94a/b GPU core (NVidia 2009a). This is a low-middle class video card, sold at around 20 000 HUF (available at any online store at around 100 €). It

- has 505 million transistors,
- a core clock rate of 650 MHz,
- 64 stream processors, each running at 1 625 MHz,
- memory bandwidth of 57,6 GB/sec, and
- a theoretical *processing rate* of 312 GFlops!

The last feature is perhaps the most important for us at first glance: it tells that the GPU is (theoretically) able to perform more than 300 billion floating point operations per second!

2.3. The usage of CUDA in developing GP-GPU programs

Even if the video card and the concept of the program to be run on the GPU are provided, it is still a question how we can in fact code a program to run on GPU.

In the first era of GP-GPU, programs had to be written to "mimic" the graphical operations, making development very complicated. (Programs have to be written "as if" they were performing graphical operations, as GPUs only understood such (i.e. graphics-related) commands.) However, GPU manufacturers soon realised the power of GP-GPU and began developing architectures which more naturally supported the development of non-graphical programs for GPUs. (This tendency be-

came so pronounced, that lately even such "video" cards have been released that were specifically designed for GP-GPU, lacking even video output.)

NVidia's such architecture is called *Compute Unified Device Architecture* (NVidia 2009b), or CUDA for short. This is not only a programming language, but a complete environment for GP-GPU. Naturally it also includes a programming language, called "C for CUDA" which can be used to develop programs for GPU. This is the approach we employed for coding our MC-simulational program; an excerpt from the code is shown on Figure 1.

Figure 1. Excerpt from the CUDA program performing MC-simulation

```
iState1 = iState + 1;
iStateM = iState + MT_MM;
if(iState1 >= MT_NN) iState1 -= MT_NN;
if(iStateM >= MT_NN) iStateM -= MT_NN;
mti  = mti1;
mti1 = mt[iState1];
mtiM = mt[iStateM];

x    = (mti & MT_UMASK) | (mti1 & MT_LMASK);
x    =  mtiM ^ (x >> 1) ^ ((x & 1) ? config.matrix_a : 0);
mt[iState] = x;
iState = iState1;

//Tempering transformation
x ^= (x >> MT_SHIFT0);
x ^= (x << MT_SHIFTB) & config.mask_b;
x ^= (x << MT_SHIFTC) & config.mask_c;
x ^= (x >> MT_SHIFT1);

x2=((float)x + 1.0f) / 4294967296.0f;

/*--- Box-Mueller transzformacio ---*/
r = sqrtf(-2.0f * logf(x1));
phi = 2 * PI * x2;
x1 = r * __cosf(phi);
x2 = r * __sinf(phi);
/*--- Box-Mueller transzformacio ---*/

/*--- Fleishman polinomialis transzformacio ---*/
x1=params[StartTid+blockIdx.x].a+params[StartTid+blockIdx.x].b*x1+params[StartTid+blockIdx.x].c*x1*x1+params[StartTid+blockIdx.x].d*x1*x1*x1;
x2=params[StartTid+blockIdx.x].a+params[StartTid+blockIdx.x].b*x2+params[StartTid+blockIdx.x].c*x2*x2+params[StartTid+blockIdx.x].d*x2*x2*x2;
/*--- Fleishman polinomialis transzformacio ---*/
```

Source: own creation

2.4. On our testing framework

From this point on, we will refer to our program as "*framework*". This is to emphasize that the most important part of our work is not the coding of statistical tests, but the coding of the MC-simulation robustness-checking module. (This also foreshadows that we took huge effort to write a *modular program*, i.e. a program in which the MC-simulational part and the description of the statistical test itself is *separated*. We will discuss this question in detail in the next section.)

In its current version, our program includes two statistical tests (Student's two independent sample *t*-test, and the Mann-Whitney *U*-test, as we will later discuss), which may be used as demonstrational subjects to the MC-simulational module. (Results from these investigations will be shown in Section 5.)

Another advantage of the "framework" approach is that new statistical tests might be *added* by a description in standard C language – no CUDA or parallel computing knowledge is needed at all.

3. Technical details on the coding of our testing framework

As we explained in the previous section, our investigations were not purely theoretical: we also developed a "testing framework" as a program which can be used to check the robustness of statistical tests with the described Monte Carlo simulation method. In this section, we will review a few considerations that cropped up during the actual implementation of this program.

3.1. Development environment

We developed the program under Microsoft's Visual Studio 2008 integrated development environment (Microsoft 2009).

As we have already mentioned, NVidia's CUDA architecture uses the so-called "C for CUDA" programming language, which is a minimal extension to the standard C language to support the coding of the GPU (NVidia 2009b). This can be comfortably handled within the MS Visual Studio environment, where coding, debugging and compiling can be performed at the same interface.

3.2. Random number generation

The key idea of Monte Carlo methods is to use random inputs (in our case: random samples) to explore the behaviour of the system (in our case: the statistical test). Thus, it is necessary to generate random numbers (lots of random numbers actually, as this has to be repeated many times) in our program.

Random number generators (RNGs) have a library-wide literature – for a relevant review, see for example (Gentle 2004) – and random number generation in a parallel environment is an even more non-trivial question.

We used the RNG employed in the CUDA's own Software Development Kit (SDK), the so-called Mersenne Twister algorithm (Podlozhnyuk 2007). It is one of the most widely used RNGs nowadays, mathematically guaranteed to be equidistributed up to 623 dimensions (Matsumoto–Nishimura 1998), passing all modern randomness test, including the diehard (Matsumoto–Nishimura 1998) and the even more stringent TestU01 (McCullough 2006). This – together with its speed – makes Mersenne Twister especially fit for Monte Carlo simulations.

However, in itself, it is not suitable for use in parallel environment. This was resolved by the so-called DCMT algorithm, which is also employed (Matsumoto–Nishimura 2000).

The Mersenne Twister with DCMT produces uniformly distributed numbers. As we will see in the next subsection, we need random numbers from standard normal distribution for our purpose, so we applied the well known Box-Mueller transform (Ketskeméty 2005).

3.3. Fleishman's polynomial transformation method

As we have already pointed it out, the departure from the assumptions of the test will be quantified by the level of non-normality in our demonstrational example (logically, as the assumption was normality in that example). To quantify non-normality, we will apply the most widely used indicators: *skewness* and *kurtosis* (i.e. the third and fourth standardized central moments, but see Subsection 3.6 as well).

Note that distributions can be viewed as points on the skewness/kurtosis space (or plane, to be more specific).

If we recall the procedure of MC-testing the robustness (Section 2.3), it is immediately obvious that we will need to generate samples from distributions having arbitrary kurtosis and skewness. (In other words we have to iterate through the points (distributions) of the skewness/kurtosis space, and generate samples at every point.) This is a nontrivial problem, as the widely used distributions either have given kurtosis and/or skewness (like normal or exponential distribution) or only one parameter[4] governs kurtosis and skewness (like gamma or lognormal distribution).

There are many solutions for this problem. We might well go back to Karl Pearson when reviewing the history of this question, as the system of Pearson distributions (developed in the early 20th century) can be viewed as an answer to this problem. Since then, many other solutions have been developed, more or less fit for our purpose. (As an example: the Pearson distribution is largely unfit, as it involves distributions that have diverse algebraic structure, so the generating algorithm would have to change between essentially different distributions when iterating through the skewness/kurtosis space.)

The solution we employ now was published in 1978 by Allen I. Fleishman (Fleishman 1978). As it might be unknown to the reader, we will discuss it in detail.

The key idea behind Fleishman's solution is the following. When generating samples from a standard normal distribution we can control neither the mean, nor standard deviation, nor the skewness, nor the kurtosis of the distribution. However, if we add a constant to every generated sample, we can arbitrarily set the mean of the distribution. Similarly, by multiplying the samples we can arbitrarily set the standard deviation of the distribution. (With this, we essentially switched from standard normal distribution to general normal distribution.) But the multiplication and the adding of a constant is simply a linear transformation, a subclass of polynomial transformation. The crucial point is to note that when the transforming polynomial has an order of zero (i.e. adding of a constant), we can only set the mean (i.e. the first moment), when it has an order of one (i.e. multiplication and adding of a constant), we can set the mean and the standard deviation (i.e. the first and second moments). One would expect that by transforming the sample with a polynomial having

[4] Obviously, in this case the two indicators cannot be set independently, to the contrary: setting one of them automatically defines the other.

an order of two (i.e. quadratic term), we would be able to arbitrarily set the first three moments, that is, mean, standard deviation and skewness.

To investigate this, Fleishman analytically derived the first four standardized central moment of the distribution generated with the $Y = a + bX + cX^2 + dX^3$ transformation, where $X \sim N(0;1)$. (Obviously, he used a polynomial having an order of three, as the aim was to set the first four moments arbitrarily.) He obtained the following equations for the mean, variance, and μ_3 and μ_4 indicators of skewness and kurtosis:

$$\mu = a + c$$

$$\sigma^2 = b^2 + 6bd + 2c^2 + 15d^2$$

$$\mu_3 = 2c\left(b^2 + 24bd\right) + 105d^2 + 2$$

$$\mu_4 = 24\left(bd + c^2\left[1 + b^2 + 28bd\right] + d^2\left[12 + 48bd + 141c^2 + 255d^2\right]\right)$$

Now we can answer the previous question: this power transformation method can be used to generate a distribution with arbitrary μ_3 skewness and μ_4 kurtosis if and only if the above system of equations can be solved for the μ_3 / μ_4 skewness/kurtosis in question. (Mean might be chosen to 0 and variance to 1, without loss of generality. Our figures will follow this convention, i.e. we will only plot skewness and kurtosis on them; mean will be assumed to be 0, variance to be 1, everywhere.)

It can be demonstrated that as a result we obtain a skewness/kurtosis plane that may be generated (i.e. where the above system of equations may be solved) being quite close to the theoretically possible[5] one.

Another favourable aspect of Fleishman's method is that it only requires random numbers from standard normal distribution – exactly what we obtain from our random number generator.

Clearly, the only complicated part is the solution of the above system of equation. Although it is a complex, non-linear system of equations, and can be solved only numerically (not analytically), it can luckily be solved off-line: the necessary coefficients can be calculated before the MC-simulation. This is important, as the solution of the equation is time-consuming. To sum up, the correct approach is to solve these equations for every possible skewness/kurtosis combination we will need, store the results and then give them to the simulation routine as a constant.

We used the GNU Scientific Library (GSL) to solve the non-linear system of equations. It is a free C library designed to support mathematical/scientific computa-

[5] There are skewness/kurtosis combinations which do not represent a distribution at all. In other words, correct probability distributions cannot have arbitrary skewness and kurtosis, namely $\mu_4 \geq \mu_3^2 + 1$ stands for every probability distribution. See (Ferenci 2009a) for more detail.

tions, developed under the GNU project, and licensed under GNU GPL (GNU 2009). We employed version 1.13.

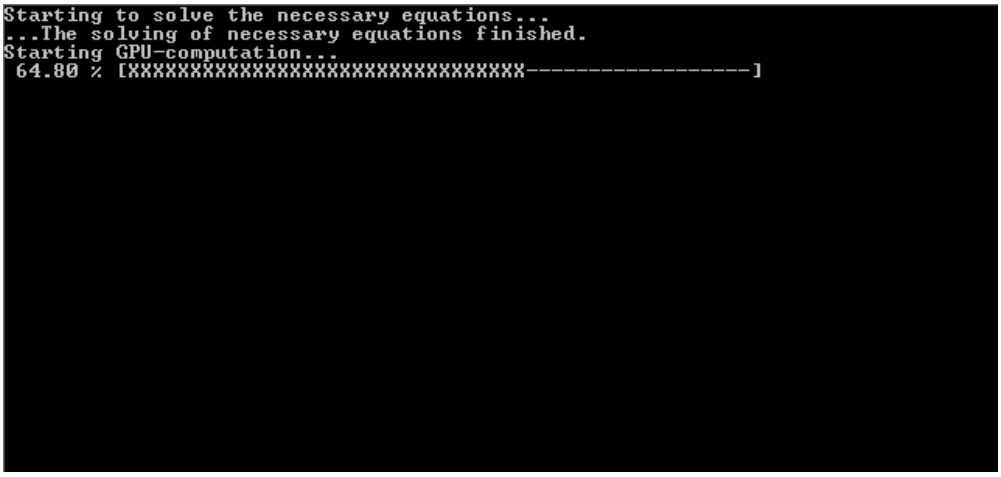
We used `gsl_multiroot_fsolver` to solve the system of equations[6] with initial values set to (0.0,1.0,0.5,0.5). (These were chosen to ensure that in normal case, the solution would be in fact the trivial transformation.) We used a simple solver algorithms (not using the derivatives, obviously), the so-called discrete Newton algorithm (`gsl_multiroot_fsolver_dnewton`). The solution was naturally iterative; we defined to stopping criteria: the number of iteration steps exceeds 1 000 (this was a backup, normally never active), and the residual is less then 10^{-5}.

3.4. User interface

As our program is computation-oriented with minimal user interaction, we designed a rather simple character-based user interface. (This also made coding easier, as we could write the program as a simple console application.)

The user interface has a *dual purpose*. First, it continuously informs the user on the progress of the computations. (This is especially important when calculations take more than minutes.) The user receives information on the current activity, and – in case of the most time-consuming GPU-computation – is also informed on the progress of that activity (with a "percentage complete" type feedback). The screenshot on Figure 2 shows this state of the user interface.

Figure 2. Screenshot of our program during the calculation phase



Source: own creation

[6] Multiroot stands for multidimensional root-finding.

The other aim of the user interface is to provide summary statistics when the calculation is finished. This is shown on Figure 3.

These statistics detail the parameters of MC-simulation (checked skewness and kurtosis levels, sample size etc.), the number of random number generations based on this and the time needed to complete the simulation. Based on the two latter information, the program displays the estimated speed of the computation in hypotheses testing/sec. It also provides a minimum estimate[7] for the speed of random number generation (in random numbers generated/sec).

Figure 3. Summary statistics provided by our program at the end of calculations

```
Starting to solve the necessary equations...
...The solving of necessary equations finished.
Starting GPU-computation...
100.00 % [XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX]
...GPU-computation finished.
Starting to write results to storage device...
...The writing of results to storage device finished.

Summary of generation:
20 levels of skewness (interval: 0.0 - 4.0, step size: 0.20)
20 levels of kurtosis (interval: 0.0 - 10.0, step size: 0.50)
10 is the sample size
10000000 sample-pair in each skewness/kurtosis combination
That is altogether 20 * 20 * 10 * (10000000 * 2) = 80000000000 random number gen
erations...

... and 20 * 20 * 10000000 = 4000000000 hypotheses testing.

The 80000.00 million random number generation and the 4000.00 million hypotheses
 testing was done in 388.898 seconds (that is approximately 0h6m).
That means the speed is: 10.29 million hypothesis testing/sec
Hence, random number generation itself had the minimum speed of: 205.71 million
random number generation/sec

End of program, press any key to exit!_
```

Source: own creation

No result is displayed on the console: the program saves every numerical result to file on the hard disk.

3.5. Visualisation

As we previously mentioned, the program saves the results as plain-text files. Although these can be viewed with any text editor (and their content interpreted), it makes analysis much simpler if we visualise their content. In other word, further (post)processing is needed, which we fulfilled with R scripts (R 2009) that we have written specifically for the purpose. We used R version 2.9.1; every figure showing results in this paper was generated this way.

[7] While the speed of hypothesis testing is a real, trustworthy indicator (as the random number generations form a part of hypothesis testing in this context), the speed of random number generation is well underestimated this way (as the time also includes calculations not related to random number generation, e.g. the performing of the statistical test).

3.6. On the measuring of kurtosis

Kurtosis – as a concept – is a complex abstraction, and its measuring is not at all unambiguous. (See (Hunyadi 2009) on this question.)

Classically the fourth standardized central moment (i.e. $\frac{E\left[\left(X-E[X]\right)^k\right]}{\sigma^k}$ for $k=4$) is used to measure the kurtosis of a distribution. This takes the value of 3 for normal distribution, so sometimes the so-called excess kurtosis is also introduced:

$$\frac{E\left[\left(X-E[X]\right)^4\right]}{\sigma^4}-3$$

This way, positive (excess) kurtosis indicates a leptokurtic, negative indicates a platykurtic distribution.

For our goals, these definitions are not fortunate. As we will generate distributions with substantial skewness, we cannot forget about the theoretical minimum of skewness. (See Footnote 5) When skewness equals just 5, the minimum possible kurtosis is 26 (compare to the minimum value of 1, in case of distributions that are not skewed at all). Should we use these kurtosis-measures, graphical representation would be overly complicated, as kurtosis would span over – literally – magnitudes. Interpretation would be also rendered more difficult, as the results for the higher values of skewness would be "shifted up". (Quite drastically, due to the quadratic nature of the theoretical minimum.)

A natural solution would be to subtract the theoretical minimum (determined utilizing the skewness) from every kurtosis value, in other words to switch from "kurtosis" to "kurtosis above the theoretical minimum". Formally this suggests a new kurtosis-definition that has the form $\mu_4-\left(\mu_3^2+1\right)$. (Graphically, this means to subtract the μ_3^2+1 parabola from every μ_4 coordinate.)

However, instead of that we chose to use the

$$\mu_4-\left(\mu_3^2+3\right)$$

new kurtosis definition. We will name this "FK-kurtosis".

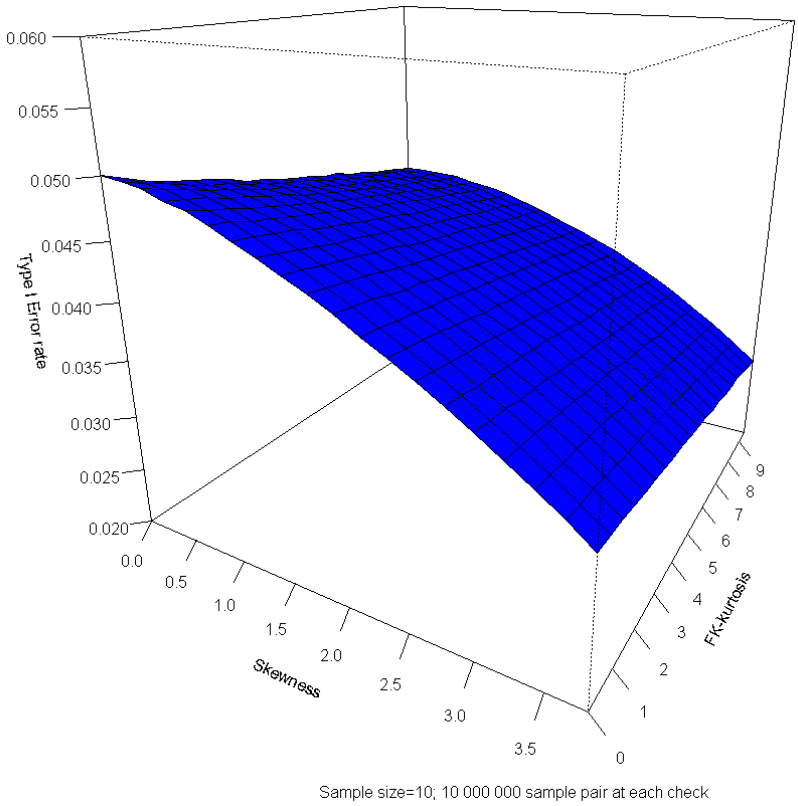
This has two advantages over the previous one. First, those distributions that have a positive FK-kurtosis can be surely generated by Fleishman's method, as it has problems[8] only with those distributions that are close to the theoretical minimum. Second, we ensure this way that the "zero skewness, zero kurtosis" point belongs to the normal distribution (just as with excess kurtosis).

[8] I.e. the necessary system of equations cannot be solved, see Subsection 4.2.

4. Results

Figure 4 shows the central result for our demonstrational example: the robustness of the Student's *t*-test for the violation of the normality assumption.

Figure 4. Robustness of the *t*-test for non-normality



Source: own creation

On the *z*-axis, we indicated the empirically found Type I Error Rate, while the *x*-*y* plane is a representation of the skewness/kurtosis space. (Using FK-kurtosis, as described in Subsection 3.6.)

One can clearly see that in case of normal distribution (zero skewness, zero excess kurtosis) the empirically found Type I Error Rate is almost exactly 0.05, the significance level. This confirms that the test is valid if its assumptions are met.

It is interesting to see how the empirically found Type I Error Rate deviates from the specified 0.05 as the samples become more and more non-normal. (To be

more precise: the populations become more and more non-normal from which the samples are obtained.) This suggests that the test is not robust for the violations of the normality assumption: it loses its validity in case of samples where that assumption is not met. But we can be even more specific: it can be clearly seen that the test is far more sensitive for skewness than for kurtosis.

These results are all consistent with the literature data (Vargha 1996).

Note how *smooth* the figure is! This illustrates that the extremely high number of sample-pairs generated at each skewness/kurtosis point diminished the statistical fluctuation (due to the sampling error[9]) to almost zero. This is why very high computing performance was needed, what could be reached only with supercomputers, grid computing etc. traditionally, but what became widely available with our method based on GP-GPU.

As we always emphasized, the above results are uninteresting themselves. (These have been known virtually since the *t*-test has been introduced; long before not only MC-simulation, but also computers.) What is important: the *performance* we could reach when generating the results. (Because these are not scientific results nowadays, but our aim was to develop a framework for testing – that's the reason why we always called it only a demonstrational example. What we have to analyze is the framework itself, not the obtained results.) So now we will have a closer look on performance data.

4.1. Performance of the MC-testing

In the demonstrational example, we performed tests with populations having skewness from 0.0 to 4.0 (with 0.2 step size) and excess kurtosis from 0.0 to 10.0 (with 0.5 step size). It is important to note that we tested every possible combination of them (in other words, we completely "mapped" the skewness/kurtosis space, we iterated through every possible combination); hence, we had to perform $20 \cdot 20 = 400$ checking of validity.

For each checking we generated 10 million sample pairs. This meant $400 \cdot 10 = 4000$ million, or 4 billion hypothesis testing. Taking into account that the sample size was 10 (and sample-pairs were generated), it meant $4 \cdot 10 \cdot 2 = 80$ billion random number generation.

And this all was done in 389 seconds – again: on our 100€, middle-class, available-from-every-shop video card. In terms of speed, this means in excess to *10 million hypotheses testing per second*! (Purely the RNG itself has to produce a speed greater than 200 million random numbers generated a second.)

According to our measurements (performed by executing the very same simulation on both GPU and CPU), the GP-GPU computation is about *38 times faster*

[9] As such, this could be quantified, but the standard deviation due to sampling would be likely much smaller than even a single pixel.

than traditional CPU computation – even though the CPU in the comparison was a modern, dual-core processor.

This performance of GP-GPU is well enough for every common task, even if multiple parameters describe the departure from the assumptions and we have to iterate on them in a combinatorial way (i.e. using every possible combination; just as in our example).

4.2. On the flexibility of our framework

We already emphasised that *flexibility* was one of the primary priorities when designing our framework; now we will show two demonstrations for this principle.

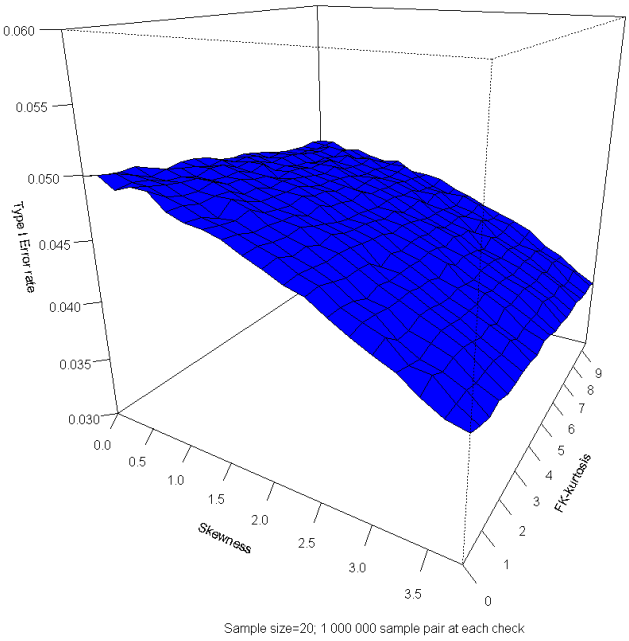
Perhaps the most important manifestation of this effort is the modularity of our program: we tried to isolate the description of the statistical test and the MC-testing parts as much as possible. This has two advantageous aspects: first, one can change the parameters of the testing by simply altering a few constants in the program; second, it is possible to add other statistical tests to be tested with leaving every other part of the program unchanged.

This means that our program is *highly scalable*: virtually any other test can be added (and, as we will see in 4.2.2 in more detail, there is no CUDA knowledge needed for this). Thus, we can use the framework to test any statistical test, and obtain results more interesting than our demonstrational example.

4.2.1. The effect of Central Limit Theorem

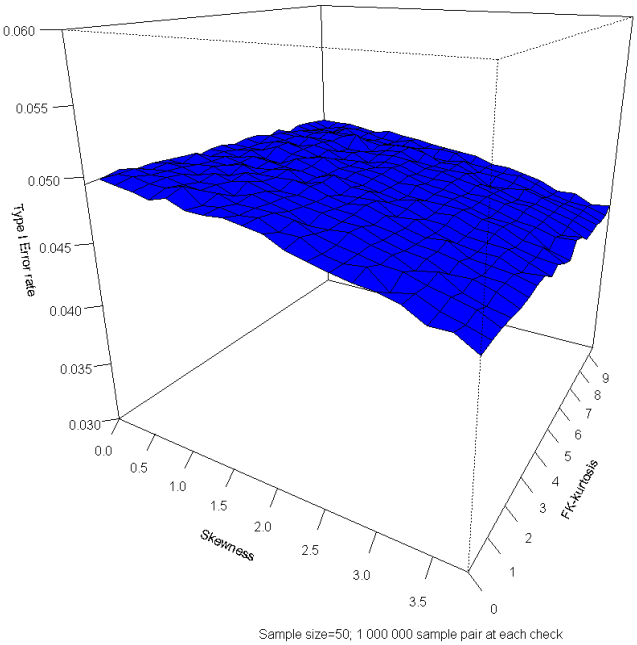
To demonstrate how simple it is to change the parameters of an MC-simulational testing, we re-run the test on the same example, but with sample sizes raised to 20, 50 and 100. Results are shown on Figures 5, 6 and 7.

Figure 5. The robustness of the *t*-test for non-normality with sample size 20



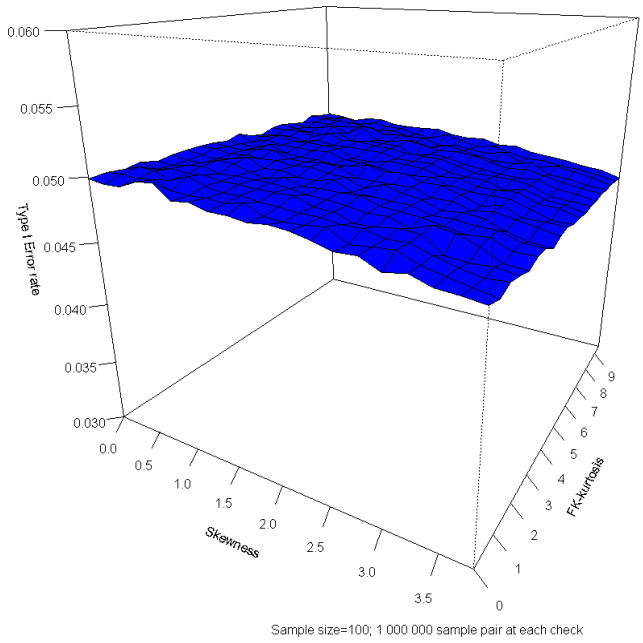
Source: own creation

Figure 6. The robustness of the *t*-test for non-normality with sample size 50



Source: own creation

Figure 7. The robustness of the *t*-test for non-normality with sample size 100



Source: own creation

One can clearly see the effect of the Central Limit Theorem: as it is well known from probability theory, with increasing sample sizes the sample mean follows more and more normal distribution, even if the population distribution was non-normal. As the test statistic of the *t*-test is based on the sample mean, this result in a higher robustness (increasing with sample size). At sample size=*100*, the test was virtually completely valid at every explored non-normality.

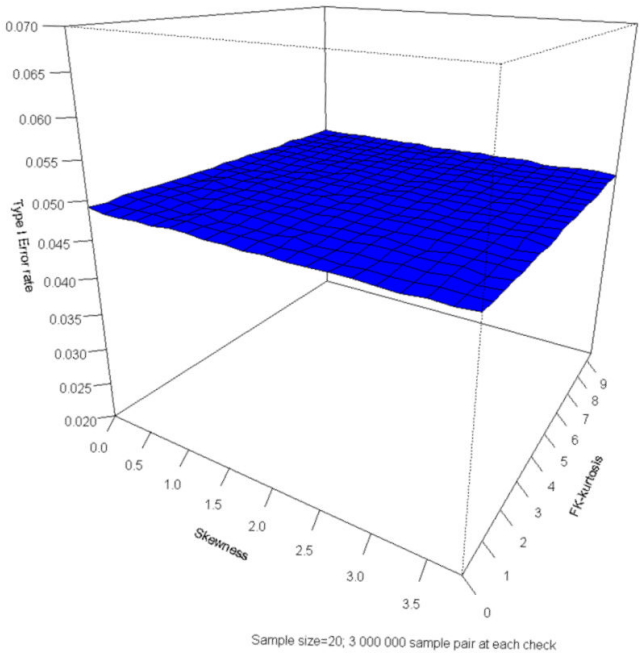
Again, not this "result" is interesting itself (it has also been known since the introduction of the *t*-test), but the way we obtained it: only a single number had to be changed in the program code for the above tests!

4.2.2. Expansion with other statistical tests

Perhaps even more important is the feature that our program can be very simply extended with other statistical tests. The adding of a statistical test to the framework only requires the specification of the test (as this part is completely separated from the code responsible for the MC-simulational testing).

To demonstrate this, we added the Mann-Whitney *U*-test (or Mann-Whitney-Wilcoxon test) to our framework, and then run exactly the same MC-simulational testing. Results are shown on Figure 8.

Figure 8. The robustness of the Mann-Whitney *U*-test for non-normality



Source: own creation

It is also important that this specification has to be done in standard C language – *no* CUDA or parallel computing knowledge is needed. (Parallelity is coded in other parts of the program.) The result is that other statistical tests can be added to our program with ease.

It can be clearly seen the Mann-Whitney *U*-test remains valid regardless of non-normality: it is robust from this point of view. This again empirically confirms the theoretical knowledge: if the PDFs of the populations have the same shape (they are just shifted) the non-parametric Mann-Whitney *U*-test is essentially distribution-free.

5. Conclusions and possibilities for further development

5.1. Summary

In this paper we discussed the question of testing the robustness of statistical tests, and especially the using of the so-called Monte Carlo (MC) method for that end. We have shown that this is an extremely computation-intensive method, which traditionally required supercomputers, grids etc., which made it available only to a few researchers.

However, a new approach, called GP-GPU can harness the extreme performance of modern – even low and middle class – video cards which can be magnitudes higher than that of CPU, for appropriate programs – just like MC-simulation. Environments for developing programs for GP-GPU are available (e.g. NVidia's CUDA).

Based on this idea, we developed a CUDA program for MC-testing the robustness of statistical tests. This can be run even with cheap, widely available video cards in ordinary PCs, despite that, it provides very high performance.

The true power of our work lies in the fact that the program we developed acts as a framework: virtually any statistical test may be included and tested – for which, no CUDA knowledge is needed at all.

Finally – as a demonstration – we performed the analysis of the well known Student's *t*-test; and – using this as a starting point – we also exhibited the main advantages of our framework.

5.2. Possibilities for further development

There are many ways to continue and improve this work. In this subsection, we will show those that seem to be the most promising.

There are many parameters we can experiment with – even in the case of the overly simple *t*-test, there is sample size and variance; from which, we have only investigated the effects of sample size. In case of more complex statistical tests, there might be far more parameters that might be changed.

The fact, that we can almost freely set the parameters of sample-generation might be used in other, essentially new ways, in addition to changing sample size and other simple descriptors. We might, for example, generate samples that deliberately do not meet the investigated test's null hypothesis – hence we can examine the power of the test. By generating samples representing different alternative hypotheses, we can also plot the approximate power function of the tests.

Fleishman's method (although having many advantageous properties) is not without problems. There are many alternatives (like Tukey's Generalized λ Distribution) which might be tried for the same end.

Finally, our framework might be in fact supplemented with statistical tests, i.e. a "test-bank" might be constructed by adding new and interesting tests to the already coded two. Such test-bank would be useful not only for research but also for education.

References

Cserny, L. 1997: *Mikroszámítógépek*. LSI Oktatóközpont, Budapest.

Ferenci, T. 2009a: *Kismintás biostatisztikai vizsgálatok néhány módszertani kérdése*. BCE TDK dolgozat, Budapest.

Ferenci, T. 2009b: *Kiskorú magyar populáció obesitassal összefüggő paramétereinek biostatisztikai elemzése*. BME VIK Diplomaterv, Budapest.

Fleishman, A. I. 1978: A Method for Generating Non-normal Distributions. *Psychometrika*, Vol. 43, Issue 4, 521-532. p.

Gentle, J. E. 2004: *Random Number Generation and Monte Carlo Methods*. Springer, New York.

GNU 2009: *GSL-GNU Scientific Library*. http://www.gnu.org/software/gsl/, 2009-12-31.

Hald, A. 1998: *A History of Mathematical Statistics from 1750 to 1930*. Wiley-Interscience, New York.

Hagerdoorn, H. 2009: *The history of Guru3D.com Part II*. http://www.guru3d.com/article/the-history-of-guru3dcom-part-ii, 2009-12-31.

Hunyadi, L. 2001: *Statisztikai következtetéselmélet közgazdászoknak*. Központi Statisztikai Hivatal, Budapest.

Hunyadi, L. – Vita, L. 2006: *Statisztika közgazdászoknak*. AULA Kiadó, Budapest.

Hunyadi, L. 2009: A negyedik mutatóról. *Statisztikai Szemle,* Vol. 87, Issue 3, 262-286. p.

Ketskeméty, L. 2005: *Valószínűségszámítás*. Műegyetemi Kiadó, Budapest.

Lee, P. M. 2009: *Bayesian Statistics: An Introduction*. Wiley, New York.

Matsumoto, M. – Nishimura, T. 1998: Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulation*, Vol. 8, Issue 1, 3-30. p.

Matsumoto, M. – Nishimura, T. 2000: Dynamic Creation of Pseudorandom Number Generators. In Niederreiter H. – Spanier J. (eds.): *Monte Carlo and Quasi-Monte Carlo Methods 1998*. Springer Verlag, Berlin, 56-69. p.

McCullough, B. D. 2006: A Review of TESTU01. *Journal of Applied Econometrics*, Vol. 21, Issue 5, 677-682. p.

Microsoft 2009: *Microsoft Visual Studio*. http://msdn.microsoft.com/hu-hu/vstudio/default(en-us).aspx, 2009-12-31.

NVidia 2009a: *GeForce 9600 GT Specifications*. http://www.nvidia.com/object/product_geforce_9600gt_us.html, 2009-12-31.

NVidia 2009b: *NVidia CUDA Programming Guide Version 2.3.1*. http://developer.download.nvidia.com/compute/cuda/2_3/toolkit/docs/NVIDIA_CUDA_Programming_Guide_2.3.pdf, 2009-12-31.

Owens, J. D. – Luebke, D. – Govindaraju, N. – Harris, M. – Krüger, J. – Lefohn, A. E. – Purcell, T. J. 2007: A Survey of General-Purpose Computation on Graphics Hardware. *Computer Graphics Forum*, Vol. 26, Issue 1, 80-113. p.

Podlozhnyuk, V. 2007: *Parallel Mersenne Twister*. http://developer.download.nvidia.com/compute/cuda/sdk/website/projects/MersenneTwister/doc/MersenneTwister.pdf, 2009-12-31.

R 2009: *The R Project for Statistical Computing*. http://www.r-project.org/, 2009-12-31.

Rubinstein, R. Y. – Krose, D. P. 2008: *Simulation and the Monte Carlo Method.* Wiley, Hoboken.

Vargha, A. 1996: Az egymintás t-próba érvényessége és javíthatósága. *Magyar Pszichológiai Szemle*, Vol. 36, Issue 4-6, 317-345. p.

Research on information propagation analyzing odds in horse racing

Shingo Ohta[1] *–Masato Yamanaka*[2] *–Joe Sato*[3] *–Takaaki Nomura*[4]

We focus on the odds in horse racing to study the information propagation to get an idea for fluctuation on information propagation. Analyzing past data of horse racing, and constructing a mathematical model of winning probability, we find a correlation between odds and results of races.
As a first step, we focus on two ways of betting called „WIN" and „EXACTA." We confirm that the winning probability derived from odds of each horse is mostly in accord with actual data. Then, comparing the result with a stochastic model of winning probability constructed with EXACTA odds, we find out fluctuations between them. Finally, we consider where is the origin of the fluctuations.

Keywords: information propagation, stochastic model, horse racing

1. Motivation and Introduction

Our motivation is to find a "better way" to share information. Especially, we focus on the case that people share the information. The "better way" means the way to convey information more correctly, more rapidly. To find it, we must take into account the "fluctuations" caused by human errors, since the "fluctuations" cause information errors. Therefore the information considered in this research should be formed by numerous people. In this sense, the horse racing is desirable to analyze.

[1] Shingo Ohta, Bachelor of Science, Ph.D student, Saitama University, Faculty of Science, Japan.
[2] Masato Yamanaka, Ph.D, post doctoral fellow, University of Tokyo, Institute for Cosmic Ray Research, Japan.
[3] Joe Sato, Ph.D, Associate Professor, Saitama University, Faculty of Science, Japan.
[4] Takaaki Nomura, Ph.D, post doctoral fellow, Saitama University, Faculty of Science, Japan.

2. The analysis of *WIN* odds

2.1. Odds and WIN

There are several ways to bet in horse racing. For the first step, we focus on *WIN*, which is the simplest way to bet. Before explaining our analysis, we explain briefly "odds" and "*WIN*".

The odds are the values which change according to betting, and dividends are defined by these values. WIN is a way to bet that people guess which horse will win. The odds therefore reflect popularity of each horse. WIN odds of a horse is defined as the ratio between total money bet on a race and the money bet on a horse:

$$\frac{\text{total money bet on a race}}{\text{money bet on a horse}}.$$

2.2. χ^2 test of goodness-of-fit

In the analysis of WIN odds, what we want to know is whether the WIN odds reflect strength of each horse.Then we consider that how well the results reflect the expectations. We define N(O) as the number of horses which have odds O and P(O) as the probability that the horses which have odds O win. Then we consider that "with what probability P(O), N horses with odds O win?"

For this analysis, we construct a stochastic model and test the goodness-of-fit between the theoretical value of the number of winner, N(O)P(O), and the observed value of number of winner, n(O). Note that we can get data of the odds O, N(O) and n(O). To test the goodness-of-fit, we should define the criterion of the fit.

We adopt the χ^2 test of goodness-of-fit for this analysis.We split whole region of odds, give a number to the each odds bin like O_i, and sum up the random variable corresponding to each odds bin. Then we derive the following χ^2 by assuming that N is large enough to satisfy the Stirling's formula. Using this criterion, we test the goodness-of-fit with a significance level:

$$\chi^2 = \sum_{i}^{\#\text{ of bin}} \frac{(n(O_i) - \bar{n}(O_i))^2}{\bar{n}(O_i)(1 - P(O_i))} \; ; \; \bar{n}(O) = N(O)P(O) \; .$$

2.3. Stochastic model

As the last preparation for the analysis, we set a stochastic model of winning probability. We assume that people know a winning probability of horses correctly. Then, it is equal to a share of a bet, namely, it is given by

$$P(O_i) = \frac{0.788}{O_i - 0.1} .$$

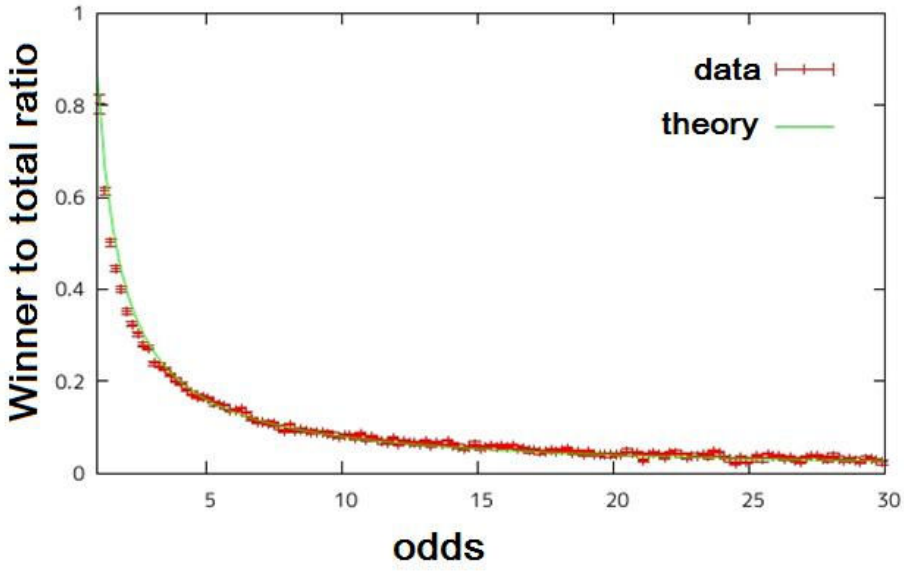
The right hand side is originally a share of a bet derived according to the definition of dividends by Japan Racing Association(JRA).

Note that the width of each odds bin should be wide enough. To analyze more precisely, there should be enough data in each odds bin. In Figure 1, we divided odds region from odds equal 1 to odds equal 31 into 150 bins. Then, we draw the theoretical value $P(O_i)$ and plots the actual results $n(O_i)/N(O_i)$, where the horizontal axis means the odds of horses, and the vertical axis means the probability of winning. At a glance, they coincide with each other very clearly.

2.4. Quantitative check of the stochastic model

To test the goodness-of-fit, we show the χ^2 values for each year from 2003 to 2008 with corresponding significance level (Table 1). We thus confirm the good fitting quantitatively.

Figure.1 The result of WIN odds analysis(degrees of freedom are 150).



Source: own creation

Most of data plots are in accord with the theoretical line.(JRA official data 1986-2009, free database soft PC KEIBA Database for JRA-VAN Data lab)

Table.1 χ^2 values from data in each year from 2003 through 2008. We apply the usual notations, * for significant at 10%, ** significant at 5%, *** significant at 1%.

year	chi-square value with significance level
2003	150.524 (*)
2004	180.915 (**)
2005	150.702 (*)
2006	187.789 (***)
2007	160.555 (*)
2008	171.232 (*)

Source: own creation

3. Comparison with EXACTA odds

3.1. What is EXACTA

Next, we compare the winning probability constructed from WIN odds with that of EXACTA. This is a way to bet that people guess the winner and the next in order. The reason we choose EXACTA is that the process to guess is almost same as that of WIN. Though the two processes are similar, there may be a little difference between them. For instance, we may expect that even if two horses' WIN odds are small, it is not always true that its EXACTA correspondense.

3.2. Stochastic model and indicator of the "fluctuation"

As we did in the analysis of WIN odds, we construct a stochastic model for the analysis of EXACTA odds. We have to construct a probability that horse α win and β finish next.

One of such a probability is given by WIN odds, O_α and O_β:

$$P(p_\alpha, p_\beta) = P(\alpha)P(\beta|\alpha) = p_\alpha \frac{p_\beta}{1-p_\alpha} \; ; \; p_\sigma = \frac{0.738}{O_\sigma - 0.1} \; (\sigma = \alpha, \; \beta)$$

The right equation is the stochastic model for WIN odds. The p_α is the probability for the horse α win, and the fact that the horse β is the second winner is same as the

fact that the horse β win without horse α. Therefore the probability is given by a conditional probability $p_\beta/(1- p_\alpha)$, and we get the above equation.
The other probability is given by EXACTA odds, $O_{\alpha\beta}$:

$$P_{\alpha\beta} = \frac{0.738}{O_{\alpha\beta} - 0.1} .$$

We assume that "people know the probability that horse α win and the horse β is the next", as WIN odds. Then, the stochastic model for EXACTA odds is similar with that for WIN odds. The difference is the numerator factor, which comes from the difference of definition between two types of dividends.

Next, to compare these values, we define the indicator of the "fluctuations". We denote two types of probabilities as follows:

$$P_{\alpha\beta} = \frac{0.738}{O_{\alpha\beta-0.1}} \longrightarrow P^a_{\text{actual}},\ P(p_\alpha, p_\beta) = p_\alpha \frac{p_\beta}{1-p_\alpha} \longrightarrow P^a_{\text{theo}}.$$

We give a serial number „a" for all conditions of horses α and β from 1 to m. In each race the number of combination is given by m_i and hence $m=\Sigma_{i=1}^{M} m_i$ with total M races. Thus the super script „a" runs from 1 to m. Then, we define the correlation coefficients as the indicator of the "fluctuations":

$$r = \frac{\sum_{a=1}^{m}(P^a_{\text{actual}} - \bar{P}_{\text{actual}})(P^a_{\text{theo}} - \bar{P}_{\text{theo}})}{\sqrt{\sum_{a=1}^{m}(P^a_{\text{actual}} - \bar{P}_{\text{actual}})^2}\sqrt{\sum_{a=1}^{m}(P^a_{\text{theo}} - \bar{P}_{\text{theo}})^2}},$$

where

$$\bar{P}_{\text{actual}} = \sum_{a=1}^{m} P^a_{\text{actual}}/m\ ,\ \bar{P}_{\text{theo}} = \sum_{a=1}^{m} P^a_{\text{theo}}/m,$$

and also check the slope of regression line for these data,

$$b = \frac{\sum_{a=1}^{m}(P^a_{\text{actual}} - \bar{P}_{\text{actual}})(P^a_{\text{theo}} - \bar{P}_{\text{theo}})}{\sum_{a=1}^{m}(P^a_{\text{actual}} - \bar{P}_{\text{actual}})^2}.$$

3.3. A fluctuation between two kinds of probabilities

In Figure 2, we plot $(P^a_{\text{actual}}, P^a_{\text{theo}})$ for all races in year 2007. The horizontal axis is the probability given by EXACTA odds and the vertical axis is that of given by WIN odds. Note that if there is a perfect coincidence between them all data points

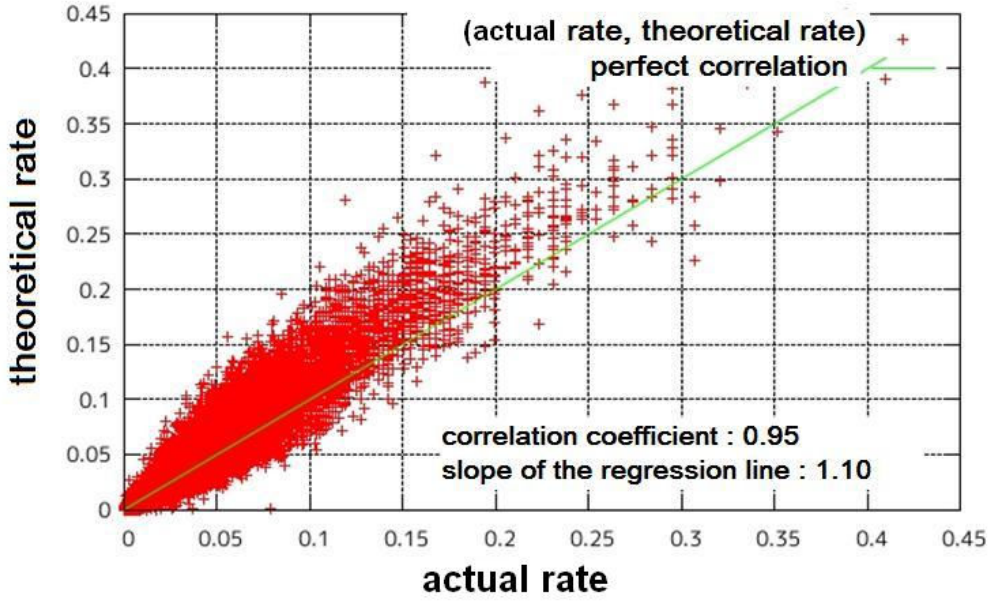
must lie on the green line. We find that the data points are dense along the line and most of them are along the line. Indeed, they have large correlation coefficient, that is, they have quite strong correlation.

However, in whole region, they have a little larger slope than that of the green line, in other words, we find that they have different distributions though they should have same ones because of their propaties. Especially, in the region with high probability, most of the data points are over the green line. In lower probability region, the situation is quite opposite.

Figure 3 is a magnified figure in the region with low probability. Contrast to whole region, the regression line has a little smaller slope than that of green line. This means that the pairs of horses with high EXACTA odds look more attractive than expected by WIN odds, which represent the strength of each horse. This is the fluctuation caused by people to bet who are eager to get much money at one bet.

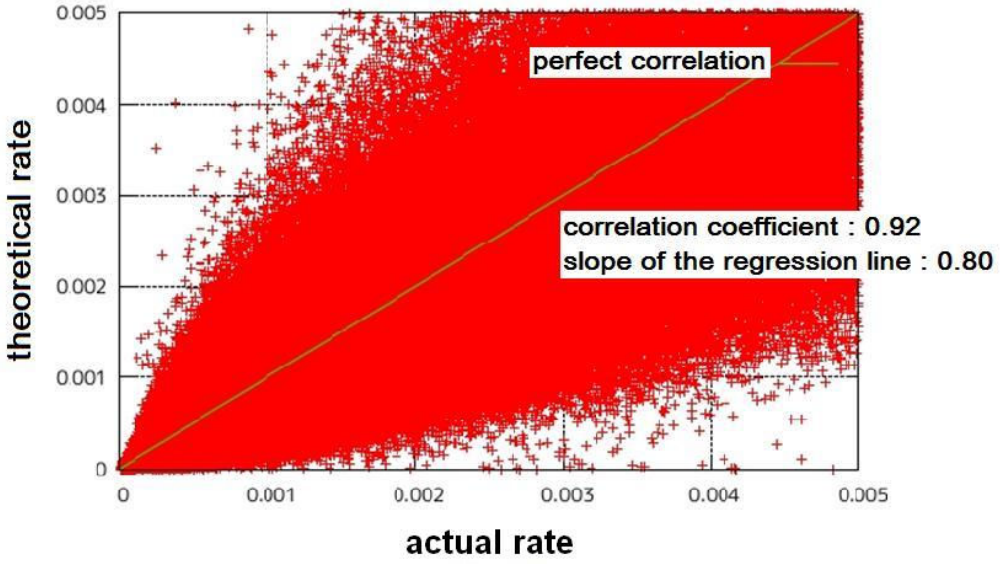
Hence we conclude that there is definitely the fluctuation between WIN odds and EXACTA odds.

Figure 2 The comparison between two kinds of probabilities we saw in section 3.2(for all races in 2007).



Source: own creation

Figure 3 The magnified figure of Fig.2 in low probability region. The slope of regression line is lower than that of perfect correlation line.



Source: own creation

4. The origin of fluctuations

4.1. Stochastic model and Indicator of the "fluctuations"

In the last step, we focus on the origin of the fluctuations. From the previous two analyses, we saw that the expectations in EXACTA deviate from the WIN expectations though they should be same. Thus we investigate where is the origin of the deviations.

For the investigation, we compare two types of probabilities.One is the probability we saw in the WIN odds analysis,

$$P_\alpha = \frac{0.788}{O_\alpha - 0.1} ,$$

which contains the information purely about the expectation of winner.
The other is defined as the sum of probabilities for EXACTA,

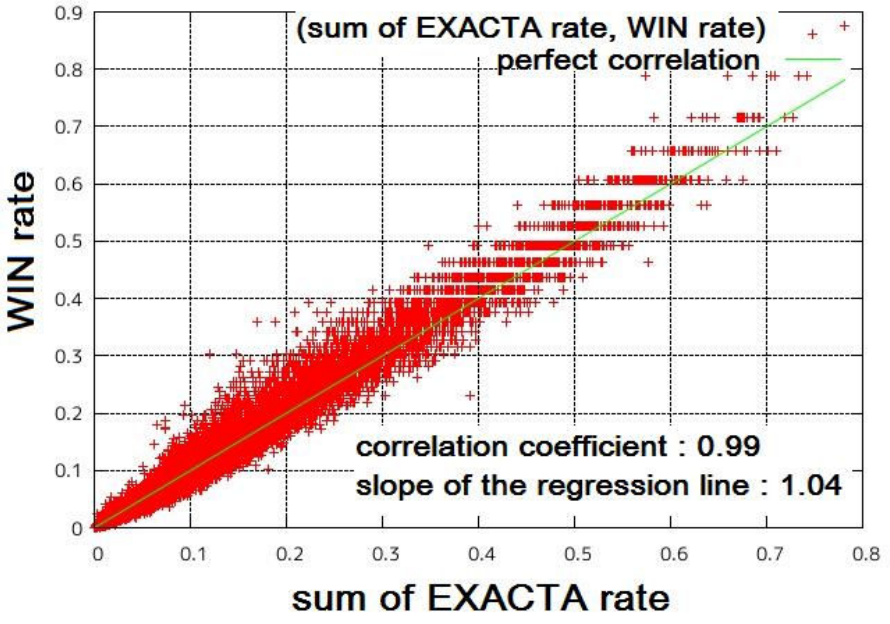
$$P_{1\alpha} = \sum_{\beta \neq \alpha} P_{\alpha\beta} \ ; \ P_{\alpha\beta} = \frac{0.738}{O_{\alpha\beta} - 0.1} \ ,$$

which contains the information of the expectation for the second horse. Then, comparing these two values, we can find out the deviations of the EXACTA odds from the WIN odds. In this analysis, we adopt the correlation coefficient and the slope of regression line as the indicators of deviations as in the previous analysis.

4.2. The Deviation from the probability for WIN odds

In Figure 4, we plot $(P_{1\alpha}, P_{\alpha})$ for all races in year 2007. The horizontal axis is the probability given by the EXACTA odds, and the vertical axis is that of given by WIN odds. There is quite strong correlation between these two values. However, as we can see, they are not exactly same each other. Therefore the main origin of the deviation in the previous analysis is seen in this analysis. Thus we conclude that the deviation is in the expectation of the second horse.

Figure 4 The comparison between two kinds of winning probability(for all races in 2007).



Source: own creation

5. Summary

In the first analysis, we confirmed that the result of each race reflects the WIN odds. In the next, we confirmed the existence of the "fluctuations" in EXACTA odds. Then, from these results, we investigated the origine of the deviations of the EXACTA odds from the WIN odds. Finally, we concluded that there is fluctuation in the expectation of the second horse.

References

[1] K. Park and E. Domany, Power law distribution of dividends in horse races, Europhy Lett, 53(4)pp.419-425.

[2] T. Ichinomiya, Physica A, Volume 368, Power-law distribution in Japanese racetrack betting, Issue 1, 1 August 2006, Pages 207-213

[3] D. A. Harville. Assigning Probabilities to the Outcomes of Multi-Entry Competitions, pp.213-217 in Efficiency of racetrack betting markets 2008 edition D. B. Hausch, V. SY Lo and T. Ziemba

[4] Plackett, R. L. (1975). The analysis of permutations. Apl. Statist., 24, 193-202

Floating utility criterion in a problem of optimal budget distribution

Eugeniy Ozhegov[1]

This paper contains the research of neuroeconomics results such as formulation and analysis of Ultimatum game (Sanfey et al. 2003) and neuromarketing (Renvoisé-Morin 2007). As a result the rational behavior of consumer during the decision-making of consume object prejudiced. In particular the axiom of reflexiveness of the rational utility theory was disproved. That axiom maintains that the fixed set of goods is not worse that itself. A conclusion that consumer choice based on the utility criterion depends not only on the set of goods but on the consume environment was made. The hypothesis of irrational behavior allowed to formalize floating utility criterion and correlation between the basket of products utility and consume environment during the consumer decision-making. Based on floating utility criterion the problem of optimal consumer's budget distribution in conditions of integral utility maximization on limited time interval and consideration of the predicted environment factors value posed. Then the problem of intertemporal consumer choise for floating criterion was posed. The solution analysis of that problems had allowed to draw a conclusion of a significant influence of the predicted environment factors value exactness on an optimal solution and a dependence of that exactness on a consumer satisfaction.

Keywords: utility, floating criterion, decision-making, optimality

1. Introduction

Rationality of classical economic theory is often questioned. Different areas of behavioral economics give different interpretation of the rationality of individual decision-making.

Risk theory formulates such irrationality in the form of the Allais paradox, the effects of equal treatment and equal difference, is trying to formalize the heterogeneity of individual behavior with the factor called "propensity for risk".

Neuroeconomical science, standing at the crossroads of economics and neuropsychology, explores the reasons of the economic decision-making depending on the activity of various parts of the brain. The origin of this area of science came through the formulation and investigation of the "Ultimatum game", as proposed in 2003 by Alan G. Sanfey in his article "Neural basis of decision-making in the

[1] Evgeniy Ozhegov, student, Perm State University, Faculty of Informational systems and mathematical methods in economics, Russian Federation

ultimatum game» (Sanfey et al. 2003). Meaning of the Game is: one of the participants must share, for example, $10 for himself and the second party, which means he can offer any amount from $0 to $10. The second party, after a proposal by a certain amount, must decide to accept it or not. If he agrees to take this amount of money, both parties remain the same amount of money, which they agreed. If the second party refuses, then they both remain with nothing. Rational economic theory suggests that any offer greater than zero must be taken because any positive amount of money is better than nothing.Empirical results showed that offers with amount of the $ 2 or $ 3 was rejected. Researchers of the game found a correlation between the refusion of the second party, during an unfair division, and activation of insula - the brain area associated with negative emotions. Also, they found a correlation between the agreement to accept the amount and activation in the dorso-lateral prefrontal cortex, which is connected with our thinking and planning actions.

One example of the capitalization of knowledge about the irrationality of decision-making is Neuromarketing (Renvoisé-Morin 2007) - a purely practical discipline that studies the decision-making to purchase consumer products during an influence of a variety of emotional factors, such as advertising, packaging, etc.. Illustration of correlation between the certain decisions and the emotional influence can be shown by the example of the American consumer, which in terms of consumption utility sees no difference between Coke and Pepsi. This result was obtained during the experiment, in which consumers were blindfolded and drank the appropriate drinks, while he did not know which one he is drinking at the moment. The activity of the brain during the consumption of different beverages were indistinguishable. If the consumer knew that he is drinking Coke at the moment, the utility which he received, expressed in activity in his brain parts was more than the utility of Pepsi by 10 times. This effect is achieved through aggressive advertising of the first drink and its other characteristics not related to the ingredients of product consumed. The result of this study is allowed not to speak on the formed consumer preferences, which described by mainstream marketing and studied in a rational utility theory, but a preference, depending on the consumption context. The consumption decision-making depends not only on the product itself, but also on external conditions under which consumption takes place. Thus, a certain set of products under different external conditions may be different levels of consumer utility. This statement refutes the first axiom of rational preferences (the axiom of reflexivity - a specific set of goods is not worse than himself) and gives rise to further study and formalization of floating utility criterion.

2. Problem of optimal budget distribution

Assume that the utility function of rational consumer is given in standard form $U: R^n \to R^1, \quad X = \{x_i\}, \ i = \overline{1,n}, \ U(X) = U(x_1, \ldots, x_n)$. On the utility of

consumption of these products have a significant influence *m* environmental factors. Let specify that factors as $S = \{s_j\}, j = \overline{1,m}$. The influence is specified in the form of the operator $K: R^n \to R^{n+m}$ or $K = \{k_{ij}(s_j)\}$, where $k_{ij}(s_j)$ is the functional coefficient of the utility sensitivity of the good x_i to a external factor s_j. Thus, the utility function takes the form: $U(K[X]) = U(X,S) = U(\{[\prod_{j=1}^{m} k_{ij}(s_j)]x_i\})$ for the case of the multiplicative effect of environmental factors on the utility from consumption of the product. In this case, if the factor s_j has no effect on the good x_i, then the value will be $k_{ij} = 1$ for every value of s_j. Utility function, written in this form will be called a floating utility criterion of the consumption basket of products.

Consider the application of this function to the problem of the optimal budget distribution - this is the problem of optimal distribution of the consumer funds in terms of maximizing the integral utility of consumer basket for a limited time interval.

Utility function is given in the form $U(\{\prod_{j=1}^{m} k_{ij}(s_j)x_i\})$. There is $x_i = x_i(t)$ – consumption of the *i*-th product at time *t* and $s_j = s_j(t)$.

$I(t)$ – the cash flow at time *t*.

$I(t) \geq \sum_{i=1}^{n} C_i(t)$ – budget restriction, where $C_i(t) = p_i(t)x_i(t)$ – the cost of consumption of the *i*-th product. In this problem defined condition non-borrowing and non-crediting of funds period *t* for other periods.

The solution of the problem requires a description of the matrix *K*, in the form of a rectangular matrix *m*n* functional coefficients of the utility sensitivity of the consumer goods *i* from the external environmental factors *j*. Methods of specification of functional coefficients are not listed, but they may be obtained from the research methods of neuroeconomics, the theory of risk, the theory of adaptive preferences, expert - it depends on the subject area of the problem and of tools explorer.

Another important parameter of the problem are the values of functions $s_j(t)$. Assume that they are known, or set a forecast value for the period *T*.

The last parameter required to solve the problem is the set of forecast of the dynamics of prices of products $p_i(t)$.

Thus, the problem becomes:

$$\int_0^T U(\{\prod_{j=1}^{m} k_{ij}(s_j(t))x_i(t)\})dt \to max \ (*)$$

$$I(t) \geq \sum_{i=1}^{n} C_i(t)$$

$$x_i(t) \geq 0, \forall i = \overline{1,n}, \forall t \in [0,T]$$

Consider the solution to this problem (*) a certain example:

$$U(X) = a_0(a_1 x_1^{\alpha_1} + a_2 x_2^{\alpha_2})^{\beta}$$

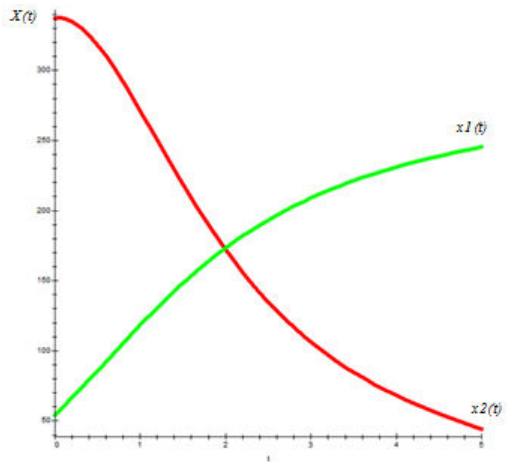
$$a_0 = 1, a_1 = 2, a_2 = 3, \alpha_1 = 0.3, \alpha_2 = 0.6, \beta = 1$$

$$K = \begin{pmatrix} -0.002s_1 + 1.5 & 1 \\ 1 & -0.01s_2 + 0.01e^{s_2} \end{pmatrix}$$
$$S = \begin{pmatrix} 180 + 180\sin\left(\frac{2t}{\pi}\right) \\ 1 + 0{,}6t \end{pmatrix}$$
$$i(t) = 200\ln(t+1) + 500$$
$$P(t) = \left(\{p_i^0(1 + \frac{t}{T}\pi_T)\}\right)$$
$$T = 5, \pi_T = 0.1,\ p_1^0 = 1,\ p_2^0 = 3$$

Solution to this problem is reduced to the optimization of the utility function for every *t* (in accordance with the additivity property of the integral) in the restriction to the budget equality to costs, because that is clear that the integral utility value will be optimal in case of Pareto-efficient budget distribution at each time moment.

Thus, the solution of the problem will be the vector of functions $F(t) = \{f_i(t)\}$, defined on the interval *[0, T],* reflecting the optimal trajectory of the consumption basket of goods for each *t.* The solution of the problem (*) looks like:

Figure 1. Solution *X(t)* for the problem (*)



Source: own creation

The optimum value of accumulated utility for the period *[0,T]* is $\int_0^T U(t)dt = 147{,}8171$.

Consider the sensitivity of the problem (*) solution to the exactness of the forecast trajectory set of the external environmental factors.

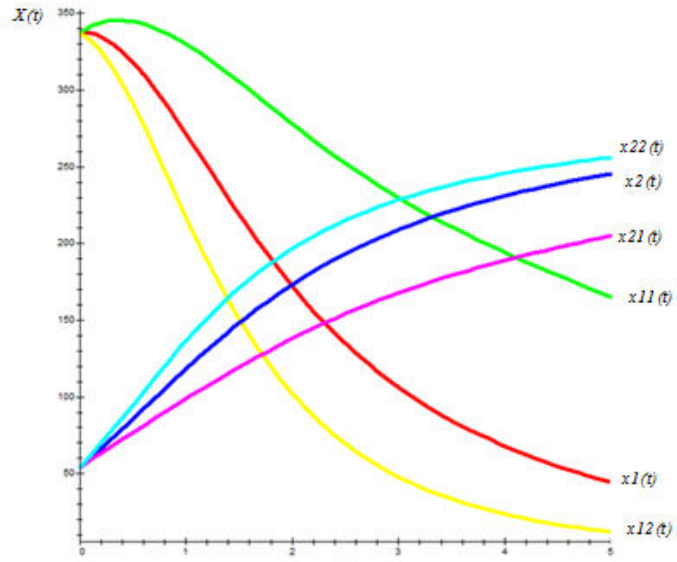
Define the forecasts of the trajectory $\tilde{S}$, different from S:

$$\tilde{S}_1 = \begin{pmatrix} 180 + 180\sin\left(\frac{2t}{\pi}\right) \\ 1 + 0{,}3t \end{pmatrix}$$

$$\tilde{S}_2 = \begin{pmatrix} 180 + 180\sin\left(\frac{2t}{\pi}\right) \\ 1 + 0{,}9t \end{pmatrix}$$

Obtain the optimal solution found in the assumptions of correct prediction. This decision will characterize the optimal trajectory of the consumption basket of goods in the sense of a set exactness of the forecast. These trajectories will look like: *X(t) x1(t) x2(t) x12(t) x21(t) x11(t) x22(t)*

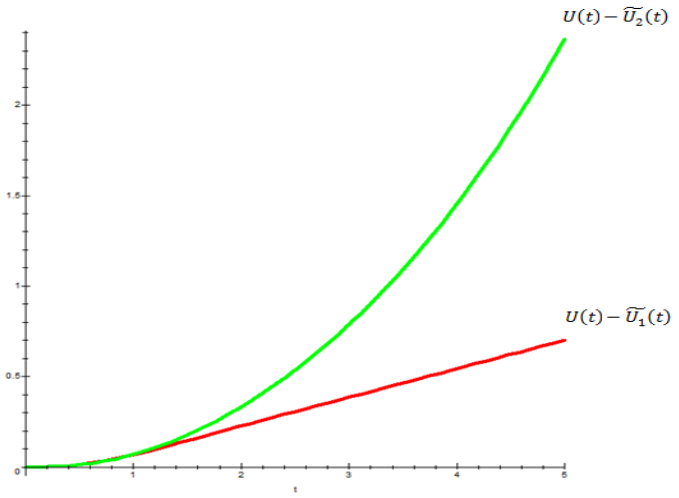
Figure 2. Sensitivity analysis of solutions *X(t)*



Source: own creation

Based on the obtained solution it seems to be possible to give a conclusion that the decision-making on the basis of not exact prediction leads to a deviation value of basket of goods utility received as predicted from the optimum obtained for the actual values of the context of consumption:

Figure 3. The deviation from the optimal solution for predictive



Source: own creation

Thus, the need to specify the forecast values of environmental factors most approximate to the reality increases the planning exactness of the trajectories of consumption and minimizes the deviation from the optimal solution.

3. Problem of utility maximization for intertemporal choise

Consider a modification of the problem (*) for a model of intertemporal choice, i.e. the possibility of borrowing funds from the budget of period *t* in period *τ*. Formally, it will looks like in the period *t* there is the funds receiving in size *I(t)*, function *α (t, τ)* describes the proportion of funds allocated for implementation of consumption in the period *τ*, then the remaining funds will amount to:

$\hat{I}(t) = \int_{\tau=0}^{T} \alpha(\tau, t)(1 + \frac{r}{T}|t - \tau|)^{t-\tau} I(\tau) d\tau$, where *r* is the money market interest rate.

So, for the model of the intertemporal choice, integral utility maximization problem becomes:

$$\int_0^T U(\{\prod_{j=1}^{m} k_{ij}\left(s_j(t)\right) f_i(t)\}) \to max \ (**)$$

$$\hat{I}(t) \geq \sum_{i=1}^{n} C_i(t)$$

$$x_i(t) \geq 0, \forall i = \overline{1,n}, \forall t \in T$$

$$\int_{\tau=0}^{T} \alpha(t,\tau)d\tau = 1, \forall t \in T$$

The problem (**) is trivial in the sense that the optimal solution will be achieved only if $\alpha(t,\tau) = \alpha^*(\tau,t) = \begin{cases} \infty, t = t^* \\ 0, t \neq t^* \end{cases}$, where t^* is determined from the condition $U(t^*) = \max_{t \in T} U(X, S, \hat{I})$

Simply put, the budget constraint is determined by the condition of crediting all income received in the period t^* in which the maximum utility value at the point achieved.

Thus, $\hat{I}(t)$ takes the form $\hat{I}(t) = (\int_{\tau=0}^{T} i(\tau)\left(1 + \frac{r}{T}|t-\tau|\right)^{t-\tau}d\tau)\varphi(t-t^*)$, where $\varphi(t-t^*) = \begin{cases} 1, t = t^* \\ 0, t \neq t^* \end{cases}$

The problem solution will looks like $X^*(t^*)$, where:

$$\Sigma_{i=1}^{n} p_i(t^*)X_i(t^*) = \hat{I}(t^*)(I)$$
$$U(X^*(t^*)) = \max_{X(t^*) \in (I)} U(X(t^*))$$

Consumption of goods will occur only during the period t^*, in the remaining periods of the interval T consumption is equal to 0. Obviously that pose the problem in this form does not make sense as to sustain an individual life requires a specific set of benefits, other than 0, which yields the minimum required utility. In this situation, it makes sense to introduce in the problem (**) a restriction on the minimum required value $\underline{U}$, then the problem becomes:

$$\int_0^T U(\{\prod_{j=1}^{m} k_{ij}(s_j(t))x_i(t)\})dt \to max \ (***)$$
$$\hat{I}(t) \geq \sum_{i=1}^{n} C_i(t)$$
$$x_i(t) \geq 0, \forall i = \overline{1,n}, \forall t \in T$$
$$\int_0^T \alpha(t,\tau)d\tau = 1, \forall t \in T$$
$$U(t) \geq \underline{U(t)}, \forall t \in T$$

Consider the solution for the problem (***):

To maximize the integral utility required to provide the interval T consumption $\underline{U}$, respectively, to solve the auxiliary problem in terms of $\underline{x_i}(t)$ where $\underline{I}(t) = p_i(t)\underline{x_i}(t)$:

$$\underline{I}(t) \to min, \forall t \in T$$
$$U(\underline{x_i}(t)) \geq \underline{U}(t), \forall t \in T$$

$$\underline{x_i}(t) \geq 0, \forall i = \overline{1,n}, \forall t \in T$$

The obtained solution of the problem to minimize the budget will be used for solving the problem of maximizing the integral utility in terms of $\overline{x_i}(t)$ and $\overline{\alpha}(\tau, t)$ with the restriction on the minimum utility value, then the problem (***) becomes:

$$\int_0^T U(\{\prod_{j=1}^{m} k_{ij}\left(s_j(t)\right)(\overline{x_i}(t) + \underline{x_i}(t))\})dt \to max \quad (****)$$

$$\bar{I}(t) \geq \sum_{i=1}^{n} C_i(t)$$

$$\bar{I}(t) = \int_{\tau=0}^{T} \overline{\alpha}(\tau, t)\,(I(\tau) - \underline{I}(\tau))(1 + \frac{r}{T}|t - \tau|)^{t-\tau} d\tau$$

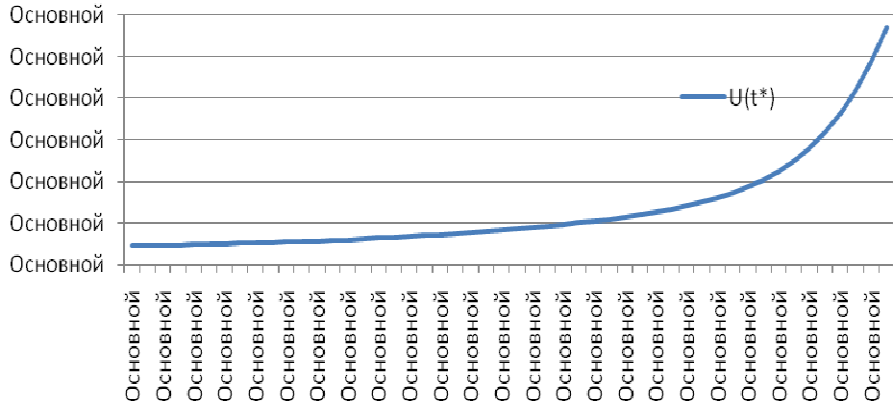
$$\overline{x_i}(t) \geq 0, \forall i = \overline{1,n}, \forall t \in T$$

$$\int_{\tau=0}^{T} \overline{\alpha}(t, \tau)d\tau = 1, \forall t \in T$$

The problem (****) solution is analogous to the solution of the problem (**) and will set consumption only at the point t^*, the remaining points of the interval T consumption is equal to 0. The general solution will be $\underline{x_i}(t) + \overline{x_i}(t)$, the optimal value of the integral value will be a $U\left(\{\underline{x_i}(t) + \overline{x_i}(t)\}\right)$. Thus, the final solution allows to achieve the minimum required value, while at t^* will be observed jump in accumulated utility to a value of $U\left(\{\underline{x_i}(t^*) + \overline{x_i}(t^*)\}\right)$

For the problem (**) found the optimal consumption for each t in the case of lending of all incoming funds in period t^*. For each solution is calculated accumulated value on the interval [0, T]:

Figure 4. Graph of the solution of (**) in the form of *U(t*)*



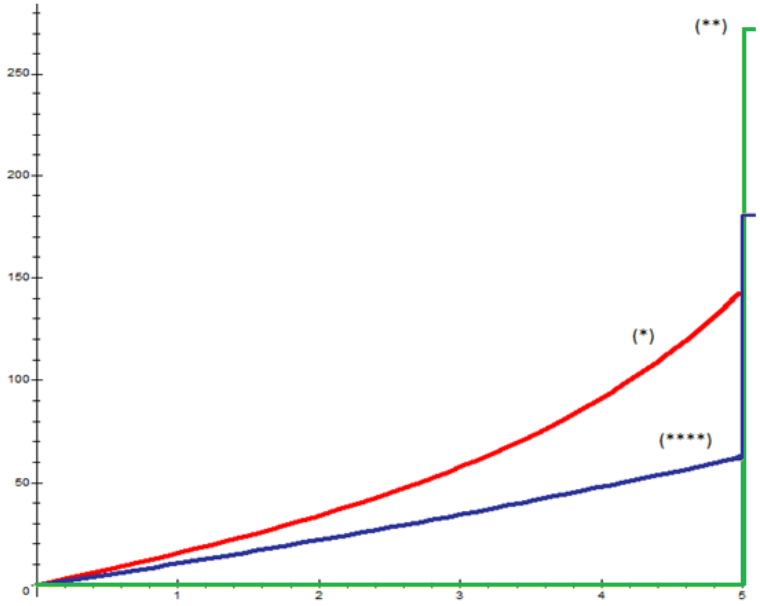
Source: own creation

Is thus seen that the maximum accumulated value is achieved if all the funds will be used for consumption in period $t^* = T = 5$. This is due to the budget growth rate for $t \to T$ and the growth rate for K, which is anticipatory for growth rate of prices. The value of accumulated utility for the problem (**) optimal solution was $\int_{t=0}^{t} U(t, \hat{I}(t))\, dt = 285{,}0304$.

Founded the solution for the problem (****) in terms of problem (1) with set of constraint for $\underline{U(t)} = 10 + t$. The solution in the form of accumulated utility will be a $\int_0^T U(\{\underline{x_i}(t^*) + \overline{x_i}(t^*)\}) = 181{,}5412$.

Consolidated graph of the solutions of problems (*),(**),(****) in the form of accumulated integral utility for the period t:

Figure 5. Comparative analysis of solutions of problems (*), (**), (****) in the form of an accumulated utility



Source: own creation

Thus, can be made a conclusion that the optimal value achieved in the problem (**) but the problem with the minimal utility value restriction most approximates the reality in studied problem class and the optimal solution is achieved in the restriction in the minimal required utility.

References

Ozgehov, E. 2009a: *Floating utility criterion in the problem of optimizing the allocation of budgetary transfers.* Perm State University*, 1-19. p.*

Ozgehov, E. 2009b: Practical aspects of neuroeconomics. *Economics and management: actual problems and seek solutions*, 45-46. p.

Renvoisé, P. – Morin C. 2007: Neuromarketing: Understanding the "buy buttons" in your customer's brain. *SalesBrain LLC*, 78-81. p.

Sanfey, A. G. – Rilling, J. K. – Aronson, J. A. – Nystrom, L. E. – Cohen, D. J. 2003: The Neural Basis of Economic Decision-Making in the Ultimatum Game. *Science Mag*, 300, 1755-1758. p.