REGIONAL STATISTICS, 2012, VOL. 2: 3–12 DOI: 10.15196/RS02101

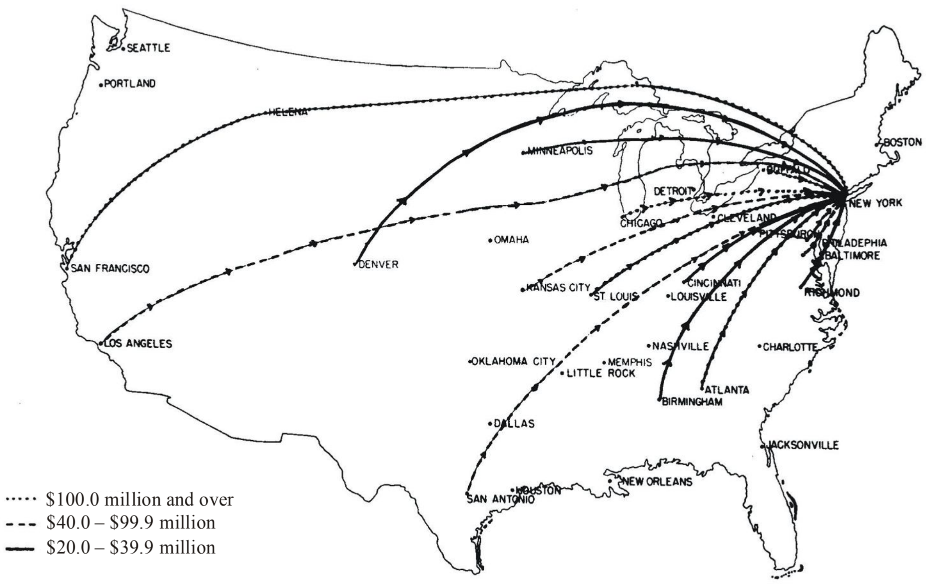
WALDO TOBLER

Looking at Some Data from Isard's "Methods of Regional Analysis"

In his path-breaking book "Methods of Regional Analysis" Walter Isard presented several sets of data. Some of these are fictitious, designed to illustrate concepts. Others contain real empirical values. Among the latter are several tables. I have used two of these, not in a research mode, but rather as exploratory visualizations. Of the two matrices one is 36 by 36 in size, the other is 47 by 47. I have used these tables in their original form without attempting any transformations such as biproportional adjustment to equalize marginals, or eigenvector analysis, or partitioning into symmetric and skew symmetric parts to estimate potentials (Figure 1). Nor have I attempted to reduce the resolution of the data by combining, or by omitting, rows or columns from the matrices.

Figure 1

Net inflows through the Interdistrict Settlement Fund to the New York Federal Reserve Office, June–July 1954



Source: Isard 1960 p. 149.

Table 1

Net commercial and financial flows between Federal Reserve zones, week ended October 13, 1954

(Millions of dollars; plus sign indicates net inflow to area at top of column)

Towns	Boston	New York	Buffalo	Philadelphia	Cleveland	Cincinnati	Pittsburgh	Richmond	Baltimore	Charlotte	Atlanta	Birmingham	Jacksonville	Nashville	New Orleans	Chicago	Detroit	St. Louis
Boston	–	+26	+1	+3	+1	+1	+4	+1	+1	+3	0	0	+1	+1	0	+8	+6	–1
New York	–26	–	+3	–25	+33	–3	+27	–4	–6	+24	–36	–19	+10	–9	+19	+70	+77	–66
Buffalo	–1	–3	–	0	+2	0	+1	0	0	0	0	–	0	–1	0	+6	–3	0
Philadelphia	–3	+25	0	–	–3	–3	–6	–4	–3	–4	–5	–1	–2	–1	–1	+7	–5	–1
Cleveland	–1	–33	–2	+3	–	+3	+5	+2	+1	–1	–1	0	+1	+1	0	+10	–5	+3
Cincinnati	–1	+3	0	+3	–3	–	–1	–1	0	0	0	0	0	–1	0	+8	–3	0
Pittsburgh	–4	–27	–1	+6	–5	+1	–	–5	+5	0	–5	+1	0	+1	+1	+7	+1	–1
Richmond	–1	+4	0	+4	–2	+1	+5	–	+8	–15	+6	0	–2	0	0	+7	+3	+2
Baltimore	–1	+6	0	+3	–1	0	–5	–8	–	–1	0	0	–1	0	0	+3	–1	0
Charlotte	–3	–24	0	+4	+1	0	0	+15	+1	–	+9	+1	0	0	0	–4	0	+1
Atlanta	0	+36	0	+5	+1	0	+5	–6	0	–9	0	–6	–11	–4	–12	+5	+3	+2
Birmingham	–	+19	–	+1	0	0	–1	0	0	–1	+6	–	0	–2	–2	+1	0	+2
Jacksonville	–1	–10	0	+2	–1	0	0	+2	+1	0	+11	0	–	+1	+1	–3	0	+1
Nashville	–1	+9	+1	+1	–1	+1	–1	0	0	0	+4	+2	–1	–	+1	+1	0	+2
New Orleans	0	–19	0	+1	0	0	–1	0	0	0	+12	+2	–1	–1	–	+5	+1	+3
Chicago	–8	–70	–6	–7	–10	–8	–7	–7	–3	+4	–5	–1	+3	–1	–5	–	–	–9
Detroit	–6	–77	+3	+5	+5	+3	–1	–3	+1	0	–3	0	0	0	–1	–	–	0
St. Louis	+1	+66	0	+1	–3	0	+1	–2	0	–1	–2	–2	–1	–2	–3	+9	0	–
Little Rock	0	–1	–	0	0	0	0	0	–	0	0	–	0	0	–1	–1	0	+4
Louisville	–1	–30	0	0	–2	+3	0	0	0	–1	–1	0	–1	0	0	+3	0	+9
Memphis	+2	–11	–	0	0	0	0	0	0	–7	0	–2	–1	–1	+1	–1	+2	+15
Minneapolis	0	+34	–4	0	–4	+1	0	+1	0	0	0	0	0	0	–1	–15	+3	+4
Helena	0	–1	0	0	0	0	–	–	–	–	–	0	–	–	–	–1	0	0
Kansas City	0	+21	+2	+2	+3	+1	–2	–1	0	0	0	–1	0	0	–1	+10	+1	+5
Denver	–1	+5	–	0	+1	0	0	0	0	0	0	–	–	–	+1	–2	+1	+2
Oklahoma City	0	+12	–	+1	0	+1	+1	0	0	0	+3	0	–	–	–4	–21	0	+5
Omaha	0	+3	0	0	0	+1	0	0	0	0	0	–	–1	–	0	–1	0	+1
Dallas	–3	–46	0	0	+1	+2	0	0	0	0	+2	0	–2	–1	–3	+7	+2	+10
El Paso	–1	–1	–	0	0	–	–	0	–	0	–	–	–	–	0	+1	0	+1
Houston	–1	–31	0	–3	0	0	–4	0	0	0	–1	0	–1	+2	+4	+3	0	+1
San Antonio	0	+6	–	+1	0	0	0	0	0	0	0	0	0	0	+1	+6	0	+1
San Francisco	+14	–23	+2	+1	–1	+1	–4	–5	+1	0	–1	0	0	0	+5	+22	0	+2
Los Angeles	–3	+56	+1	+2	+2	+1	+1	+1	0	+1	–1	0	0	0	+1	+15	+11	+3
Portland	–1	+5	0	–3	0	0	0	–1	0	–	0	0	0	–	0	+1	+1	–1
Salt Lake City	0	–2	–	0	0	0	0	0	–	–	–	–	–	–	–	–2	0	+1
Seattle	0	+7	0	+1	0	0	0	0	0	0	0	–	0	–	–	–2	+2	+1

Source: Isard 1960, pp. 152-153.

Note: Careful examination of Isard's table and the maps reveals that they came from different sources and relate to different time periods. Thus the flows mapped here differ from those illustrated in the book.

Relating to the first of my examples, Isard has four maps of the movement of funds between Federal Reserve cities. These are illustrated on pages 149, 154–156 of the book. He also presented a table of similar monetary transfers between the thirty-six Federal Reserve cities in the United States (Table 1). Curiously none of the four maps shown is made using the information from this table. But he argued, rightly in my opinion, that such tables and maps should be used more widely. To date I have not seen much evidence that this suggestion has been followed. An examination of recent issues of the *Annals of Regional Science* convinces one of this fact. Isard thus recognized that insufficient use is often made of the human power of inference from visualizations in research. As an exercise to demonstrate the usefulness of this idea I have used a simple and freely available computer program [1](Figure 2) to create several maps from the 36 by 36 table. In order to do this it was only necessary to look up the latitudes and longitudes of the cities on the internet and to find a computer compatible outline map of the United States. The preparation of such maps is then very simple and alternate versions of the maps can be created quickly. The computer program allows one to interactively hover (with a mouse) over a location to read its name, and to hover over the flow symbol to be presented with the numerical value. Thus exploratory research use can be made of the maps. This interaction of course is not possible with the printed versions of the maps (Figure 3).

Figure 2

All fiscal transfers during a week (of 1,260 possible)

Figure 3

Net transfers greater than 2 107 dollars*

1 http://www.csiss.org/clearinghouse/FlowMapper/

The 36 by 36 matrix yields a possible 1,260 movements so that some parsing is desirable. Such tables are not random, but contain considerable structure. The majority of the volume is contained in a minority of the flows so that deleting flows below the mean tends to leave 75% of the movement volume but requires only 20% of the movement activity. This is a far better strategy than, for example, reducing the resolution by lumping all of the thirty six cities into the twelve Federal Reserve Districts and thereby reducing the 1,260 values to 132 with the attendant loss of geographic detail.

Examples of maps from one set of data hardly constitutes a research project, but inclusion of such fiscal-movement maps can enhance most studies, and a sequence from several time periods is certain to be enlightening, especially if rendered within an animation.

Figure 4

Above average transfers to New York City

Figure 5

Above average transfers from New York City

Another important data array is available in an input-output table. As pointed out by Isard "… the *interregional input-output* approach is most prominent, both in terms of accomplishment and recognition. It represents a fruitful method for depicting and investigating the underlying processes, which bond together the regions of a system and all the separate facets of their economies" (Isard 1960 p. 310.). A graphical exercise, similar to that given above, has been undertaken with respect to the New England input/output table from the "Methods" book. In this case there are forty-seven industry groupings and the flows are represented by a 47 by 47 table. The industries are quickly given coordinate locations by a metric multidimensional scaling program (Borgatti 1996). This works somewhat as follows. Consider two industries to be "close" if the quantity of exchanges between them is large. Then construct a diagram in which "close" industries are placed near each other. Do this by positioning a first industry at random. Take the industry "closest" to this and place it on a circle of chosen radius near the first one, thereby establishing a scale for the diagram. The next industry, the third industry, then lies at the intersection of two circles drawn about the first two, with appropriately scaled radii. There is a slight ambiguity and arbitrariness here because two circles can intersect in two places. Choose one of the options and insert the third industry, and then proceed to locate the next (fourth) industry using more circles. The degrees of freedom rapidly decrease and the choices become more difficult. Mathematically the fitting via circles can be replaced by an iterative procedure that places industries into the diagram so that "close" industries are also close in the diagram (see Tobler 1996). The contacts between industries are then drawn as flows between these locations using the same mapping computer program. Curiously, the scaling algorithm has placed most of the industries on the perimeter of a circle with only a few industries in the interior. The suspicion is that the starting configuration, the Gower algorithm was used, began with all items on a circle. The non-metric version, the Torsca algorithm, gave very similar results.

The few interior-lying industries seem to dominate the table. The "Household" category is especially prominent. The important industries exchanging goods and services are easily identified when using the interactive computer program, as are the magnitudes involved. Close visual examination of the original input-output matrix would permit recognition of these major flows, especially when studied by the original compiler of the table. Nevertheless, without such close scrutiny a casual user could find the graphical version(s) more useful and especially rapid. This technique has, to the best of my knowledge, not been used very often, if at all, with input/output tables.

The 47 by 47 matrix could contain as many as 2,162 entries, if the diagonal is ignored, so that a parsing by size is again appropriate, as is seen clearly in the diagrams. Only a little over a dozen entries are significant. As is the case with the movement maps, one example does not constitute a research project but a series of such diagrams, even an animation, can enhance the study of changing regional structure. An opportunity to study regional dynamics thus becomes available. An interesting experiment might be provided by applying the technique to the Leontiev inverse, or to the input/output coefficient matrix.

Table 2

New England: input requirements and commodity

Industry (Sector) Purchasing / Industry (Sector) Producing	1. Agricultural and Fisheries	2. Food and Kindred Products	3. Tobacco Manufactures	4. Textile Mill Products	5. Apparel	6. Lumber and Wood Products	7. Furniture and Fixtures	8. Paper and Allied Products	9. Printing and Publishing	10. Chemicals
1. Agriculture and Fisheries	244.5	534.2	7.6	571.2	1.2	8.3	–	1.2	–	57.9
2. Food and Kindred Products	53.6	174.3	0.1	16.5	0.6	*	*	4.1	*	32.8
3. Tobacco Manufactures	–	–	8.1	–	–	–	–	–	–	+
4. Textile Mill Products	1.4	0.1	–	358.0	246.0	0.1	15.4	5.9	1.6	0.6
5. Apparel	1.0	7.2	–	–	124.2	–	0.3	2.8	–	1.4
6. Lumber and Wood Products	3.3	2.9	0.2	4.9	0.1	47.2	20.8	36.8	0.1	2.2
7. Furniture and Fixtures	–	–	–	3.3	–	–	0.4	0.7	–	–
8. Paper and Allied Products		16.1	0.6	21.4	1.6	0.2	0.8	358.0	70.3	15.8
9. Printing and Publishing	–	1.4	–	0.6	–	–	–	–	49.8	0.8
10. Chemicals	18.7	51.5	0.2	219.8	9.0	1.1	3.4	25.2	6.3	127.0
11. Products of Petroleum and Coal	10.3	2.1	*	8.2	0.3	3.2	0.1	8.7	0.2	15.5
12. Rubber Products	2.7	0.3	–	3.6	1.1	0.4	0.3	1.2	0.2	+
13. Leather and Leather Products	–	–	–	0.6	3.4	0.2	0.4	–	0.3	–
14. Stone, Clay and Class Products	1.5	9.0	–	0.3	–	0.6	1.8	3.9	–	12.3
15. Iron and Steel	0.1	0.1	–	–	0.1	0.4	5.2	–	–	0.2
16. Nonferrous Metals	–	–	–	–	–	0.1	0.9	–	0.9	9.0
17. Plumbing and Heating Supplies	–	–	–	–	–	–	–	–	–	–
18. Fabricates Structural Metal Products	–	–	–	–	–	–	0.3	–	–	–
19. Other Fabricated Metal Products	1.9	19.3	0.1	*	0.4	1.5	7.1	2.3	0.1	6.2
20. Agricultural, Mining and Construction Machinery	1.3	–	–	–	–	–	–	–	–	–
21. Metalworking Machinery	–	–	–	–	–	–	–	–	–	–
22. Other Machinery (except electric)	–	0.5	–	9.6	1.3	0.6	0.6	1.9	2.3	+
23. Motors and Generators	–	–	–	–	–	–	–	–	–	–
24. Radios	–	–	–	–	–	–	–	–	–	–
25. Other Electrical Machinery	–	–	–	–	–	–	–	–	–	+
26. Motor Vehicles	2.5	0.1	–	–	–	+	–	–	–	–
27. Other Transportation Equipment	0.2	–	–	–	–	–	–	*	–	+

(Continued the next page)

(Continued)

Industry (Sector) Purchasing / Industry (Sector) Producing	1. Agricultural and Fisheries	2. Food and Kindred Products	3. Tobacco Manufactures	4. Textile Mill Products	5. Apparel	6. Lumber and Wood Products	7. Furniture and Fixtures	8. Paper and Allied Products	9. Printing and Publishing	10. Chemicals
28. Professional and Scientific Equipment	–	–	–	–	–	–	0.1	0.8	2.1	0.6
29. Miscellaneous Manufacturing	0.1	0.4	–	1.1	16.2	*	0.9	2.1	–	1.4
30. Coal, Gas and Electric Power	1.4	6.9	+	28.8	2.3	1.0	1.0	17.0	1.9	9.0
31. Railroad Transportation	9.9	19.5	0.2	25.8	3.8	6.2	2.9	30.9	4.4	13.7
32. Ocean Transportation	1.6	4.5	+	3.6	0.7	0.4	*	2.2	*	2.1
33. Other Transportation	12.5	13.0	0.2	21.7	1.6	6.0	2.2	16.1	1.6	4.5
34. Trade	30.6	14.8	0.4	62.6	23.4	2.6	3.2	24.3	2.0	8.3
35. Communications	+	1.5	+	2.5	1.2	0.4	0.3	1.1	2.5	1.1
36. Finance and Insurance	5.4	5.1	+	5.5	1.5	3.3	1.0	2.5	1.5	0.9
37. Rental	53.9	3.2	–	6.9	6.1	0.8	0.9	3.6	4.0	1.6
38. Business Services	0.2	18.9	1.0	19.5	6.1	0.8	3.1	3.0	3.8	20.3
39. Personal and Repair Services	8.3	4.2	*	0.8	0.2	1.8	0.2	0.6	1.3	0.5
40. Medical, Educational and Nonprofit Organizations	–	–	–	–	–	–	–	–	–	–
41. Amusements	–	–	–	–	–	–	–	–	–	–
42. Scrap and Miscellaneous Industries	–	–	–	6.6	–	–	–	34.5	–	5.3
43. Undistributed	–	73.1	1.3	120.3	83.0	38.0	17.7	27.7	39.6	83.2
44. Eating and Drinking Places	–	–	–	–	–	–	–	–	0.1	–
45. New Construction and Maintenance	4.5	4.2	+	10.7	1.0	0.5	0.4	5.8	1.0	1.7
46. Government	18.3	40.3	1.0	175.6	23.8	14.6	6.0	68.5	21.8	36.5
47. Households	431.7	222.3	3.8	902.8	254.3	110.6	57.3	297.9	197.2	164.1
Total Gross Outlays	*921.5*	*1250.7*	*25.0*	*2612.9*	*814.6*	*251.1*	*154.9*	*991.4*	*416.8*	*636.9*

Source: Isard 1960 pp. 604–607. (Also see the discussion of the data in the table on pages 603–617.)

Figure 6

The complete I/O table in graphic form

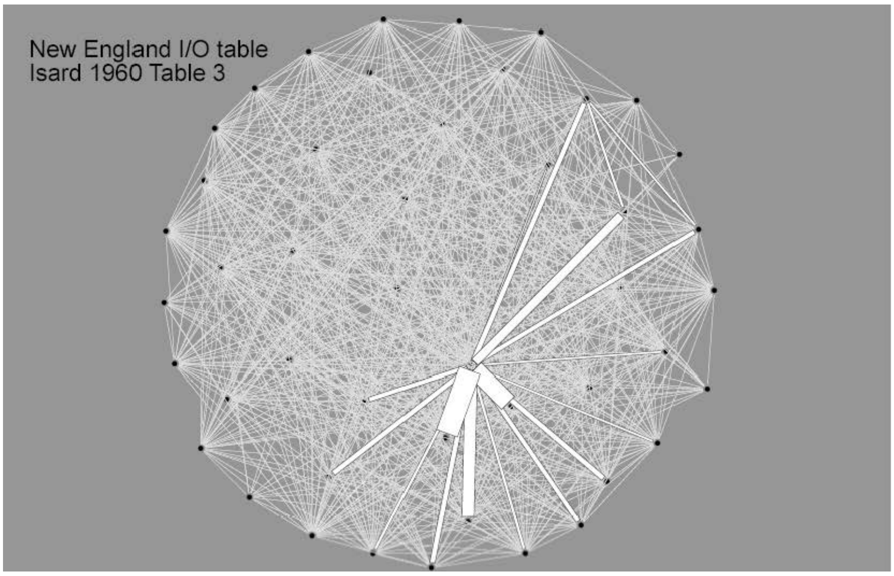


Figure 7

All exchanges above the average

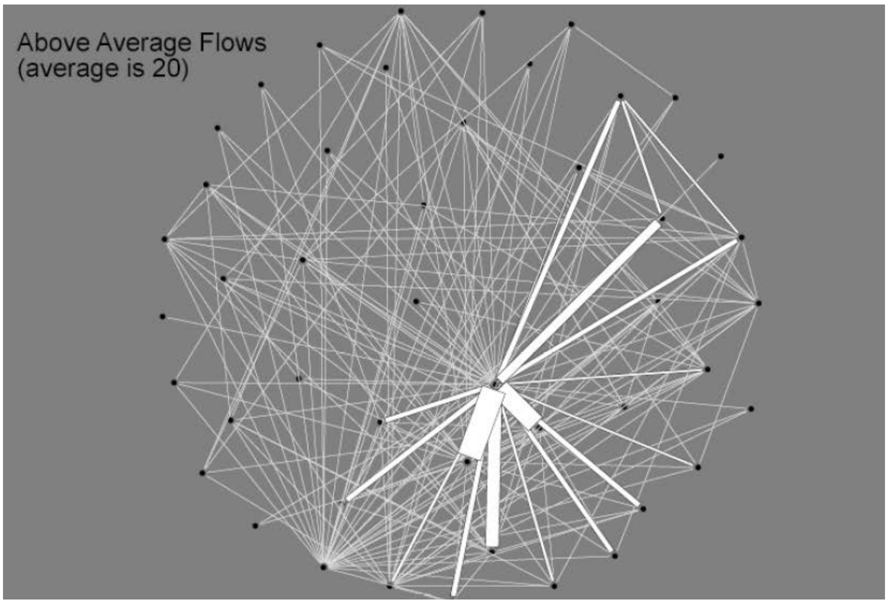


Figure 8

*Above 5 * 108 dollars*

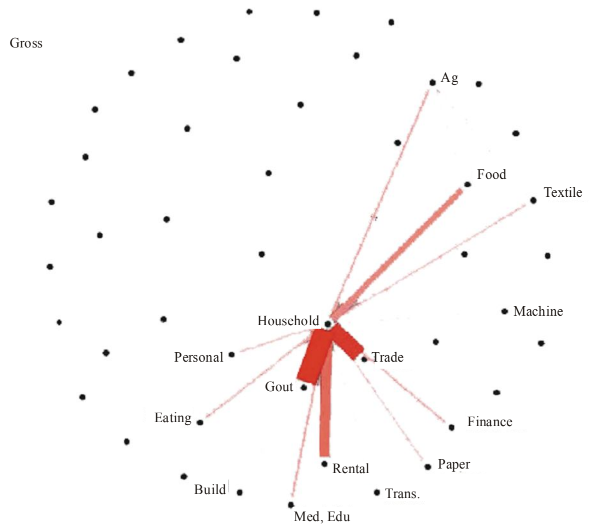
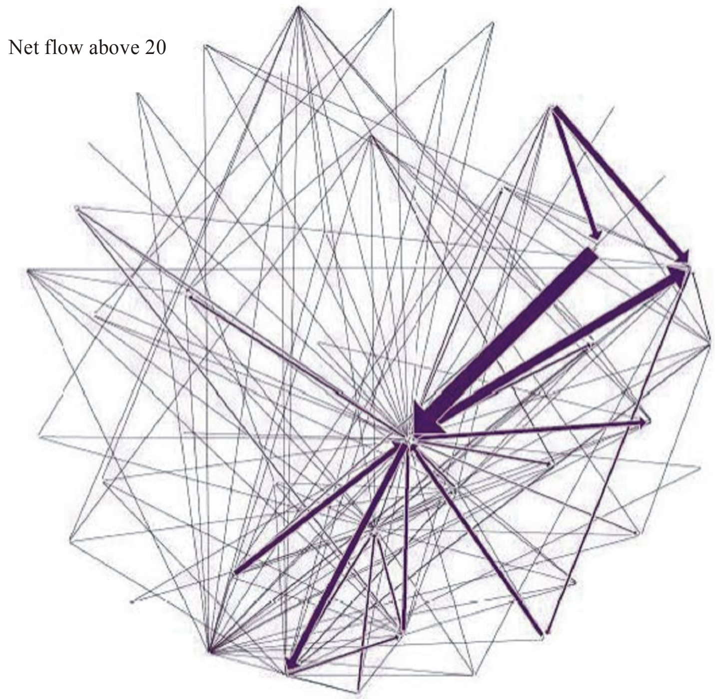


Figure 9

*Net exchanges above 2 * 107 dollars*



REFERENCES

Tobler, W. (2005): Using Asymmetry to Estimate Potential. *Solstice* 16 (2) http://www-personal.umich.edu/~copyrght/image/solstice/win05/Sunbelt%20XXV%20talk%20ed.pdf
Isard, W. (1960): *Methods of Regional Analysis*. MIT Press, Cambridge.
Tobler, W. (1981): A Model of Geographic Movement. *Geographical Analysis* 13 (1) 1–20.
Borgatti, S. P. (1996): *ANTHROPAC 4.0*. Analytic Technologies. Natick, Mass.
Tobler, W. (1996): A Graphical Introduction to Survey Adjustment. *Cartographica* 32 (1) 33–42.

Keywords: monetary flow, input-output tables, visualization.

Abstract

Isard has argued for greater presentation of empirical data in regional studies by displaying tables and geographical maps. In this note his suggestion is followed by displaying two sets of computer maps using tables given in his classic 1960s book.

REGIONAL STATISTICS, 2012, VOL. 2: 13–26 DOI: 10.15196/RS02102

KARIMA KOURTIT – PETER NIJKAMP – ANDREA CARAGLIU
– CHIARA DEL BO

Spatial Contextual Impacts on Business Performance of Dutch Firms

Aims and scope

The literature on industrial innovation shows a remarkable pattern. In earlier contributions (see e.g. Schumpeter 1934) much attention has been given to the institutional and market conditions (for instance, competition) that acted as drivers for innovations by entrepreneurs. Innovations were mostly seen as necessary instruments to survive in a competitive economy.

In the 1980s much attention was also given to the economic impacts of industrial innovation, e.g. on labour markets, on regional disparities, or on economic growth. However, at the same time it was recognized that innovation is not ‘manna from heaven’, but the result of our deliberate actions, in both the public and the private sector. Against the background of endogenous growth theory, much attention was also paid to the impact of R&D, education, training, and subsidies on the innovation intensity of entrepreneurs. Entrepreneurial innovation was no longer seen as a ‘fate’, but as an opportunity to be created by active entrepreneurs.

In the past two decades, the regional and urban context of innovation has received much interest. The popularity of concepts such as ‘regional innovation systems’ or ‘open innovation systems’ has prompted an avalanche of studies on the knowledge-based society, on the high-tech economy, and in more recent years on the creative economy (see e.g. Kourtit et al. 2011). The spatial dimension of the innovation society has increasingly been put at the centre of scientific research. In particular, the socio-economic, technological, institutional, and cultural conditions of cities and regions have been given much attention from the perspective of drivers or facilitators of economic growth. In parallel, the local socio-economic impacts of innovation have also become a centrepiece of recent scientific research in innovation.

The present study aims to highlight the drivers and impacts of innovative activities in the high-tech sectors from the perspective of local and regional dimensions. To that end, an extensive dataset on Dutch high-tech firms has been created and investigated. The attention will in particular be focused on a triple-layer analytical framework: the spatial context, the industrial context, and the entrepreneurial context.

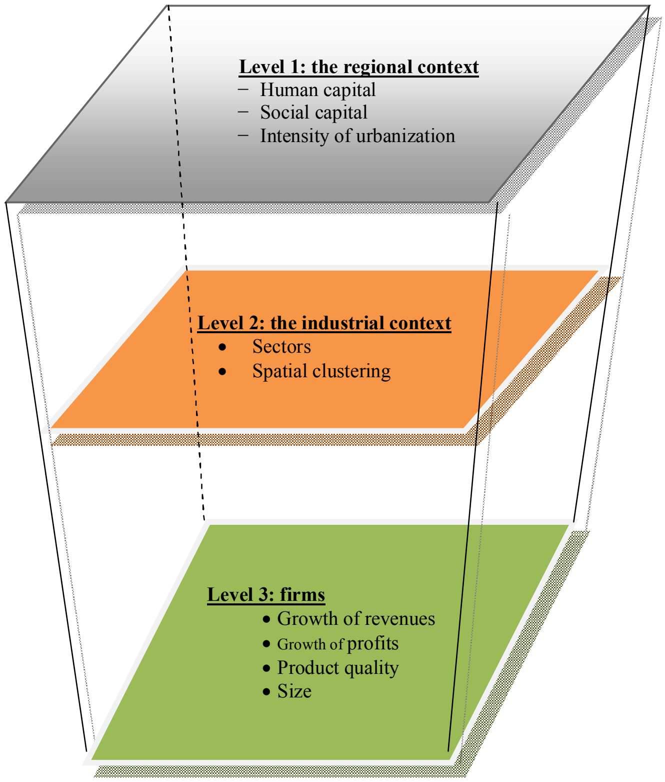
This paper is organized as follows. After this introductory section, we will provide in Section 2 more information on the empirical database. Next, Section 3 will describe the regional and industrial context, where human capital, urbanization, and social capital will extensively be discussed. Section 4 will then analyse the regional variation in the determinants of the economic performance of the firms concerned, while Section 6 will offer concluding remarks.

Description of the sample and methods for data collection

The present paper aims to offer an evidence-based framework for studying the relationship between the performance of high-tech firms and spatial context in the Netherlands. To that end, data was collected via 244 interviews among managers, CEOs, directors etc. working for 61 Dutch high-tech firms. This individual micro data set is next merged with regional attributes and characteristics, collected at the COROP[1]/NUTS3 level. To systematize our data set a triple-layer structure is designed, with a regional level, an industrial level, and a firm level (see Figure 1).

Figure 1

A multi-level approach to firm performance



Source: Authors' elaborations.

1 COROP ("*Coördinatiecommissie Regionaal Onderzoeksprogramma*") regions are statistical areas whose data are collected by the Dutch Central Bureau for Statistics (CPB). At the European level, and, therefore, for EUROSTAT purposes, they correspond to the NUTS3 level.

The above presented multi-level approach to firm performance was inspired by the recently developed '*Flying Disc'* multilevel model by Kourtit and Nijkamp (2012). According to these authors, the architecture of the 'Flying Disc' model of firms' business performance in geographical space 'serves as a strategic navigation instrument that maps out main directions in a comprehensive micro-meso framework, which includes an integrated set of essential locational factors (inputs) in core geographical zones, as well as linkages that determine a firm's micro-business performance (outputs)'. This framework addresses dynamic force fields that occur in cities and regions. This framework is structured to reflect the functional mechanism through which firms, in terms of competitive behaviour, optimize their spatial economic performance. This leads essentially to a conceptual complex and multilevel model. Figure 1 shows that a decomposition of these force fields leads to new and functionally distinct constructs, where each construct is mutually connected with the others. Thus, our multi-level approach is based on a ramification of forces and regional and industrial determinants, that altogether make up an integrated system, which may matter in achieving a higher business performance and success. Each type of force and construct can be associated with a collection of relevant research challenges in order to trace the regional key drivers of the urban agglomeration dynamics that exert a decisive impact on the XXPQP (maximum contribution to productivity, quality, and profitability) of firms. This performance framework is largely similar to the XXQ concept; see Nijkamp (2008). Clearly, this framework can be transformed into an operational measurement model (using actual data) to evaluate and rank the comprehensive performance of innovative firms in the high-tech sector, provided detailed assessments of geographical and urban determinants are available from reliable statistics.

The regional and industrial context

Introduction

In this section, we will organize the dataset in a form that is appropriate for statistical analysis. The firm-level data described in the previous section have been collected by interviewing firms in 18 COROP (NUTS3) Dutch regions, out of 39. The localization of the regions where the sampled firms belong to is shown in Figure 2.

The regional sample covers 54 per cent of the total Dutch population, and 56 per cent of total value added.[2] In the regions belonging to the sample, labour productivity is on average 2 per cent higher than the Dutch average. Within the sample of regions where the interviews took place, labour productivity differentials are much larger: the most productive region is 75 per cent more productive than the Dutch average, while the least productive suffers from a 40 per cent productivity gap with respect to the same mean value.

2 Source of the figures: authors' calculations, on the basis of 2009 NUTS3 level EUROSTAT data.

Figure 2

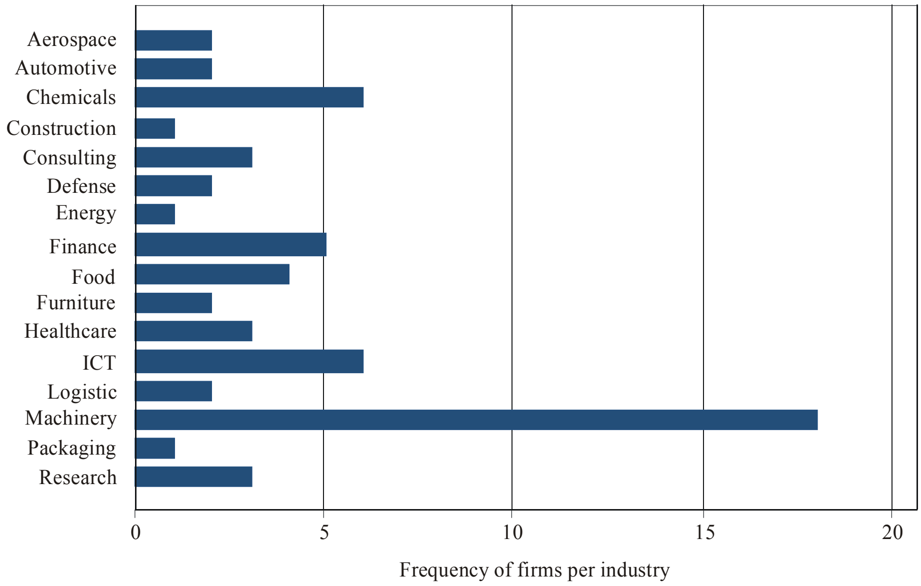
COROP (NUTS3) regions in the Netherlands where at least one firm has been interviewed: the spatial context

Source: Authors' elaborations.

All firms interviewed belong to medium-high and high-tech industries (Figure 3). In particular, a large majority of firms are active in the high-tech machinery sector, with ICTs (Information and Communication Technologies) and chemicals as the second most common sectors (see also Kourtit et al. 2013).

Figure 3

Frequency of the interviewed firms per industry



Source: Authors' elaborations.

Because of the increasing relevance of the regional context for firms' performance (Section 1), the present section presents an overview of the main elements characterizing the regional context where the interviewed firms are active. In particular, Section 3.2 discusses the regional endowment of human capital; Section 3.3 shows how strongly each region hosting at least one interviewed firm is urbanized; and Section 3.4 finally, discusses the wealth of social capital of the regions of the sampled firms. These three elements represent the three core characteristics differentiating regions where the interviews have been carried out, and provide, along with the industrial dimension, the framework for firms' activities and performance. Although in the present paper these dimensions are only analysed and described, they can be used as the ideal setting of an econometric evaluation of firms' performance.

Human capital

At all spatial scales, the endowment of human capital has been found to be remarkably growth-enhancing, in particular through the positive impact on firm performance (Becker 1964, Mincer 1974). Substantial empirical evidence of this claim is also available at the firm level (Crook et al. 2011, for instance, represent a recent convincing meta-analysis on this issue), stressing the importance of an educated and high-skilled workforce for firms' performance.

More recently, because of the increasingly widespread availability of an educated labour force in most Western countries, complex forms of human capital, which move beyond the pure level of education-based definitions provided by earlier literature, have been proposed and analysed. Recent contributions (Wößmann 2003, Vandenbussche et al. 2006, Caragliu et al. 2012) posit that high-level professions, creative capital, and urban knowledge capital are increasingly relevant in determining urban performance.

In order to properly capture the level of human capital in Dutch COROP regions, a measure comprising both traditional (viz. level of education) as well as more recent (i.e. the share of high-level professionals, the wealth of creative capital, and the urban knowledge capital) approaches to human capital is built, by aggregating human capital indicators via a Principal Component Analysis.[3] The resulting human capital indicator is mapped, in Figure 4 for the regions in the sample.

Figure 4 clearly shows a north-west-south east axis of high-human capital regions, with a strong concentration of high values within the *Randstad* conurbation.[4] More remote areas, such as the country's north-east and the regions bordering Germany, present instead a less human-capital rich environment. This issue can be further analysed by inspecting the regions' urbanization level (Section 3.3).

Intensity of urbanization

Figure 5 represents the intensity of urbanization for each area within COROP regions.[5] According to the classical definition provided in Hoover (1936), in fact, firms located in denser urban areas enjoy productivity increases because of cost savings, external both to the firm and to the industry; these are labelled as *urbanization economies*. Indeed, the map represented in Figure 5 clearly matches the information commented on in Section 3.2: firms located in the highly urbanized Randstad area also enjoy, *ceteris paribus*, a better access to human capital, which is typically concentrated in large metropolitan areas. Whether both these productivity advantages are beneficial to the firms' focus on short- and long-term performance indicators will be examined in Section 4.

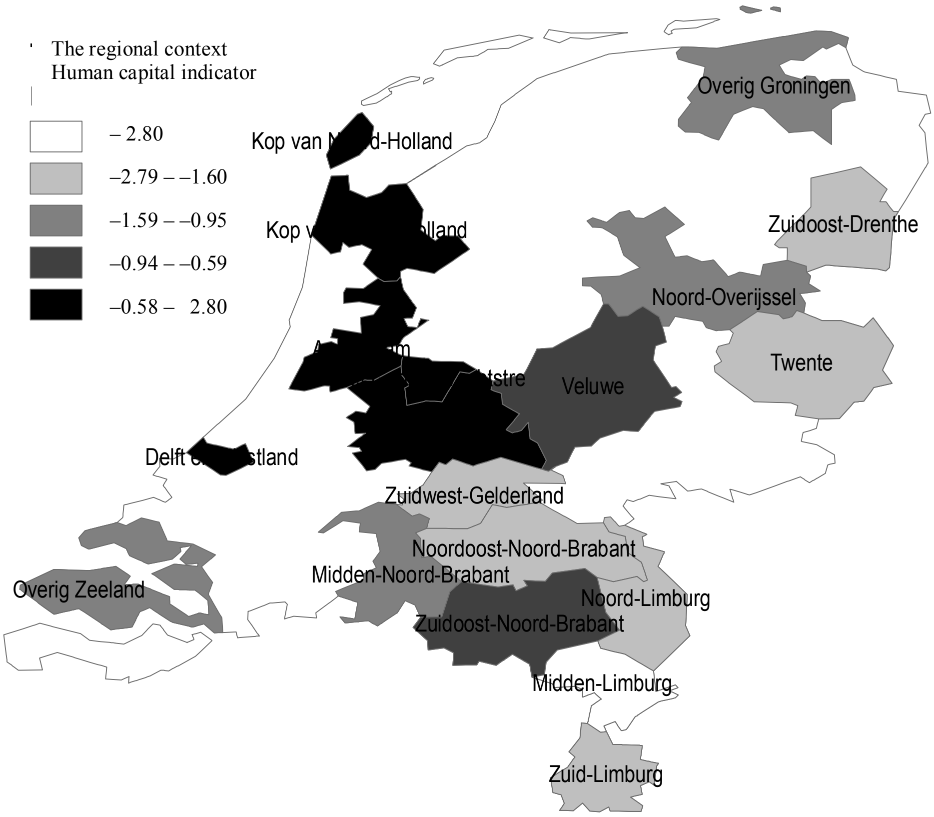
3 The details of the performed PCA and a thorough description of the human capital indicators used can be found in Kourtit et al. (2013).

4 The Randstad is one of the largest conurbations in Europe. It comprises the four largest Dutch cities (viz. Amsterdam, Rotterdam, The Hague, and Utrecht), and, according to the 2011 Randstad Monitor, it hosts more than 6.8 million people, which represents about 40 per cent of the total Dutch population.

5 The data here employed classify areas according to a five-tier classification (from top to bottom, *very strongly urbanized* through *not urbanized*). Data are collected by aggregating the density of registered addresses in 500m –sided squares. These data, along with the COROP/NUTS3 boundaries, are represented in Figure 5. For more details on the methods for data collection, see also CBS (2008).

Figure 4

The regional context: human capital



Source: Authors' elaborations.

Social capital

Social capital (Putnam 2000, Putnam et al. 1993, Fukuyama 1995, Bourdieu 1983, among many others) refers to the set of norms, networks, and institutions forming the glue of a society. Since firms located in a region with a high level of social capital are expected to reap the benefits in terms of productivity advantages (Bosma et al. 2004), the regional endowment of social capital represents a crucial asset for firms' performance. Figure 6 maps social capital values in the 18 analysed COROP regions.[6]

6 The variable being portrayed in Figure 6 results from a PCA performed on trust, norms, and networks COROP regions indicators as explained in Kourtit et al. (2013).

Figure 5

The regional context: level of urbanization

Source: Authors' elaborations.

Figure 6 displays a concentration of high social capital values in South-Eastern COROP regions, and in regions with relatively low levels of urbanization, as exemplified in Figure 5.[7] Whether firms enjoy more relevant productivity advantages from high levels of social of human capital, or from being located in a more urbanized area, requires further research.

Classifying regional variations in the determinants of firm performance

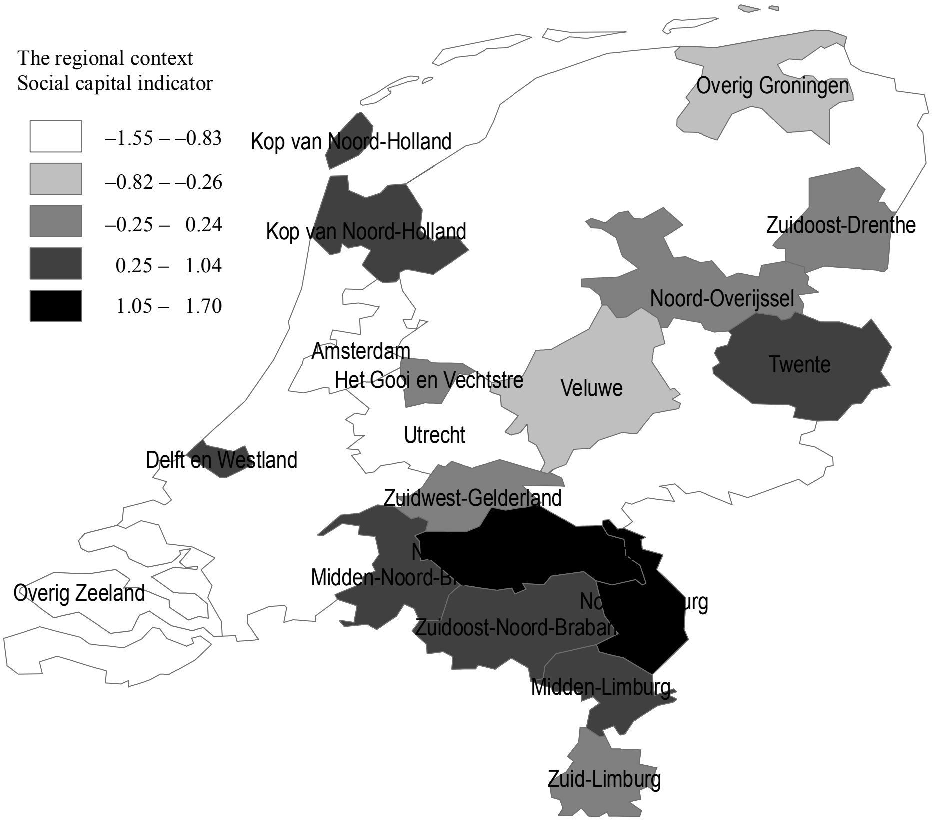
Section 3 has provided a clear description of the relevant spatial clustering of region-specific firm performance determinants in our sample. An a-spatial positive association between firm performance indicators and such characteristics suggests that firm

7 On this issue, also see Section 3.3.

performance, or objectives, could also be spatially clustered. In this Section, we inspect such association more formally.

Figure 6

The regional context: social capital



Source: Authors' elaborations.

Table 1 begins by presenting the extent to which firm objectives, as stated in the questionnaires, referred to in Section 2, are clustered in COROP regions. In the last row, we also present the ratio of the highest to the lowest mean value recorded for all firm objectives indicators considered. We focus in particular on the following goals or objectives for the firm: growth of revenues, growth of profits and product quality, and we report the mean value of the importance of each objective at the firm level (on an increasing scale from 1 to 5) in each COROP region.

The evidence presented suggests that while profits (second indicator, Table 1) represent a common goal, with a similar relative importance for most firms, irrespective of their location, the growth of revenues (first indicator, Table 1), and product quality (third indicator, Table 1) present relevant spatial clustering. Focusing on product quality, this final goal for firms' objective function is characterized by wide spatial differences. This is further corroborated, from results in the last row of Table 1, by a relevant difference (quantified by a ratio of 5) between the region in which such objective is least important (viz. Zuidwest-Gelderland), and the regions in which it is of maximum relevance (i.e., the NUTS3 regions of Overig Groningen and Veluwe).

Table 2 classifies the same firm objectives presented in Table 1 by firm size. Interestingly, it appears that smaller firms not only present a higher focus on short-term objectives, such as the mean growth of revenues or profits (as suggested in several previous contributions; see for instance Audretsch, 1991), but also, although less significantly so, they consider investing in increased product quality a relevant objective in their activities.[8]

Table 1

Spatial clustering of firm objectives in Dutch COROP regions

Name of COROP/NUTS3 region	NUTS3 code	Mean growth of revenues	Mean growth of profits	Mean product quality
Overig Groningen	NL113	4.00	4.00	5.00
Zuidoost-Drenthe	NL132	4.00	4.00	3.00
Noord-Overijssel	NL211	4.00	4.00	3.00
Twente	NL213	3.00	4.00	3.00
Veluwe	NL221	4.00	4.00	5.00
Achterhoek	NL222	3.00	3.00	2.00
Arnhem/Nijmegen	NL223	4.00	4.00	3.00
Zuidwest-Gelderland	NL224	2.00	2.00	1.00
Utrecht	NL310	2.96	2.83	2.94
Kop van Noord-Holland	NL321	4.00	4.00	4.00
Amsterdam	NL326	2.86	3.14	3.00
Het Gooi en Vechtstre	NL327	3.50	4.00	3.50
Agglomeratie Leiden	NL331	1.00	4.00	4.00
Delft en Westland	NL333	3.00	3.00	3.00
Rijnmond	NL335	2.50	2.50	2.00
Overig Zeeland	NL342	3.00	3.00	4.00
Midden-Noord-Brabant	NL412	4.00	3.33	3.33
Noordoost-Noord-Brabant	NL413	4.00	4.00	4.00
Zuidoost-Noord-Brabant	NL414	3.23	3.38	3.54
Noord-Limburg	NL421	3.00	4.50	3.50
Midden-Limburg	NL422	3.00	4.00	3.00
Zuid-Limburg	NL423	2.50	4.00	3.00
Ratio of highest to lowest		*4.00*	*2.25*	*5.00*

Source: Authors' elaborations.

8 Remarkable statistical significance is associated with the different relevance of revenues growth, (the p-value of the classical t-tests for differences in group means is equal to 0.06). All other t-tests turn out to be marginally not significant.

Table 2

Firm objectives and firm size

Firm size	Mean growth of revenues	Mean growth of profits	Mean product quality
Small	3.26	3.41	3.14
Large	2.79	3.36	3.11
Ratio of highest to lowest	*1.17*	*1.01*	*1.01*

Source: Authors' elaborations.

Table 3 classifies the three-abovementioned firm objectives according to the regions' level of human capital, measured by the indicator described in Section 3.1. In this case, and somewhat unexpectedly, firms located in regions with a lower than average level of human capital pay more attention to profits growth and product quality, and significantly so.[9]

Table 3

Firm objectives and the region's human capital

Wealth of human capital	Mean growth of revenues	Mean growth of profits	Mean product quality
Low	3.28	3.62	3.38
High	2.97	3.19	2.91
Ratio of highest to lowest	*1.10*	*1.14*	*1.16*

Source: Authors' elaborations.

This piece of evidence can be usefully confronted with the same type of test performed by considering the intensity of urbanization of the region where the firm is located. From results presented in the vast literature explaining how skilled workers concentrate in urban areas (e.g. Moretti 2004, among countless others), one would expect more urbanized areas to host firms committing more to all the three objectives, above all product quality.

Table 4 shows instead that strongly urbanized areas host firms which pay the lowest attention to all three objectives. The highest focus on revenues growth is registered for areas with an intermediate degree of urbanization, which typically represent companies active in mature industries, looking for residential locations with relatively lower rent levels, in order to reap the productivity benefits stemming from large-scale production. On the contrary, firms located in denser urban areas may actually look for different objectives, mostly related to knowledge access, as suggested also by the knowledge spillovers (Anselin et al. 1997) and the milieu innovateur (Camagni 1991) literature. Such statements can be further crosschecked in Figure 7 below.

9 The p-value associated to the classical t-tests for differences in group means, is equal to 0.04 in both cases.

Table 4

Firm objectives and the region's level of urbanization

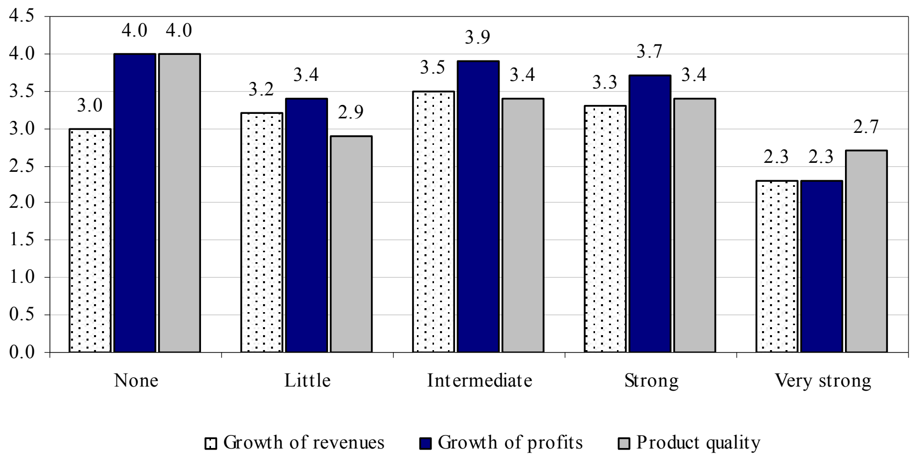
Level of urbanization	Mean growth of revenues	Mean growth of profits	Mean product quality
Not urbanized	3.00	3.00	4.00
Little	3.21	3.43	2.86
Intermediate	3.50	3.88	3.38
Strong	3.31	3.69	3.38
Very strong	2.33	2.33	2.66
Ratio of highest to lowest	*1.50*	*1.66*	*1.50*

Source: Authors' elaborations.

The apparently contradictory results obtained here, viz. that firms located in moderately urbanized areas and in areas with a lower than average human capital, seem to be significantly more interested in profits and revenues growth, and in product quality, deserves further research efforts.

Figure 7

Firm objectives and the region's level of urbanization



Source: Authors' elaborations.

Finally, Table 5 presents similar statistics, with a classification of firm objectives by levels of regional social capital.

Table 5

Firm objectives and the social capital

Wealth of social capital	Mean growth of revenues	Mean growth of profits	Mean product quality
Low	2.87	3.23	2.97
High	3.37	3.57	3.30
Ratio of highest to lowest	*1.17*	*1.11*	*1.11*

Source: Authors' elaborations.

In this case, the results are in line with previous findings on the positive association between social capital (in the form of stronger weak ties, higher trust, and more intense sociability). For all three firm objectives analysed here , the results from the plant-level sample observed suggest that firms pay uniformly more attention to profits and revenues growth, and product quality, as the level of social capital of the region in which they are located increases.

This last result is however slightly at odds with the previously described findings on urbanization levels and human capital. In fact, social capital is consistently found to be positively associated (with a one-to-one relationship) with human capital (Coleman 1988). Societies with a more educated labour force seem to enjoy higher levels of trust, while the reverse is also true (i.e., societies with better weak ties/stronger bonds and higher trust tend to invest more time and money in accumulating human capital).

The picture stemming from the analysis of the industrial and regional environments where the firms interviewed are active clearly calls for a deeper analysis of the causal direction of the relations depicted here. Particular attention should finally be paid to the level of analysis where each impact between firm performance determinants, profit and revenues growth, and product quality achievements, takes place.

Concluding remarks

Spatial dynamics is the outcome of a complex socio-economic and technological force field, with a complex ramification of micro-based motives (e.g. entrepreneurial performance strategies) and macro-oriented (or meso-oriented) contextual drivers. The study of the nexus firms-regions calls for a thorough statistical data collection and advanced statistical analyses.

The present analysis, using the Netherlands as a case study, aims to identify the determinants and success factors of competitive innovation strategies of high-tech firms. Their locational behaviour (either the search for new locations or the decision to stay where they are) appears to be a key factor for their performance. Clearly, location does not refer here to the specific locality of a firm, but also to the broader urban regions and spatial networks in which a firm has to operate. Our study offers an evidence-based illustration of the great importance of local and regional conditions for firms' strategies and behaviour.

It should be added that there is much scope for a further enhancement of our insights into the drivers and effects of spatial dynamics. Systematic database collection at a micro (firm) level, conceptualization of the various forces at work (in a multi-level configuration), and development of advanced statistical-econometric tools that are fit for purpose are indispensable for new insights into a complex space-economy.

REFERENCES

Anselin, L. – Varga, A. – Acs, Z. (1997): Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics* 42 (3): 422–448.

Audretsch, D. (1991): New firm survival and the technological regime. *Review of Economics and Statistics* 73 (3): 441–450.

Becker, G. S. (1964): *Human capital: a theoretical and empirical analysis, with special reference to education*, Chicago (IL): University of Chicago Press.

Bosma, N. – van Praag, M. – Thurik, R. – De Wit, G. (2004): The Value of Human and Social Capital Investments for the Business Performance of Startups. *Small Business Economics* 23(3): 227–236.

Bourdieu, P. (1983): Ökonomisches kapital, kulturelles kapital, soziales kapital. In: Kreckel, R. (ed.) *Soziale Ungleichheiten (Soziale Welt, Sonderheft 2)*. pp. 183–98. Otto Schartz & Co., Goettingen.

Camagni, R. (1991): Local milieu, uncertainty and innovation networks: towards a new dynamic theory of economic space. In: Camagni, R. (ed.) *Innovation networks: spatial perspectives*. Belhaven Press, London.

CBS (Centraal Bureau voor de Statistiek) (2008): *Bevolkingskernen in Nederland, 2008*. Retrieved online on Oct 12, 2012 at the URL http://www.cbs.nl/nl-NL/menu/themas/dossiers/nederland-regionaal/publicaties/geografische-data/archief/2012/2010-bevolkingskernen-in-nederland-2008.htm.

Coleman, J. S. (1988): Social capital in the creation of human capital. *American Journal of Sociology* 94 (Supplement): S95–S120.

Crook, T. R. – Todd, S. Y. – Combs, J. G. – Woehr, D. J. – Ketchen , D. J. (2011): Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. *Journal of Applied Psychology* 96 (3): 443–456.

Hoover, E. M. Jr. (1936): The measurement of industrial localization. *The Review of Economics and Statistics* 18 (4): 162–171.

Fukuyama, F. (1995): *Trust: the social virtues and the creation of prosperity*. Free Press, New York (NY).

Kourtit, K. – Nijkamp, P. – Stough, R. (eds.) (2011): *Drivers of Innovation, Entrepreneurship and Regional Dynamics*. Springer, Berlin.

Kourtit, K. – Nijkamp, P. (2012): In Search of Creative Champions in High-Tech Spaces; A Spatial Application of Strategic Performance Management. *Journal of Regional Science* (forthcoming).

Kourtit, K. – Nijkamp, P. – Caragliu, A. – Del Bo, C. (2013): Impacts of Multi-Level Capital Resources on Local Business Performance. In: Kourtit, K. – Nijkamp, P. – Stimson, R. (eds.) *Applied Modelling of Regional Growth and Innovation Systems*. Springer, Berlin. (forthcoming).

Mincer, J. (1974): *Schooling, experience and earnings*. Columbia University Press, New York (NY).

Moretti, E. (2004): Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics* 121 (1-2): 175–212.

Nijkamp, P. (2008): XXQ Factors for Sustainable Urban Development: A Systems Economics View. *Romanian Journal of Regional Science* 2 (1): 1–34.

Putnam, R.D. (2000): *Bowling alone: the collapse and revival of American community*. Simon & Schuster, New York (NY).

Putnam, R.D. – Leonardi, R. – Nanetti, R.Y. (1993): *Making democracy work: civic traditions in modern Italy*, Princeton University Press, Princeton (N.J).

Schumpeter, J. (1934): *The Theory of Economic Development*. Harvard University Press, Cambridge MA.

Vandenbussche, J. – Aghion, P. – Meghir, C. (2006): Growth, distance to frontier and composition of human capital. *Journal of Economic Growth* 11 (2): 97–127.

Keywords: innovation, high-tec, clusters, performance.

Abstract

This paper offers a statistical analysis of the regional contextual drivers and aspects of innovative firms in the high-tech sector in the Netherlands. Data are collected by means of 244 interviews among actors working for 61 Dutch high-tech firms. This individual micro data set is next merged with regional attributes and characteristics, collected at the COROP (NUTS3 level) in the Netherlands in order to analyse the relationship between the economic performance of individual firms and the broader regional economic environment.

REGIONAL STATISTICS, 2012, VOL. 2: 27–44 DOI: 10.15196/RS02103

IMRE LENGYEL – IZABELLA SZAKÁLNÉ KANÓ

Competitiveness of Hungarian Urban Micro-regions: Localization Agglomeration Economies and Regional Competitiveness Function*

Introduction

Increasing regionalization represents one of the most spectacular processes of the economies that develop and transform as a result of globalization processes; while the (relative) importance of national economies is decreasing, the economic role of regions and cities seems to grow. Global competition has also intensified spatially, especially with the growing importance of the agglomeration economies. Interregional competition, which refers to the competition of regions and cities for scarce resources, educated human labour, investments etc., is increasingly prevalent (Enyedi 2009).

It appears to be generally accepted in regional science these days, that there is some sort of competition among regions, but this may be characterized by different attributes such as the competition among corporations or countries (Batey–Friedrich 2000, Chesire 2003, Malecki 2002). Capello states (2007a, xviii): 'Regions compete on absolute rather than comparative advantage". The results of interregional competition are similar to those of the competition among countries: welfare (living standard) improves in the successfully competing regions, employment and incomes (wages) are high, new investments take place, talented young people and successful businessmen migrate there, etc. (Malecki 2004, Polenske 2004). Successfulness in competition, or in other words, competitiveness has been one of the key concepts over the past two or three decades partly due to the sharpening of global competition (Camagni 2002).

Today territorial competitiveness covers both economic growth and economic development. This complex point of view is well demonstrated by the fact that Capello (2007a) emphasizes the connections between territorial competitiveness and local development, as well as regional growth (both for endogenous and exogenous) in her book entitled 'Regional economics'. However, while theoretical approaches of econometric regional growth between 1960 and 1990 were based on increasing productivity and individual welfare indicators as described by traditional neoclassical models, the shift in the 1990s resulted in a definite turn towards strengthening competitiveness (Capello 2007b). In territorial endogenous growth theories, regional growth is the result of partly independent mechanisms (Capello 2007b, pp. 757–758): a competitive process, a socio-relational process, a territorial and spatial process, an interactive process, and an endogenous process.

* This research was supported by the TÁMOP 4.2.1/B-09/KONV-2010-0005 project of the Hungarian National Development Agency.

The modes of improving regional competitiveness and regional economic development strategies are heavily dependent on the type of the given region. This is because regions in different phases of their development are in different positions when it comes to interregional competition. Porter et al (2008) classified these phases as: resource-driven stage, investment-driven stage, and innovation-driven stage. These categories are especially important in understanding regional development in transition economies, where regions are hardly in the innovation-driven phase (Lengyel–Cadil 2009, Lengyel–Leydesdorff 2010, Lengyel 2009b). However, based on agglomeration advantages Budd-Hirmis (2004) points out that metropolitan regions with urbanization agglomeration economies are competing with more emphasis on their comparative advantages, while regions of localization agglomeration economies tend to compete on competitive advantages. McCann (2008) considers that size of regions is a strong influential factor when it comes to the organization of clusters, which play a very important role in interregional competition: pure agglomeration (urban), industrial complex (local but not urban), and social networks (local but not urban).

The next section of this paper covers the pyramidal model of regional competitiveness. This model is a logical systematization for measuring endogenous regional growth and the factors influencing it; the model will be used to introduce the regional competitiveness function (RCF). After introducing the theoretical model, we are going to investigate the competitiveness of Hungarian urban micro-regions (LAU1) with a population of above 50 thousand citizens. Our statistical analysis to underline the classification of micro-regions by competitiveness types is based on the multivariate linear regression analysis.

Pyramidal model and regional competitiveness function (RCF)

Three major issues emerged in the debates aiming at the interpretation of competitiveness (Barkley 2008): *(1) how to define regional competitiveness and its factors; (2) what indicators should be used to measure it;* and *(3) how can regional competitiveness be improved?* These three questions usually lie in the background of other professional debates too; while representatives of regional science concentrate on the first one, the regional economist on the second one, the experts of regional policy tend to focus on the third one.

There were a number of attempts to define the new notion of competitiveness according to new global competition conditions in the mid-1990s. The *standard notion of competitiveness* in the European Sixth Regional Periodic Report of EU (EC 1999): '*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*'. In the European Competitiveness Report (EC 2008, p. 15): "Competitiveness is understood to mean a sustained rise in the standards of living of a nation or region and as low a level of involuntary unemployment, as possible." In other words 'high and rising standards of living and high rates of employment on a sustainable basis' (EC 2001).

Porter (2007) suggests using prosperity, measured by standard of living and inequality for *measuring regional competitiveness*. Prosperity, defined by per capita income is decomposed into two factors: labour productivity and labour utilization.

Factors influencing labour productivity are skills, capital stock, and total factor productivity. Factors of labour utilization are working hours, unemployment, and workforce participation rate (population age profile).

Kitson, Martin, and Tyler (2004) use three indicators for measuring competitiveness: regional productivity, employment, and standard of living. They also claim that competitiveness is influenced by hard and soft elements as well. The bases of the regional competitive advantage are: productive capital, human capital, social-institutional capital, cultural capital, infrastructural capital, and knowledge/creative capital. The region-specific economic and social qualities, like social capital, knowledge/creative capital, and territorial capital are gaining more and more in importance (Camagni 2009, Lengyel I. 2009a). Thus, regional competitiveness studies are increasingly influenced by theories of endogenous growth and development.

Stimson, Stough, and Salazar (2009) suggest a new conceptual model framework for *regional endogenous development*. The dependent variable of endogenous growth is measured by two indicators, on one hand by the change of employment or income, on the other hand by an employment-based location quotient (LQ). Explanatory variables include, among others, resource endowments (estimated by 13 indicators) and market fit (measured by 4 indicators). Their model includes several indicators for leadership quality, as well as institutions and entrepreneurship.

The standard notion of competitiveness obtained in this way cannot be used, however, to identify factors responsible for regional competitiveness or areas, which are to be strengthened or developed by regional development policies and programmes for improved competitiveness. Since the notion of competitiveness can be seen as refining that of economic growth, it can often be observed that proposals for improved competitiveness combine traditional means of economic development with methods based on endogenous development.

The standard definition refers to "relatively high income". This can be measured by means of the per capita GDP and the GDP growth rate. A high employment level is in turn indicated by the rate of employment. These two indicators can be measured independently from one another, but per capita GDP can also be expressed as follows, respectively:

$$\frac{GDP}{total.population} = \frac{GDP}{employment} \times \frac{employment}{working - age.pop.} \times \frac{working - age.pop.}{total.population}.$$

This formula suggests that measuring regional competitiveness can be traced back to two interdependent economic categories (Lengyel 2004):

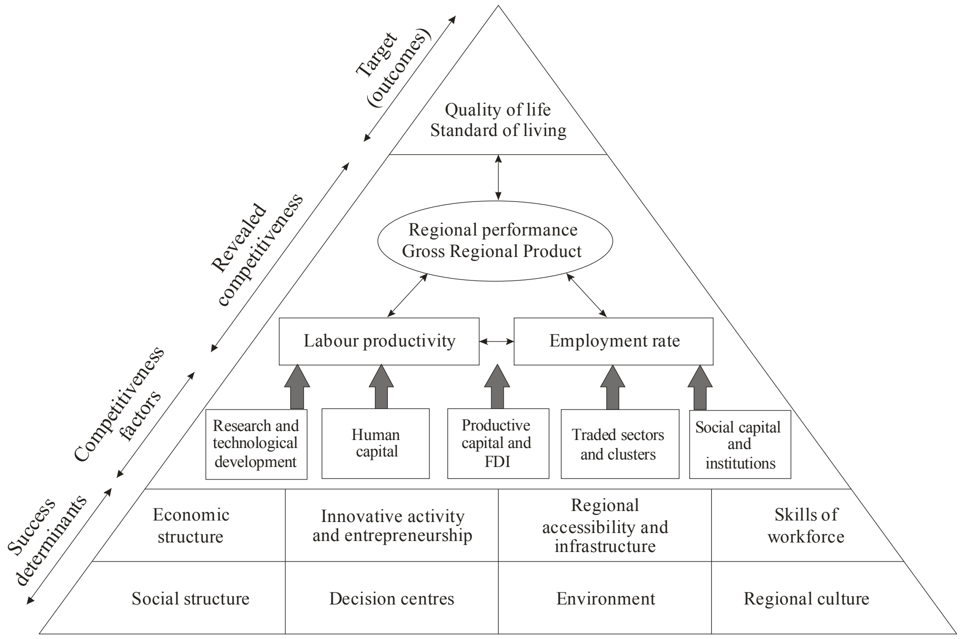
$$\textit{Regional income} \cong \textit{Labour productivity} \times \textit{Employment rate}.$$

Hence the *substance of regional competitiveness*: the economic growth in the region, which growth is generated by both a *high level of labour productivity* and a *high level of employment*. In other words, competitiveness means *economic growth driven by high productivity and a high employment rate*.

Our study reviewing the competitiveness of Hungarian micro-regions is built on the pyramidal model since it is coherent with the above-mentioned findings, and is established on the basis on the inputs- outputs- outcomes relationship (Figure 1). The target (outcomes) is the standard of living; the prosperity of any region depends on its competitiveness. Outputs are the *revealed competitiveness* indicators: per capita Gross Regional Product, labour productivity, and employment rate. Sources of competitiveness, inputs influencing regional competitiveness can be divided into two groups of *direct* and *indirect* components. Of particular importance are *competitiveness factors* with a direct and *short-term influence* on economic output, labour productivity, and employment rates. Nevertheless, social, economic, environmental and cultural processes and parameters, the '*success determinants*', with an indirect, *long-term impact* on competitiveness are also to be taken into account.

Figure 1

The renewed pyramidal model of regional competitiveness



Source: Based on Lengyel, I. (2000, 2004).

Three levels can be distinguished with regard to the objectives of regional development programming and the various characteristics and factors influencing competitiveness:

- *Revealed competitiveness* (or basic categories) (*ex post* indicators, output): these categories measure competitiveness and include income, labour productivity and employment rate.
- *Competitiveness factors* (*ex-ante* factors): input factors with an immediate impact on revealed competitiveness categories. These can be used to influence regional competitiveness by means of institutions in short-term programming periods.
- *Success determinants* (social, economic, and environmental backgrounds): input determinants with an indirect impact on basic categories and competitiveness factors. These determinants take shape over a longer period and their significance reaches beyond regional policy-making.

The *pyramidal model of regional competitiveness* seeks to provide a systematic account of these means and to describe the basic aspects of territorial competitiveness. ‘This model is useful to inform the development of the determinants of economic viability and self-containment for geographical economies’ (Pike et al. 2006, p. 26). ‘This is an aggregate notion, in a regional context, labour productivity is the outcome of a variety of determinants (including the sort of regional assets alluded to previously). Many of these regional factors and assets also determine a region’s overall employment rate. Together, labour productivity and employment rate are measures of what might be called ‘revealed competitiveness’, and both are central components of a region’s economic performance and its prosperity (as measured, say, by GDP per capita), though obviously of themselves they say little about the underlying regional attributes (sources of competitiveness) on which they depend’ (Gardiner–Martin–Tyler 2004, p. 1049).

Competitiveness factors of the *renewed pyramidal model* include such constituents of endogenous development theory like social capital and regional specialization, besides traditional factors of production like capital, labour, and technology:

- RTD – *Research and technological development (RTD)*: rapid introduction of innovations and new technologies creates competitive advantages. Innovation may come from outside the region (e.g. technological transfer), but the competitiveness of the region is most effectively advanced by successful R&D activities, innovations and their fast and wide-ranging distribution. The introduction of innovations and creation of patents may be effectively advanced by knowledge-intensive businesses.
- HUM_CAP – *Human capital*: population of active age, size and age structure of the workforce are important growth factors. However, the education level of the workforce is also important, especially the rate of people holding a tertiary degree.
- CAP_FDI – *Productive capital and FDI*: capital is indispensable for improving the competitiveness of a region. Investments from outside the region, especially foreign direct investments, usually create new sectors, markets, new technologies, and new jobs. It also improves labour productivity and can also encourage technological transfer.
- TS_CLUST – *Traded sectors, entrepreneurship, and clusters*: a strong traded (export-oriented) sector is an important source of competitiveness, which may become even more competitive by clustering. Flexible regional specialization may be furthered by entrepreneurship and small and medium-sized enterprises (SMEs). Innovative SMEs are flexible and can quickly adapt to market changes, they are principally responsible for generating employment in the region.
- SOC_CAP – *Social capital and institutions*: economic prosperity also presupposes efficient cooperation among firms, governmental and non-governmental institutions. Successful companies also depend on the level of administrative services and public institutions. Social capital is particularly important: trust, reliability, readiness to cooperate, etc.

In order to investigate the relations between indicators of revealed competitiveness (RC) and competitiveness factors, we intend to introduce the *Regional Competitiveness Function* (RCF):

$$RC = f\,(RTD, HUM\_CAP, CAP\_FDI, TS\_CLUST, SOC\_CAP).$$

The basic premise of the study is thatwe assume that there is a relationship between competitiveness factors and revealed competitiveness. Causality is to be determined by multivariate regression. Our dependent variable is revealed competitiveness measured by a calculated index, while the five competitiveness factors are explanatory variables.

RCF is an extension of traditional regional growth concepts from the latest work on endogenous growth research. The importance of traded sectors and regional specialization is pointed out by Porter (2003, 2008), Stimson, Robson, and Shyy (2009), while Acs and Szerb (2007), Fischer and Nijkamp (2009) emphasize the significance of SMEs and entrepreneurship, and Varga (2006, 2007) stresses the importance of innovation and knowledge spillover. Sociological research alludes to the importance of social capital (and territorial capital), brought to the attention of regionalists by Camagni (2009), Faggian and McCann (2009), Florida (2002) and Glaeser (2008).

The weight of each RCF competitiveness factor in measuring revealed competitiveness was assessed during our study of Hungarian micro-regions. This assessment excluded the success determinants of the pyramidal model, because we assume that the RCF is mainly useful for describing short-term relationships.

Background of competitiveness studies in Hungary

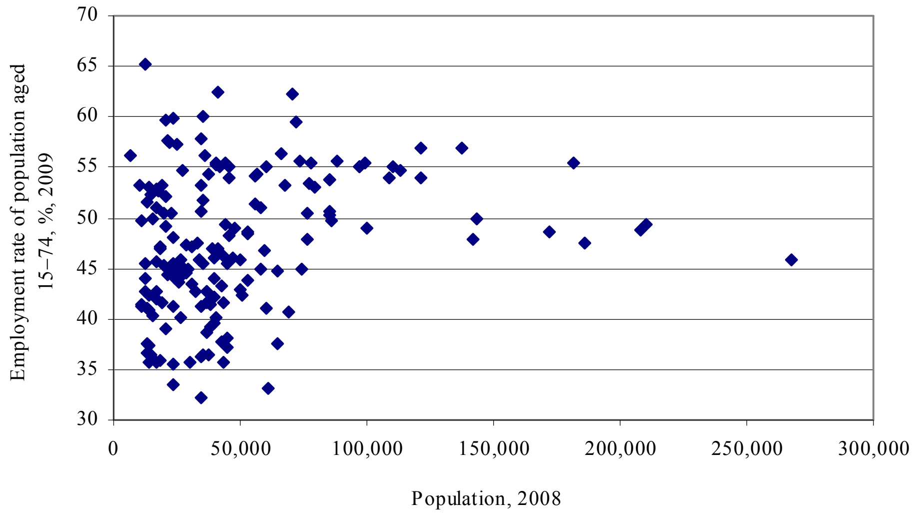
Regional competitiveness studies tend to be relative, i.e. we mostly compare the competitiveness of the chosen regions to each other. It is recommended to choose nodal regions, because workforce commuting, business relationships, etc. rarely adhere to the spatial distribution of normative regions. It is difficult to gather reliable statistical data about nodal (functional) regions, thus Level LAU1 micro-regions were chosen this time. We assume that, except for Budapest, micro-regions are able to provide a good assumption of workforce commute zones (Lukovics 2009, Szakálné Kanó 2011).

In 2008, Hungary consisted of 7 regions (NUTS 2), 19 counties (NUTS 3) and the capital, as well as 174 micro-regions (LAU 1). Statistical data usable for competitiveness investigations are available for these territorial levels. All LAU1 micro-regions have a town centre.

The indicators of *revealed competitiveness* (GDP per capita, employment, labour productivity) show a broad distribution in LAU1 micro-regions. Examining *employment rates* by micro-regions based on their populations, one may get a very diversified distribution (Figure 2). Employment rates in micro-regions with less than 70 thousand inhabitants (four fifth of micro-regions) are distributed evenly, mostly between 35% and 60%. In those 31 micro-regions with more than 50 thousand inhabitants in their town centres (so-called *urban micro-regions*), employment rates vary between 45 and 55% (in Budapest it is 56.6%). It can be established that the critical mass, population as employees and consumers, as well as more sophisticated business and other urban services are crucially important factors in the development of employment (Bajmócy–Szakálné Kanó 2009).

Figure 2

Employment rate and population of LAU1 micro-regions

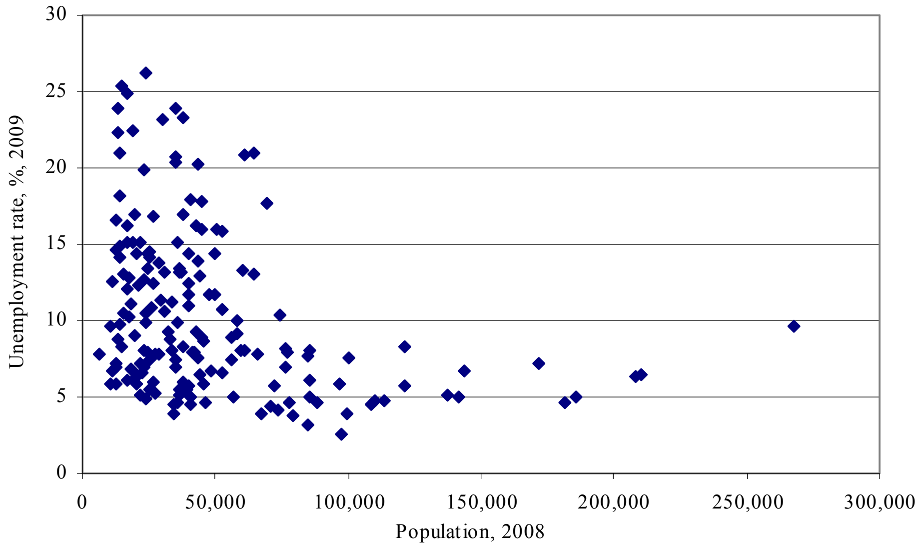


Source: Calculations of authors based on National Employment Office (http://kisterseg.afsz.hu/index.php) and KSH Territorial Statistical Yearbook.

Note: Without Budapest.

Figure 3

Unemployment rate and population of micro-regions



Source: Calculations of authors based on National Employment Office (http://kisterseg.afsz.hu/index.php) and KSH Territorial Statistical Yearbook.

Note: Without Budapest.

Unemployment rates have an opposite relationship (Figure 3). With this indicator, an important milestone can also be seen at 50 thousand urban inhabitants: more populated micro-regions have unemployment rates of 5 to 10%, while less populated micro-regions have between 7 and 28%. No influence on this situation can be seen in micro-regions with less than 50 thousand urban inhabitants, as these have a similar distribution as larger ones.

Our empirical study includes urban micro-regions, with more than 70 thousand inhabitants (and more than 50 thousand urban inhabitants), potentially able to show *localization agglomeration advantages*. The groups of 174 micro-regions, according to agglomeration economies:

– *Budapest* (population of 2 million): urbanization agglomeration economies (Jacobs' externalities),
– *31 micro-regions with urban centre, as urban micro-regions* (at least 50,000 population of urban centres, sum total 3.6 million population): localization agglomeration economies (Marshall' externalities),
– *142 small (rural type) micro-regions* (sum total 4.4 million population).

Budapest was intentionally left out of this study due to its vastly different characteristics. To sum up, urban micro-regions with potential localization agglomeration economies were studied by using the pyramidal model.

Empirical testing of the Regional Competitiveness Function

Our empirical study included the competitiveness of 31 urban micro-regions. Goals of the investigation:

– comparison of these micro-regions by competitiveness, ranking, establishment of region types,
– to show how the indicators and indicator groups used influence regional competitiveness.

Our study adheres to the logical construction of the pyramidal model. Revealed competitiveness indicators show recently achieved competitiveness as ex-post indicators. Competitiveness factors point out their contribution to revealed competitiveness. On the other hand, these show 'capabilities' and future possibilities as ex ante indicators: by developing these, we can see how the competitiveness of micro-regions might change in the near future.

Difficulties were liable to occur during the database creation process, because several theoretical categories (like social capital) are not straightforward to operationalize, and it is difficult to obtain reliable and authentic data for all Hungarian micro-regions (Bajmócy–Lukovics–Vas 2010). Computer analysis was done with SPSS-18.[1]

The basic idea of our study: we assume that there is a relationship between competitiveness factors and revealed competitiveness. Causality is to be determined by multivariate regression. Our *dependent variable* is revealed competitiveness measured by a calculated index, while the 5 competitiveness factors are explanatory variables.

1 Micro-regional competitiveness indicators and database were collected by Miklós Lukovics, Zoltán Bajmócy and György Málovics, thanks for their help.

Our multivariate linear regression model:

$$RC= \beta_0 + \beta_1 RTD + \beta_2 HUM\_CAP + \beta_3 CAP\_FDI + \beta_4 RS\_CLUST + \beta_5 SOC\_CAP + \varepsilon.$$

The indicators used were set up based on the pyramidal model (see *Appendix 1*):

- revealed competitiveness (RC) is calculated by 3 indicators,
- competitiveness factors are described by a total of 34 indicators: RTD (5 indicators), HUM_CAP (9 indicators), CAP_FDI (6 indicators), RS_CLUST (6 indicators), SOC_CAP (8 indicators).

To test RCF, we first calculated the value of revealed competitiveness; afterwards we analysed it with multivariate linear regression to determine how far competitiveness factors are able to explain the value of revealed competitiveness.

(a) Revealed competitiveness

Micro-regions may show enormous distortion due to data localization, therefore it might be misleading to calculate GDP, also major companies are calculated as being a one-point business at their headquarters' location. Therefore we concluded that 3 out of the revealed competitiveness (PIT_INH: taxable income per capita; GVA_EMPL: gross value added per employee; EMPL_RATE: employment rate) shall undergo principal component analysis to determine the principal component (RC), which shall be used later on as the dependent variable*:*

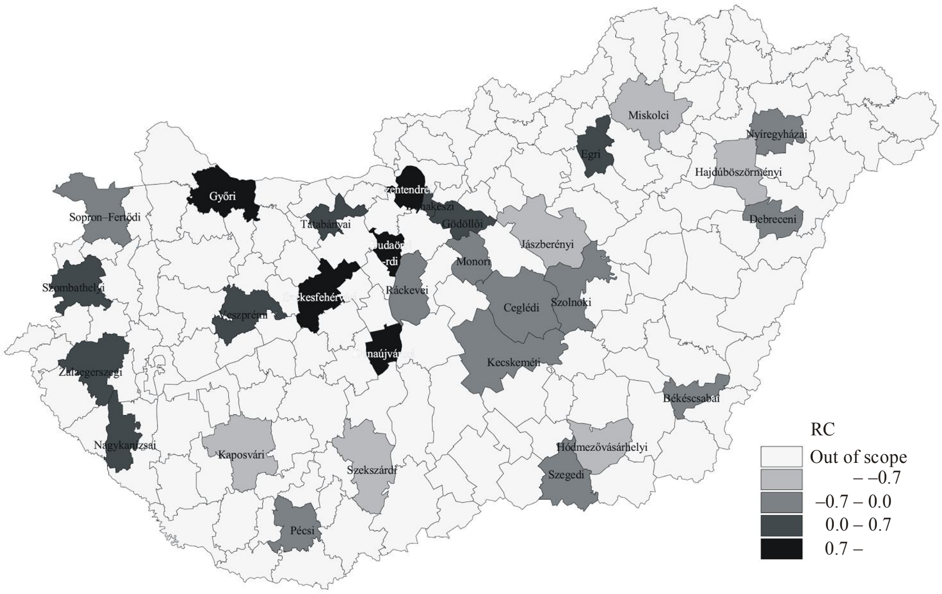
- RC contains 60.7% of the 3 indicators,
- commonalities: PIT_INH 0,835; GVA_EMPL 0,5; EMPL_RATE 0,485.

Based on principal component analysis we found four types of Hungarian urban micro-regions (Figure 4):

- the most competitive 6 micro-regions are found in Transdanubia (Dunaújváros, Győr, Székesfehérvár) with significant foreign-owned manufacturing capacities, as well as in the western agglomeration of Budapest,
- the second type (8 micro-regions) includes all other Northern Transdanubian micro-regions with some further micro-regions to the east of Budapest,
- the third type (11 micro-regions) includes other county capitals, with poor economy and human capital, as well as in the southern agglomeration of Budapest,
- while the least competitive 6 regions are found in the southern and eastern part of the country with some rural settlements.

Figure 4

Types of micro-regions by revealed competitiveness



(b) Relationship between competitiveness factors and RC

The analysis included the effect of the 5 competitiveness factors of the renewed pyramidal model on the dependent variable (RC). Each competitiveness factor was based on 5 to 9 indicators, therefore we performed factor analysis within the indicator group in order to compress information and establish 1 to 2 factors per indicator group:

– RTD (research and technological development): one single factor, including 68% of information,
– HUM_CAP (human capital): two factors, one containing 36.8% (HUM_CAP1), the other 33.6% (HUM_CAP2) of the information,
– CAP_FDI (productive capital and FDI): one single factor, including 68 % of the information,
– TS_CLUST (regional specialization and clusters): two factors, one explaining 39.3% (TS_CLUST1), the other 36.1 % (TS_CLUST2) of the information,
– SOC_CAP (social capital and institutions): two factors, one explaining 31.6% (SOC_CAP1), the other 30.0% (SOC_CAP2) of the information available.

The above-mentioned 8 factors were used in multivariate linear regression, where RC was considered a dependent variable and the forward method resulted in 2 factors: CAP_FDI and SOC_CAP2. These two factors account for 85.2% (R^2=0.852) of the dependent variable's (RC) standard deviation.

The model created:

$$RC_i = + 0{,}452\ CAP\_FDI_i - 0{,}615\ SOC\_CAP2_i + e_i.$$

The regression model provides adequate explanation for the dependent variable:

- there is no multicollinearity to observe, VIF=1.308,
- residuals show a normal distribution,
- there is no heteroscedasticity to observe.

Table 1

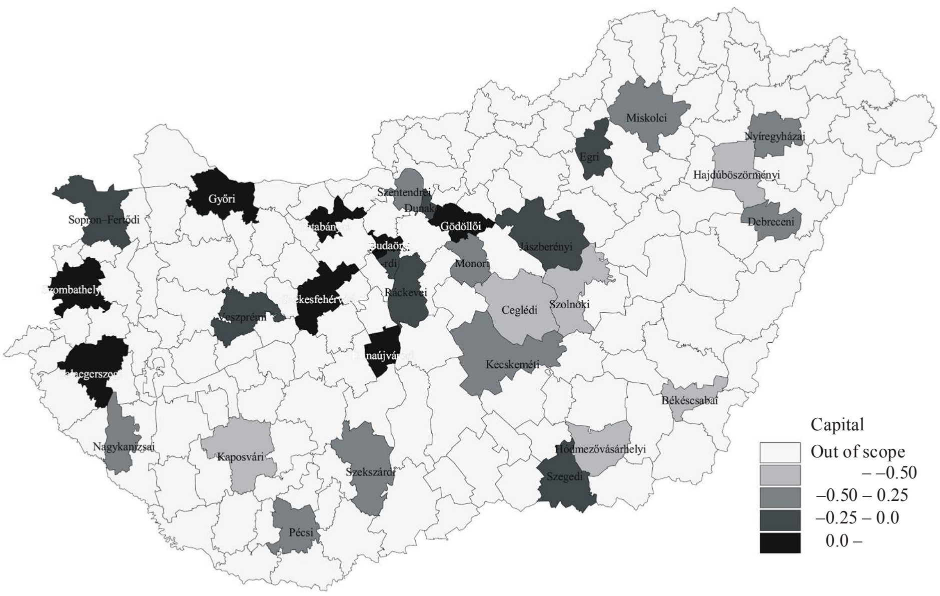
Indicators having major influence on the competitiveness of micro-regions

CAP_FDI	Component	SOC_CAP2	Component
CONS-INH	0.773	PAYER-PIT	–0.653
SHARE-INH	0.936	POOR	0.858
FDI-INH	0.963	CULT	0.029
FDI-CAP	0.962	DIS-PENS	0.731
FDI-EMPL	0.944	DIPL-LOCAL	–0.041
FDI-REV	0.950	CRIME	0.039
		UNEMPL-RATE	0.835
		NONGOV	0.075

Based on these results, these two factors explain the competitiveness of micro-regions (Table 1). The first factor (CAP_FDI) only includes positive variables: a foreign direct investment, total assets of enterprises (CONS-INH) and paid-in capital of enterprises in the micro-region (SHARE-INH). In the second factor (SOC_CAP2): the proportion of personal income taxpayers increases, while poverty rate, unemployment rate and disability pensioners reduce competitiveness.

Figure 5

Micro-regions by CAP_FDI factor



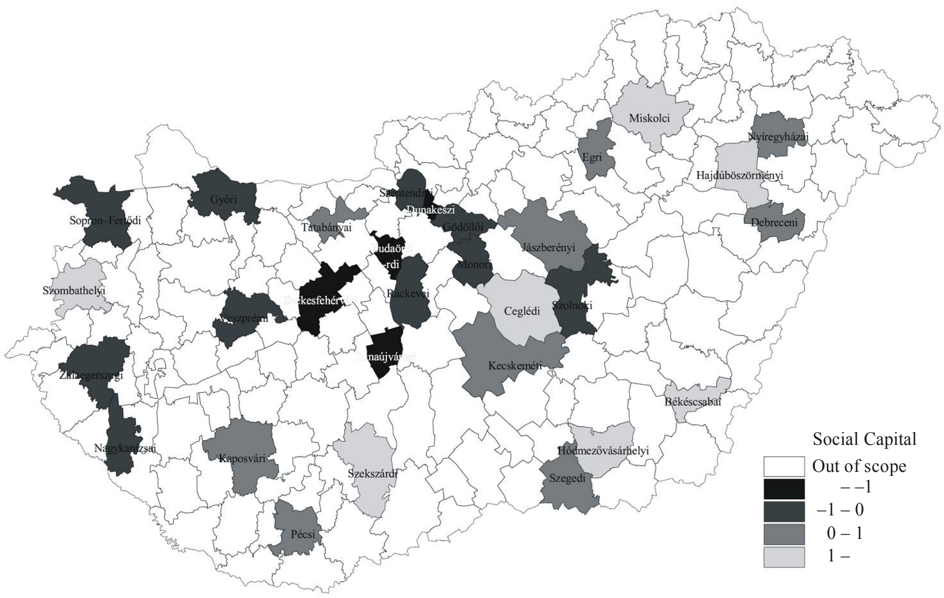
Micro-regions may be classified based on *productive capital and FDI* and even their spatial distribution may be determined (Figure 5):

- most competitive 8 micro-regions, similarly to revealed competitiveness, are found in manufacturing centres of Western Transdanubia and in smaller centres around Budapest,
- the next category (8 micro-regions) is also dominated by regions around Budapest, but a few major cities also appear from other regions of the country,
- the third group (9 micro-regions) is characterized by country capitals from everywhere around the country,
- while the least competitive 6 micro-regions are found in the south and the east part of Hungary.

Classification of micro-regions based on *social capital* is similar to the previous ones (Figure 6). Social capital is quite strong around the capital and in western parts of the country, while it is practically missing in other regions. It has to be noted, that variables included in the factor, like unemployment rate, poverty rate, number of disability pensioners under retirement age, etc. not only describe social capital, but may also be linked to human capital.

Figure 6

Micro-regions by SOC_CAP_2 factor



(c) Relationship between RC and the factors created from the indicators

There may be multicollinearity among the indicator groups of the five competitiveness factors. Therefore we used a different methodology to review and test the relationship between the RC dependent variable and each of the 34 indicators considered: we performed factor analysis on the 34 indicators to generate independent factors. These factors were used in multivariate linear regression. This was especially beneficial because it enabled us to test the structure of the pyramidal model. However, it bears the disadvantage that one has to find an explanation afterwards for each factor based on the indicators included.

Factor analysis was performed for 34 variables with 4-5-6-7-8 factors; obviously, the higher the number of factors better explains the standard deviation (Table 2). We performed multivariate linear regression in each case, and found the best alignment for 5 factors:

– there is no multicollinearity to observe,
– residuals show a normal distribution,
– there is no heteroscedasticity to observe.

Table 2

Factor weights for 34-indicator factor analysis

Factors	4 factors	5 factors	6 factors	7 factors	8 factors
1	23.58	22.26	22.15	22.31	22.34
2	21.53	20.76	20.46	20.19	20.30
3	16.13	16.47	14.61	14.91	14.82
4	9.85	9.58	8.95	8.89	8.66
5	–	8.15	8.75	8.78	6.56
6	–	–	6.42	4.98	5.52
7	–	–	–	4.45	4.89
8	–	–	–	–	4.38
Total	*71.09*	*77.22*	*81.34*	*84.51*	*87.47*

These five factors account for 81.1 % (R^2=0.811) of the dependent variable's (RC) standard deviation. Our calculations resulted in the following multivariate linear regression model:

$$RC_i = +\ 0{,}213\ F1_i + 0{,}665\ F2_i + 0{,}421\ F3_i + 0{,}301\ F4_i + 0{,}236\ F5_i + e_i.$$

The interpretation is complicated by the fact that each indicator may be present in more than one factor; therefore, it is recommended to consider components with an absolute value greater than 0.5 (Table 3).

Table 3

Factor components

Factor1		Factor2		Factor3		Factor4		Factor5	
DIPL_EMPL	0.887	FDI_EMPL	0.940	MIGR	0.885	EXP_GVA	0.794	CULT	0.692
SERVICES	0.876	FDI_REV	0.939	BIRTH_MORT	0.795	CLUST_PROP	0.787	SEC_EMPL	0.677
SELF_EMPL	0.863	FDI_INH	0.933	VITALITY	0.694	PAYER_PIT	0.656		
MANAG_EMPL	0.850	FDI_CAP	0.931	PATENT_OUT	0.595	EXP_INH	0.636		
DIPL-LOCAL	0.817	SHARE-INH	0.918	SME_INH	0.660				
KIMS	0.791	CONS-INH	0.725	KIBS	0.569				
NONGOV	0.716	EXP_INH	0.626	YOUNG_INH	0.527				
R&D_INH	0.594	KIBS	0.559	POOR	–0.518				
CRIME	0.515	SME_INH	0.505	ENTRE	–0.520				
SCHOOL	–0.752			DIS_PENS	–0.650				
				UNEMPL_RATE	–0.688				

Indicators of the pyramidal model's competitiveness factors appear in several calculated factors as components (Table 4). The pyramidal model's research and technological development competitiveness factor (RTD) is only linked to one factor; we attribute this to the fact that among the studied 31 micro-regions, there is research and development only in a few university towns. Indicators of human and social capital appear in several factors, especially because these are difficult to operationalize.

Table 4

Relationship between the competitiveness factors and the calculated factors

Competitiveness factors	Factor1	Factor2	Factor3	Factor4	Factor5
Research and technological development	x				
Human capital	x		x		x
Productive capital and FDI		x			
Regional specialization and clusters		x		x	
Social capital and institutions	x		x		x

Revealed competitiveness is most broadly influenced by the Factor2, dominated by productive capital and FDI, as well as regional specialization (entrepreneurship). This factor expresses one of the elements to the pyramidal model (Productive capital and FDI), complemented by a few indicators of other elements. Spatial distribution of micro-regions based on this factor shows exact conformity with Figure 5.

Factor1 contains research and technological development, human capital and social capital indicators (Figure 7). Micro-regions that are strong on this factor are distributed quite evenly around the country; usually in university towns, sometimes even being the centres of less developed regions. Compared to previous results it is salient that highly competitive micro-regions of Transdanubia show weak competitiveness on human capital and RTD values.

Factor4 is linked to the pyramidal model's regional specialization and clusters element. This indicates the spatial distribution of Hungarian manufacturing industries (Figure 8). It is interesting to see that manufacturing industries with export capabilities are located in Northern Transdanubia and beyond the daily commute zone of Budapest's agglomeration.

The RCF was tested for 31 Hungarian micro-regions based on the pyramidal model. In our opinion, both analyses rendered useful results for regional policy-makers and for fine-tuning the model itself.

Figure 7

Factor1: research and technological development, human capital, and social capital and institutions

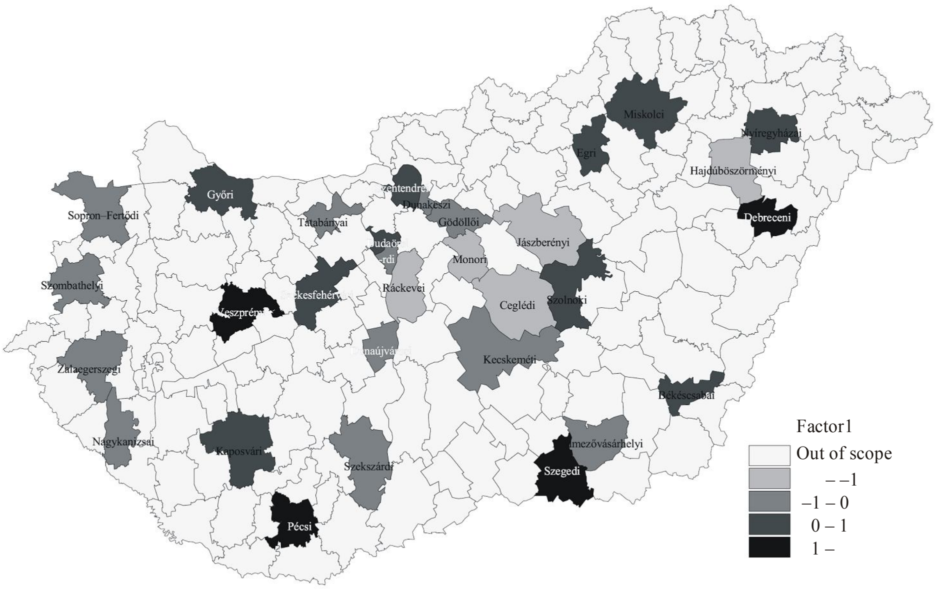
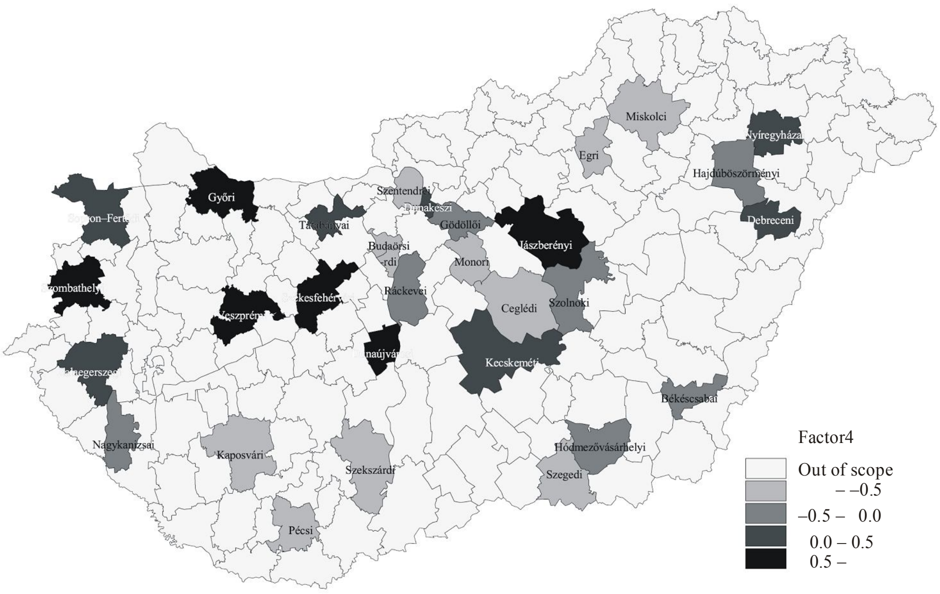


Figure 8

Factor4: regional specialization and clusters



Summary

The aim of this study was to apply the pyramidal model of regional competitiveness and perform a study of LAU1 micro-regions with potential localization agglomeration economies. The pyramidal model rests on endogenous growth factors, and it reflects on competitiveness advantages and disadvantages besides measuring competitiveness itself. Influencing factors of competitiveness have been modelled by the Regional Competiveness Function, created by multivariate linear regression models.

Hungary has shown slow economic growth for about a decade, and employment figures have been falling behind the EU-average. These factors together demonstrate that the Hungarian economy is lacking competitiveness. Data shows that the area around Budapest is still growing dynamically, well exceeding the EU-average, while other parts of the country are able to stagnate at best. Regional differences in the country are enormous, among the major ones in the EU. Our research was based on the question as to why these provincial regions are unable to gain more competitiveness.

The aim of our empirical study was to analyse those provincial LAU1 micro-regions, which have an urban population of at least 50,000. The Regional Competitiveness Function was estimated in two ways, because in our opinion, both methods are useful and are able to amend each other in regional competitiveness studies. In the future, however, it will be more beneficial to examine nodal regions, which are a much better representation of business and institutional relationships.

Our empirical results show a good representation of Hungarian region types in their specific developmental phases:

- *Budapest and micro-regions around it*: this region, housing about 3 million inhabitants, is developing quickly, offering wide-ranging urbanization advantages.
- *Manufacturing micro-regions*: significant FDI and export, high employment, weak RTD and human capital. These regions are located at the north-western border and are well integrated into the EU economy, however, their labour productivity is low and foreign-owned companies do not have a wide supply base in the region. These are remote controlled regions unable to vitalize their own economies, because their human capital and innovation capacity required for higher value-added products and services and innovation is quite weak.
- *University towns*: excellent human capital and state-financed RTD, but a low level of export capabilities in the business sector, low levels of productive capital, labour productivity, and employment. These micro-regions are distributed around the country. They are unable to vitalize the economy of their broader region because there are no significant enterprises in the region.
- *Stagnated urban micro-regions*: weak human capital, low levels of export capability, usually encircled by rural settlements.

The weak performance in the Hungarian economy is partially an outcome of inadequate regional policy. There is an enormous need for decentralized territorial development in order to strengthen the competitiveness of provincial urban regions, which should also enable them to execute bottom-up development strategies more strongly adhering to the unique characteristics of each micro-region.

There is still a long way to go towards the establishment of a Regional Competitiveness Function. The road is full of conceptual and methodological barriers. However, there is an explicit need for a better understanding of regional development in less prospective European countries. We believe that the synthesis of endogenous growth theories and regional competitiveness studies would benefit a more refined framework for empirical analyses to do this. The potential outcome is a better policy framework.

REFERENCES

Acs Zoltán – Szerb László (2007): Entrepreneurship, Economic Growth and Public Policy. *Small Business Economics* 28 (2-3): 109–122.

Barkley, D. L. (2008): Evaluations of Regional Competitiveness: Making a case for case studies. *The Review of Regional Studies* 38 (2): 212–143.

Batey, P. – Friedrich, P. (2000): *Aspects of Regional Competition*. In Batey, P.– Friedrich, P. (eds): *Regional Competition*. pp. 3-33. Springer, Berlin.

Bajmócy Zoltán – Szakálné Kanó Izabella (2009): Measuring the Innovation Performance of Hungarian Subregions. In: Bajmócy Zoltán – Lengyel Imre (eds) *Regional Competitiveness, Innovation and Environment*. pp. 99-121. JATEPress, Szeged.

Bajmócy Zoltán – Lukovics Miklós. – Vas Zsófia (2010): A subregional analysis of universities' contribution to economic and innovation performance. *Transition Studies Review* 17 (1): 134–150.

Budd, L. – Hirmis, A. K. (2004): Conceptual framework for regional competitiveness. *Regional Studies* 38 (9): 1015–1028.

Camagni, R. (2002): On the Concept of Territorial Competitiveness: Sound or Misleading? *Urban Studies* 39 (13): 2395–2411.

Camagni, R. (2009): Territorial capital and regional development. In: Capello, R.–Nijkamp, P. (eds) *Handbook of regional growth and development theories*. pp. 118–132.Edward Elgar, Cheltenham.

Capello, R. (2007a): *Regional economics*. Routledge, London and New York.

Capello, R. (2007b): A forecasting territorial model of regional growth: the MASST model. *The Annals of Regional Science* 41 (4): 753–787.

Chesire, P.C. (2003): Territorial Competition: Lessons for (Innovation) Policy. In: Bröcker, J. – Dohse, D. – Soltwedel, R. (eds) *Innovation Clusters and Interregional Competition*. pp. 331–346 Springer, Berlin.

EC (1999): *Sixth Periodic Report on the Social and Economic Situation and Development of Regions in the European Union*. European Commission, Luxembourg.

EC (2001): *Second Report on Economic and Social Cohesion*. European Commission, Brussels.

EC (2008): *European Competitiveness Report 2008*. European Commission, Brussels.

Faggian, A. – McCann, P. (2009): Human capital and regional development. In: Capello, R.–Nijkamp, P. (eds) *Handbook of regional growth and development theories*. pp. 133–151. Edward Elgar, Cheltenham.

Enyedi György (2009): Competitiveness of the Hungarian Regions. *Hungarian Geographical Bulletin* 58 (1): 33–48.

Fischer, M. – Nijkamp, P. (2009): Entrepreneurship and regional development. In: Capello, R.–Nijkamp, P. (eds) *Handbook of regional growth and development theories*. pp. 182–198. Edward Elgar, Cheltenham.

Florida, R. (2002): *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*. Basic Books, New York.

Gardiner, B. – Martin, R. – Tyler, P. (2004): Competitiveness, Productivity and Economic Growth across the European Regions *Regional Studies* Special Issue 38 (9): 1045–1067.

Glaeser, E. L. (2008): *Cities, Agglomeration, and Spatial Equilibrium*. Oxford University Press.

Kitson, M. – Martin, R. – Tyler, P. (2004): Regional Competitiveness: An elusive yet key concept? *Regional Studies* 38 (9): 991–999.

Lengyel Balázs - Cadil, V. (2009): Innovation policy challenges in transition countries: foreign business R&D in the Czech Republic and Hungary. *Transition Studies Review* 16 (1): 174–188.

Lengyel Balázs – Leydesdorff, L. (2010): Regional innovation systems in Hungary: The failing synergy at the national level. *Regional Studies* 45 (5): 677–693.

Lengyel Imre (2000): A regionális versenyképességről (On regional competitiveness). *Közgazdasági Szemle* 47 (12): 962–987.

Lengyel Imre (2004): The Pyramid Model: Enhancing Regional Competitiveness in Hungary. *Acta Oeconomica* 54 (3): 323–342.

Lengyel Imre (2009a): Knowledge-based local economic development for enhancing competitiveness in lagging areas of Europe: The case of the University of Szeged. In: Varga Attila (ed.) *Universities, Knowledge Transfer and Regional Development: Geography, Entrepreneurship and Policy*. pp. 322-349. Edward Elgar, Cheltenham–Northampton.

Lengyel Imre (2009b): Bottom-up Regional Economic Development: Competition, Competitiveness and Clusters. In: Bajmócy Zoltán – Lengyel Imre (eds) *Regional Competitiveness, Innovation and Environment*., pp. 13–38 JATEPress, Szeged.

Lukovics Miklós (2009): Measuring Regional Disparities on Competitiveness Basis. In: Lengyel Imre – Bajmócy Zoltán (eds) *Regional Competitiveness, Innovation and Environment*. pp. 39–53. JATEPress, Szeged.

Malecki, E. J. (2002): Hard and Soft Networks for Urban Competitiveness. *Urban Studies* 39 (5-6): 929–945.

Malecki, E. J. (2004): Joyckeing for position: What it Means and Why It Matters to Regional Development Policy When Places Compete. *Regional Studies* 38 (9): 1101–1120.

McCann, P. (2008): Agglomeration economics. In: Karlsson, C. (ed) *Handbook of Research on Cluster Theory*. pp. 23-38. Edward Elgar, Cheltenham.

Pike, A. – Rodrígues-Pose, A. – Tomaney, J. (2006): *Local and regional development*. Routledge, New York.

Polenske, K. R. (2004): Competition, Collaboration and Cooperation: An Uneasy Triangle in Networks of Firms and Regions. *Regional Studies* 38 (9): 1029–1043.

Porter, M. E. (2003): The Economic Performance of Regions. *Regional Studies* 37 (6-7): 549–578.

Porter, M. E (2007): *Competitiveness: implications for Central Europe and the Czech Republic*. Paper presented in Prague, 22 October.

Porter, M. E (2008): *Regional Competitiveness in a Global Economy*. The Summit for American Prosperity, The Brookings Institution, Washington, 11 June.

Porter, M. E.–Delgado, M.–Ketels, C.–Stern, C. (2008): Moving to a New Global Competitiveness Index. In: *The Global Competitiveness Report*. pp. 45–63. World Economic Forum, Davos.

Stimson, R. – Robson, A. – Shyy, T-K. (2009): Modeling regional endogenous growth: an application to the non-metropolitan regions of Australia. *The Annals of Regional Science* (43 (2): 379–398.

Stimson, R. – Stough, R. R. – Salazar, W. (2009): *Leadership and Institutions in Regional Endogenous Development*. Edward Elgar, Cheltenham.

Szakálné Kanó Izabella (2011): A gazdasági akitivitás térbeli eloszlásának vizsgálati lehetőségei. *Statisztikai Szemle*, 89 (1): 77–100.

Varga Attila (2006): The Spatial Dimension of Innovation and Growth: Empirical Research Methodology and Policy Analysis. *European Planning Studies* 14 (9):. 1171–1186.

Varga Attila (2007): Localized knowledge inputs and innovation: The role of spatially mediated knowledge spillovers in Hungary. *Acta Oeconomica* 57 (1): 1–20.

Keywords: regional competitiveness, pyramidal model, endogenous regional development, regional competitiveness function.

Abstract

Nowadays, more and more scholars of regional science are interested in the role of agglomeration economies in the knowledge-based economy. This issue can be dealt with from different points of view: different development types of functional or nodal regions with the examination of the factors influencing regional competitiveness.

In this paper, we outline our analytical framework: the renewed pyramidal model of regional competitiveness. The renewed pyramidal model is a logical systematization for measuring endogenous regional development and the factors influencing it; the model shall be used to introduce the regional competitiveness function (RCF). After introducing theoretical model and new function, we are going to investigate the competitiveness of Hungarian urban microregions (LAU1), where firms potentially enjoy localization agglomeration economies. The statistical analysis to underline the classification of microregions by competitiveness types is based on multivariate linear regression models.

REGIONAL STATISTICS, 2012, VOL. 2: 45–60 DOI: 10.15196/RS02104

JUDIT GÉBERT – GYÖRGY MÁLOVICS – ZSUZSA FÁSKERTI

The Limits of Well-being Measurement at Sub-regional Level

Introduction

To characterise the well-being of different regions, certain indices of the System of National Accounts (SNA) are used by the national and international literature. However, it is already evident that the informational base of the SNA and especially its most popular indicators (e.g. GDP, GNP) is very narrow. This base does not contain even basic information about the state of the natural environment, the effect of economic processes on nature, and several other aspects of social well-being (Sen 1999, van den Bergh 2007, Stiglitz et al. 2009).

Consequently, many organisations and researchers have recently attempted to create systems of indicators to represent more aspects of general well-being and sustainable development. As a result of this tendency in research projects, more than 500 sets of indicators aiming to operationalise sustainable development are mentioned in a recent study (Böhringer-Jochem 2007). Besides this fact, every significant international organisation (like the EU, UN, OECD, etc.) has its own indicator system for sustainability. In 2003, the UN, the European Commission, the IMF, the OECD and the World Bank published a proposal for a framework for integrated environmental and economic national accounts (Eurostat 2007, Hungarian Central Statistical Office 2008, OECD 2003, UN 2007, UNECE et al. 2008). These sustainability indicator sets all have their 'well-being' pillar, namely they all contain indicators which aim to provide information on the present level of social well-being. In our paper we concentrate on this, well-being measurement.

Recently, a committee led by Nobel laureate economists published a report on the insufficiency of the dominant welfare indicators and the desirable criteria for an alternative indicator system (Stiglitz et al. 2009). These facts suggest that there are valid arguments to use alternative indices in political decision-making in order to obtain a better understanding of the well-being of societies.

In our paper, we aim to establish a system of well-being indicators for the sub-regional level. Thus, our work contributes to a better understanding of present social well-being at this level of analysis.

Our paper consists of three parts. In the first, we describe the theoretical background and the methodology. On the basis of contemporary well-being theories, we create an overall well-being typology. This typology includes the ideal well-being measures for the sub-regional level. After that, we describe the process of the operationalisation of the well-being measures and specifically those tools which are available for the sub-regional level to determine the level of prosperity in Hungary. Lastly, the average well-being rates in Hungary are analysed at sub-regional level with the help of multivariate statistical methods. In this part of the paper, we introduce our results in a national comparison. In the final section of the paper, we draw conclusions from our results and from our theoretical underpinning. Our main findings are that (1) well-being measurement at a sub-regional level reveals tendencies which are masked by traditional economic and competitiveness analysis; and (2) a measurement exercise like ours has serious limitations, i.e. there is a significant measurement gap between theoretical models of well-being and their operationalisation, most of all because of the current state of data availability.

Methodology and theoretical background

As the first step of our research we overviewed the current literature on well-being theories – utility-based approaches (Hausmann-McPherson 1997), basic resources (or needs) (Rawls 1971, Streeten 1979) and the capability approach (Sen 1993, 1995, 1999, Nussbaum 2011, Comim et al. 2008). The main components of these approaches can be seen in Table 1.

Table 1

Theories about well-being

Utility-based approach	Basic needs, resources	Capability approach
Income Unemployment	Rawls (1997): Basic rights and liberties Powers of offices and positions of responsibility Income and wealth Social bases of self-respect Streeten (1979): Means for bare survival Means for continued survival Means for productive survival Non-material needs	Sen (1999): Real opportunities Nussbaum (2011): Life Bodily health Bodily integrity Senses, imagination and thought Emotions Practical reason Affiliation Other species Play Control over one's environment

We do not intend to choose between the different theories, which all have their own theoretical underpinning. Instead, we synthesise these contemporary theories of well-being in a new theoretical framework and use the notion of ‘dimension’, which was suggested by Alkire (2002). We define dimension as follows: *a dimension is any of the component aspects of well-being* (Alkire 2002). So a dimension is a kind of aspect or viewpoint that captures a part of well-being. While one single dimension cannot give a full picture of well-being, there is no requirement that dimensions of well-being must be free of overlap. Our aim with this notion is to interpret well-being according to the results of contemporary welfare economics from the broadest possible base of information.

According to the aforementioned well-being theories, we compiled the following list of dimensions (Table 2).

Table 2

Dimensions of well-being

1. Financial goods	All of the theories about well-being admit that we need certain material and financial goods to survive. But both the basic resources approaches and the capability approach emphasise that financial and material goods are just means to achieve well-being and not goals in themselves.
2. Rights for freedom	Some of the authors on well-being (Rawls 1971, Nussbaum 2011) attribute significance to basic human rights for freedom (like freedom of movement, free choice among a wide range of occupations, freedom to take part in decision-making, etc.), which cannot be alienated from both intrinsic and instrumental reasons.
3. Physical well-being	To live a long life, people prima facie need health. Things which are needed for survival and for being healthy belong to this dimension, like drinking water, food, shelter, sanitation, etc. According to the capability approach, health is the most important means of achieving valuable goals in life (Sen 1999).
4. Relationship with family/friends	Having relationships with other members of society, playing with them, having emotions about them are among the most important capabilities according to Nussbaum (2011), but this is mentioned in basic goods theories as well.
5. Quality of environment	The quality of the environment is neglected by the mainstream economic theories. However, this important factor is not just intrinsically valuable, but it is also the basis for other elements of well-being like clean drinking water, healthy food, safe environment, etc.
6. Leisure activity	Being able to play is mostly emphasised by Nussbaum. According to the capability approach, life is more valuable if people have a wide variety of leisure activities, like cinema and theatre going, sports, civil activities, etc.
7. Labour	Labour, as one of the factors of production, has an important role in economic theories and competitiveness analysis. However, having a job also gives people self-esteem, thus it is another important dimension (Sen 1999).
8. Forms of social care	Forms of social care are services which are provided by the community. These services cannot be created by individuals on their own.
9. Education	Education has prior importance in widening the capabilities of individuals. More educated people produce more value-added products and services, and have a better chance of getting a job (Sen 1999).
10. Information access	According to the capability approach, to live a valuable life, people need to get information about the world in general, such as about job opportunities, political actions, etc. (Sen 1999). But Rawls also emphasises the right to be informed (1971).
11. Basis of social self-respect	The basis of social self-respect means recognition by social institutions that gives a sense of self-worth to citizens. People need certain goods to be full members of society, like appropriate clothing, communication devices, etc. (Rawls 1971).

Source: Own illustration.

In the second step of our research, we established our set of indicators. To grasp the sub-regional dimension of the statistical indicators of well-being, our database relied on data from TeIR and, most of all, on the data of the Hungarian Central Statistical Office. Exceptions are income data, which originate from the tax authority. The resultant database of potential well-being indicators consists of 58 welfare indicators, all of which can be linked to a dimension of well-being (Appendix 1).

The concept of dimension according to Alkire (2002) is then divided into three levels in order to classify potential well-being indicators. These are: (1) Theoretical main dimension, (2) Dimension and (3) Sub-dimension. The theoretical main dimensions refer to the 11 dimensions based on theory. The other two levels (dimension and sub-dimension) represent a more detailed conceptual analysis, virtually a group of indicators. These were created by using factor analysis.

The indicators we actually used for our analysis were selected after a multi-step process. All of the 58 indicators were classified into one of the theoretical main dimensions on the basis of well-being theories. Unfortunately, we did not manage to find indicators for 5 of the previously defined 11 theoretical well-being dimensions. Thus indicators were available in the following six theoretical main dimensions:

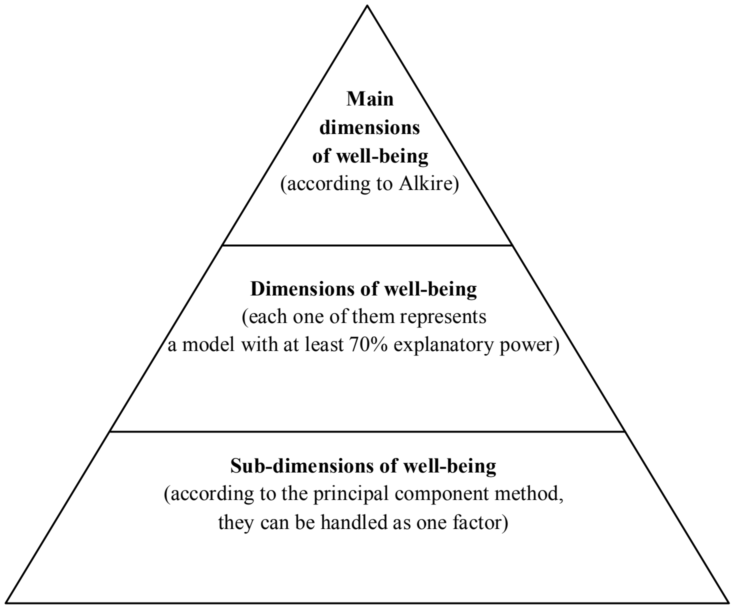
1. Financial goods
2. Physical well-being
3. Leisure activities
4. Labour
5. Forms of social care
6. Education

We standardised our variables, and created specific and rate indicators to be able to compare the results of the sub-regions. After standardising the variables, we made different factors (sub-dimensions) by principal component analysis from the standardised variables in the theoretical main dimensions. Our set of indicators consists of three levels, as mentioned above, which can be seen in Figure 1. The theoretical main dimensions are lined up according to the theoretical literature and are in accordance with Alkire's dimensions of well-being. The sub-dimensions represent one given aspect of well-being and they can be handled as single factors in our model because of their empirical connection.

We created the level of dimensions in our analysis, which is between the level of main dimensions and sub-dimensions (Figure 1). From a theoretical perspective, this level reflects the constituents of well-being, which are narrower than the theoretical main dimensions, but broader than the sub-dimensions. Moreover, when analysing the indicators in one dimension by the principal component method, we found that the indicators within one dimension were part of the same model with at least 70% explanatory power, thus one dimension can be considered as one single unit. Accordingly, the actually used set of indicators consists of standardised variables, which are connected strongly and unequivocally to the factors (or main components) from the principal component method. Thus we have 47 indicators, classified into 23 sub-dimensions (factors).

Figure 1

Levels of indicators



Source: Own illustration.

In the third step of our analysis we specified the output of the measurement exercise. The actually used indicators were the basis for two kinds of output. The first are factors or sub-dimensions themselves. According to the process interpreted above, we saved the factor values and used them later in the evaluation. To use this process, it is a basic requirement that the artificial variables or factors should have real – close to reality – meaning (see Table 3).

The second kind of output is dimensions. Dimensions were made not from the belonging factors, but from the variables themselves related to the factors with a weighting process. The factor analysis attributes a communality value to each variable. These values show the weight of the variable in the hypothetical or artificial variable. Therefore this is one kind of weighting which is offered by the applied method.

As a consequence, the value of one dimension is the average of the included standardised variables where the weights are the square root of the communalities. It is important to emphasise that not every main dimension was divided into dimensions (some of the main dimensions had 70% explanatory power in themselves).[1] During the analysis, the main dimensions with high explanatory power were counted as one of the

1 For instance, the theoretical main dimension of Financial goods was kept as a single dimension, because the three indicators in them (Inland incomes, Poverty rate, Sub-regional Gini index) have more than 70% explanatory power (see Appendix 1).

dimensions of well-being. Thus in the analysis for well-being at sub-regional level, we had the following 11 dimensions:

– Financial goods (3 indicators)
– Physical well-being (4 indicators)
– Leisure activities – culture (7 indicators)
– Leisure activities – recreation (4 indicators)
– Labour (1 indicator)
– Forms of basic social care (5 indicators)
– Forms of social care – healthcare (4 indicators)
– Forms of social care – childcare (6 indicators)
– Forms of social care – basic education (5 indicators)
– Forms of social care – safety (5 indicators)
– Education (3 indicators)

Table 3

Main dimensions, sub-dimensions and their connection to actual meaning

<table>
<tr><th>Theoretical main dimension</th><th>Dimension</th><th>Sub-dimension (factor)</th></tr>
<tr><td rowspan="2">Financial goods</td><td rowspan="2">Financial or material well-being</td><td>Income deprivation</td></tr>
<tr><td>Income distribution</td></tr>
<tr><td rowspan="2">Physical well-being</td><td rowspan="2">Physical well-being</td><td>Cancer-type illness</td></tr>
<tr><td>Other long-term illness</td></tr>
<tr><td rowspan="5">Leisure activities</td><td rowspan="3">Culture</td><td>Opportunities for public culture</td></tr>
<tr><td>Opportunities to visit cinema/theatre</td></tr>
<tr><td>Opportunities to visit museums</td></tr>
<tr><td rowspan="2">Recreation</td><td>Availability of recreation facilities in town</td></tr>
<tr><td>Quality of recreation facilities</td></tr>
<tr><td>Labour</td><td>Labour</td><td>Unemployment</td></tr>
<tr><td rowspan="12">Forms of social care</td><td rowspan="3">Forms of basic social care</td><td>Sewage system</td></tr>
<tr><td>Sanitation</td></tr>
<tr><td>Availability of post offices in town</td></tr>
<tr><td rowspan="2">Healthcare</td><td>Availability of basic health care services</td></tr>
<tr><td>Hospitals</td></tr>
<tr><td rowspan="3">Child care</td><td>Quality of kindergartens</td></tr>
<tr><td>Availability of kindergartens</td></tr>
<tr><td>Endangered minors</td></tr>
<tr><td rowspan="2">Basic education</td><td>Availability of primary schools</td></tr>
<tr><td>Quality of primary school education</td></tr>
<tr><td rowspan="2">Safety</td><td>Crime</td></tr>
<tr><td>Safety of travelling</td></tr>
<tr><td>Education</td><td>Education</td><td>Availability of education</td></tr>
</table>

Results

To analyse the well-being situation of sub-regions, we made clusters along the examined dimensions and made groups from sub-regions in similar situations. The result of our cluster analysis is shown in Tables 4 and 5.

Table 4

Cluster centres

Well-being dimensions	Clusters			
	1	2	3	4
	Final cluster centres			
Financial goods (inverse)	–0.110	0.509	–0.341	–0.363
Physical well-being (inverse)	0.061	0.279	–0.363	–0.337
Leisure activities – culture	0.069	–0.320	0.133	1.537
Leisure activities – recreation	–0.084	–0.183	0.323	–0.060
Labour (inverse)	–0.330	1.279	–0.784	–0.730
Forms of basic social care	–0.062	–0.418	0.565	–0.570
Forms of social care	–0.179	–0.400	0.745	–0.877
Forms of social care – childcare	0.061	–0.454	0.367	0.200
Forms of social care – basic education	0.062	–0.058	–0.187	2.463
Forms of social care – safety	–0.087	–0.270	0.473	–0.937
Education	0.142	–1.026	0.861	–0.143

Table 5

Size of the clusters

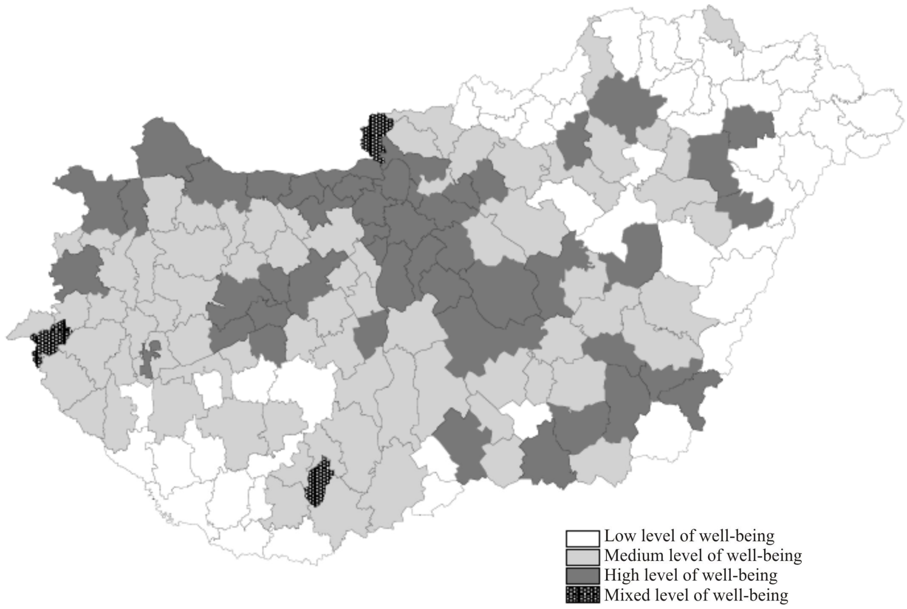
Cluster	Number of sub-regions in the cluster
1	73
2	50
3	48
4	3
All	174

As we can see, there are only three sub-regions in the fourth cluster (Őriszentpéter, Pécsvárad, Szob). Here well-being seems to be mixed and diverse. These sub-regions prosper in some of the dimensions (Financial goods, Physical well-being, Leisure activities, Labour, Basic education) compared to other sub-regions in the country. However, in other dimensions (Recreation, Education) these regions are slightly below average, while well-being is poor as regards some dimensions (Forms of basic social care, Healthcare, Safety).

The other three clusters are much more homogenous. In the first cluster, we find sub-regions where the employment situation is relatively good and overall well-being is about average. The exception is the dimension of Healthcare, because here the cluster is below average. Sub-regions in the second cluster have a relatively low level of well-being, while for those in the third cluster the level is relatively high. In the third cluster every dimension has higher values than those in the second cluster, except Basic education. The same is true for the first and third cluster. This simple relationship between the clusters becomes more complicated and less homogenous when we analyse the variance between the values. In the case of the second cluster, the variance of Physical well-being around the centre of the cluster is high; the sub-regions in this cluster vary greatly compared to the value attributed to the centre of the cluster. In the case of the third cluster, the same can be said about the dimensions of Leisure activities – recreation and Forms of social care – Healthcare.

Figure 2

Clusters of well-being at sub-regional level



Source: Own illustration.

If we analyse the results in a national comparison (see Figure 2), we find that the sub-regions with relatively high well-being are located around a north-west and south-east axis. The regions around Lake Balaton with high well-being and some of the bigger cities in the eastern part of Hungary (Eger, Miskolc, Nyíregyháza, Debrecen) are outside this axis. It seems that the distribution of sub-regions with higher well-being is influenced by the route of highways, because there is a significant spatial concentration alongside the most important roads. It is important to emphasise that there is a concentration of sub-regions with high well-being in the agglomeration of Budapest. In contrast, we find sub-regions with the lowest well-being in the north-eastern and south-western parts of Hungary, primarily in the North Hungary and South Transdanubia regions.

Discussion

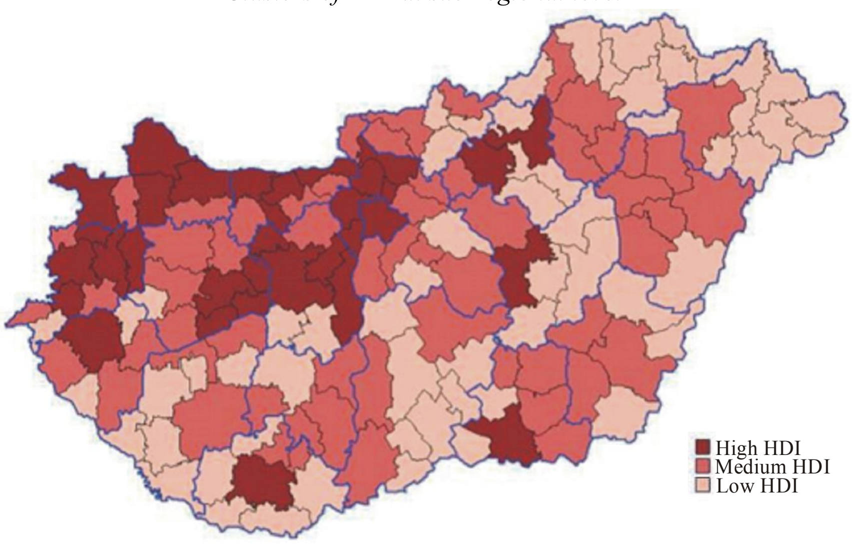
As aforementioned, our main aim was to create a synthesis between different theories of well-being in our study. We tried to include as many aspects of well-being as possible. Consequently, we included much more information in our research than conventional welfare or competitiveness analyses, which usually use only price-based SNA indicators and/or employment/unemployment levels. Although we also used income-related indices and unemployment rates, these are not the only indicators but rather only one of the components in our complex analysis of well-being. We claim that this type of multidimensionality and the resultant wide informational base are the main advantages of our research perspective.

Therefore, it is small wonder that our outcome shows differences compared to the results of the traditional economic or competitiveness measurement.[2] Not only do the single dimensions differ (which are obviously different, because they measure different things), but also our analysis gives a more detailed and diverse picture of sub-regional well-being compared to SNA and competitiveness results (Lukovics–Kovács 2011) or HDI results (Garami 2009).

Although here we do not have the opportunity to give a detailed analysis of these differences, we illustrate them by a brief comparison. Figure 3 shows sub-regional HDI results (Garami 2009). By comparing the two maps (Figures 2 and 3), one can see that they give a somewhat different picture of sub-regional well-being in Hungary. While a north-west south-east axis is present in our results, the HDI analysis implies a better developed north-western part of Hungary – which is more nuanced in our analysis shown in Figure 2.

Figure 3

Clusters of HDI at sub-regional level



Source: Garami 2009.

During our research we realised that we had to face serious limitations and possible biases if we wanted to keep the broad informational base of our analysis. We think that these limitations and their consequences on scientific work regarding well-being

2 For a detailed discussion of our results on sub-regional well-being see Málovics et al. 2010.

measurement (in Hungary) are at least as important from a scientific point of view as our results related to Hungarian sub-regional well-being are. The reason is that biases similar to the ones we encountered are likely to occur with any well-being measurement using a wide informational basis. Therefore we discuss these biases below in detail.

The biases are: (1) the arbitrariness of choosing dimensions and statistical tools; (2) the operational gap between the foundational level and the practical level of our research; and (3) the constraints caused by data availability.

(1) It can be argued that our research suffers from arbitrariness in two ways. First of all, there is the uncertainty of the statistical methods used. Both the results of factor analysis and clustering depend very much on the decisions of the researcher. In factor analysis the researcher decides which factor should be filtered out, and how to rotate or iterate the variables. The same can be said about clustering: the characterisation of the different clusters depends on the researcher's choices.

The other form of arbitrariness, often mentioned in the literature of contemporary welfare economics as well (Alkire 2002, Alkire at al. 2008), is that the choice of dimensions is arbitrary and depends only and exclusively on the evaluation of the researcher[3]. It is also debatable whether the dimensions we created are constituents of well-being or not. Also, there may be aspects which are missing from our study but may be deemed important by another evaluation. However, as Sen (1999) argues, we can avoid the overgeneralisation of an arbitrary view of well-being if we make these limitations and decisions explicit and do not claim that the results are general, exclusive and objective.

(2) A significant limitation of our research results from the operational gap between the foundational or conceptualised level and the empirical level of analysis in the study. After creating the list of dimensions, we had to face the fact that some of them were difficult or even impossible to quantify or measure. Five main dimensions had to be ruled out (Rights for freedom, Information access, Relationship with family/friends, Basis of social self-respect and Quality of environment) because they were considered as immeasurable (Berg-Schlosser 2004). Although there are techniques to operationalise these notions, like rights for freedom (see for instance Berg-Schlosser 2004), there are no indicators available on them for the Hungarian sub-regions at the moment. Thus to measure them would involve a whole new research project demanding significant financial resources.

(3) The last significant constraint (bias) of our research is data availability. We realised that databases which are available in Hungary at the moment are of poor quality if one is to operationalise sub-regional well-being. This is true for even those dimensions of well-being where we managed to find some related data (indicators). Although some of the dimensions – like Financial status or Forms of social care – are theoretically easier to quantify than for example Rights for freedom or Information access, even for these dimensions we found a very low number of indicators to represent them.

3 This arbitrariness is true even if the selected dimensions have theoretical underpinning. First of all, the list of dimensions of well-being is - and will always be - incomplete. One can always add another important dimension. Thus, choosing the relevant dimensions incorporates the decision of the researcher. Second, dimensions are normative notions and not prescriptive in any sense.

For instance, in the case of basic education, we measured the quality of education with the available five indicators: primary school performance, the number of classes per 1 000 pupils, the number of full-time primary school teachers per 1 000 pupils, the number of computers and the number of primary schools per settlement. It is not hard to see that, for example, to measure the quality of education to a better extent we could include other indices as well, e.g. performance indices revealing skills gained from education (Stiglitz et al. 2009).

As we chose our indicators from existing databases, our research was constrained by already existing measures. As a consequence, our study has become less theory-driven since we could not measure well-being in a way contemporary well-being theory would suggest. Instead, because we relied on available indicators, our research is rather information/data-driven. This means that we had to adjust our measurement to the data currently available and rely less on theory. As a result, our study might show a significantly biased and/or limited picture of well-being from a theoretical point of view. However, after reviewing the literature (van den Bergh 2007, Stiglitz et al. 2009, Comim at al. 2008, Robeyns 2006) it seems that basically almost all well-being measures and indicator systems face similar problems.

Summary

In our study, we analysed the level of well-being in the sub-regions of Hungary. At the foundational level we started with conceptualising the notion of well-being according to the theories in contemporary welfare economics. Next we identified different dimensions on three levels of analysis to explain even more aspects of well-being and thus deal with its complexity. After that, we chose indicators from already existing databases to cover these dimensions. We used factor analysis and clustering to scrutinise the well-being level of the sub-regions.

Besides providing different results to earlier Hungarian economic and well-being analyses, our main contribution is a new theoretical conceptualisation of well-being measurement previously unknown in the Hungarian literature. We do not just aggregate different indicators but assign them to different theoretical dimensions of well-being. The advantages of such an analysis are twofold. First, this concept may help to understand the complexity of well-being. Second, it helps to categorise indices into theoretically based broader well-being categories.

Both of these advantages have consequences for policy-making. Local (or sub-regional) policy-making is generally strongly driven by price-based SNA indicators and/or unemployment levels (Bajmócy 2011) although the well-being of a sub-region is a more complex phenomenon. Therefore our indicator system enables the informational base of policy-making to be broadened. A common problem with such exercises is that they imply a tension between theoretical grounding and practical applicability (Steinbuka et al., n.d.). While it is important for political decision-makers to have simple and easy-to-understand indicator sets, this understandability and simplicity may result in a reduced theoretical grounding. Our indicator set is a first attempt to overcome this dilemma. By creating an indicator set based on welfare economic theory, we provide a theoretical grounding. By introducing three levels of aggregation, we create a reduced set of relatively easily understandable well-being dimensions. Thus we contribute to practical applicability while retaining a theoretical grounding.

Our main conclusions are twofold. First, our research reveals a much more detailed picture about sub-regional well-being than SNA or competitiveness studies do. One of the main novelties of our research is the broadening of the information base of sub-regional well-being measurement. Second, we state that well-being measurement in general is quite constrained at the moment, and this is especially true of the opportunities for well-being measurement in Hungary at a sub-regional level. The reasons for this are: (1) the arbitrariness of choosing dimensions and statistical tools; (2) the gap between well-being theory and practical measurement; and (3) constraints resulting from data availability. It seems to us that statistical data availability has to improve considerably for researchers to be able to give a comprehensive, theoretically sound and non-data-driven picture of sub-regional well-being in Hungary.

REFERENCES

Alkire, S. (2002): Dimensions of Human Development. *World Development* 30 (2): 181–255.

Bajmócy Zoltán (2011): *Bevezetés a helyi gazdaságfejlesztésbe*. JatePress, Szeged.

Berg-Schlosser, D. (2004): The Quality of Democracies in Europe as Measured by Current Indicators of Democratization and Good Governance. *Journal of Communist Studies and Transition Politics* 20 (1): 28–55.

Böhringer, C. – Jochem, P.E.P. (2007): Survey: Measuring the immeasurable — A survey of sustainability indices. *Ecological Economics* 63 (1): 1–8.

Comim, F.– Qizilbash, M. – Alkire, S. (eds.) (2008): *The Capability Approach - Concepts, Measures and Applications*. Cambridge University Press, Cambridge.

Dasgupta, P. (2001): *Human Well-Being and the Natural Environment*. Oxford University Press, Oxford.

Eurostat (2007): *Measuring progress towards a more sustainable Europe.* Internet: http://ec.europa.eu/sustainable/docs/estat_2007_sds_en.pdf.

Garami Erika (2009): A humán erőforrás területi különbségei. Az emberi fejlődés indexének hazai alkalmazhatósága. *Területi Statisztika* 49 (3): 281–298.

Lukovics Miklós – Kovács Péter (2011): A magyar kistérségek versenyképessége. *Területi Statisztika* 51 (1): 52–71.

Hanley, N. (2000): Macroeconomic measures of sustainability. *Journal of Economic Surveys* 14 (1): 1–30.

KSH (Hungarian Central Statistical Office) (2008): *A fenntartható fejlődés indikátorai Magyarországon - Sustainable development indicators in Hungary.* KSH, Budapest.

OECD (2003): *OECD Environmental Indicators: Development, Measurement and Use*. Reference paper. Internet: http://www.oecd.org/dataoecd/7/47/24993546.pdf.

Hausman, D. M. – McPherson, M. S. (1997): *Economic analysis and moral philosophy*. University Press, Cambridge.

Málovics György – Bajmócy Zoltán – Fáskerti Zsuzsa – Gébert Judit (2010): Kistérségek jóléti helyzete és ennek innovációs képességgel való kapcsolata a Dél-Alföldi régióban. In: Bajmócy Zoltán (szerk.): *A Dél-alföldi régió innovációs képessége. Elméleti megközelítések és empirikus elemzések*. (CD Book), pp. 419-468. SZTE Gazdaságtudományi Kar, Szeged.

Nussbaum, M. (2011): *Creating capabilities: the human development approach.* Belknap Press of Harvard University Press, Cambridge.

Rawls, J. (1971): *A Theory of Justice*. Oxford University Press, Oxford.

Robeyns, I. (2006): The Capability Approach in Practice. *The Journal of Political Philosophy* 14(3): 351–376.

Sen, A. K. (1993): Capability and Well-Being. In.: Nussbaum, M. – Sen, A. K. (eds.) *The Quality of Life.* Oxford, Oxford University Press.

Sen, A. K. (1995): *Inequality Reexamined*. Cambridge University Press, Cambridge.

Sen, A. K. – Williams, B. (ed.) (1996): *Utilitarianism and beyond.* Cambridge University Press, Cambridge.

Sen, A. K. (1999): *Development as Freedom.* Oxford University Press, Oxford.

Šteinbuka, I. – Wolff, P. (2007): *Indicators and better policy-making: the case of sustainable development.* http://epp.eurostat.ec.europa.eu/portal/page/portal/sdi/files/LISBON%20ISI%20AUG%202007%20REV2.PDF.

Stern, N. (2006): *Stern Review on the Economics of Climate Change.* http://www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/ stern_review_report.cfm.

Stiglitz, J. E. - Sen, A. K. - 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.

Streeten, P. (1979): From growth to basic needs. *Financial Development* 19(3): 28–31.

UN (2007): *Indicators of Sustainable Development:Guidelines and Methodologies.* United Nations, New York.

van den Bergh, J. C. J. M. (2007): *Abolishing GDP.* TI Discussion Paper, 07-019/3. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=962343.

Keywords: well-being, sub-regions, measurement.

Abstract

Well-being measurement has been an intensively discussed topic in recent economic literature. It has become clear that the informational base of the System of National Accounts (SNA) is far too narrow to include many aspects of well-being. Many economists emphasise that new economic measures which are more sensitive to the social and environmental aspects of decision-making are needed (van den Bergh 2007, Stiglitz et al. 2009) in order to make social decision-making itself sensitive to such aspects. It is small wonder that we encounter a great many sustainability and well-being indicators in economic literature (Böhringer-Jochem 2007).

In our paper, we establish a model for the well-being measurement of sub-regions based on contemporary welfare economics. Based on our model we operationalise and measure well-being at a sub-regional level within Hungary using statistical data. Our research shows that (1) well-being measurement at a sub-regional level reveals tendencies which are masked by traditional economic and competitiveness analysis and (2) well-being measurement has serious limitations, i.e. there is a significant measurement gap between the theoretical models of well-being and their operationalisation based on statistical data sets.

Appendix 1

	Numerator	Reference year	Denominator
	Material welfare		
1	Inland incomes	2007	Population of the sub-region
2	Poverty rate (Number of inhabitants with incomes less than 60% of the compulsory minimum wage)	2007	Number of personal tax payers
3	Sub-regional Gini index*	2007	
	Physical well-being		
4	Malignant tumours of lungs and bronchial tubes	2008	100 000 inhabitants
5	Malignant tumours	2008	100 000 inhabitants
6	Diseases of the respiratory system	2008	100 000 inhabitants
7	Diseases of the digestive system	2008	100 000 inhabitants
8	Malignant breast tumours	2008	100 000 inhabitants
9	Diseases of the circulatory system	2008	100 000 inhabitants
10	Infant mortality (Infant death)	2008	1 000 births
	Cultural services		
11	Seating capacity of cinemas	2008	100 000 inhabitants
12	Number of cinema performances	2008	100 000 inhabitants
13	Number of museums	2008	100 000 inhabitants
14	Number of museum exhibitions	2008	100 000 inhabitants
15	Number of creative cultural communities	2008	100 000 inhabitants
16	Number of regular cultural activities	2008	100 000 inhabitants
17	Number of municipalities with cultural institutions	2008	Number of municipalities in the sub-region
18	Number of municipalities with a public library	2008	Number of municipalities in the sub-region
19	Number of cultural institutions	2008	100 000 inhabitants
	Recreation services		
20	Number of municipalities with sports halls, sports grounds	2008	Number of municipalities in the sub-region
21	Number of municipalities with baths/swimming pool	2008	Number of municipalities in the sub-region
22	Number of playgrounds, sports grounds and picnic areas	2008	100 000 inhabitants
23	Surface of playgrounds, sports grounds and picnic areas	2008	1 000 inhabitants
	Labour opportunities		
24	Number of persons seeking employment for over 180 days	2008	Population aged 15–59
	Forms of social care – Basic forms		
25	Number of flats connected to public drainage	2008	Number of flats in the sub-region
26	Number of municipalities with public sewage disposal	2008	Number of municipalities in the sub-region
27	Number of flats connected to drinking water system	2008	Number of flats in the sub-region
28	Number of flats involved in regular waste collection	2008	Number of flats in the sub-region

(Continued on the next page)

(Continued)

	Numerator	Reference year	Denominator
29	Number of municipalities with post office(s)	2008	Number of municipalities in the sub-region
30	Number of inhabitants involved in water delivery because of unsatisfactory quality of drinking water from the aspect of public health	2008	10 000 inhabitants
31	Number of public wells	2008	1 000 km^2 within municipality boundaries
	Forms of social care – health		
32	Number of municipalities with family doctor	2008	Number of municipalities in the sub-region
33	Number of municipalities with outpatient medical attendance	2008	Number of municipalities in the sub-region
34	Number of municipalities with pharmacy	2008	Number of municipalities in the sub-region
35	Number of functioning hospital beds	2008	100 000 inhabitants
36	Number of family doctor services	2008	100 000 inhabitants
37	Number of paediatric services	2008	100 000 inhabitants between age 0–17
38	Number of pharmacies	2008	100 000 inhabitants
	Forms of social care – child care		
39	Number of places in kindergartens (including special education)	2008	1 000 children of kindergarten age
40	Number of kindergartens (including special education)	2008	1 000 children of kindergarten age
41	Number of municipalities with kindergartens	2008	Number of municipalities in the sub-region
42	Number of municipalities with day care	2008	Number of municipalities in the sub-region
43	Number of children (from age 0-17) placed under child protection	2008	Number of children from age 0 to 17
44	Number of endangered children (from age 0 to 17)	2008	Number of children from age 0 to 17
45	Number of children applying for day care but rejected due to the lack of day care places	2008	1 000 children of day care age
	Forms of social care – basic education		
46	Number of municipalities with primary school	2008	Number of municipalities in the sub-region
47	Number of primary schools (including special education)	2008	1 000 children of primary school age
48	Number of primary school classes (including special education)	2008	1 000 children of primary school age
49	Number of full-time teachers (including special education)	2008	1 000 children of primary school age
50	Number of computers in primary schools (including special education)	2008	1 000 children of primary school age

(Continued on the next page)

(Continued)

	Numerator	Reference year	Denominator
	Forms of social care – security		
51	Number of assaults	2008	100 000 inhabitants
52	Reported crimes	2008	1 000 inhabitants
53	Accidents caused by vehicles	2008	100 000 inhabitants
54	Number of casualties and serious road accidents	2008	100 000 inhabitants
55	Number of people seriously injured or killed in road accidents	2008	100 000 inhabitants
	Education		
56	Average number of completed years of schooling of inhabitants over 7 years old	2001	
57	Number of inhabitants who did not complete the first year of primary school	2001	Inhabitants over 7 years old
58	Number of inhabitants with primary qualification only	2001	Inhabitants from age 18 to 24

REGIONAL STATISTICS, 2012, VOL. 2: 61–73 DOI: 10.15196/RS02105

TAMÁS DUSEK

Bidimensional Regression in Spatial Analysis*

Bidimensional regression is a method developed by Waldo Tobler for comparing the degree of resemblance between two two-dimensional configurations of points or surfaces. In case of spatial analysis and any other research focusing on two- or multidimensional configurations, a need arises to compare these with each other. Bidimensional regression makes it possible in such a way that it transforms one of the configurations of points being in different coordinate systems into the coordinate system of the other with the proper degree of displacement, rotation or rescaling. Between the points of configurations transformed into a joint coordinate system this way, it is possible to determine the degree of the local and global similarity or dissimilarity of the configurations.

After the antecedents in the sixties and seventies, Waldo Tobler published his study introducing this method in 1994[1] (Tobler 1961, 1965, 1978, 1994). In his demonstrative example, Tobler compared a medieval map[2] of Britannia with a modern map. He mentioned further application options, for example comparisons between the faces of father and son, shapes of leaves, skill forms of the Australopithecus and the Homo sapiens as well as signatures. The possible applications of this technique cover such a large range that the procedures should become as well known as the technique of ordinary regression. (Tobler 1994, 187) Lloyd and Lilley also used this method to analyse the Gough map.

In case of point configurations generated with a multidimensional scaling, if they represent a matrix containing non-air kilometre distances between different geographical points, as well as at cognitive maps, a need arises to compare it with geographical coordinates. Because of the differences (displacement, rotation and scaling) between the geographical and the multidimensional scaling (hereinafter: MDS) coordinate systems, the configurations cannot be directly compared, however, bidimensional regression makes it possible just by the objective fitting of the two coordinate systems with the smallest possible defect. In connection with the comparability of time and geographical spaces, Ewing wrote in 1974 that there was no such method in existence (Ewing 1974, 165); however, the study of Ahmed and Miller of 2007 had already used this method.

This paper presents the main characteristics of the bidimensional regression mostly by the help of a demonstrative example. Since unidimensional, bivariate regression is a well

* This study was made with support of the Bolyai János Research scholarship. The study is the edited version of the presentation delivered at the session of "Analytical methods with space parameters" of the HAS RSC Research Methodology Subcommittee on 28 September 2010.

1 In addition to the study of Waldo Tobler of 1994, the study of Friedman and Kohler of 2003 focuses on the general issues of this method. In the demonstrative example of this latter study, a cognitive map is compared with a topographic map. However, there is no description of this method in the books of spatial statistics or general statistics.

2 Gough map.

known and widely used analytical method, I will dwell on the resemblances and differences between the bidimensional regression and the unidimensional regression.

Signs and abbreviations

In a unidimensional regression, the closeness of the relationship between different characteristics belonging to the same observational units is under investigation, while in case of a bidimensional regression, the degree of the relationship between two configurations that can be mutually associated with each other. In a unidimensional regression, the possibility to associate data is obviously ensured by the fact that they belong to the same observational units. In case of a bidimensional one, in geographical applications, the observational units are the same, but the distances between them may be different (e.g. air kilometre distance and time distance). In non-geographical applications, some kind of procedure is applied to make the points of two point patterns with the same number of points, e.g. the coordinates of specific points of two leaves, mutually connectible. The configurations may represent real point configurations and plains represented by points.

The point configurations to be compared may be displayed in two or one coordinate system(s) (figure 1.a and 1.b). The b part of the figure is confusing to some extent as the starting-point, the scaling and the direction of the coordinate systems may be arbitrary. To be more explicit, one part of the task is to ensure the possibility to display configurations of different coordinate systems in a single coordinate system. If X_i, Y_i are the coordinates of the independent configuration, A_i, B_i the coordinates of the points of the dependent configuration, then A_iB_i will be the coordinates of the independent configuration displayed in the coordinate system of the dependent configuration (Table 1). If we would like to express MDS coordinates in geographical coordinates, then the MDS configuration must be chosen as an independent configuration and the geographical configuration as a dependent one. In the knowledge of the relationship between the two configurations, on the one hand, we would like to determine the average degree of displacement, rotation and scaling; on the other hand, it is desirable that we would be able to give the transformed coordinates of those points where there was no observation.

Figure 1.a

The point configurations to be compared are consisted of points that can be mutually associated

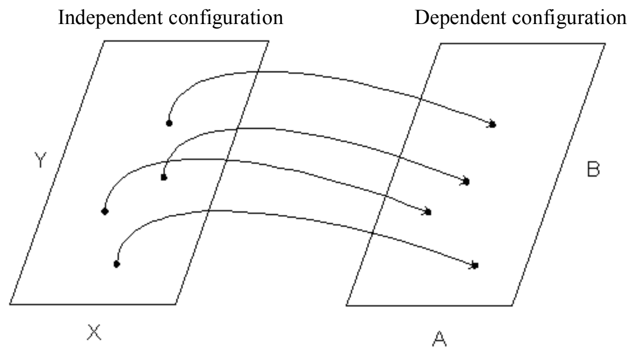
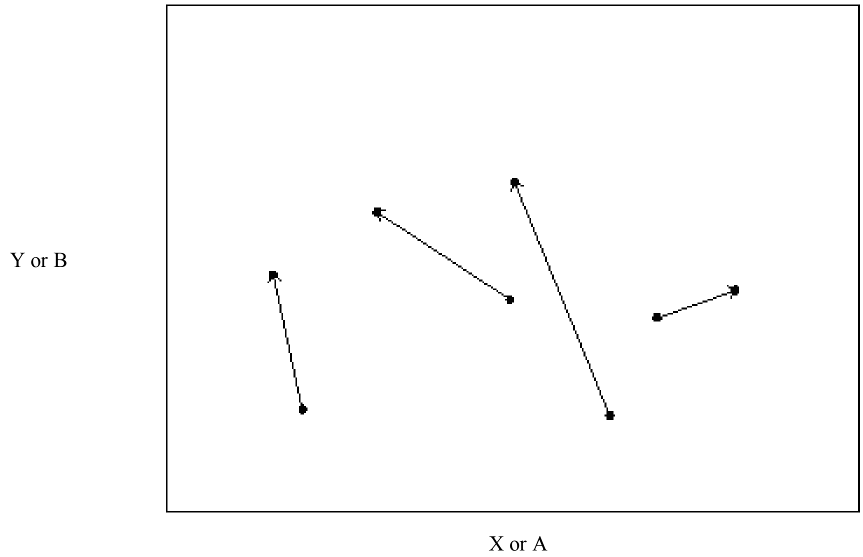


Figure 1.b

Joint coordinate system showing the point configurations to be compared



Source: Own chart based on Tobler (1994).

Table 1

Marking the coordinates of two point configurations

XY (Independent configuration)	AB (dependent configuration)	$\hat{A}\hat{B}$ (independent configuration in the coordinate system of the dependent configuration)
$X_1\ Y_1$	$A_1\ B_1$	$\hat{A}_1\hat{B}_1$
$X_2\ Y_2$	$A_2\ B_2$	$\hat{A}_2\hat{B}_2$
$X_3\ Y_3$	$A_3\ B_3$	$\hat{A}_3\hat{B}_3$
.		
.		
.		
$X_n\ Y_n$	$A_n\ B_n$	$\hat{A}_n\hat{B}_n$

In a unidimensional regression, we would like to minimize the $\sum(y_i - \hat{y}_i)^2$ amount, while in a bidimensional regression the $\sum(X_iY_i - \hat{A}_i\hat{B}_i)^2$ amount. Based on the method of transforming coordinates, Waldo Tobler mentioned four possible solutions: Euclidian, affine, projective and curvilinear transformation; of these, the first three are linear functions, i.e. what is a straight line in one of the coordinate systems will be a straight line in the transformed coordinate system as well. In the Euclidian version, x and y coordinates change to the same degree (proportion), in the affine one the x and y coordinates may change in different proportions, while in the projective one the scale, the shape, and the rotation may change from point to point. This study is restricted only to how the Euclidian version is determined. The calculation of the three other ones would be more complicated, while in the projective and the curvilinear ones, the interpretation of the results would be more difficult as well.

Computation and interpretation of the parameters of the Euclidian regression

Equations relating to the computation of the Euclidian version can be seen in Table 2. Of the markings, x and y are the coordinates of the independent configuration, a and b the coordinates of the dependent configuration, while $\hat{a}$ and $\hat{b}$ the coordinates of the independent configuration inside the coordinate system of dependent configuration. Of the four parameters of the first (matrix-algebraic) equation, α1 is to determine the degree of the horizontal displacement, while α2 the degree of the vertical displacement. If we draw a parallel between the one and the two dimensional regressions, then these two coefficients will correspond to the β1 (constant) parameter of the one dimensional regression. β1 and β2 are to determine the scale difference (Φ) and the angle of the rotation (Θ) in a way that can be seen in the first and second equalities. If Φ=1, then there is no scale difference between the two configurations, if Φ>1 then it means the zooming of XY and the reduction of that at Φ<1. If Θ=0, then there is no need to rotate the XY coordinate system, if it is negative, then it means a clockwise rotation. Since the arcus tangent equation can be interpreted only between -90 degrees and +90 degrees, 180 degrees must be added to Θ, if $\beta_1<0$. Φ is the β_1 parameter of the one-dimensional regression, Θ is the specific parameter of the two dimensional case.

Table 2

Equations of a bidimensional Euclidian regression

1. Equation of the regression	$\begin{pmatrix}\hat{A}\\ \hat{B}\end{pmatrix}=\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}+\begin{pmatrix}\beta_1 & -\beta_2\\ \beta_2 & \beta_1\end{pmatrix}*\begin{pmatrix}X\\ Y\end{pmatrix}$
2. Scale difference	$\Phi=\sqrt{\beta_1^2+\beta_2^2}$
3. Rotation	$\Theta=\tan^{-1}\left(\frac{\beta_2}{\beta_1}\right)$
4. Calculation of β_1	$\beta_1=\frac{\sum(a_i-\bar{a})*(x_i-\bar{x})+\sum(b_i-\bar{b})*(y_i-\bar{y})}{\sum(x_i-\bar{x})^2+\sum(y_i-\bar{y})^2}$
5. Calculation of β_2	$\beta_2=\frac{\sum(b_i-\bar{b})*(x_i-\bar{x})-\sum(a_i-\bar{a})*(y_i-\bar{y})}{\sum(x_i-\bar{x})^2+\sum(y-\bar{y})^2}$
6. Horizontal displacement	$\alpha_1=\bar{a}-\beta_1*\bar{x}+\beta_2*\bar{y}$
7. Vertical displacement	$\alpha_2=\bar{b}-\beta_2*\bar{x}-\beta_1*\bar{y}$
8. Correlation based on random errors	$r=\sqrt{1-\frac{\sum\left[(a_i-\hat{a}_i)^2+(b_i-\hat{b}_i)^2\right]}{\sum\left[(a_i-\bar{a})^2+(b_i-\bar{b})^2\right]}}$
9. Breaking down the sum of squares	$\sum\left[(a_i-\bar{a})^2+(b_i-\bar{b})^2\right]=\sum\left[(\hat{a}_i-\bar{a})^2+(\hat{b}_i-\bar{b})^2\right]+\sum\left[(a_i-\hat{a}_i)^2+(b_i-\hat{b}_i)^2\right]$ SST=SSR+SSE
10. Calculation of $\hat{a}$	$\hat{a}=\alpha_1+\beta_1(X)-\beta_2(Y)$
11. Calculation of $\hat{b}$	$\hat{b}=\alpha_2+\beta_2(X)+\beta_1(Y)$

Source: Based on Tobler (1994) and Friedman–Kohler (2003). A key to these signs and abbreviations can be found in the text.

The greater the similarity between the two point configurations, the greater the degree of the bidimensional correlation (r). If as a result of displacement, rotation, and scaling, the coordinates of the points will correspond to each other, the indicator will reach the maximum value at one. The minimum value of the correlation is zero, which means that all points of one of the point configurations have the same coordinates. At this time, the centre of gravity of the two configurations will be the same, but the distance between them will be the same as the distance from the centre of gravity of that configuration, which is not clustering around one point. It may happen that we would like to disregard one of the three transformations, in that way the calculations relating to the similarity between the two configurations might also be implemented.

In principle, the breaking down of the sum of squares is performed in the same way as in case of the unidimensional one, the abbreviations are also the same (SST: total sum of squares, SSR: the sum of squares explained by the regression, SSE: residual sum of squares, which is not explained by the regression). The practical and interpretational difference is indicated by the fact that in a bidimensional regression, the difference indicates the distance from the centre of gravity of the analysed points and not from the average of a quantitative attributum variable.

In tables 3 and 4 as well as in figure 2, an example can be seen showing a comparison between the geographical distances and the railway kilometre distances of five settlements. Railway kilometre distances determine a non-Euclidian space, a two dimensional scaling was applied to their bidimensional approximation. (For example, it shows that as Mosonmagyaróvár may be reached by only a roundabout route first of all from Sopron and, to a smaller extent, from Szombathely and Pápa, therefore it will be farther away from the previously mentioned three settlements in the space of the railway network). The average of the MDS coordinates is zero, they have no unit of measurement, and the geographical coordinates are the coordinates of the uniform national projection given in terms of kilometres. The value of Φ of 62.41 gives the exact difference of their scale. MDS coordinates must be rotated by 19.347 degrees in an anticlockwise direction (value of Θ). The degree of the horizontal shift is 504.59 (α_1), and 247.96 is that of the vertical one (α_2).

Table 3

Two-dimensional regression between the space of the railway network and the geographical space

Settlement	MDS-coordinates		Geographical coordinates		MDS-coordinates	
	X	Y	A	B	$\hat{A}$	$\hat{B}$
Győr	0.350	0.055	543.90	260.70	524.10	258.40
Mosonmagyaróvár	0.578	0.354	516.60	280.00	531.30	280.70
Pápa	0.273	–0.540	530.00	223.10	531.80	222.00
Sopron	–0.600	0.456	465.30	262.20	459.80	262.4.0
Szombathely	–0.601	–0.330	467.10	213.80	476.00	216.20
Average	*0.000*	*0.000*	*504.58*	*247.96*	*504.60*	*247.94*

Table 4

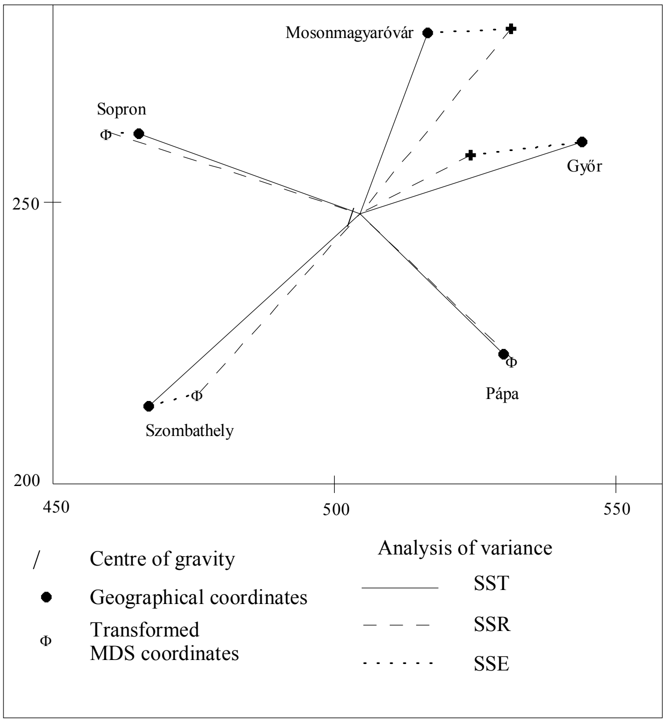
Results of the two dimensional regression (concerning data of table 3)

r	r^2	α_1	α_2	β_1	β_2	Φ	Θ	SST	SSR	SSE
0.956	0.913	504.59	247.96	58.877	20.673	62.41	19.347	8462.9	7730	732.9

The utility of the correlation and the determination coefficients is the same as in the case of the unidimensional regression, however, as it is well known in this latter case, it should be cautiously interpreted since its value may be influenced by many kinds of factors. The comparison of random errors of observations shows those points, which are most responsible for the deviations, though the warning that the influential observations are not necessarily the most outstanding ones is also valid in this case. The shift in the direction of the centre of gravity means that the average distance of the given point from the other points is smaller in the independent configuration, than in the dependent configuration. For example, in figure 2, the average distance between Győr and Szombathely, regarding its proportions, is smaller, while that of the other three settlements is greater in the space of the railway network, than based on air kilometre distances. (The railway network distances, in an absolute way, are also greater in case of Győr and Szombathely than the air kilometre distances, but they increased at a lower rate compared with the three other settlements.)

Figure 2

Geographical distances of settlements and their transformed MDS coordinates representing railway network distances



The independent point configuration with its original coordinates cannot be seen in figure 2, as the great differences of the scale and the displacement would not make possible the joint representation of XY and AB. The relative configuration of XY and AB and the proportions between the distances of certain points correspond to each other. The results gained with the transposition of dependent and independent configurations can be seen in table 5. The closeness of the correlation do not change, the absolute values of the transformations will be different; instead of extension, a reduction corresponding to the degree of the reciprocal value of the extension, shifts of opposing direction and of different degree and a rotation of opposing direction (but of the same degree) will be necessary.

Table 5

Results of a bidimensional regression
(to the data of table 3, but XY and AB are transposed)

r	r^2	α_1	α_2	β_1	β_2	Φ	Θ	SST	SSR	SSE
0,956	0,913	–8,171	–0,978	0,014	–0,00485	0,01464	–19,347	1,9851	1,8132	0,1719

Comparison with other procedures

Several other methods exist to compare two dimensional configurations: various shape indices, separate regressions by dimension, correlation between the distance matrices of points. The advantage of the two-dimensional regression as opposed to all existing methods, is that this is the only method that takes into account all information relating to the configuration of the formations. The separate regression by dimension is a unidimensional method, it does not take into account that XY and AB express the coordinates of certain points in an inseparable way. Regressions by dimension are also sensitive to rotation (or the lack of that). The regression between the distance matrices, though it provides another indicator for the degree of similarity, is not appropriate to determine scale difference, rotation and displacement.

The coordinates of the configurations to be compared may derive from a bidimensional scaling. However, the different local and global indicators (the Stress measures) of the goodness of the two-dimensional scaling do not reveal a deviation in relation to the geographical space, but the differences of the original distance matrix and the two-dimensional distance matrix. Therefore, these two methods may be connected inside the same analysis, but in a mutually supplementing and not substituting manner.

Toolkit of visual presentation and an example on the application of this method

This method is not included in any statistical or geostatistical software. However, the coefficients themselves can be easily computed in Excel based on the formulas. Mapping is much more labour intensive. The software, originally developed by Waldo Tobler, then redeveloped by Guerin, which can be downloaded from the following website: http://www.spatial-modelling.info/Darcy-2-module-decom-paraison is a help in this. The coefficients also can be calculated with this; however, the applicability of the visual display is significantly reduced by the fact that the figures are non-editable and non-inscribable.

I will show through examples the tools of the visual display relating to this method including some further minor questions. Concerning our 23 towns (Budapest and towns of county rank with the exception of Érd), table 6 is to show the results of the relationships between coordinates received with bidimensional scaling using geographical coordinates and corresponding distance matrices determined in different ways. Of the options offered by the Elvira internet time table search engine, I took the shortest travel time as a railway time distance, the public road distance matrices were made and put at my disposal by Péter Tóth, who used a piece of Microsoft MapPoint 2009 software (Széchenyi István University). Public road data are relating to July 2008, the railway data to November 2009.

Table 6

Results of bidimensional regressions between geographical coordinates and coordinates determined by multidimensional scaling from other different distance matrices (22 towns of county rank and Budapest)

Denomination	r	r^2	α_1	α_2	β_1	β_2	Φ	Θ
Railway network distance	0.974	0.949	649.0	199.5	188.3	40.2	192.6	12.06
Railway network time distance	0.941	0.885	649.0	199.5	184.9	28.3	187.0	8.71
Railway cost distance	0.973	0.947	649.0	199.5	191.0	29.9	193.3	8.91
Public road distance	0.990	0.980	649.0	199.5	192.8	32.7	195.5	9.63
Public road time distance	0.975	0.950	649.0	199.5	–190.6	–28.7	192.8	188.57

Remark: Railway distances: November 2009, public road distances: July 2008.

The relatively high level of correlations indicates that at this level there is no such significant distortion in the railway, public road network and time space, which may significantly alter the neighbourhood relationships of the settlements. If we analyse the more detailed network of 142 settlements, then the picture is modified at the time distances (Table 7). In the selection of the 142 settlements, the settlement size and the railway network location played a role. In those configurations, which are made up by many points, because of the crowdedness of the chart, it is more difficult to see the shifts, but in case of a good fitting, it is easy to interpret the overall view. However, many points and bad fitting jointly result in such charts, which can be interpreted in a more difficult way, though the more interesting single shifts may be stressed in this case as well and the spatial segment especially responsible for the bad fitting might also be identified.

Table 7

Results of two-dimensional regressions between geographical coordinates and coordinates determined by multidimensional scaling from other different distance matrices (142 Hungarian settlements)

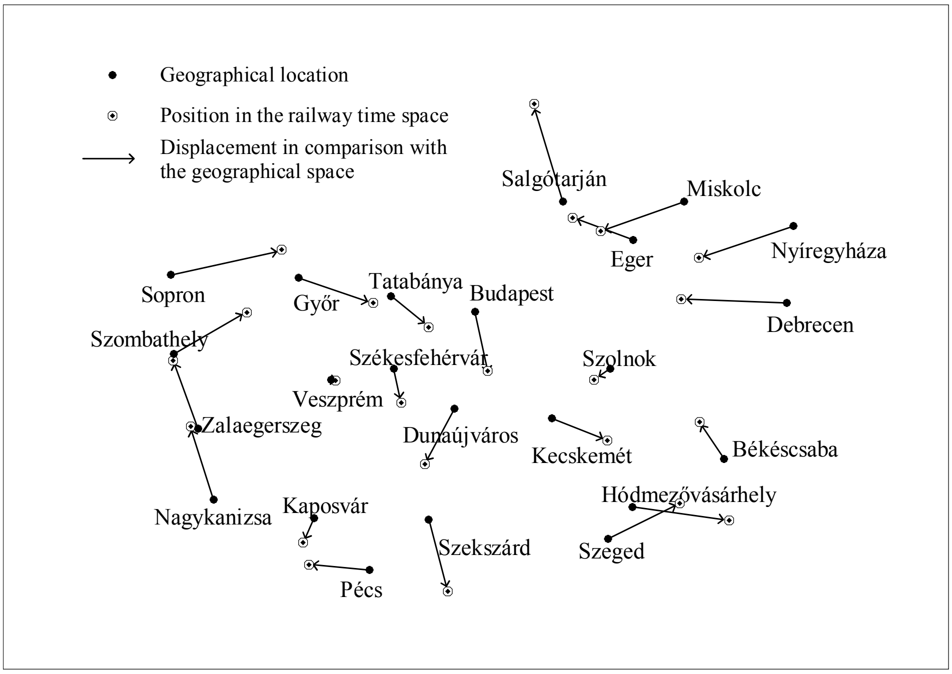
Denomination	r	r^2	α_1	α_2	β_1	β_2	Φ	Θ
Railway network distance	0.984	0.968	687.0	216.0	–187.4	–60.5	197.0	197.9
Railway time distance	0.781	0.610	687.0	215.9	–155.3	–33.7	158.9	192.2
Public road time distance	0.446	0.199	687.0	216.0	–84.4	–30.1	89.6	199.6

Remark: Railway distances: November 2009, public road distances: July 2008.

The geographical and railway time spatial position of the examined settlements can be seen in Figure 3. Graphical display is much richer in information than the quantitative display of the coordinates and their differences, because it shows the size and the direction of the change concerning all settlements. For example, it can be seen that Budapest shifted in the direction of the centre of gravity, because its railway accessibility is better than its otherwise favourable, near-central geographical location. Nyíregyháza, Debrecen, Miskolc, Győr, Sopron, Tatabánya significantly shifted in the direction of the centre of gravity, and, to a smaller extent, the same can be said of Szombathely and Békéscsaba. The settlements of Southern Great Plain and Southern Transdanubia moved farther away from each other, because of the non-appropriate east-west interconnection of the south of the country. The time spatial position of Dunaújváros and Salgótarján is significantly worse than their geographical positions. Gábor Szalkai draw the same conclusions through the use of detour indices concerning the differences of the domestic railway time space and the geographical space (Szalkai 2001, Szalkai 2004), however, the method of the detour index is not appropriate in itself to indicate those differences that can be determined from direction vectors.

Figure 3

Location of towns of county rank and Budapest based on geographical coordinates and railway time spatial coordinates



Remark: Figure 3–5 were made with MapInfo based on computed, modified coordinates.

At the same time, the figure contains only those points that take part in the regression, so it is more difficult to read the information, similarly to that as if only the estimated points would be displayed in a unidimensional two-variable regression instead of the straight line of the regression. Interpolation makes it possible to display those points that do not take part in the regression if we know the shift of those points that take part in the regression. The mesh fitted on the coordinate system of the dependent configuration and the interpolated, modified position thereof further generalize the information derived from those points that take part in the regression. In parts a and b of figure 4, in addition to the original and the interpolated mesh, those points that take part in the interpolation (towns of county rank) can also be seen. However, it is very rare that there are no linear geographical objects that are well known from the topographical maps, for example in case of Hungary county borders, the position of which may be interpolated as well (figure 5). This figure spectacularly displays the direction and degree of the more significant differences between the geographical space and the railway time space.

Figure 4.a

Geographical coordinates of a grid

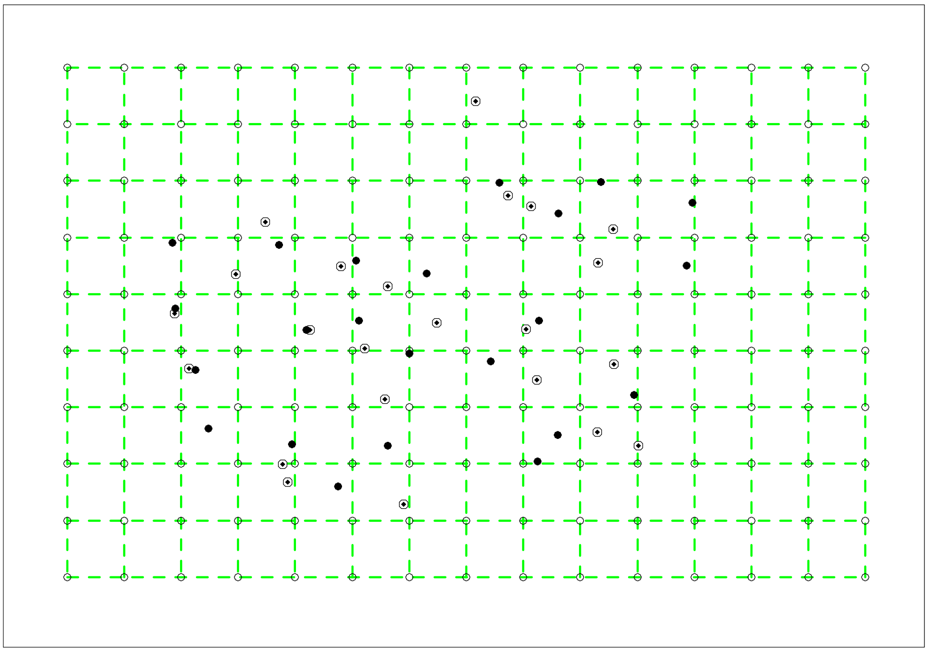


Figure 4.b

Warped grid: time space coordinates of the grid

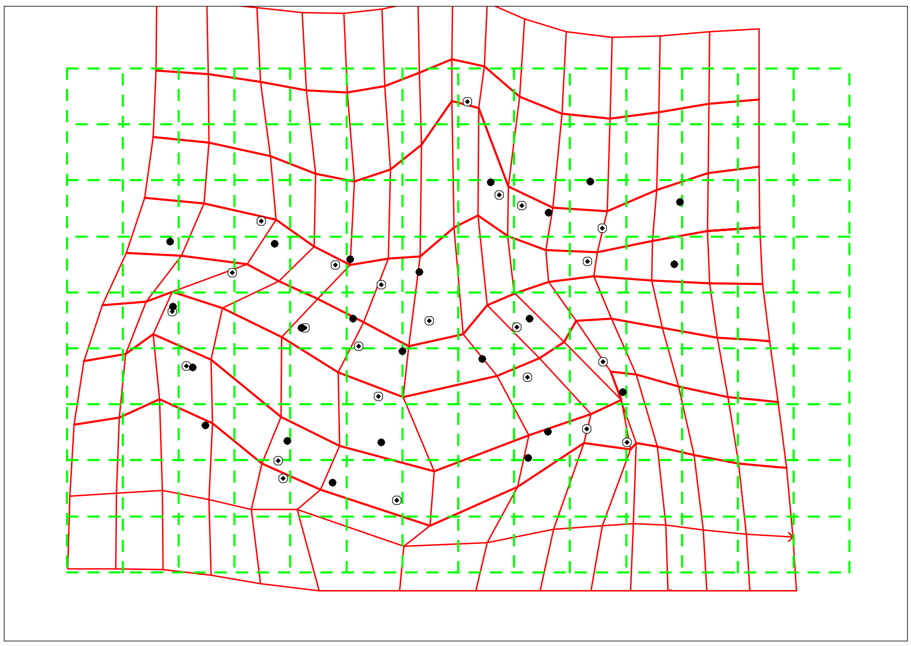


Figure 5

Interpolated position of county borders in railway time space

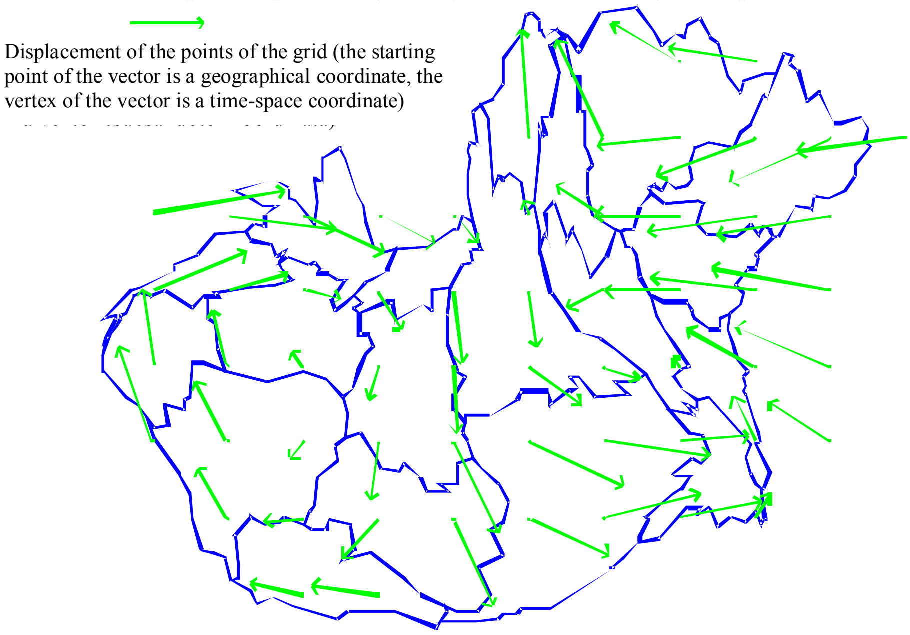
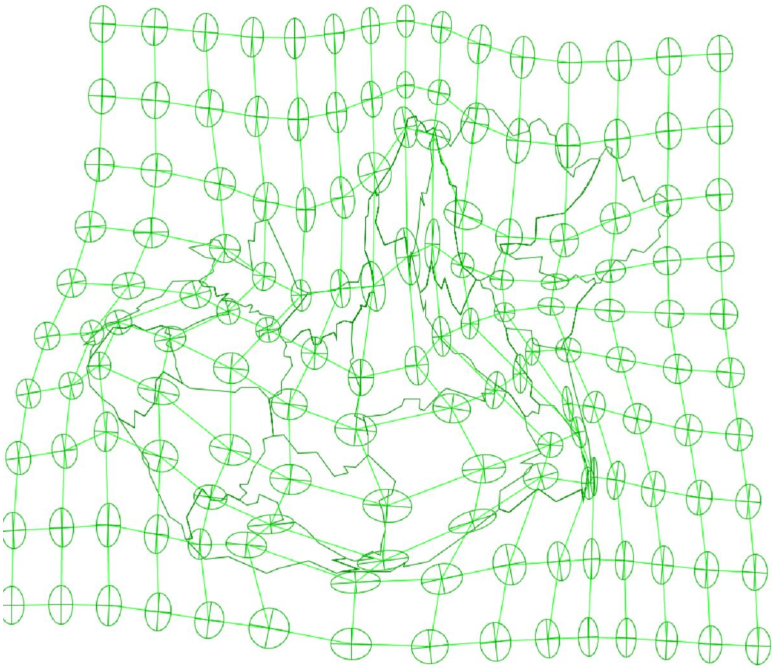


Figure 6

Distortion of the railway time space in comparison with the geographical space



Remark: This figure was made with the Darcy software (by B. Guerin). This software can be downloaded from the following website: http://www.spatial-modelling.info/Darcy-2-module-de-comparaison.

In the cartography, the distortion of the geographical projection may be expressed in a graphic manner with the help of the Tissot's indicatrix. Waldo Tobler presented, as an analogy of this, that the degree of the local deviation of the bidimensionally-compared configurations can be displayed with distortion ellipses (Tobler 1994). These distortion ellipses can be seen in the nodes of the interpolated grid on figure 6. If the area of the ellipse is greater than average, then in the given spatial segment, the time space will be more elongated than the geographical space. The main axis of the ellipse will shift in the direction of the greatest distortion.

Concluding thoughts

As it can be seen from the presented calculations and display options, bidimensional regression provides basic data for the visualization, interpretation, measurement, and analysis of the non-Euclidian spaces, and it is not only a procedure of comparing configurations with each other, which is much more effective and rich in information compared with the other known methods. If the coordinates are already available (the determination of these can be very laborious if the distance data needs to be collected on an individual basis), the calculation of correlation and regression is not complicated; in the visual display of the results only the display of the interpolated coordinates is more labour intensive. Then again, in the relation to this method, a number of questions of detail might be discussed, among them it should be stressed, the development of indicators that can be used to characterize the influential observations, the connection with the multidimensional scaling and the presentation of the affine transformation as well as the interpolation procedures. The present study's intention is to draw attention to the existence of this method in the first place, with the aim to facilitate its penetration and practical application.

REFERENCES

Ahmed, N. – Miller, H. J. (2007): Time-space Transconfigurations of Geographic Space for Exploring, Analyzing and Visualizing Transportation Systems. *Journal of Transportation Geography* 15 (1): 2–17.

Ewing, G. (1974): Multidimensional Scaling and Time-space Maps. *Canadian Geographer* 18 (2): 161–167.

Friedman, A. – Kohler, B. (2003): Bidimensional Regression: Assessing the Configural Similarity and Accuracy of Cognitive Maps and Other Two-Dimensional Data Sets. *Psychological Methods 8(4):468–491.*

Lloyd, C. D. – Lilley, K. D. (2009): Cartographic Veracity in Medieval Mapping: Analyzing Geographical Variation in the Gough Map of Great Britain. *Annals of the Association of American Geographers* 99 (1): 27–48.

Szalkai Gábor (2001): Elérhetőségi vizsgálatok Magyarországon. *Falu–Város–Régió* 3 (10): 5–13.

Szalkai Gábor (2004): A közlekedéshálózat fejlesztésének hatása az elérhetőség változására. Magyar Földrajzi Konferencia CD-kiadványa

Tobler, W. (1961): Map Transconfigurations of Geographic Space. PhD dissertation, University of Washington, Seattle

Tobler, W. (1965): Computation of the Correspondence of Geographical Patterns. *Papers of the Regional Science Association* 15 (1): 131–139.

Tobler, W. (1978): Comparisons of Plane Forms. *Geographical Analysis* 10 (2): 154–162.

Tobler, W. (1994): Bidimensional Regression. *Geographical Analysis* 26 (3): 186–212.

Keywords: bidimensional regression, regional analytical methods, multidimensional scaling, railway time space.

Abstract

Bidimensional regression is a method developed by Waldo Tobler for comparing the degree of resemblance between two two-dimensional configurations of points or surfaces. It is an extension of linear regression where each variable is a pair of values representing a location in a two-dimensional space. Bidimensional regression numerically compares the similarity between two-dimensional surfaces through an index called bidimensional correlation. The aim of the study is the general description of the method, and to present examples of its application with the help of some real data, that of the Hungarian railway time space.

REGIONAL STATISTICS, 2012, VOL. 2: 74–89 DOI: 10.15196/RS02106

ÁRON KINCSES – GÉZA TÓTH

Geometry of Potential Models*

Introduction

We aim to rethink and systematize accessibility potential models, and to investigate what they measure and under what conditions, what characteristics they possess by way of analogy with a gravitational field and what features they "inherit".

The question is what has potential, what a particular potential value means, and why we use this value to describe social space. What conclusions can be drawn from a potential model? To answer this latter question, the values forecast by the models were compared with traffic data.

Potentials measure the position of a spatial domain compared to the rest of the areas, and the impacts of the mass distribution of the particular spatial division at the same time. In the present paper we make an effort to filter out these effects, and to disaggregate potentials.

Space in geometry and in physics, gravity

Geometric models

In mathematics, three theoretically different geometric analysis methods can be distinguished: axiomatic, group-theoretical and differential-geometric.

The axiomatic construction of geometry begins with the selection of the object set (e.g. pairs of numbers or points in Euclidean space). In this set basic elements, basic relations (e.g. adaptation) are then defined. Subsequently, general statements are generated in support of which the particular system is introduced. These are axioms. If axioms are true in an object set, then this object set is referred to as a model of the geometry provided by the geometric system. So axioms – howtever strange this may be – need to be justified in constructing a model. The model and "reality" can be connected by this moment.

Geometric space in physics

Model creation, the description of space by models is not a characteristic of geometry or mathematics only. It is right to ask in physics as well what geometric relations can be applied to describe physical phenomena.

* The study is the edited version of the presentation made at the "Analysis methods using spatial parameters" session of the Research Methodology Sub-committee of the Regional Science Commission of the Hungarian Academy of Sciences, held on 28 September 2010.

The research was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

According to János Bolyai "*the law of gravitation also seems to be closely related to the size, structure and quality of space.* This means the recognition that there should be an internal coherence between physical gravitational field and the geometrical structure of space (Oláh-Gál 2008).

When detailing his theory, Einstein relied on the non-Euclidean concept of space, and used its interpretation and symbols further developed by B. Riemann more than two decades after Bolyai. L. Infeld, one of the colleagues of Einstein writes the following: "The interpretation of the gravitational field as geometric space is one of the greatest and most revolutionary achievements ever in the history of physics. A world without masses, electrons and electromagnetic space is an empty world, a false idea. However, if masses, charged particles and electromagnetic space emerge, then the gravitational field emerges too. If the gravitational field emerges, then our world bends. Its geometry is Riemann's geometry and not the Euclidean one (Gábos 2004)."

Gravitational force, force field, field strength, potential

According to the law of universal gravitation or the gravitation law of Newton (1686) any two point-like bodies mutually attract each other with a force directly proportional to the product of the masses of the bodies and inversely proportional to the square of the distance between them:

$F = \gamma \cdot \frac{m_1 \cdot m_2}{r^2}$, where γ is the proportionality factor, the gravitation constant (independent from place and time).

If the radius vector drawn from mass point 2 to mass point 1 is indicated by $\vec{r}$, then the unit vector pointing from 1 to 2 is $-\vec{r}/r$, and so the gravitational force on mass point 1 by mass point 2 is:

$$\vec{F}_{1,2} = -\gamma \cdot \frac{m_1 \cdot m_2}{r^2} \cdot \frac{\vec{r}}{r}.$$

The negative sign in the formula expresses that body i attracts body j (Budó 1970).

Generally, in case of a gravitational field from any body or any system of bodies, the force on a body with mass m is proportional to m: $\vec{F} = m\vec{K}$.

The $\vec{K}$ vector quantity of the gravitational field, which – disregarding the dimension – means a force on a body with unit mass, is referred to as (gravitational) field strength. Field strength usually depends on place and maybe time t: $K = K(x, y, z, t)$.

The fundamental concept of the force field comes from Faraday (around 1840), who replaced the concept of "distance action" concerning electronic and magnetic force effects between distant bodies as well as in vacuum, accepted until that time, with the principle of proximity effect or spatial effect. Accordingly, force effects between separate bodies are always mediated by space: the direct reason for the force by a body in place A on a body in place B is that the field strength by the body in place A differs from zero in place B.

A force field can be completely defined if field strength $\vec{K}$ can be given by direction and size in every point of the particular domain. Many force fields, however, including the gravitational field, can be described much more simply as well, by only one scalar function instead of three, the “potential”. The potential is related to field strength as work or potential energy to force. So the value of the potential in some point P of the gravitational field:

$$U_P = -\int_0^P K_s ds,$$

i.e., not taking into consideration dimension, equals work against gravitation forces while taking a body with unit mass from “zero point” O (in any way) to point P (Budó 1970).

Accessibility potential models

Potential model in regional analyses

The regional analysis tools of spatial interactions are potential models, which show the position advantage of each area compared to other areas, quantifying the advantage ensured by them (Schürmann–Spiekermann–Wegener 1997). According to other definitions accessibility “shows the character of spatial interaction”, and “is the attraction of a node, taking into account the mass of other nodes and the costs of access to them on the network” (Bruinsma–Rietveld 1998). There is no generally agreed definition of accessibility; many different indicators with differing methodologies are used in empirical studies (for example Ingram 1971, Morris–Dumble–Wigan 1978, Handy–Niemeier 1997). On the one hand, the main task of accessibility tests is to ensure a measuring tool appropriate for evaluating the accessibility to all origin and destination points and on the other hand, for explaining the differences measured in accessibility (Chapelon 1997).

Both gravitation and potential models are based on the fact that the behaviour of spatial groups of people is determined by certain laws, and these laws are the same as physicists’ laws determining the behaviour of groups of molecules. Human beings are certainly not like molecules but the behaviour of people and bodies are similarly subject to gravitation law. On the basis of this analogy the investigation of people’s behaviour based on physical laws is also referred to as “social physics” (Carrothers 1956). Therefore a common feature of the models is that potential interaction between two settlements, areas etc. is inversely proportional to the distance between them. The other similarity is that any person in the examined settlements generates the same amount of interaction as anybody else. That is, the amount of interaction between two areas is directly proportional to the masses of units in the analysis.

In literature on accessibility, indicators are divided basically into three groups. There are models based on infrastructure, activities and usefulness. Accessibility potential models belong to the group based on activities (for more details see Tóth–Kincses 2007). In this study, potential models based on gravitation analogy are examined in more detail. In these cases – similarly to Newton’s law of gravitation –interaction between masses in social space (population, economic volume) is usually described by the function of the value directly proportional to mass and inversely proportional to the power of the distance between them (for example, these types of models are applied in the analyses of attraction zones).

If two points in space, i and j are given, with masses P_i and P_j respectively, and the difference between them is d_{ij}, then the following hypothesis can be made for the intensity of interaction between them (G):

$$G = c \cdot \frac{P_i \cdot P_j}{f(d_{ij}^k)}$$, where c and k are constants.

With this formula, space can be divided, on every point of which it can be decided which of two close mass points has a more intensive impact on it.

Relations, interactions, as in inter-connections in physics, are not limited to pairs of points (or masses) in society either, many other points have an impact on each point. Mass points generate space around themselves, thus creating force field. Social space – on the analogy of gravitation (electric, magnetic) field – is attempted to be approximated by potential models. The general form of a potential in a particular point of social space:

$$T_i = \sum \frac{P_j}{f(d_{ij})} ,$$

where P_j is the active mass assigned to point j in the examined space, and d_{ij} is the distance between points i and j.

Some characteristics of potential models

Selection of area of analysis

It seems that every point of Earth has an impact on the potential of all other points. Certainly it does not mean that researchers take into account the data of all territorial units when making calculations, but one should be aware that selection affects the form of potential surface (Lukermann and Porter 1960, Houston 1969). An additional essential criterion is that the area of analysis should form a relatively closed socio-economic system.

Territorial division

Territorial division is an important issue from the point of view of potential analyses too. The problem originates primarily from making calculations not at the level of individuals in statistical analyses but applying the characteristics of groups of individuals using some administrative or statistical groupings. The difficulty in compiling data limits the selection of levels as well. If data are available, it is worthwhile to make calculations with different numbers and sizes of territorial units, for the problem of modifiable territorial units also means a relevant aspect of analysis in this respect (Dusek 2004). By applying smaller territorial units, one can have a more detailed, while in case of larger units, a smoother potential surface.

Mass factor

The interaction ability of the different areas principally depends on the scale of their socio-economic activities. To ensure that the potential can appropriately show the interaction ability of the different areas a mass factor should reasonably quantify the level of the particular activity. Its selection differs from one task to another. In the most often applied approach, population is used as mass of analysis, either unweighted or weighted by some socio-economic factor (e.g. qualifications, income). There are investigations in which population is replaced by the volume of retail sales or the income of households. In the models, mass can mean almost any extensive quantity describing social space. Discussions on the application of mass factors have lost importance since there is a close correlation between most of such factors, so their selection has a relatively small impact on the calculated potential. The selection and application of the factor of distance may be much more important (Houston 1969).

Distance

First of all, the introduction of the factor of distance in socio-geographic analyses is due to the fact that spatial separation hinders the co-operation among the many different areas; consequently, its quantification in some way is recommended. Naturally, the most simple application of the model is the use of linear distances. However, in respect of accessibility indicators, the distance, costs or time of travel by some mode of transport is always considered. The distance to be covered between two points is referred to as the spatial resistance factor (Tóth–Kincses 2007).

Distances between the particular "masses" are also considered differently in the different models. Several approaches are known which apply the reciprocal value of distance or some power of that (see among others Hansen 1959, Davidson 1977, Fortheringham 1982). Thus, there are models applying squared, exponential (Wilson 1971, Dalvi–Martin 1976, Martin–Dalvi 1976, Song 1996, Simma–Vritic–Axhausen 2001, Schürmann–Spiekermann–Wegener 1997), Gaussian (Ingram 1971, Guy 1983) and log-logistic (Bewley–Fiebig 1988, Hilbers–Veroen 1993) resistance factors. The approximations of distances with different functions are owing to the effort made to achieve the most favourable fit when examining spatial structures. For even better approximations these non-linear resistances are even transformed (examples include the Box-Cox transformation, which makes the residuals of the regression homoscedastic[1], and transforms them by approximating them to normal distribution (Box–Cox 1964)). Although these models are based on a gravity analogy, the form of the potential is different, and the meaning of factors in the formula is not always the same either. However, there is a determined relationship between field strength and the potential:

$$\vec{K} = -\mathrm{grad}U \Rightarrow \qquad K_x = \frac{\partial U}{\partial x}; \quad K_y = \frac{\partial U}{\partial y}.$$

So one can work with different types of potentials (from that induced by the analogy of gravitation), but in this case the effects of force are also different among the sources of space. These models differ from one another in that attraction forces remain above a pre-set threshold value within different distances. However, the extent to which they describe the real balance of forces between social masses is already another question.

Generally, it can be said that if the force field remains central (i.e. the effects of force depend only on distance besides mass), to describe space it is then not necessary to know

1 Equal variance with two samples.

the value and direction of field strength in every point of space. In case of central forces (resulting from irrotationality) space can be described by a single scalar function, the potential. In this case there is no need for vectors to be added even if there are several sources (masses), the values can be added mathematically as scalars.

Self-potential

In the analysed space the degree of the location-dependent potential in a particular point of space does not only depend on the distance of masses from it and their size but also on the size of the force field the particular point can generate around itself (Frost–Spence 1995, Bruinsma–Rietveld 1998). Consequently, inherent, internal and external potentials are distinguished in such potential analyses (Nemes Nagy 1998, 2005). The distinction between these latter two factors stems from the distinction between the forces of the area of analysis in a narrower sense and of the space influencing it from outside. So the potential is calculated by adding these three factors.

When calculating the self-potential of an area, it is assumed that it is not only transport from one territorial unit to another that can be a factor improving accessibility but also that of transport within the different areas/settlements. That is, it can be stated that the different products/services do not need to be transported to other areas if they can be sold within the particular area too. Leaving out of consideration the role of self-potential may lead to misleading results. It is easy to see that the accessibility of central settlements of agglomerations and settlement groups in Hungary would always be lower in such cases than for the rest of the settlements in the settlement groups of large towns.

When calculating self-potential – similarly to other potential analyses – the area of a particular area is taken into account (possibly inner area instead of the administrative one). Considering the area as a circle, the area's radius is calculated, which is regarded as proportional to distances on public roads within the different settlements, so it is also referred to as its own distance. This distance is used in models applying linear distances, while in those applying network distances this distance is recalculated in support of some average speed/costs etc. and substituted in the formula.

Calculation of potential

Scalar addition requires linear superposition among the different members in the definition of potential. There is no interaction among the different effects, they do not amplify or attenuate one another but each has its solitary impact independently from the others (there are only two-body forces and no multi-body forces), and then these independent members are added. Larger masses do not "suppress" the attraction impact of other areas and are independent from that according to the formula. Though this is a very important feature in physics, it is uncertain that social space also has this characteristic.

The location-dependent accessibility potential can be calculated as the sum of inherent and internal potentials according to the following formula:

$$P_i = \sum_i A_i = SA_i + \sum_i BA_i \ ,$$

where ΣA_i is all the accessibility potentials of area i, SA_i is the own and BA_i is the internal potential. Furthermore, there are approaches which take into account "external potentials" in addition to the area of analysis.

In most accessibility analyses the value of the potential is not calculated in all points of space, instead, the calculation is made for towns, and the resulting data are extrapolated using methods of geographical information science to the areas the centres of which are the particular towns (Baradaran–Ramjerdi 2001). This approach is slightly strange in physics, where the potential describes space created by masses and is a function that assigns a number to each point in space. That is, the potential is a point characteristic describing space, not masses (population or income of settlements or micro-regions etc.).

Test of model creation, relation between model and real space

In describing the characteristics of potential models, different model frameworks can be constructed based on differing definitions of sets of objects (points in space), and sources, masses, basic elements (definition of lines with linear distance or distance on public roads) and basic relations of space.

In the foregoing, the structure of the different models was discussed, avoiding the questions (based on geometric interpretation) of how realistically the potential structure, generated with the definition of its basic elements and relatios, defines the space, and to what extent the volumes of spatial flows can be "confronted" with the values of the models, i.e. whether or not the axioms are fulfilled. Namely, the consequences drawn from the models can be applied to real social space only in this case. Axioms here fulfil the role of a bridge between real life and the models.

The data of the Hungarian Public Roads Non-profit PLC show the annual average daily traffic passing through a section of public road. National pubic road cross-section traffic counts are implemented – in line with international practice – by sampling procedure. This method of counting makes it possible – with knowledge of fluctuations in traffic over time – to determine average daily traffic in some cross-section from relatively few data (from small samples, from the results of counts lasting for a short time), with appropriate accuracy and reliability.

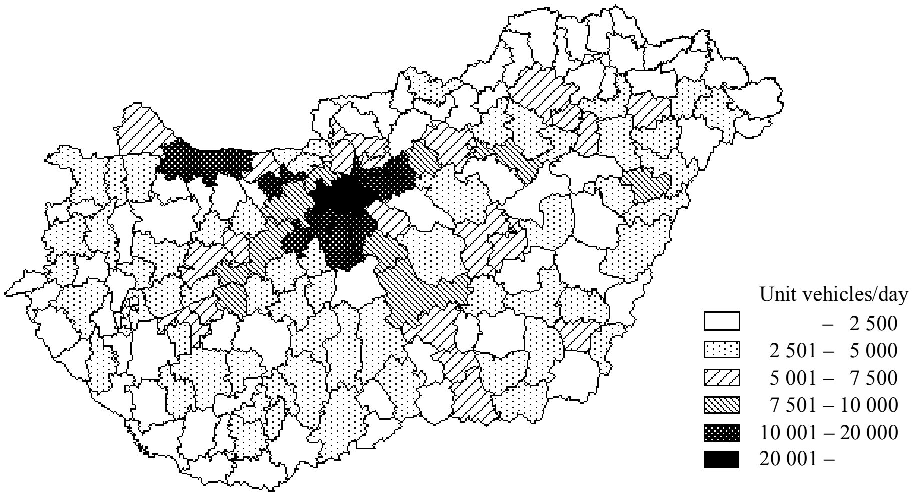
The essence of national cross-section traffic counts is that counts are carried out at a large number of stations, based on samples, distributed all over the year, on 5 different occasions (covering a period from 6 a.m. to 6 p.m.). The averages obtained from these values (g_x) multiplied by intra-day (a_x), day (b_i) and month (c_i) factors describing the regularity of traffic are the most reliable spatial values (with a probability of p=95) of average daily turnover calculated for the year as a whole:

$$ADT = \frac{1}{n} * \sum_{i=1}^{n} g_x * a_x * b_i * c_i,$$

where n is the number of days counted, g_x is traffic observed during a count lasting x hours, a_x is the intra-day factor (traffic counted in a particular part of the day relative to traffic over 24 hours), b_i is the day factor (a number belonging to each day of the week, which changes daily traffic into a monthly average value), and c_i is the month factor (a number belonging to each month of the year, to transform monthly average traffic into annual average traffic).

Figure 1

Average daily traffic calculated for a year as a whole, 2008



Traffic (flow) data for 2004 and 2008 were compared using twelve different potential models. Incomes and resident population were applied as mass factors. The models applied linear (c_1), square (c_2), „e-ad" (c_3) and „e-ad" Box-Cox (c_4), Gaussian (c_5) and log-logistic (c_6) resistance factors (details of potential models: Tóth–Kincses 2007).

Table 1

Dimensions of analysis

Dimension	Notes
Source	Accessibility is calculated and interpreted from the point of view of all people in our investigation, and the many different social groups, as well as the differing travel purposes of different travellers are not distinguished.
Purpose	The purpose to be achieved is quantified by the population and income of the particular micro-region. This "mass" factor (component) quantifying the purpose to be achieved is included in the applied models, without adjustment.
Resistance	The spatial resistance factor in the present case means theoretical accessibility time measured between the centres of micro-regions on public roads, in minutes. The applied resistance factor can be linear, square, exponential, Box-Cox, Gaussian and log-logistic.
Restrictions	When using roads between two micro-regions, restrictions on the particular section are the maximum speed for the type of the road.
Geographic scope	On defining the area of analysis, the territory of Hungary was taken into consideration. Undoubtedly, although accessible targets outside Hungary also affect domestic potentials, external impacts had to be left out of consideration, since road network maps of an appropriate level of detail were available to us only for Hungary.
Mode of transport	The aspects of passenger and freight transport were not distinguished in the analysis.
Spatial level	The basic spatial level of our research is the level of micro-regions (LAU 1 level).
Dynamics	Population, income and public road network on 1 January 2004 and 2008 were taken into account in the research.

The models involved in the investigation were the following potentials (the calculations were also made with models allowing for the agglomeration impact, achieving similar results):

$$C_1 = \frac{p_i}{c_{ii}} + \sum_j \frac{p_j}{c_{ij}} \qquad C_2 = \frac{p_i}{c_{ii}^2} + \sum_j \frac{p_j}{c_{ij}^2}$$

$$C_3 = \frac{p_i}{e^{\beta C_{ii}}} + \sum_j \frac{p_j}{e^{\beta C_{ij}}} \qquad C_4 = \frac{p_i}{e^{\beta\left(\frac{c_{ij}^{\lambda}-1}{\lambda}\right)}} + \sum_j \frac{p_j}{e^{\beta\left(\frac{c_{ij}^{\lambda}-1}{\lambda}\right)}}$$

$$C_5 = \frac{p_i}{\frac{-c_{ii}^2}{w * e^u}} + \sum_j \frac{p_j}{\frac{-c_{ij}^2}{w * e^u}} \qquad C_6 = \frac{p_i}{1+e^{a+b\ln c_{ii}}} + \sum_j \frac{p_j}{1+e^{a+b\ln c_{ij}}},$$

where c_{ij} indicates the time needed to traverse the distance between micro-regions i and j on road, while p_i the relevant social mass of micro-region i, and a, b and w are constants of the particular spatial structure.

In the table which depicts how the potential models fit together, the results of models calculated with income and population are presented, using data from 2004 and 2008. It can be shown that somewhat better fitting can be achieved with income than with population data, although the difference is not significant. On the basis of our micro-regional tests, models applying the log-logistic resistance factor (C_6) can be considered as the best accessibility potential model. It is to be noted, however, that one cannot be sure that using a different territorial division would have led to this result.

Table 2

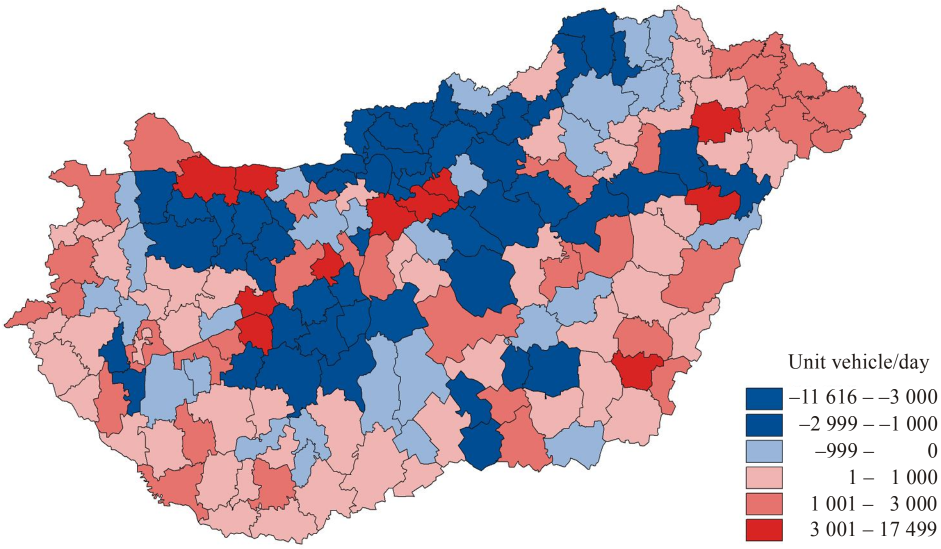
Fitting of models – based on gravitation analogy – to traffic data (R^2)

Denomination	C_1	C_2	C_3	C_4	C_5	C_6
Population, 2004	0.43	0.26	0.55	0.52	0.19	0.63
Income, 2004	0.42	0.24	0.56	0.53	0.18	0.73
Population, 2008	0.45	0.45	0.56	0.52	0.13	0.69
Income, 2008	0.46	0.45	0.58	0.55	0.11	0.72

Regression values – with a few exceptions – indicate moderate correlation. There are no large disparities between the explanatory powers of models. It can be seen that conclusions drawn from potential models need to be treated with caution, for the correlation between them and social space is not strong enough to do so in specific cases. In the following, using the best fitting model, using residuals, it was examined where significant spatial disparities existed between potential space and the flows.

Figure 2

Differences between traffic and trends calculated from potential (C_6)



One can see that in the blue areas of the country traffic is lower than the value expected by the model, while in the area of the capital, close to large towns and next to the border and Lake Balaton, the volume is larger than what could be expected by the potential. This is also logical, as these areas can be highlighted target areas, which has not been integrated into any of the models.

Hence these can be the following steps of improving the potential models so that the models can also become better and better in a mathematical sense, and the conclusions drawn from them can be applicable to the whole social space.

Relation between space and masses, disaggregation of potential

The topology of geometry of the accessibility potential showed that whatever models are used, it is a common feature of them that they measure the impacts of spatial structures, spatial division, of the location of the different spatial domains and of the distribution of masses at a time. The location of a spatial domain is basically defined by geographical position, which is somewhat modified by accessibility (depending on mode of transport). That is, in case of a particular potential value, it cannot be stated whether that results from favourable/unfavourable (settlement, spatial) structure or position, or the location of masses, spatial size or the impact of own mass. In this chapter we aim to disaggregate these impacts, to describe parts as a proportion of total potential values, and to present spatial disparities.

According to our idea on the gravitation field of social masses, therefore, an arbitrary division of space is given (settlement, micro-regional structure etc.), and then a distribution of masses according to this division (masses being either quanta or tokens "assigned" to the particular spatial structure). The value of the potential in a particular point is defined by the sum of these two impacts (internal potential) and that of the impact of own mass and own spatial size (self-potential).

In an arbitrary point of space, by the effect of the potential caused only by the division of space we mean the value which would be generated if masses were equal in all delimited spatial units. The mass distribution impact in an arbitrary point of space is the difference between the value of internal potential and spatial structure potential in the particular point. The impacts of spatial size and own mass can be interpreted in an analogous way in case of self-potentials:

$$U_i^{total} = U_i^{mass\_distribution} + U_i^{spatial\_structure} + U_i^{own\_mass} + U_i^{spatial\_size}$$, where

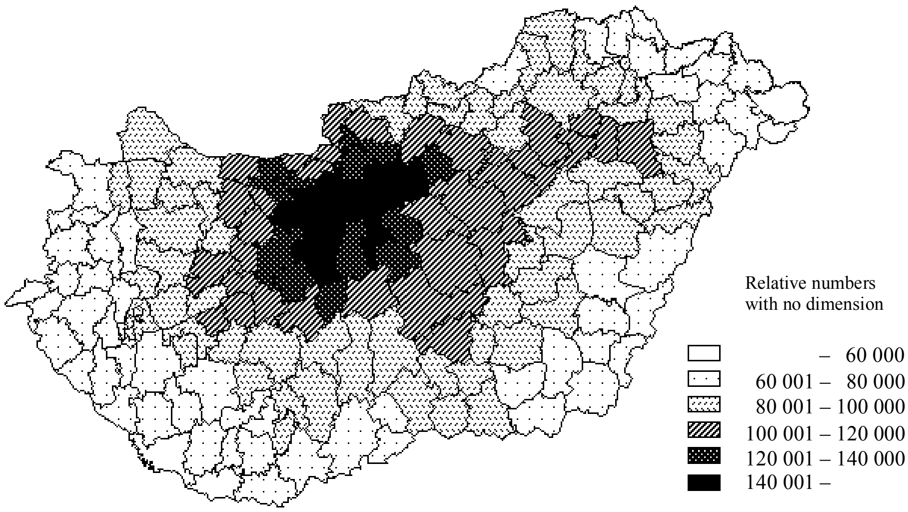
$$U_i^{spatial\_structure} = \sum_j \frac{\frac{\sum_{k=1}^{n} m_k}{n}}{f(d_{ij})}, \quad U_i^{mass\_distribution} = U_i^{internal} - U_i^{spatial\_structure},$$

$$U_i^{spatial\_size} = \frac{\frac{\sum_{i=1}^{n} m_i}{n}}{f(d_{ii})}, \qquad U_i^{own\_mass} = U_i^{self} - U_i^{spatial\_size}.$$

In the next example, the starting point of calculations was the micro-regional data series of resident population in Hungary (1 January 2008). The division of the above potential was made by the model applying linear resistance factor, using distances on public roads. The impacts of masses beyond Hungary and the cross-border effect of internal masses were not taken into account here either.

Figure 3

Potential values of resident population accessibility by micro-regions



The most important result of the split of the potential is that total potential depends on spatial division to the highest extent. This – as already mentioned – depends on the topographical situation of areas on the one hand and on the other, the accessibility factor modifying it, due to which the picture of the role of spatial division does not grow concentrically towards the borders but is somewhat distorted. The proportion of spatial division within total potential is in the 55% and 119% range. With areas along the border – having low total potential in national comparison – primarily determined by this factor, it can be stated that these micro-regions have a disadvantage already owing merely to their location, which cannot really be offset by changes in either accessibility or mass distribution or own mass. Location advantages have their impact mainly in the central part of the country.

The case of mass distribution is completely different. There are areas from where smaller-than-average masses are accessible relative to the total potential of the particular area. This means that the structural location of these micro-regions would result in higher potential values, but the distribution of masses, unfavourable to them, induces a negative effect. Examples include micro-regions with low populations, situated in a bloc along the south-western border of Hungary, or Budapest, the total potential of which is lowered by nearly 7% by the distribution of accessible masses. The effect of mass distribution varies between 43% and –21% at the national level. In essence, the role of mass distribution gradually declines moving away from the micro-regions of the Budapest agglomeration." It is worth to observe that the role of mass distribution is only positive in the Győr micro-region out of regional centres, i.e. the volume of masses available from here is very considerable. However, the rest of the regional centres primarily stand out from the potential space because of their own mass, and the further development of these is substantially hindered by relatively small masses being accessible from them.

Within the self-potential of areas the least important factor as measured against total potential is the area size of the particular area. Its share of total potential ranges between 1% and 6%. The role of own mass is positive in 45 cases from the point of view of the total potential of a particular micro-region, and negative in respect of the rest. The proportion of own masses is between 34% and –4% of the total potential of micro-regions. The most positive shares can be seen for Budapest and major cities, while the negative ones are seen in case of predominantly border-side micro-regions with low populations.

Figure 4

Role of spatial division as a proportion of potential values of resident population accessibility by micro-regions

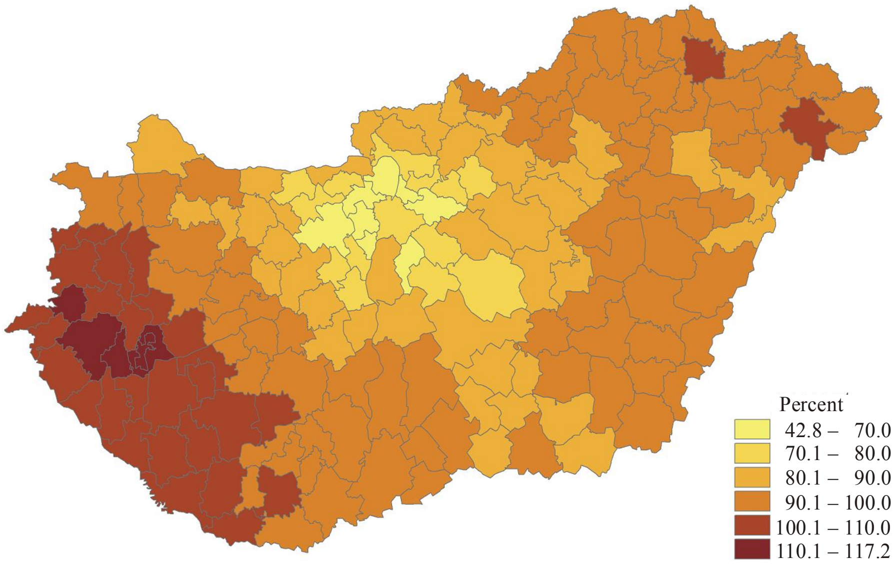


Figure 5

Mass distribution as a proportion of potential values of resident population accessibility by micro-regions

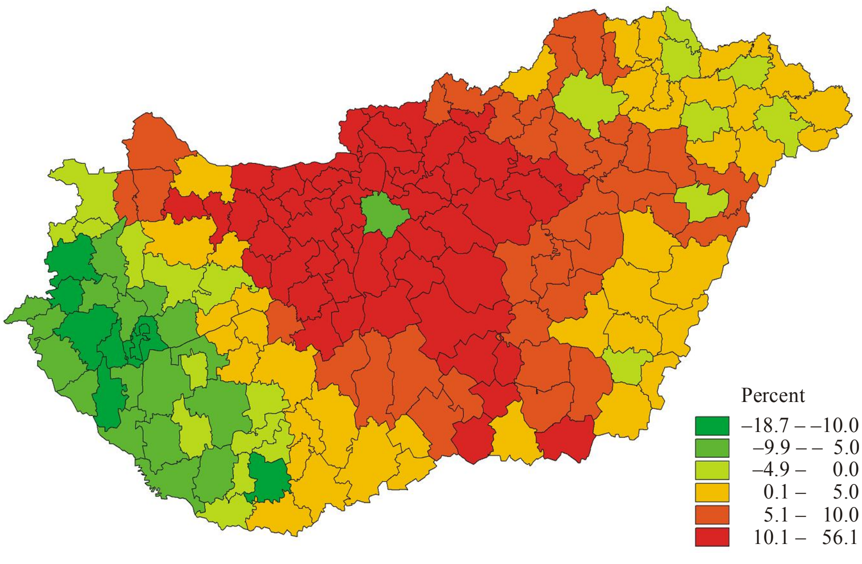


Figure 6

Role of area size as a proportion of potential values of resident population accessibility by micro-regions

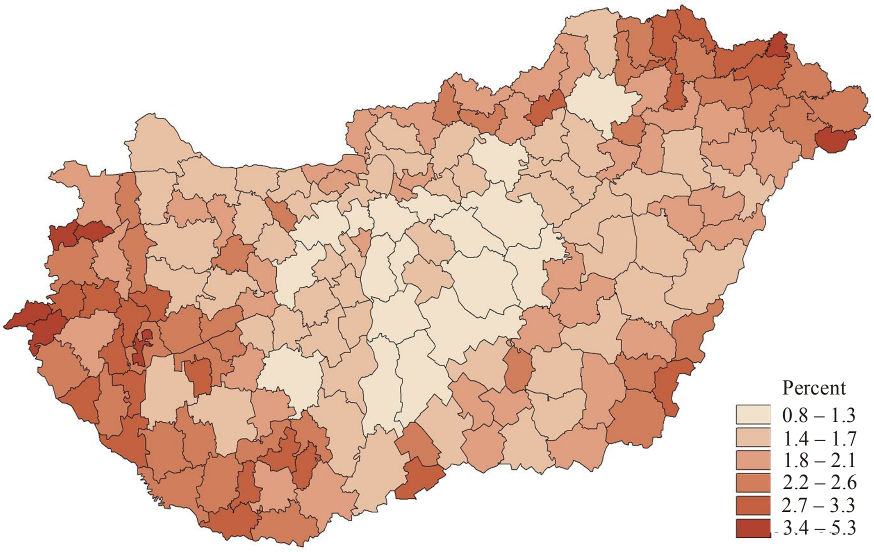
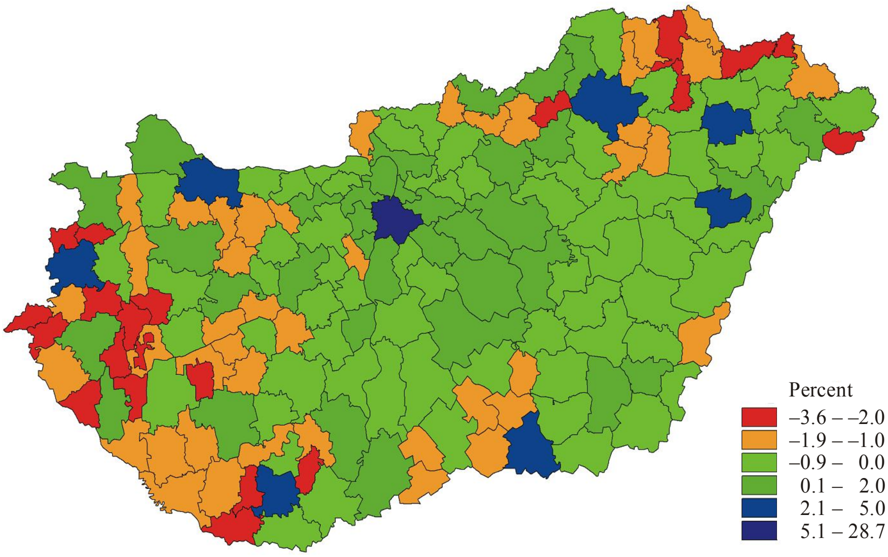


Figure 7

Own mass of areas as a proportion of potential values of resident population accessibility by micro-regions



Summary

By discussing the geometry of accessibility potential models and in support of gravity analogies, we attempted to make a short review of the application and applicability of the models. We aimed to present what may be concluded from different potential models. To answer this question, the values of the models were compared to traffic data. We found that regression values showed moderate correlation. There are no large disparities between the explanatory powers of the models. We stated that the conclusions drawn from the potential models needed to be treated with caution, for the relationship between social space and the models is not strong enough for this in particular cases.

According to the analysis of the potential models, it is not possible to state in case of a particular potential value if it results from favourable/unfavourable (settlement, spatial) structure or position, or the location of masses, area size or the impact of own mass. Therefore we attempted to disaggregate these effects of the potentials by the mathematical method detailed in the study. We pointed out that total potential depended on spatial division to the highest extent, but the impact of mass distribution as well as of own mass may also be important influential factors.

REFERENCES

Baradaran, S. – Ramjerdi, F. (2001): Performance of accessibility measures in Europe. *Journal of Transportation and Statistics* 4 (2/3): 31–48.

Bewley, R. – Fiebig, D. G. (1988): A flexible logistic growth model with applications to telecommunications. *International Journal of Forecasting* 4 (2): 177–192.

Box G. E. P. – Cox D. R. (1964): An Analysis of Transformations. *Journal of the Royal Statistical Society Series B (Methodological)* 26 (2): 211–252.

Bruinsma, F. R. – Rietveld, P. (1998): The Accessibility of European Cities: Theoretical Framework and Comparison of Approaches. *Environment and Planning A* 30 (3): 499–521.

Budó Ágoston (1970): *Kísérleti Fizika I.* Nemzeti Tankönyvkiadó, Budapest.

Carrothers, G. A. P. (1956): An historical review of the gravity and potential concepts of human interaction. Journal American Institute of Planners, 22 (2): 94–102.

Chapelon, L. (1997): *Offre de transport et aménagement du territoire: évaluation spatio-temporelle des projets de modification de l'offre par modélisation multi-échelles des systèmes de transport*. Thèse de doctorat: Aménagement: Tours: Laboratoire du CESA.

Dalvi, M. Q. – Martin, K. M. (1976): The measurement of accessibility: some preliminary result. *Transportation* 5 (1): 17–42.

Davidson, K. B. (1977): Accessibility in transport/land-use modelling and assessment. *Environment and Planning A* 9 (12): 1401–1416.

Dusek Tamás (2001): A területi mozgóátlag. *Területi Statisztika* 41 (3) 215–229.

Fortheringham, A. S. (1982): A new set of spatial-interaction models: the theory of competing destinations. *Environment and Planning A* 15 (1): 15–36.

Frost, M. E. – Spence, N. A. (1995): The rediscovery of accessibility and economic potential: the critical issue of self-potential. *Environment and Planning A* 27 (11): 1833–1848.

Gábos Zoltán (2004): A klasszikus gravitációelméletről. *Fizikai Szemle* 54 (12): 397–401.

Guy, C. M. (1983): The assessment of access to local shopping opportunities: a comparison of accessibility measures. *Environment and Planning B: Planning and Design* 10(2): 219–238.

Handy, S. L. – Niemeier, D. A. (1997): Measuring Accessibility: An Exploration of Issues and alternatives. *Environment and Planning A* 29 (7): 1175–1194.

Hansen, W. G. (1959): How accessibility shapes land use. *Journal of the American Institute of Planners* 25 (2): 73–76.

Hilbers, H. D. – Veroen, E. J. (1993): *Het beoordelen van de bereikbaarheid van lokaties. Definiëring, maatstaven,* toepassingen beleidsimplicaties. INRO-VVG, 09, TNO Inro, Delft.
Houston, C. (1969): Market potential and potential transportation costs: an evaluation of the concepts and their surface patterns in the U.S.S.R.. *Canadian Geographer*, 13 (3): 216–236.
Ingram, D. R. (1971): The Concept of Accessibility: A Search for an Operational Form. *Regional Studies* 5 (2): 101–107.
Lukermann, F. – Porter, P. W. (1960): Gravity and potential models in economic geography. *Annals of the Association of American Geographers* 50 (4): 493–504.
Martin, K. M. – Dalvi, M. Q. (1976): The comparison of accessibility by public and private transport. *Traffic Engineering and Control* 7 (12): 509–513.
Morris, J. M. – Dumble, P. L. – Wigan, M. R. (1978): Accessibility Indicators for Transport Planning. *Transportation Research A* 13 (2): 91–109.
Nemes Nagy József (1998): *Tér a társadalomkutatásban – „Ember–Település–Régió”*. Hilscher Rezső Szociálpolitikai Egyesület, Budapest.
Nemes Nagy József (ed.) (2005)*: Regionális elemzési módszerek*. Regionális Tudományi Tanulmányok, 11., ELTE, Budapest.
Oláh-Gál Róbert (2008): Bolyai János egyik leghosszabb fizika tárgyú kéziratáról. Fizikai Szemle 58 (9): 302.
Schürmann, C. – Spiekermann, K. – Wegener, M. (1997): *Accessibility Indicators*. Berichte aus dem Institut für Raumplanung, 39., Institute of Spatial Planning, University of Dortmund, Dortmund.
Simma, A. – M. Vrtic – K. W. Axhausen (2001): *Interactions of travel behaviour, accessibility and personal characteristics: The Case of Upper Austria*. Presentation, European Transport Conference, September 2001, Cambridge.
Song, S. (1996): Some Tests of Alternative Accessibility Measures: A Population Density Approach. *Land Economics* 72 (4): 474–482.
Tóth Géza – Kincses Áron (2007): Elérhetőségi modellek. Tér és Társadalom 21 (3): 51–87.
Wilson, A. G. (1971): A family of spatial interaction models, and associated developments. *Environment and Planning A* 3 (1): 1–32.

Keywords: regional analysis methods, potential model, spatial relations.

Abstract

With the help of the description of geometry of potential methods and the gravity models, we tried to briefly explain these models' application and applicability. Our aim was to show what may be concluded from different potential models. To answer this question, traffic data were compared to the values predicted by the models. We found that the values of the regression showed a moderately strong correlation. We could determine that the conclusions drawn from the potential models should be treated with reservations.

According to the analysis of the potential models it is not possible to determine in case of a potential value whether it results from positive/negative (municipal, spatial) structure or position, or the location of masses, area size or the impact of its own mass. Thus, the mathematical method described in this study tried to separate the potentials. We found that accessibility potential was dependent on spatial subdivision to the highest degree, but important factors were the effect of mass distribution as well as of own mass.

REGIONAL STATISTICS, 2012, VOL. 2: 90–107 DOI: 10.15196/RS02107

JÁNOS PÉNZES

Changes in the Spatial Income Structure of North-eastern Hungary After the Change of Regime

Introduction

The market economic transition after the change of regime significantly transformed the spatial structure of development in Hungary. In the first half of the 1990s, the meltdown of the socialist economy and the deriving crisis phenomena, appearing to a different extent by region, basically determined the spatial socio-economic trends. At the same time, it demonstrated which regions were able to weather the market economic transition and were able to attract new investments due to their favourable geographical and market position. North-eastern Hungary, covering the three counties in Northern Great Plain and the three in Northern Hungary, suffered badly as a result of the socio-economic changes after the transition, but the transformation of the spatial structure of development was not at all homogenous.

The present analysis seeks an answer to the question on how micro-level changes in income can be detected in the observed area, and within the examined spatial structure, which settlements were the relative winners or losers of the processes after the transition. Another aim of the study is to throw light on spatial relationships, and to formulate general conclusions, especially about the development trends in the underdeveloped, peripheral as well as in the, even in national terms, developed regions of the observed area. The basic method of the analysis was the application of the potential model, in the course of which, the effective application of the method was also examined.

Transformation of the spatial income structure in Hungary

Instead of the complex socio-economic and infrastructural approach of development, the study only concentrated on one single indicator, the taxable personal income, hereinafter referred to only as income for the sake of simplicity. Income data on a settlement level are available in the databases of the Ministry of Finance (National Tax and Customs Administration) and the Hungarian Central Statistical Office from the introduction of personal income tax in 1988. Income is a complex indicator used widely in regional studies, as it is one of the most important indicators of living conditions and prospects of the population. In addition, due to personal income tax, it is a significant measure in the redistribution of GDP, and an important income source for the general and local governments (Ruttkay 1997). However, its significance has decreased during recent years. At the same time, it has been a continuously available indicator with unchanged content since 1988. Due to its complexity, it has significant explanatory authority in the examination of the spatial structure of development and its trends.

Income as an indicator also has some limits, arising first from the fact that because of taxation rules, it involves only some 40% of the total net available cash income of the population coming from macro-data (Kiss 2007). The total income volume of the population containing government transfers also shows a much more balanced picture than the examined concept of income, which covers mostly wages and salaries. Thus, the data of taxable incomes correlate well with employment indicators; they also show a strong statistical relationship with numerous other indicators, particularly those referring to financial development (Jakobi 1999, Adler–Skultéty 2002, Kiss 2007, Pénzes 2010). This is why income formed the basis for many examinations of spatial structures and regional inequalities (e.g. Kovács 1993, Bódi–Obádovics–Mokos 1999, Major–Nemes Nagy 1999, Németh–Kiss 2007).

In line with the methodology of HCSO, the value of income per inhabitant figuring as a unit indicator in the study was computed by the mid-year permanent population data, and was called per capita income for the sake of simplicity.

The socio-economic transformation after the change of regime had an impact on the regional processes, partly generating new regional inequalities and partly strengthening earlier phenomena. Although changes are differently assessed by researchers, they similarly summarize the basic trends (Cséfalvay–Nikodémus 1991, Rechnitzer 1993, Enyedi 1996, Kozma 1998, Nemes Nagy 1998, Beluszky–Győri 1999, Faluvégi 2000):

– the dominant development of Budapest, the increase of the difference between the capital and the rural areas;

– the emergence of the 'West-East declination';

– the differentiation of development of micro-regions and settlements.

The most characteristic features of the spatial structure were also reflected in the per capita incomes. The changes in the relative income level compared to the rural average value (in the study, it is the national average without Budapest) spectacularly illustrated that the regions were partitioned into two groups in the period between 1988 and 2008 (Figure 1).

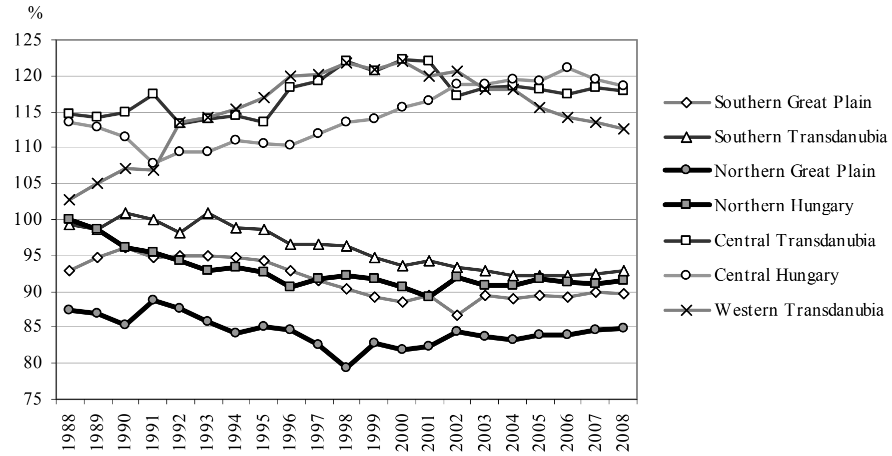
The income level in Central Hungary, Central Transdanubia, and Western Hungary has been continuously higher than the rural average since the change of regime despite the tendentious decline in the latter region since 2004. In all probability, this process will strengthen further in the light of income statistics to be published later, since the car industry and the connected sub-contractor network were significantly affected by the world economic recession, which hit the economy in Western Transdanubia hard.

The per capita income in Southern Transdanubia, Northern Hungary, Southern Great Plain and especially in the Northern Great Plain was considerably below the rural average value. The economic recession of larger extent after the change of regime is clearly shown by the regional income trends. In case of the Northern Great Plain region, it is not the declining trend that is the most striking, but the continuous and marked lagging behind the values in the other regions and the aggregated average.

The regional development trends show not only the partition of the regions into two groups, but also the fact that within the groups of regions with above average and below average incomes, the differences decreased, and so-called 'convergence clubs' evolved (Major 2001, Lőcsei 2010).

Figure 1

Income level of regions per residents compared to the rural average without Budapest between 1988and 2008



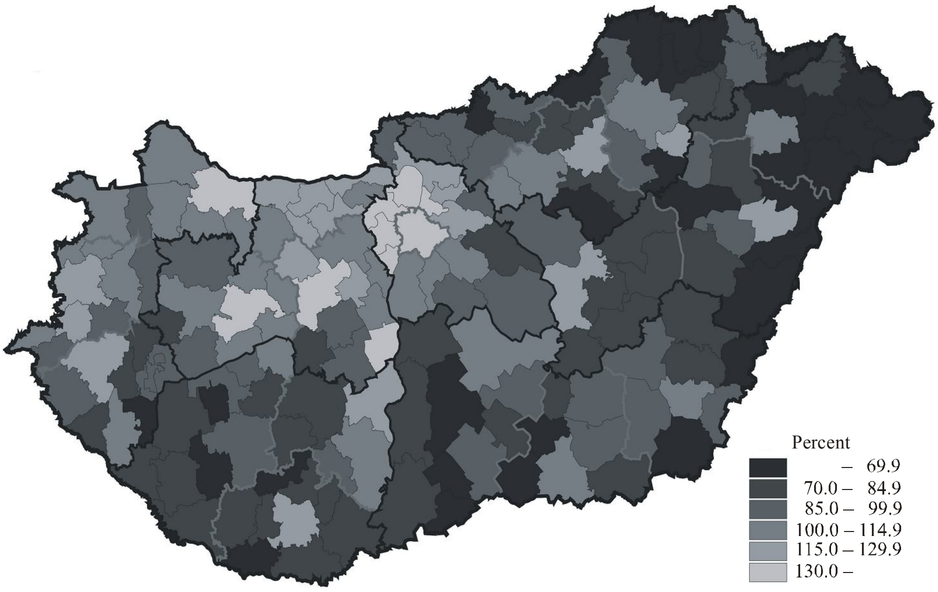
Source: Own calculation based on the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

The regional differences in income level suggest the different extent of income inequalities within regions. According to the Williamson-hypothesis basically proved in the case of Hungary, higher income levels are connected with smaller internal income inequalities (Nemes Nagy 2005, Németh–Kiss 2007, Pénzes 2010). In national terms, income inequalities showed a convergence after the turn of the millennium, but divergences could be observed as well within smaller regional units (Pénzes 2007). However, the main emphasis of this study is not on examining inequality trends but on analysing the changes in the spatial income structure and their background. It would have been possible to analyse the trends of income levels of regional units (Dusek 2005a), but, instead the focus has been on examining the incomes of 174 micro-regions compared to the rural average by detecting the changes between 1988 and 2008.

In Figure 2, the Balassagyarmat–Békéscsaba line, mentioned often in scientific literature is well demonstrated; this can be interpreted as the spatial borderline between development and lagging behind. Compared to the rural average, low income levels were characteristic of many micro-regions in Southern Transdanubia and the Southern Great Plain. In 2008, the lowest value of the country, barely more than half of the rural average, appeared in the Bodrogköz micro-region followed by the values of the Csenger, Baktalórántháza, Sarkad, Fehérgyarmat, and Abaúj–Hegyköz micro-regions. The highest value (160% of the rural average) was observed in the Budaörs micro-region preceding Budapest, and the Székesfehérvár, Dunakeszi, Dunaújváros, and Pilisvörösvár micro-regions. Characteristically, the values in the county seats, as well as those of the micro-regions in the agglomeration of Budapest, were considerably above the average. The majority of micro-regions in north-eastern Hungary were at the bottom of the income ranking.

Figure 2

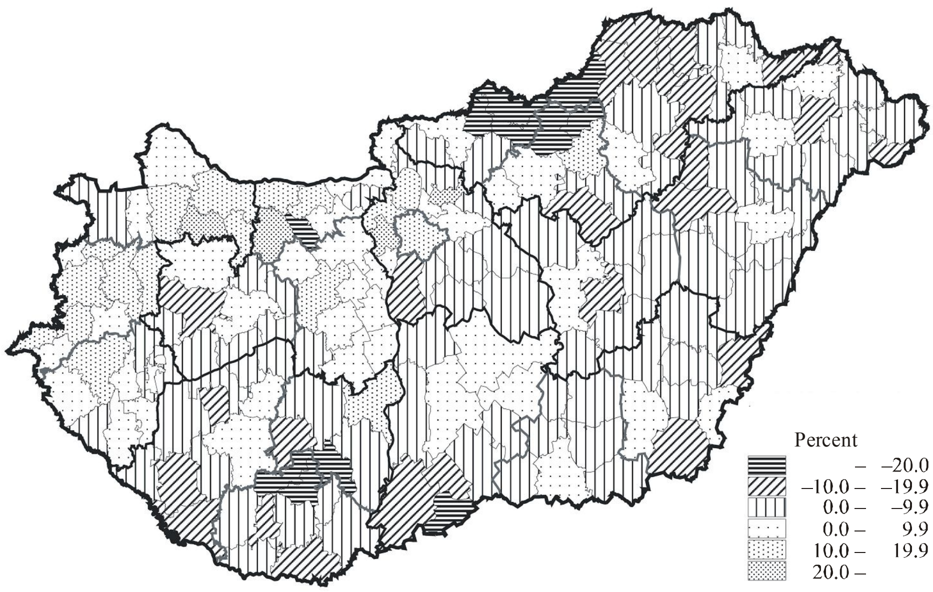
Per capita income in micro-regions compared to the rural average in 2008



Source: Own calculation based on the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

Figure 3

Changes in the per capita income level in micro-regions compared to the rural average between 1988 and 2008



Source: Own calculation based on the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

The examination of the dynamics of income data between 1988 and 2008, pointed out that the largest decline (more than 20%) in income level appeared in the former heavy industrial and mining districts of Southern Transdanubia (e.g. the Bonyhád, Komló, and Sásd micro-regions), and Northern Hungary (e.g. the Bátonyterenye, Bélapátfalva, Ózd, Pétervására, Salgótarján, and Szécsény micro-regions), afflicted by the structural crisis (Figure 3). However, there were some examples for the decline in employment and relative income caused by the structural crisis in the northern part of Transdanubia as well (in the Ajka and Oroszlány micro-regions). The northern Hungarian region was a area of important, mainly industrial developments in the socialist era, due to which, income inequalities on larger regional level diminished (Beluszky 1976). At the same time, these developments became the source of considerable micro-level inequalities, a large part of which deepened following the change of regime. The reason for this was that branches that proved to be viable after the market economic transition could weather the changes, while heavy industry also mentioned previously as well as agricultural cooperatives together with their auxiliary branches characteristic in the Great Plain and oriented mainly to the eastern markets fell apart and ceased to exist. Foreign investments, partly with privatization and Greenfield investments, appeared only to a low extent. Even these investments were concentrated in the surroundings of Budapest, in larger towns and in the vicinity of already existing motorways. At the same time, nothing could hinder the decline of regions without built-up physical infrastructure and with moderate human resources, which were often also underdeveloped earlier (Kanalas–Kiss 2006). One of the objectives of my study is to highlight the differentiated development trends of regions/settlements by a more detailed analysis of income trends in the examined area.

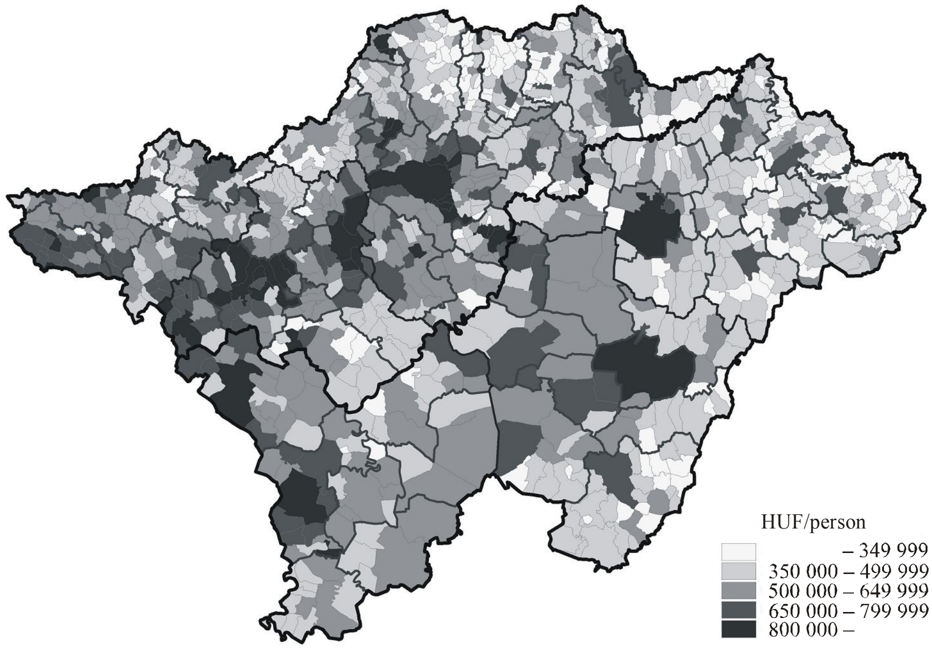
In micro-regions of north-western Hungary and the agglomeration of Budapest, except the southern part of Pest county, a relative income increase was characteristic. North-western regions, which also had a more developed infrastructure at the time of the transition and traditions in manufacturing (mainly in manufacture of machinery), and significantly, were closer to the central areas of Europe, became the main target area of foreign direct investments. In the neighbourhood of Budapest, in addition to the inward migration of wealthier individuals, the strengthening of logistic and trade activities, as well as that of production established close to the centre (e.g. manufacture of pharmaceuticals) caused the considerable rise in income level. Furthermore, micro-regions with tourism assets (e.g. Gárdony, Őriszentpéter and Tata), favourable location (e.g. Tét), and those with a considerable number of workforces commuting out (e.g. Kisbér, Pápa and Tét), were also in a better position.

Income trends in north-eastern Hungary

In the overwhelming majority of micro-regions in north-eastern Hungary, especially in the area having had significant socialist industry (even in the Tiszaújváros micro-region having basically successfully weathered the transition), and the region near the border, the relative income level declined. In the north-eastern part of the country, the income position deteriorated further in most of the traditionally underdeveloped regions after 1989.

Figure 4

Per capita income in settlements of north-eastern Hungary in 2008



Source: Own calculation based on the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

A considerable (16.3%) growth in the relative income was characteristic of the Eger micro-region. There was a perceptible increase of about 5% in the Nyíregyháza, Gyöngyös, and Debrecen micro-regions, while a slightly strengthening position was recorded in the Szolnok, Sárospatak, Kisvárda, Balassagyarmat, Mezőkövesd, and Hatvan micro-regions.

There were marked differences in the per capita income of settlements according to data of 2008 (Figure 4). Incomes were characteristically higher in county seats and settlements in their surroundings. Tiszaújváros (former Leninváros) was in the forefront of the income ranking in the country before the change of regime and for many years afterwards as well (Nemes Nagy 2001). However, after the turn of the millennium it was preceded by some hamlets in Cserehát (e.g. Teresztenye and Tornakápolna), where the development, considered 'isolated colouring factors', could be attributed only to the appearance of one or two taxpayers. It is striking how large a concentration was shown by settlements in the inner areas of Cserehát, where income level is the lowest in the country (among them, e.g. Csenyéte, Gagyapáti, Rakaca, Szakácsi, Tornabarakony were the bottom-ranked). The situation in Cserehát, which seems to be extremely contradictory, can be attributed to the settlement structure separated into small parts, where acute social problems (ageing, increasing proportion of Roma population, unemployment) as well as the appearance of some taxpayers had a more marked effect than in regions of settlements with more inhabitants (Pénzes 2010). Despite this, the area is on the whole one of the most characteristic peripheral regions of Hungary.

The income level definitely reflected settlement categories by population number in each regional frame (Table 1). The basic trend observed was that the larger the settlement, the higher the average income value. The changes between the two dates similarly reflected the relationship with the settlement size. Although the relationship and the trend of change were not fully unambiguous, the polarization by settlement size took shape between the two dates. In the regional income trends of Northern Great Plain, the polarization was a determinant characteristic after the change of regime, but its decreasing trend, in contradiction to the changes in inequalities after the turn of the millennium, did not unambiguously appear (Pénzes 2010). At the same time, when summing up the results of settlements in the two regions, the statistical relationship found between population number and per capita income was quite moderate on both dates.

Table 1

Proportion of per capita incomes of settlements compared to the rural average in Hungary by population categories, 1988–2008

Population size	Northern Great Plain			Northern Hungary			Hungary		
	1988	2008	change,	1988	2008	change,	1988	2008	change,
	percent		percentage point	percent		percentage point	percent		percentage point
– 499	52.0	48.4	–3.6	72.3	60.5	–11.8	73.5	67.3	–6.2
500 – 999	63.7	53.2	–10.5	83.5	66.3	–17.2	79.9	73.9	–5.9
1 000 – 1 999	67.1	57.6	–9.5	86.9	72.9	–14.0	84.2	80.2	–4.0
2 000 – 4 999	71.5	65.0	–6.5	89.6	81.5	–8.1	86.5	85.7	–0.9
5 000 – 9 999	80.6	73.9	–6.7	95.3	92.7	–2.7	93.5	92.0	–1.5
10 000 – 49 999	89.2	87.5	–1.8	110.9	105.9	–5.1	108.0	108.6	+0.6
50 000 –	114.2	121.0	+6.8	119.0	120.2	+1.2	120.4	126.0	+5.6
Budapest	–	–	–	–	–	–	138.6	155.2	+16.6
Rural average	*87.4*	*85.0*	*–2.5*	*100.1*	*91.4*	*–8.7*	*100.0*	*100.0*	*0.0*

Source: Own calculation based on the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

The present study intended to highlight the specific features of spatial income structure in north-eastern Hungary and the changes during the past two decades with the help of the potential model.

Characteristics of the potential model

The potential model, being a regional analysis method of physical analogy similarly to the gravitation model, is the basic method of measuring spatial interactions and examining spatial structures (Dusek 2005b). The first person who applied this method in geography was John Quincy Stewart (Stewart 1942), while its introduction in analyses in Hungary was connected to László Bene and Kálmán Tekse (Bene–Tekse 1966). Antal Papp ranked the settlements in the northern part of the Great Plain having a central function with a potential value calculated on the basis of a complex indicator (Papp 1978).

In the Hungarian scientific literature, it is an examination method used not too frequently, but generally with the help of which spatial population-economic structures were analysed (Nagy 2004, Tagai 2004, Molnár 2008, Nemes Nagy–Tagai 2009), and which was applied in case of predictive impact assessments (Tóth 2005a, Tóth 2005b, Tagai–Pénzes–Molnár 2008).

The essence of the method is that it shows so the ‘spatial strength’ of socio-economic phenomena that it takes into account in the calculation masses belonging to spatial units (the value of the observed phenomenon) and the distance between them. The results of the potential calculation are data without dimension; however, on the maps of potential difference illustrating changes, the changes compared to the values of the previous period can be expressed in percentages.

When applying the potential model, not only the own strength of spatial units, the so-called self-potential value, can be expressed with the inner potential value, but the interactive effects of volumes are taken into account as well. On the other hand, the outer potential value can be demonstrated by taking into account the volumes outside the area observed.

Accordingly, the total potential value is the sum of these three results:

Total potential = self-potential + inner potential + outer potential.

$$(1) \quad P_{saját}(A_i) = \frac{M_i}{d_{ii}^b}$$

The *self-potential* of point A_i can be computed by the *(1)* formula, where M_i is the own point mass (in this study, the total value of income), d_{ii} is the distance data associated with the spatial unit (in the simplest case the radius of a circle of the same size as the area), and *b* is the exponent.

When computing the *inner potential* of A_i *(2),* the sum of effects of other involved spatial units on A_i must be taken into account. The size of the effect depends on the other point masses and their distance from the given spatial unit. The higher its value, the larger the mass of the spatial unit and the nearer it is located in space. The square of the distance *(d_{ij})* is most frequently in the denominator.

The computation of the *outer potential* of A_i *(3)* is practically the same as that of inner potential, but here the effects of spatial units outside the observed area are taken into account. The description of the method can be found in the references (Dusek 2005b).

$$(2) \quad P_{belső}(A_i) = \sum_{j=1}^{n} \frac{M_j}{d_{ij}^b} \quad , \qquad (3) \quad \mathrm{P}_{\mathrm{külső}}(\mathrm{A_i}) = \sum_{\mathrm{k}=1}^{\mathrm{n}} \frac{\mathrm{M_k}}{\mathrm{d_{ik}^b}} .$$

Results of the potential calculation referring to north-eastern Hungary

In the course of the potential calculation for the six counties involved in the survey, income values belonging to the settlement levels were taken as a basis in order to outline a detailed picture of the spatial structure. The survey included not only a static analysis, but, by comparing the conditions in 1988 with those in 2008, a dynamic one as well. However, in the course of calculation, it was a problem that the number of settlements grew from 963 to 999 between 1988 and 2008, since numerous settlements and towns became independent in respect of public administration.

In case of potential models, the results may be strongly altered by the fact that the number of mass points, taken into account at different times, changes. As in 1988, the data of numerous settlements becoming independent later were recorded only aggregated, it would also have been reasonable to aggregate accordingly the values of 2008. However, the data of numerous settlements would not have been recorded independently this way; e.g. the data of one of the most typical suburban settlements, Bocskaikert, would have been included in the data of the former Hajdúhadháztéglás. Thus, the solution chosen was that data of settlements having become independent before the middle of the time interval, i.e. before 1998, were recorded separately; however, in their cases a data correction had to be performed. In order to avoid bias, the earliest income values of settlements published independently were converted into the real value in 1988, by decreasing the absolute income volume with the inflation.

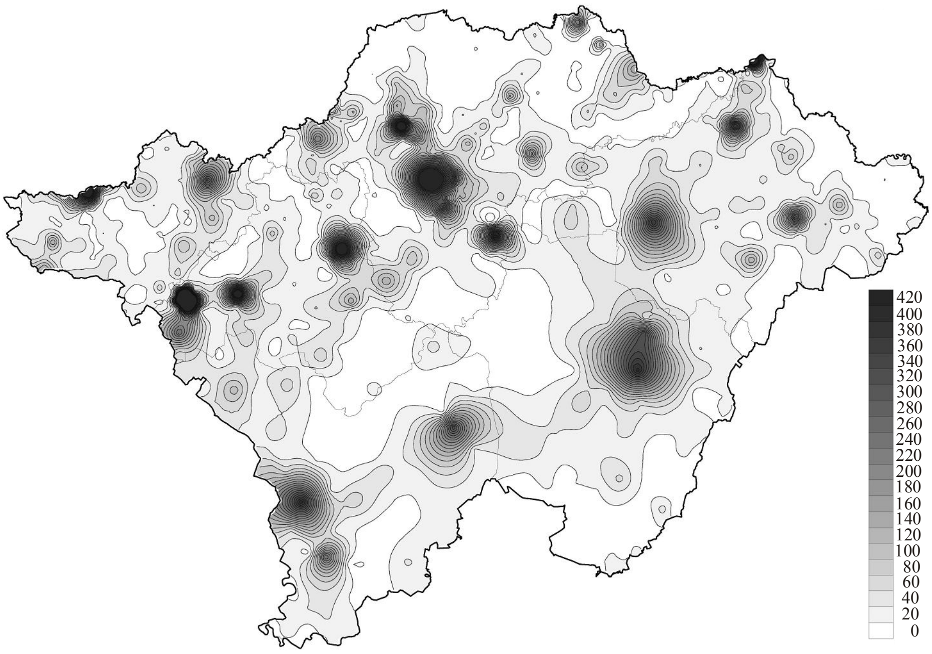
At the same time, settlements having become independent after 1998 (Berente, Farkaslyuk, Gibárt, Ipolyszög, Kerekharaszt, Nagykeresztúr, Pálosvörösmart, Rákóczibánya, Somoskőújfalu, Szarvaskő, Szorgalmatos, Tiszaszőlős), appeared only aggregated in the potential calculation.

In order to eliminate the distortion of the considerable inflation between 1988 and 2008, absolute income volumes were expressed as a percentage of the total average value of the two observed regions on both dates. The relative spatial income structure within the observed spatial system of relationships were also examined, but this did not hinder drawing conclusions. Potential values would have obviously changed if the values of north-eastern Hungary had been compared with the national average of income.

In the calculation, distance data are the accessibility of settlements from each other on public road expressed in minutes; I used their squares when applying the formula. An argument for the chosen distance concept is that the distance on public roads shows much better the system of relationships between settlements than the straight-line distance. The accessibility expressed in time meant the shortest time necessary for reaching a settlement, which makes a difference between the maximum speeds on dual carriageways and on minor roads. In order to eliminate the effects deriving from the development of the public road network and the modification of the Highway Code, I took into account the same times of accessibility for both observed dates. (The effect of the change in time of accessibility on spatial structure could be the topic of a separate study similarly to the analyses of Géza Tóth in the references.)

Figure 5

Self-potential of income in north-eastern Hungary in 2008



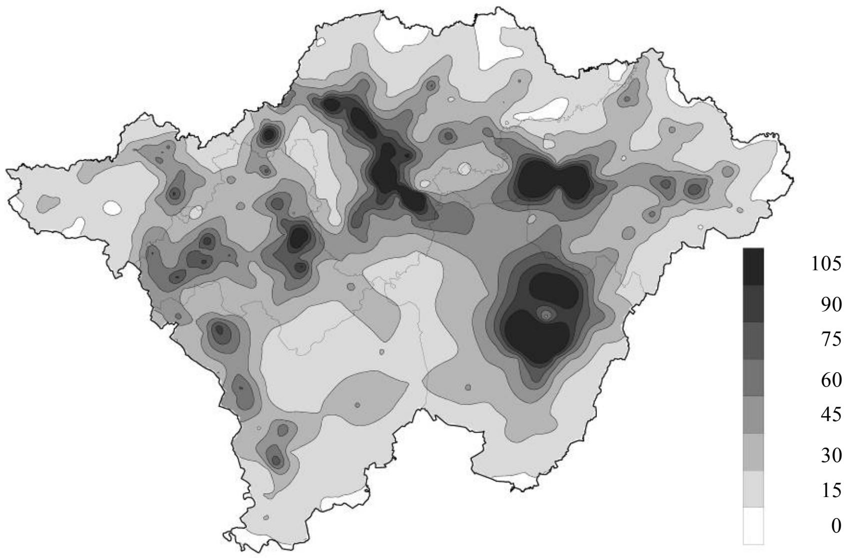
Source: Edited by Gábor Németh based on own calculation from the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

Note: Calculated for settlement level in Figures *5–10*.

According to the practice in the Hungarian and foreign methodology, when calculating *self-potential*, the distance is one third of the time of accessibility to the radius of a circle of the same size as the public administration area of settlements with a speed of 50 km/hour, so that it becomes comparable with the other potential components (Tagai 2007). The value of self-potential reflected the 'spatial strength' calculated on the basis of income, the extent of which depended on the income volume and the public administration area of the settlement (Figure 5). In the light of this, it is not surprising that the smallest settlement of the observed area, which has at the same time relatively high income volume, Petőfibánya, represented the maximum value followed by Záhony, Miskolc, Kazincbarcika, Eger, Gyöngyös, Balassagyarmat, and Tiszaújváros in 2008. Hamlets in Cserehát and Hegyköz were the last in the ranking in both of the observed years. Although there were some bias on the maps interpolated with the help of the Surfer program, the pattern taking shape suggested the regions with higher and lower 'spatial strength' of incomes in the area. Mountainous and border areas far from larger centres, as well as areas of lower settlement density in the Great Plain, especially the Middle Tisza District and the Berettyó–Körös District, appeared with a very low self-potential.

Figure 6

Inner potential of income in north-eastern Hungary in 2008



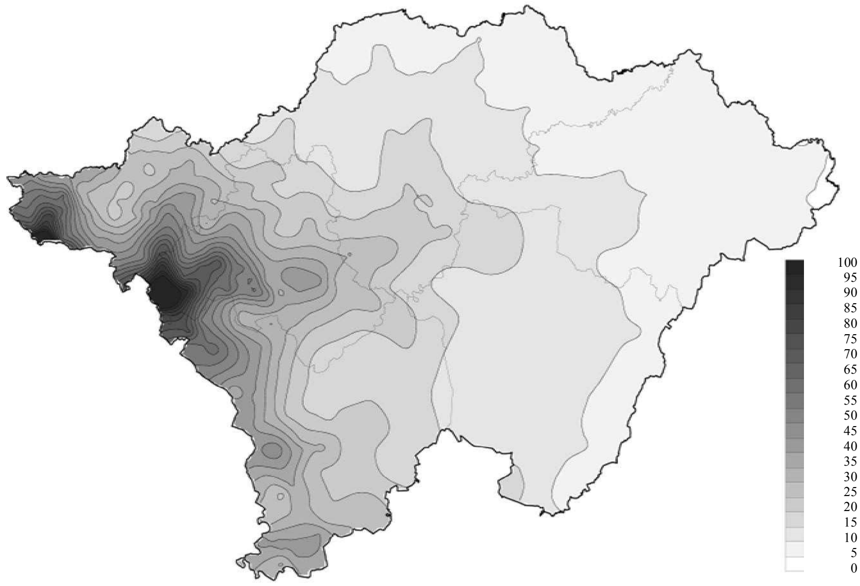
Source: Edited by Gábor Németh based on own calculation from the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

When calculating *inner potential*, I expressed the effect of the altogether 987 settlements (after correction) of the two regions on each other with the help of the method mentioned. As a starting point, the accessibility of each settlement from all other settlements in north-eastern Hungary had to be computed. The nearly one million data were extracted by the Network Analyst application of the ArcView GIS 3.3 program.[1] Consequently, the inner potential takes into account the aggregated effect of masses in the observed area in inverse proportion to the distance between them. The highest values appeared in settlements near high income volumes, first of all near county seats (Figure 6). The highest value of inner potential occurred however in Sajószöged located only a ‘stone’s throw away’ from Tiszaújváros, followed by Nyírpazony, Mikepércs, Nyírtelek, Nagytálya, Mályi, Sáránd, and Ebes in 2008. These settlements are at the same time the main target areas of people migrating out the neighbouring towns, i.e. the most significant rural areas of suburbanization (Lőcsei 2004). Inner potential was the lowest in the peripheral settlements in Szatmár, Bodrogköz, Zemplén, and Cserehát which are the furthest from the centres.

1 Special thanks to Dr Géza Tóth and Dr Szilárd Szabó for their help.

Figure 7

Outer potential of income in north-eastern Hungary in 2008



Source: Edited by Gábor Németh based on own calculation from the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

When calculating *outer potential*, I took into account the more considerable masses outside the two regions; first of all Budapest, then the county seats of the neighbouring Bács-Kiskun, Békés, Csongrád counties (Békéscsaba, Kecskemét, Szeged); additionally, a total of 95 towns of the connecting counties were included in the survey. Unfortunately, settlements beyond the borders could not be involved in the survey due to the lack of data, and this spectacularly determined the spatial structure of the outer potential. One of our earlier studies taking into consideration employment centres (Tagai–Pénzes–Molnár 2008), intended to make up for this hiatus.

Because of the above mentioned factor, the value of outer potential was typically determined by the distance from Budapest, which became spectacularly longer along the M3 motorway in eastern direction (Figure 7). In the observed time interval, this effect extended in compliance with the further construction of the motorway, although I took into account only the times of accessibility calculated for the latter date. The effect of the capital was the strongest in the south western part of Nógrád and the western part of Heves counties (in the vicinity of Rétság, Hatvan and Gyöngyös), which is also demonstrated, among others, in the number of people migrating to Budapest (Sütő 2008). In both of the observed years, the value was the highest in Szendehely. However, the favourable effect of the improving accessibility of Budapest ensured by motorways is not unambiguous in case of each region (Németh 2009). The outer potential field was influenced to a moderate degree by the values in Kecskemét and Szeged as well. The results were the lowest in those settlements of the Fehérgyarmati micro-region, which are located furthest from Budapest.

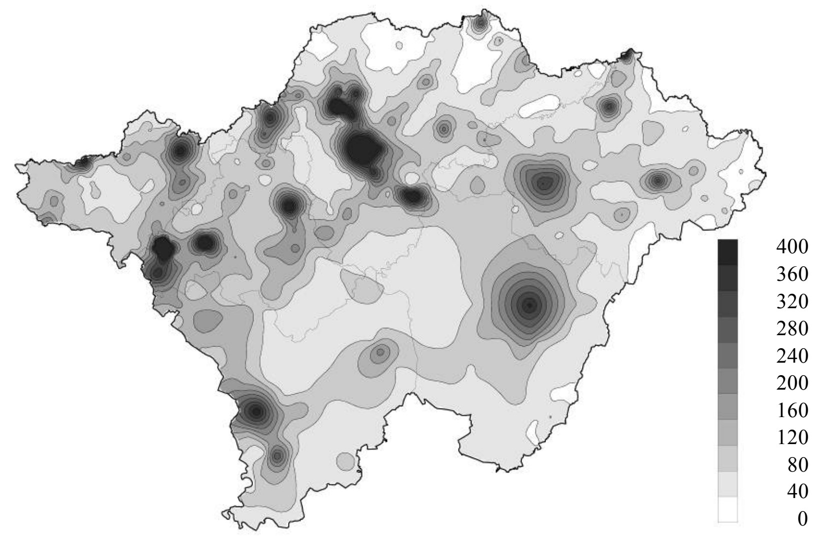
Based on own studies referred to previously, as well as on the results of studies of other authors (Kovács 1990, Süli-Zakar 1992, Baranyi 2007), it can be assumed that if mass values were also available from areas beyond the borders, the external effects would be the most significant in the region of Bihar near Nagyvárad, Szatmár, and Bereg, in the vicinity of Szatmárnémeti and Beregszász, as well as in the area of northern Cserehát located not far from Kassa (Košice).

The *total potential value* can be calculated by adding up the three components (Figures 8 and 9). As a result, Petőfibánya achieved first place with an outstandingly high value, largely due to its self-potential and the nearness of Hatvan and Budapest; it was followed by Gyöngyös, Kazincbarcika, Miskolc, Záhony, Eger, Balassagyarmat, Tiszaújváros, Mályi, Felsőzsolca, Bocskaikert, and Szolnok. Generally speaking, county seats, the towns having an important role in employment and economy, and the settlements in their vicinity affected by suburbanization had the highest potential indicator.

The foregoing findings are reflected in the lowest total potential value as well: small settlements in Szatmár, Cserehát, and Zemplén were the last in the ranking in 2008. Kishódos, Nagyhuta, Garbolc, Debréte, Tornabarakony, Pamlény, and Nagyhódos were among the bottom-ranked settlements in both of the observed years.

Figure 8

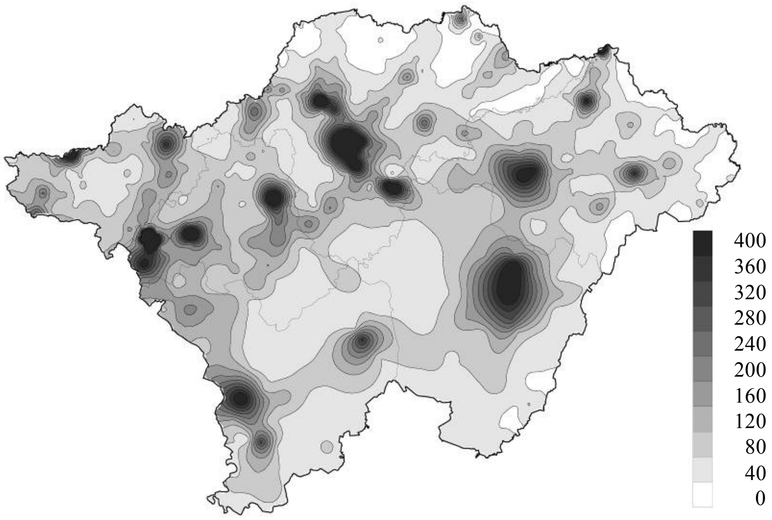
Total potential of income in north-eastern Hungary in 1988



Source: Edited by Gábor Németh based on own calculation from the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

Figure 9

Total potential of income in north-eastern Hungary in 2008



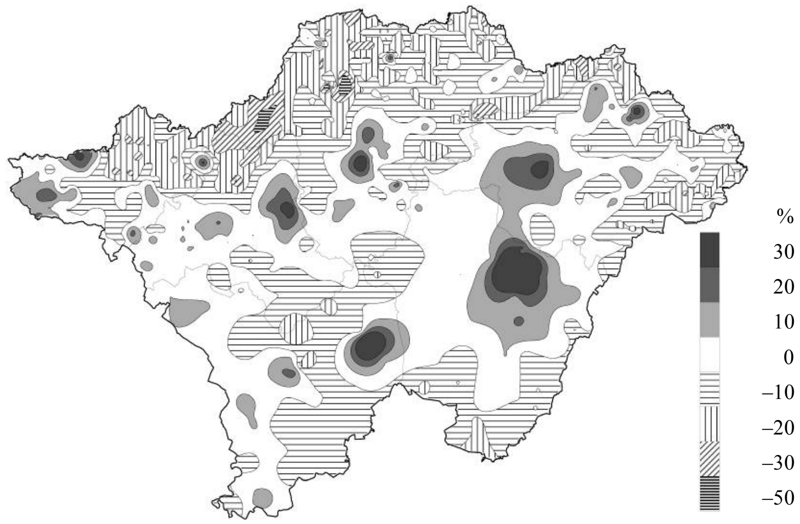
Source: Edited by Gábor Németh based on own calculation from the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

Not only the total potential values on the two dates were worth examining, but by making use of the possibilities given by the methodology, the changes having occurred could also be detected. By expressing the decrease or increase between 1988 and 2008 in percentage terms, settlements and regions with relative advantage or disadvantage can be outlined. If income volumes were compared to the national values instead of the averages of the observed area, there would hardly be any positive values. In this way, however, areas affected relatively favourably or unfavourably during the market economic transition within the observed spatial system of relationships can be determined (Figure 10). Some 60% of settlements had a decreasing total potential value.

Among settlements showing the largest increase in total potential value, many were located near large towns. The most considerable change was recorded in Bocskaikert near Debrecen; there was a quite large increase in Ostoros, Andornaktálya, and Nagytálya in the neighbourhood of Eger, and in Mályi, Kistokaj, and Arnót near Miskolc as well as in Nyírpazony, and Nyírtura connected to Nyíregyháza. But, for example, inhabitants with high income, who migrated out of Tiszaújváros, caused the increase in the values of Sajóörös and Sajószöged as well (Pénzes 2004, Ekéné Zamárdi–Pénzes 2006). At the same time, it is remarkable that some settlements having become independent in respect of public administration during the 1990s also had high values (e.g. Berekfürdő, Tiszaszalka, Mátranovák, Nagyvarsány, Téglás, but Bocskaikert having the highest value can be also classified here). Although some doubts may arise in respect of the method applied during data correction, the income position typically improved in the majority of these settlements (despite the fact that just their relatively higher values around the change of regime could not be taken into account). The majority of settlements with a larger population were characterized by increasing but not extreme total potential values.

Figure 10

Changes in total potential of income in north-eastern Hungary between 1988 and 2008



Source: Edited by Gábor Németh based on own calculation from the data of the Ministry of Finance–National Tax and Customs Administration and the Hungarian Central Statistical Office.

On the other hand, the largest decline was recorded in settlements afflicted by the cessation of mining (e.g. Izsófalva, Edelény), the meltdown of heavy industry (e.g. Ózd, Borsodnádasd), the bankruptcy of cooperative production units (e.g. Zalkod, Viss) and (also) by social problems. The outward migration and ageing was, for example, characteristic of Királd and Zalkod, a high proportion of Roma in the population was typical of Terpes, Szendrőlád, Viss, Járdánháza and Arló. The issue of the possible effects of changes in public administration also arose in case of Edelény and Ózd, since several settlements with considerable population number were separated from them, but data correction tried to eliminate this problem. Almost in each case, the largest falls in total potential value could be attributed to the dramatic decrease in the proportion of employed people.

Summary

The study intended to highlight the income trends in the two regions of north-eastern Hungary. The per capita taxable income data in Northern Hungary spectacularly fell after the change of regime, while the bottom-ranked position of Northern Great Plain did not change, but, compared to the national data, a decrease was characteristic here as well. The changes having taken place resulted in marked regional differences, the features of which reflected the acute problems of the collapse of mining and heavy industry in Northern Hungary, as well as of the decline in farming in the Great Plain. An important feature is that the further increase of disadvantaged regions, where incomes were modest earlier as well, is especially striking. At the same time, an important part of the settlements that had high income previously, retained or improved their position.

The characteristics of the spatial structure and the spatial pattern of changes were explored with the help of the potential model by comparing the data of 1988 and 2008. After data corrections, the calculation-intensive method outlined the most important features of the potential model components.

The self-potential values clearly demonstrated the spatial income structure of the observed regions, and outlined the centres of economic activities and income production, as well as the areas of lower activities.

The inner potential values reflected settlements located near larger centres and having good accessibility, the majority of which are target areas of suburbanization, and which are unambiguously winners of income restructuring in the past period. At the same time, settlements located far from centres were characterized by quite modest values.

Due to the lack of data beyond the borders, outer potential was almost entirely influenced by the distance from Budapest, which at the same time, pointed out the impact of motorways on potential values.

The total potential values, and especially their changes, clearly demonstrated the relative winners and losers of the spatial income structure having restructured after the change of regime. Most characteristically, county seats, except for Miskolc showing a decline and Salgótarján where the decline was even larger, together with the majority of settlements with a higher population number could retain or improve their relative positions. Settlements in their surroundings had increasing income potential as well, which showed the favourable socio-economic trends (first of all a higher level of employment) associated with suburbanization.

Decreasing potential values characterized the peripheral areas of north-eastern Hungary. These are the outer periphery composed of the Cserehát, Zemplén, Bodrogköz, and Szatmár-Bereg regions, the almost contiguous eastern border area, as well as the Middle Tisza District forming the more and more characteristic inner periphery, where unfortunately, not even the reviving tourism could reverse the trends (Radics–Bujdosó 2006, Dávid–Michalkó 2008). Moreover, the spectacular decline of the Bódva- and Sajó-völgy, as well as the Ózd–Salgótarján axes demonstrated the long-lasting structural crisis. The method outlined the further increase in the disadvantage of areas which were also considered underdeveloped earlier, while the advantage of developed regions typically increased in the observed period. The income changes in the settlement network of north-eastern Hungary were mainly characterized by polarization between 1988 and 2008.

REFERENCES

Adler Judit – Skultéty László (2002): Regionális foglalkoztatottsági különbségek – kiváltó okok a '90-es években III. *Comitatus: Önkormányzati Szemle* 12 (4): 59–68.

Baranyi Béla (2007): *A határmentiség dimenziói Magyarországon*. Dialóg–Campus Kiadó, Budapest–Pécs.

Beluszky Pál (1976): Területi hátrányok a lakosság életkörülményeiben – Hátrányos helyzetű területek Magyarországon. *Földrajzi Értesítő* 25 (2–3–4): 301–312.

Beluszky Pál – Győri Róbert (1999): A magyarországi városhálózat és az EU-csatlakozás. *Tér és Társadalom* 13 (1–2): 1–30.

Bene László – Tekse Kálmán (1966): *Vizsgálatok a népesség területi eloszlásának alakulásáról Magyarországon 1900–1960*. KSH Népességtudományi Kutató Csoport közleményei, 1.

Bódi Ferenc – Obádovics Csilla – Mokos Béla (1999): Adózás, jövedelemkülönbségek Magyarországon. *Területi Statisztika* 39 (2): 131–147.

Cséfalvay Zoltán – Nikodémus Antal (1991): Két századvég Magyarországon – gyorsjelentés a gazdaság regionális átrendezéséről. *Tér és Társadalom* 5 (4): 69–89.

Dávid Lóránt – Michalkó Gábor (eds.) (2008): *A Tisza-tó turizmusa*. Magyar Turizmus Zrt., Budapest.

Dusek Tamás (2005a): A kistérségek jövedelmi helyzetének alakulása 1988 és 2003 között: a változás típusai. In: (Rechnitzer János szerk.) *Átalakulási folyamatok Közép-Európában*. SZE MTDI Évkönyv, Győr.

Dusek Tamás (2005b): *Térbeli egymásrahatások, szociálfizikai modellek. Regionális elemzési módszerek*. Regionális Tudományi tanulmányok 11. ELTE Regionális Földrajzi Tanszék – MTA–ELTE Regionális Tudományi Kutatócsoport, Budapest.

Ekéné Zamárdi Ilona – Pénzes János (2006): Jövedelemegyenlőtlenségek Borsod-Abaúj-Zemplén megyében. In: Kovács Ferenc – Hevesi Attila (szerk.) *Földrajz. Tiszteletkötet Hahn György 70. születésnapjára*. A Miskolci Egyetem Közleménye. A sorozat, Bányászat, 69. kötet). Egyetemi Kiadó, Miskolc.

Enyedi György (1996): *Regionális folyamatok Magyarországon*. Ember, település, régió. Hilscher Rezső Szociálpolitikai Egyesület, Budapest.

Faluvégi Albert (2000): A magyar kistérségek fejlettségi különbségei. *Területi Statisztika* 40 (4): 319–346.

Jakobi Ákos (1999): Az anyagi jólét becsült kistérségi egyenlőtlenségei. In: *A táj és az ember geográfus szemmel*. Geográfus Doktoranduszok IV. Országos Konferenciája elektronikus kiadványa, Szeged.

Kanalas Imre – Kiss Attila (eds.) (2006): *A perifériaképződés típusai és megjelenési formái Magyarországon*. MTA RKK Alföldi Tudományos Intézet, Kecskemét.

Kiss János Péter (2007): *A területi jövedelemegyenlőtlenségek strukturális tényezői Magyarországon*. Doktori értekezés, Szegedi Tudományegyetem Természettudományi Kar, Szeged.

Kovács Csaba (1993): A települési és térségi jövedelemegyenlőtlenségek az adóköteles jövedelmek és nyugdíjak alapján. In: Enyedi György (szerk.): *Társadalmi-területi egyenlőtlenségek Magyarországon*. Közgazdasági és Jogi Könyvkiadó, Budapest.

Kovács Zoltán (1990): A határ menti területek központhálózatának átalakulása az első világháború utántól napjainkig. *Földrajzi Közlemények* 114 (1–2): 3–16.

Kozma Gábor (1998): A gazdasági élet szereplőinek térbeli preferenciái. *Falu Város Régió* 1 (9): 7–14.

Lőcsei Hajnalka (2004): A vidéki városi agglomerációk fejlődési pályája. In: *Térségi és települési növekedési pályák Magyarországon. Regionális Tudományi Tanulmányok 9*. ELTE Regionális Földrajzi Tanszék és MTA–ELTE Regionális Tudományi Kutatócsoport, Budapest.

Lőcsei Hajnalka (2010): *Területi növekedési pályák Magyarországon, 1990–2008*. Doktori értekezés, ELTE TTK, Budapest.

Major Klára (2001): *A nemzetközi jövedelemegyenlőtlenség dinamikája*. Doktori értekezés, Budapesti Közgazdaságtudományi és Államigazgatási Egyetem, Budapest.

Major Klára – Nemes Nagy József (1999): Területi jövedelem-egyenlőtlenségek a kilencvenes években. *Statisztikai Szemle* 77 (6): 397–421.

Molnár Ernő (2008): *Megújulás vagy csendes vegetálás? Az észak-alföldi kisvárosok feldolgozóipara az ezredfordulón*. Doktori értekezés, DE TTK, Debrecen.

Nagy Gábor (2004): A gazdasági távolság meghatározása potenciálmodell felhasználásával. *Területi Statisztika* 44 (1): 31–40.

Nemes Nagy József (1998): Vesztesek – nyertesek – stagnálók (a társadalmi-gazdasági változások regionális dimenziói). *Társadalmi Szemle* 53 (8-9): 5–18.

Nemes Nagy József (2001): A topon. *Comitatus: Önkormányzati Szemle* 11 (3): 63–64.

Nemes Nagy József (2005): Fordulatra várva – a regionális egyenlőtlenségek hullámai. In: Dövényi Zoltán – Schweitzer Ferenc (szerk.) *A földrajz dimenziói*. MTA Földrajztudományi Kutatóintézet, Budapest.

Nemes Nagy József – Tagai Gergely (2009): Területi egyenlőtlenségek, térszerkezeti determinációk. *Területi Statisztika* 49 (2): 152–169.

Németh Nándor (2009): *Fejlődési tengelyek az új hazai térszerkezetben. Az autópálya-hálózat szerepe a regionális tagoltságban.* Regionális Tudományi Tanulmányok 15. ELTE Regionális Tudományi Tanszék, Budapest.

Németh Nándor – Kiss János Péter (2007): Megyéink és kistérségeink belső jövedelmi tagoltsága. *Területi Statisztika* 47 (1): 20–45.

Papp Antal (1978): A Tiszántúl középső részének központjai, területi elrendeződésük, potenciáljuk. *Földrajzi Közlemények* 102 (3): 269–279.

Pénzes János (2004): A humán erőforrások centrum–periféria viszonyai a tiszaújvárosi kistérségben. In: Süli-Zakar István (szerk.) *Határon átnyúló kapcsolatok, humán erőforrások*. Kossuth Egyetemi Kiadó, Debrecen.

Pénzes János (*2*007): A területi jövedelemegyenlőtlenségek alakulása az észak-alföldi régióban. *Területi Statisztika* 47 (4): 358–384.

Pénzes János (2010): *Területi jövedelmi folyamatok az észak-alföldi régióban a rendszerváltás után*. Studia Geographica 26. Debreceni Egyetemi Kiadó, Debrecen.

Radics Zsolt – Bujdosó Zoltán (2006): Tourism as the tool of rural development in Eastern Hungary – a case study. In: Zikov, M. et. al. (eds.) *Rural areas in the modern development conditions*. Faculty for Natural Sciences and Mathematics, Institute for Geography, Skopje.

Rechnitzer János (1993): *Szétszakadás vagy felzárkózás – a térszerkezetet alakító innovációk*. MTA RKK, Győr.

Ruttkay Éva (1997): A lakossági jövedelemeloszlás területi és települési különbségei. *Comitatus: Önkormányzati Szemle* 7 (11): 11–20.

Stewart, J. Q. (1941): A Measure of the Influence of a Population at a Distance. *Sociometry* 5 (1): 63–71.

Süli-Zakar István (1992): Az államhatár társadalmi-gazdasági fejlődést akadályozó hatásának vizsgálata ÉK-Magyarország határ menti területein. *Földrajzi Közlemények* 40 (1–2): 45–56.

Sütő Attila (2008): Város és vidéke rendszerek és típusaik Magyarországon. *Falu Város Régió* 10 (3): 51–64.

Tagai Gergely (2004): *Kelet-Közép-Európa népesség- és potenciáltere*. MTA KTI Műhelytanulmányok, 26, Budapest.

Tagai Gergely (2007): A potenciálmodell erényei és korlátai a társadalomkutatásban. *Tér és Társadalom* 21 (1): 145–158.

Tagai Gergely – Pénzes János – Molnár Ernő (2008): Methods of the analysis of integration effect on border areas – the case of Hungary. *Eurolimes* 6: 150–159.

Tóth Géza (2005a): A potenciálmodell alkalmazása az autópálya-nyomvonalak területfejlesztési szempontú vizsgálatában. *Gazdaság és Statisztika* 56 (3): 3–17.

Tóth Géza (2005b): *Az autópályák szerepe a regionális folyamatokban*. Központi Statisztikai Hivatal, Budapest.

Keywords: taxable income, north-eastern Hungary, potential model, spatial structure.

Abstract

As a result of the political and economic transition, the Hungarian spatial pattern of taxable personal incomes restructured significantly. The position of north-eastern Hungary weakened, however, pronounced intra-regional inequalities arose due to the uneven collapse of the local economies. The transition reflected the group of settlements categorised by the number of settlements.

The detailed spatial income pattern of settlements and the internal and external effects influencing it were represented with the help of the potential model. The changes between 1988 and 2008 were measured by the difference of the summarized potential values. The map of differences illustrated the areas with relative advantages and disadvantages compared to the regional average value. Most of the largest towns were able to maintain or strengthen their positions giving advantage to the suburban settlements in their environs. However, the former peripheries, the mining and heavy industrial zones with a collapsed economic base, suffered the greatest relative loss over the last two decades because of the transition.

REGIONAL STATISTICS, 2012, VOL. 2: 108–128 DOI: 10.15196/RS02108

LAJOS BÁLINT

Spatial Gender Differences in Life Expectancy at Birth

The study presents the major micro-regional characteristics of life expectancy at birth between 2005 and 2009 according to the classification in force at present. The spatial pattern of life expectancy at birth by gender was described with the help of global and local autocorrelation indicators. Due to the strong relation between the regional life expectancy of men and women, the applied non-spatial cluster analysis served as a more exact methodological frame for typology. The sole purpose of the study was to describe the spatial structure of life expectancy at birth with descriptive tools.

Life expectancy at birth goes far beyond its demographic content. A number of researchers emphasize that life expectancy can be considered a stable imprint of social processes, and is – among macro-level indicators – the most important component of life quality (Sen 1998, Wilkinson 1992, Bobak and Marmot 1996, Mazumdar 2001, Dasgupta 2000). As opposed to other indicators of life quality and different composite indicators, which are often difficult to interpret, life expectancy in itself has an unambiguous content irrespective of spatial and temporal context. To live longer and to remain healthy as long as possible is one of the individual and collective aims that are naturally obvious for everybody.

In Hungary, a large number of publications dealt with the topic of regional mortality (Daróczi 1997). Until the establishment of the structure of micro-regions, researches focused mainly on county-level analyses and the differences between the districts of the capital city (Daróczi 1997, Józan and Forster 1999). Demographic and spatial epidemiological researches at a more nuanced regional scale gained ground owing to the establishment of the structure of micro-regions as well as to the more and more refined methodological tools. In the Hungarian scientific literature numerous micro-regional mortality analyses were prepared (Hablicsek 2004, Klinger 2003, 2006a, 2006b). The focus of the analyses of demographers was first of all on the compilation of regional mortality tables and the descriptive characterization of differences. Within the current framework, the emphasis is definitely put on spatiality. This study outlines the present differences in life expectancy with the help of spatial cross-sectional data.

Data and method

The micro-regional abridged mortality tables were calculated with the Chiang method separately for men and women (Chiang 1984). The detailed description of the method is well documented in numerous publications of the Hungarian scientific literature (Hablicsek 2003, Daróczi 2004). The abridged mortality tables <1, 1–4, 5–9, …, 90+ are built up from the data of age groups. The mortality and mid-year population data come from the DEMOgráfiai program of the Hungarian Central Statistical Office which creates tables. In the interest of authenticity, micro-regional life expectancies at birth contain the aggregate data of five years. The selected regional scale is the micro-regional classification in force at present[1], which divides the country into 174 disjunct regions in a compact way. The population sizes of these regions are quite different. The division of the population of the capital into districts seems a possible solution, by which the number of observations is naturally increasing, while population size differences between them are decreasing. However, changing the spatial scale leads to well-known consequences. In case of life expectancy at birth, the mean of the "sample" increases, and the results of the autocorrelation tests, especially of the local tests change. Regions with the best life expectancy are practically only the districts of Budapest and some micro-regions of the agglomeration. The main argument in favour of the unchanged traditional micro-regional scale is that the divide between the districts of the capital is nowadays much less striking than the differences between the capital and the peripherical regions in the rural areas.

Table 1

Descriptive statistics

Denomination	Men	Women
Minimum	63.54	74.54
Maximum	73.09	80.51
Range	9.55	5.97
1st quartile	67.39	76.50
Median	68.58	77.10
3rd quartile	70.02	78.17
Interquartile range	2.63	1.67
Mean	68.67	77.27
Dispersion	1.80	1.22
Variance	3.23	1.50
Skegness	–0.04	0.17
Kurtosis	–0.27	–0.56
r	0.83[a)]	
N=	174	174

a) $p < 0{,}001$.

General features

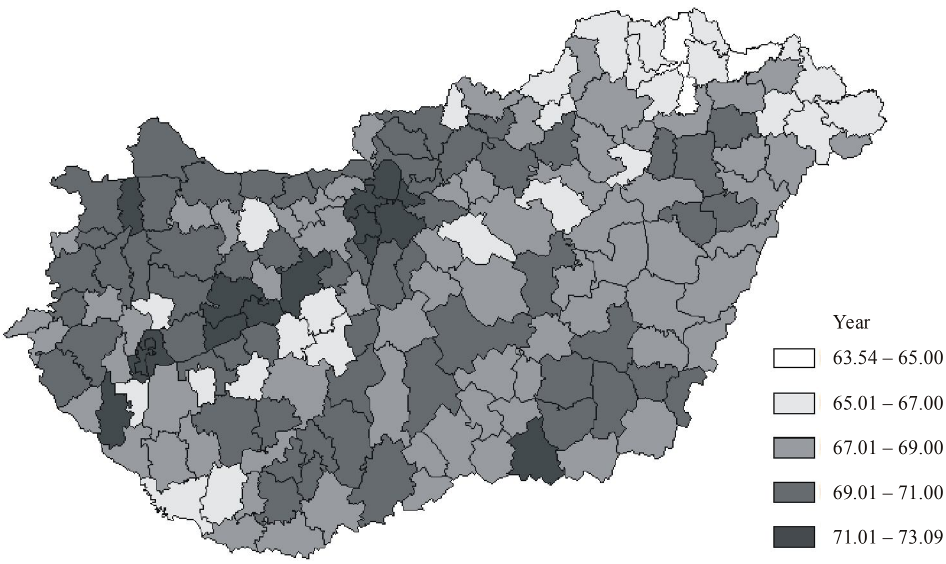
Life expectancy at birth for *men* by micro-region varied between 63.5 and 73.1 years in the observed period. The range of nearly ten years can be considered a significant difference. Micro-regions with the lowest life expectancy have wider, contiguous areas mainly in the north-eastern part of the country (typically in Borsod-Abaúj-Zemplén county) and in the region of Southern Transdanubia. Life expectancy is low in the micro-regions in the southern parts of Fejér county (Aba, Enying and Sárbogárd) as well. Furthermore, there are isolated areas with high mortality in the region Central Transdanubia. The location of areas with the most unfavourable life expectancy is similar to an hourglass rotated towards the south-west and north-east, while in the case of areas with more favourable life expectancy it is more difficult to observe such abstraction.

1 The number of micro-regions changed to 175 after the manuscript had been handed in (*the editor*).

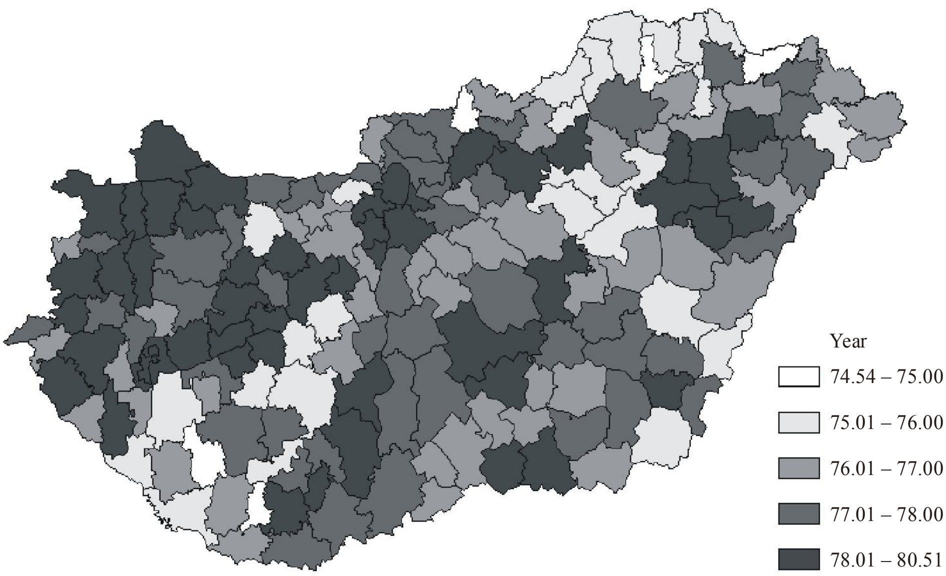
Figure 1–2

Life expectancy at birth by micro-region between 2005 and 2009

1. Men



2. Women



The areas with the highest life expectancy are in the capital and its agglomeration, as well as on the northern shore of Lake Balaton. The majority of micro-regions in Western Transdanubia have lower life expectancies in each region, first of all in the more urbanized areas.

The spatial differences in life expectancy of women by micro-region are much smaller (Table 1), as confirmed by the range (6 years) and the interquartile range as well (1.7 years as opposed to 2.4 years in case of men). However, the difference of six years between the areas with the highest and the lowest life expectancy is also considerable. The areas in the most favourable situation are – similarly as in the case of men – in Budapest and its agglomeration, in the region Western Transdanubia and on the northern shore of Lake Balaton (Figures 1–2), but, differently from men, there are such areas in the different parts of the Great Plain (e.g. in Hajdúság) as well. In the eastern part of the country women have higher life expectancy in areas centred in large cities and in their surroundings (Szeged, Debrecen, Eger, Békéscsaba).

The spatial pattern of life expectancy for men and women does not considerably differ from each other, insofar as visual impression may have some relevance. The definite correlation between the two genders is shown by the high value of the Pearson linear correlation coefficient as well ($r=0.83$, $p<0.001$).

Spatial autocorrelation

There are numerous application opportunities of spatial autocorrelation, of which I use the aspect of the explorative spatial data analysis (ESDA). I try to find an answer to the question of how far the values observed in certain locations are similar to, or different from, those of the neighbours. The present spatial structure of life expectancy is described by determining the degree of clustering of life expectancies and by presenting the location of clusters. The differences between autocorrelation tests applied for irregular polygons can be essentially attributed to the different interpretations of similarity (Waller and Gotway 2004). In the following, I present three often applied autocorrelation tests, and apply them to micro-regional life expectancies.

Moran I

The Moran *I* (Moran 1950, Cliff and Ord 1981) is built up similarly to the Pearson product-moment correlation. Moran himself used the concept of spatial correlation. The difference is the correction of spatiality by the weight matrix (W). The coefficient shows the relation of the variable with itself, i.e. with the same variable of the neighbouring locations and not the strength of the linear relation between two variables. Therefore, the Moran coefficient is an univariant and, due to the inherent feature of spatiality, multi-directional indicator. Its formula is written in the following well-known way:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

where $S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$. If the weight matrix is row-standardized, i.e. the sum of the rows is one, i.e. each existing neighbour is normed by the number of neighbours, the sum of S_0 is equal to the number of observations *(n)*. The equation is reduced:

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

where y_i is the value of the y variable in the *i-th* location, y_j is the value of the y variable in the *j-th* location, $\bar{y}$ is the mean of observations. The expected value of the Moran coefficient is $E(I) = -1/(n-1)$; according to the null hypothesis, observations are independent from each other, i.e. there is no autocorrelation. The value range of the indicator is approximately between –1 and 1. The Moran value is sensitive to the population differences in the relevant observations (Oden 1995) and the outliers, and its size depends on the selected weight matrix as well (Waller and Gotway 2004). Normality and random permutation tests, as well as saddle point approximation are available for testing the hypothesis (Bivand 2009). The value of Moran *I* of course does not change, but the significance level may alter. The normality assumption is often not true for the probability distribution of test statistics, so it is examined usually by one of the above mentioned processes, whether the observations are random or clustered.

Gear's c coefficient

The null hypothesis of Moran *I* is based on the covariance structure of the examined spatial variable $[(y_i - \bar{y})(y_j - \bar{y})]$. The assumption is that the values of locations do not differ consistently from the mean of observations. The null hypothesis of Geary c is based on another interpretation of independence, and states that the neighbouring spatial elements do not differ from each other (Geary 1954). According to the conclusion of the hypothesis, there is no consistent spatial pattern in the differences of neighbours, which are great in some places and smaller in others. Geary c can be computed by the following formula:

$$c = \frac{(n-1)\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(y_i - y)^2}{2W_{ij}\sum_{i=1}^{n}(y_i - \bar{y})^2}.$$

The symbols are the same as those used in the foregoing. The expected value of Geary c in the case of spatial independence is: $E(c)=1$. The connection between Geary c and Moran *I* is negative. As Moran is the best invariant autocorrelation test, Geary is rarely applied. The less than one indicator refers to positive autocorrelation, i.e. to the similarity (small difference) of the neighbouring observations. In the case of extreme positive autocorrelation the value of c is 0. Any value, higher than one, refers to aconsiderable difference between certain locations i.e., to negative autocorrelation. The extreme value of negative autocorrelation is 2. Since the indicator is based on the squared difference between the values of the neighbouring locations, outliers have a considerable effect on estimating autocorrelation (Fortin–Dale 2005).

Getis–Ord General G statistics

The general or global G indicator (Getis–Ord 1992) is based on the concept of the distance approach of neighbourhood, but it does not preclude the possibility of determining the weight matrix on the basis of topological relation. In the case of global G observations are generally, but not exclusively identified with a single point i.e., their centroids, whose Descartes coordinates are known. For aggregate demographic data applying the population centroid instead of geometrical procedures would be more reasonable which, in the case of micro-regions, would be equal to the coordinates of the centroid of the most populated settlement (the centre of the micro-region). Statistics examine the values of pairs of x_i and x_j points of located in d distance from each other. The formula of the indicator is:

$$G(d) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(d) x_i x_j}{\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j}, \; j \neq i$$

The global G identifies spatial relation with the product of the point pair values. In the case of a random pattern the expected value is: *E[G(d)]=W/[n(n–1)]*. If the value of G is higher than the value expected in the case of independence, then the spatial pattern is characterized by the concentration of high values. For low values we can state that the spatial pattern is much rather dominated by the low value pairs. In addition to the normality test, the significance value of global G is determined with the help of the Z value.

It is worth making some comments on the practice of the widely used G statistics published in 1992 (Getis–Ord 1992).[2] Its limit is that it can be applied only for positive values in their own (not transformed) units of measurement. Non-linear transformation considerably influences the value of test statistics. Accepting the null hypothesis, i.e. that observations are independent, does not necessarily mean a random pattern. High and low clusters may be equally present in the spatial structure. Consequently, as opposed to other global autocorrelation indicators, G statistics reveal which type of clustering is present in the observed spatial structure (Aldstadt 2010).

Results of global life expectancy test

In the case of topological weight matrices the relation was always based on first order queen contiguity. In line with symbols used in scientific literature, I applied B (binary) and W-type (row-standardized) weights, and created the binary weight matrix on the basis of the nearest 5, 10, 15 neighbours. In order to define the significance level of

2 In 1995, the authors elaborated a more flexible test, which can be applied not only for variables with positive sign but also for binary weights (Ord–Getis 1995).

autocorrelation, I used a randomization test for the Moran and the Geary values, and assumed normality for Getis statistics. The computations were made with the help of an R 13.0 program (R Development Core Team 2011).

It can be stated in general that the spatial pattern of life expectancy at birth for men shows slightly more considerable autocorrelation based on both the Moran *I* and the Geary c indicator than the same tests for women (Table 2). So, the spatial similarity of their life expectancy in the different micro-regions is also more significant. Except for the first order binary neighbourhood of women, the global G value proved to be significant in case of each weight matrix. Low values hardly differing from zero indicate that in the micro-regional pattern of life expectancy in the Hungarian spatial structure the more definite role of low values is characteristic, although it is not obvious at all based only on the impressions provided by the map.

Table 2

Results of autocorrelation tests with different weight matrices

Type of weight matrix	Moran *I*	Geary c	Global G
		Men	
B	0.371***	0.631***	0.031*
W	0.394***	0.613***	–
KNN5	0.398***	0.631***	0.029**
KNN10	0.327***	0.691***	0.058***
KNN15	0.270***	0.749***	0.087***
		Women	
B	0.237***	0.717***	0.031
W	0.264***	0.716***	–
KNN5	0.233***	0.770***	0.029***
KNN10	0.199***	0.804***	0.058***
KNN15	0.159***	0.838***	0.087***

* $p<0.05$, ** $p<0.01$, *** $p<0.001$.

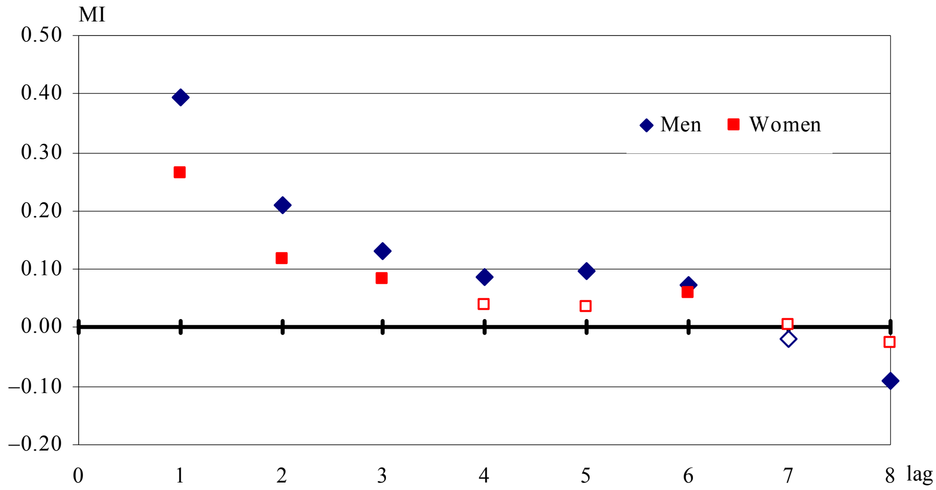
Correlogram

Characterizing spatial structure is a complex task, which involves the size, intensity and direction (isotropic or anisotropic) of spatial processes. One of the tools of characterizing the pattern is the spatial correlogram i.e, the representation of Moran *I* values as the function of distance (D=1, 2, …, d) or the different orders of neighbourhood (K=1, 2,…, k). The correlogram gives a clear visualization of the permanence of spatial dependence. The coefficients in the correlogram are usually marked by specific symbols depending on their significance; the empty sign refers to the absence of significance, while the signs of significant coefficients (α=0.05) are filled (Fortin–Dale 2005). According to common experience, autocorrelation is the largest in the case of first order or nearest neighbourhood, and dependence is gradually diminishing by the increase of distance or the higher order of neighbourhood. For those who are familiar with geostatistics, the similarity between the correlogram and the semivariogram is easily recognizable. The correlogram can be practically considered the inverse of the semivariogram.[3]

I took into account 1–8 lags of the 174 observations, i.e. the eighth possible neighbour (Figure 3). It is clear to see in the diagram that, on the one hand, the autocorrelation of men is always more significant than that of women, and, on the other hand, its intensity is strongly present up to the sixth order of neighbours, while in the case of women it practically disappears after the fourth order . At the same time, a significant autocorrelation not differing from that of men occurs again at the sixth order of neighbourhood.

Figure 3

Moran I correlogram



Note: The empty formation shows the insignificant spatial effect (the absence of spatial effect).

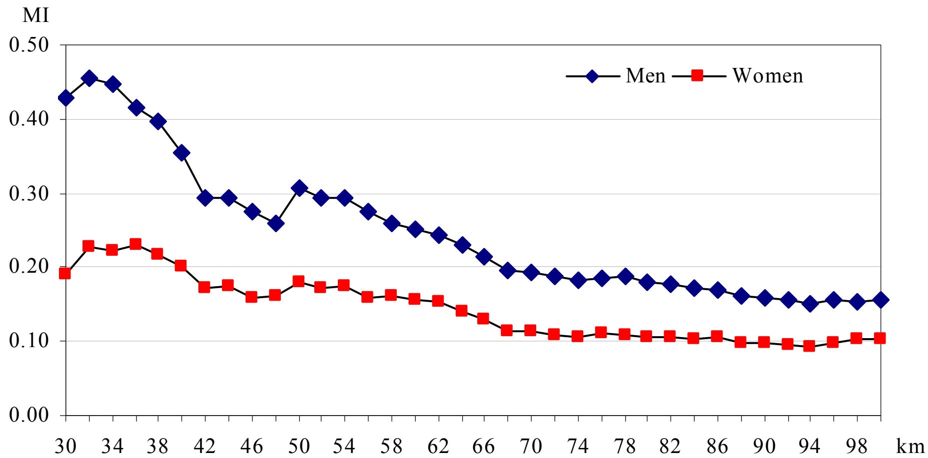
Compared to the correlogram which is based on the different orders of neighbourhood, distance-type approaches are also accepted, yet their substantive content is not at all unambiguous due to the nature of the data, the irregularity of spatial elements, i.e., the differences between distances. I defined the distance based on the Euclidean distance between the centroids of the micro-regions in an interval between 30 and 100 kilometres, for every two kilometres (Figure 4). The minimum 30 kilometre distance ensures that each observation (micro-region) has at least one neighbour. In the distance-based approach, the maximum value of Moran *I* was realized at 32 kilometres (MI = 0.447) for men and between 32 and 36 kilometres (MI = 0.22) for women. The highest value has an outstanding role, since the spatial effect is the largest here, and further additive effects are not realizable any more. After this, the value of the Moran coefficient is practically continuously decreasing but, in the case of neighbours within a

3 On the senivariogram, the distances between the observed points are on the x axis and the variances are on the y axis. In the case of similarity between two near points, their variance is less different, but by the increase of the distance, the differences are increasing as well. The fitting of the proper function shows the operation of the spatial structure.

circle of 100-kilometre radius, representing a size of a region in Hungary, a considerable and strongly significant autocorrelation can be observed.

Figure 4

Distance-based Moran I correlogram



Indicators or local autocorrelation

As opposed to the global autocorrelation tests, local approaches reveal the local features of spatial structures and describe the differences and similarities within the observed spatial structure by giving an exact answer to the questions ‘where?’ and ‘which are those?’. While global indices inform on the extent of autocorrelation in one single indicator, their local variants evaluate each observation one by one. The study by Anselin, (1995) published in the mid-1990s, was a significant leap in the field of autocorrelation of polygon-type spatial data. He marked the family of local indicators with the acronym LISA (Local Indicators of Spatial Association). Each LISA indicator is proportional to its global equivalent. In other words, global indicators can be decomposed into their local components.

Local Moran

Among the indicators of the LISA family the most widely spread is the local Moran connected to Anselin. ‚It can be defined for the i-th location as follows:

$$I_i = z_i \sum_j w_{ij} z_j \ ,$$

where z_i and zj are the differences from the mean, $(y_i - \bar{y})$ and $(y_j - \bar{y})$, w_{ij} is the local weight matrix. Due to the summation of values of j observations, only the neighbouring values are involved, i.e. $j \in J_i$. In the interest of simpler interpretation, w_{ij} is usually row-standardized, and – not necessarily, but in the general practice – w_{ii}=0. The expected value of local Moran is $E(I_i)$=–w_i/(n–1). Significant local clusters may be defined also along with the assumption of normality and randomization. The interpretation of local Moran is the same as the equivalent types of quadrants of the Moran scatter diagram. Four significant outputs are possible. There are high-high and low-low clusters, where the fix locations and their surroundings have similar values. Furthermore, there are also low-high and high-low clusters. Significant clusters may also be defined by assuming normality on the basis of its known momentums. An especially wide-spread solution is the random permutation test recommended by Anselin and the saddle point approximation of Tiefelsdorf (LLoyd 2011).

Getis–Ord local G

The Getis–Ord-statistics (Getis–Ord 1992, 1995) is less widely spread in the Hungarian practice, so its more detailed description, including its momentums, seems to be justified. The local version of the Getis–Ord-statistics can be defined as follows:

$$G_i(d) = \frac{\sum_{j=1}^{n} w_{ij}(d) x_j}{\sum_{j=1}^{n} x_j}, \qquad i \neq j,$$

where w_{ij} is the symmetrical spatial weight matrix. The value of the matrix elements is 1 if they are located within distance *(d)* defined by us and it is 0 in all other cases. According to this type of indicator, the given location is not neighbouring with itself, and the value of the weight matrix is 0. This distinguishes it from the indicator $G_i^*(d)$, where $i = j$. The $G_i(d)$ statistics actually identifies the strength of spatial association with the concentration of weighted spatial points. In the case of clustering of the above average values within a given distance the value of G_i will be high, while in the case of concentration of low values it will be low.
The expected value of the indicator can be defined as follows:

$$E(G_k) = \frac{W_i}{(n-1)}, \qquad \text{where } W_i = \sum_{j=1}^{1} w_{ij}(d), \quad i \neq j.$$

For the variance of the indicator we have to define:

$$Y_{i1} = \frac{\sum_{j=1}^{n} x_j}{n-1} \qquad \text{and} \qquad Y_{i2} = \frac{\sum_{j=1}^{n} x_j^2}{n-1} - Y_{i1}^2, \qquad i \neq j.$$

The variance is:

$$Var(G_i) = \frac{W_i(n-1-W_i)}{(n-1)^2(n-2)} \left(\frac{Y_{i2}}{Y_{i1}^2} \right).$$

The difference between the observed and the expected value of the G_i-statistics gives an answer as to whether the clustering of the high or the low value of the variable is characteristic in the surroundings of the location. The standardized Z value, $G_i(Z)$ belonging to G_i can be written as follows:

$$Z_i = \frac{G_i(d) - E[G_i(d)]}{\sqrt{Var(G_i(d))}}$$

It may be easy to fefining significance under the above conditions,, since the critical value of $G_i(Z)$ is 1.96 on a confidence level of 95%. When defining local G_i value, the key issue is to define the optimum d distance. According to the proposal of Getis (1995), defining the distance with the maximum extent of autocorrelation is the most obvious.

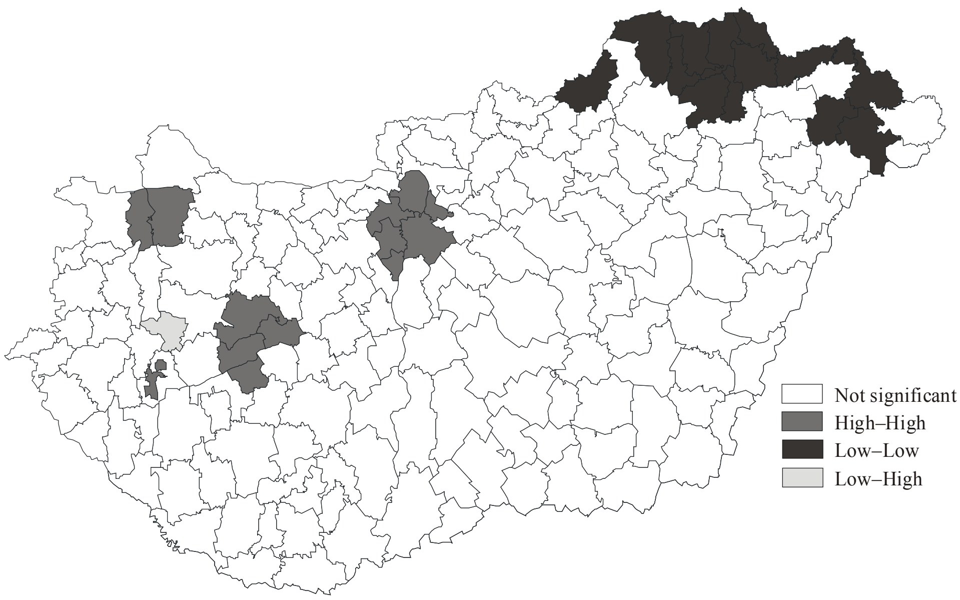
Local autocorrelation of micro-regional life expectancies

I examined the local features with the local Moran- (Anselin 1995) and Getis–Ord- (1992) statistics.Frst-order binary weight matrix based on queen contiguity was applied in both tests. Instead of mapping the local values of autocorrelation, I only wish to point out the significant clusters. In both tests the significance was defined in the same way, along with the assumption of normality, the limitations of which were indicated previously.

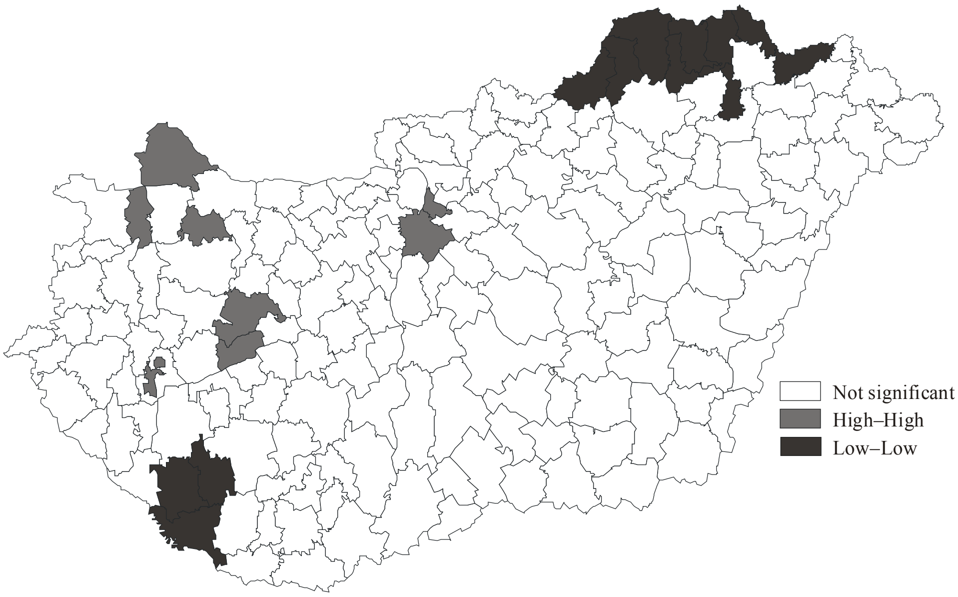
Figure 5–8

Local autocorrelation of micro-regional life expectancies
Local Moran I_i

5. Men

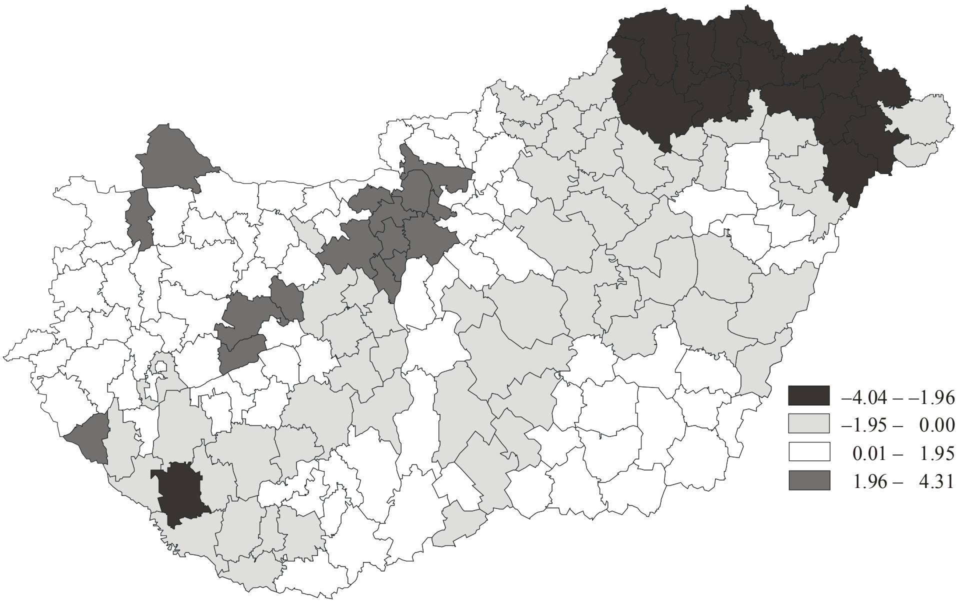


6. Women

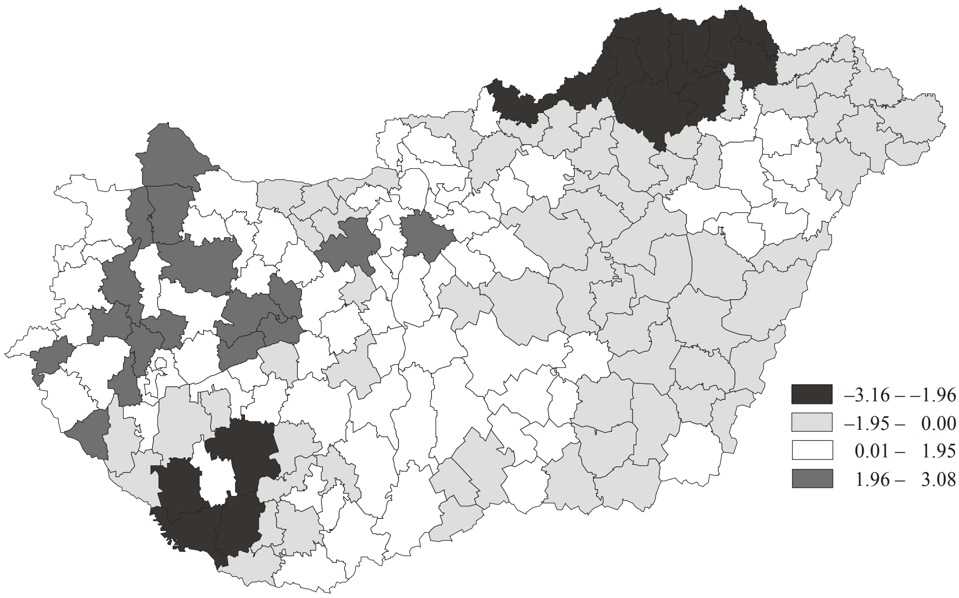


Local $G_i(Z)$

7. Men



8. Women



On the whole, the two tests led to very similar results both for men and women. The cluster of men is in the north-eastern part of the country. The cluster according to local G turned out to be somewhat more extensive than that according to the local version of Moran. It is interesting that, on the basis of the local Moran, there was no significantly low cluster in Southern Transdanubia, while in case of the Getis test, a cluster centre emerged (Nagyatád micro-region). On the basis of clustering of high values, four-four clusters can be distinguished, differing practically only in size. The areas with high life expectancy for men are mostly in the neighbourhood of the capital city, in the surroundings of Balatonfüred and Veszprém (the northern areas of Lake Balaton) and in Győr-Moson-Sopron county. According to local Moran, the Hévíz micro-region surrounded by the Keszthely micro-region, while according to Getis–Ord, the Letenye micro-region were identified as a high cluster.

The areas with low life expectancy are in the northern edge of the country in case of women as well, however, this cluster proved to be much narrower and somewhat shorter than that for men. Both local tests showed a cluster of more considerable extent in the micro-regions of Somogy in Southern Transdanubia (Nagyatád, Barcs, Kadarkút). It should be noted that according to $G_{i,}$ the Kaposvár micro-region was also a part of the low cluster. Except for Budapest, areas with longer lifetime cluster compactly as well. The cluster around Budapest is not at all as extensive as that for men, and only the centre of the cluster, i.e. Budapest proved to be significant. It seems that the effect of higher living standards is different for the two genders in the central region of the country,. While in the case of men a diffuse effect in the areas around Budapest is clearly visible, in the case of women this seems to be limited. The high clusters are in the micro-regions of Western Transdanubia. Compared to the former difference, the difference between the two testsis considerable. The number of clusters shown by local Moran is considerably lowerthan the number of those shown by local G.

Gender differences

The difference between life expectancies of men and women became obvious after the 1st world war, and since that time life expectancy of women has been higher than that of men in each more developed country (Nathanson 1984). The more favourable mortality of women can be observed in each age group; however, differences diminish in older age (Read and Gorman 2011). In older age, not only is the difference in mortality probabilities moderate, , the structure of causes does not show any significant difference, either (Cockerham 2004). The extent of gender differences considerably changed over time (Annandale 2010). In addition to the rlevant time lag, the fluctuation of differences can be partly explained by lifestyle factors and the consequent changes in the structure of causes of death (Cutler et al. 2006). Nowadays, the difference is much smaller in countries with low mortality than the peak values in the late 1960s and early 1970s. In the most developed countries differences between 4 and 7 years can be observed in general. So, for example, in Austria, where mortality is on Western European level, life expectancy at birth was 77.7 years for men and 83.3 years for women in 2009 (a difference of 5.6 years). Among the Nordic countries a smaller, 4–5 years difference is characteristic of Sweden (79.3, 83.4) and Norway (78.5, 83.4), while the difference is more considerable, 6.8 years in Finland (76.7, 83.5). Among the Western European countriesgender differences are strongly significant in France (77.8, 84.8) (HFA–DB).

In the Central and Eastern European region differences are incomparably larger: in the former Soviet region differences much above ten years can be considered usual (Cockerham 1999). In Russia the difference between men and women is exactly 12 years (62.7, 74.7) (HMD), and the situation is the same in the Ukraine as well (62.3, 74.1) (HFA–DB).

In Hungary the difference between life expectancies of men and women was 8 years (70.0, 77.9) in 2010, which can be considered a medium between Eastern European countries with extremely high and Western European countries with low mortality. Gender differences decreased in the past one and a half decades, but the lag is considerable compared even to the higher differences in the Western European.

To a certain extent, gender differences can be explained by biological (genetic, hormonal) features. The temporal fluctuation of macro-level differences is a reasonable argument for the limits to the biological explanation. In addition to biological reasons, various social, behaviour and psychological factors have a determinant role. Behaviour differences refer to the different lifestyle of men and women. Apart from some exceptions, health damaging habits (smoking, alcohol consumption) occur more often among men, their eating habits are less healthier than those of women, high-risk activity forms (injuries, accidents) occur more regularly among them and more men fall victims of fatal violent crimes too. The majority of victims of completed suicides are also men. It is also known that the participation of men in screenings and healthcare is lower than that of women.

In the spatial examination of gender differences two more important empirical trends can be defined. One of them clearly focuses on the analysis of the spatial structure, while the other is looking for an answer to the reasons for the differences. Preston (1976) intended to explain the inequality on the basis of the different (economic, social, psychological) dimensions of modernization. He stressed the different effects of the various factors on the structure of causes of death. The Swedish study examining the connection between gender differences and health status concluded a negative correlation between the health status and the different segments of equality, i.e. the equality in gender roles generated an increasing inequality in health, which was just the opposite of the most important finding of inequality literature (Backhans 2007).

The spatial connection between healthy life expectancies of men and women was examined by Groenewegen and his colleagues (Groenewegen et al. 2003). He found a relatively strong connection between the regional values for men and for women. The regional analysis of the British health authorities examining the data of the first half of the 1990s found more significant differences in the case of men than in the case of women. Gender differences were more significant in deprived areas and more balanced in areas of higher social status (Relaigh and Kiri 1997). When analysing 373 regional units in Poland, Malczewski (2009) obtained entirely different results. Life expectancies of men and women correlated only to a medium degree, and the spatial patterns of the two genders considerably differed from each other. A Scottish study examined the regional data of alcohol-related mortality by gender (Emslie and Mitchell 2009). The authors found a strong connection between the mortalities of the two genders. According to their finding, the same social factors are in the background of alcohol-related mortality in case of both genders.

Experiences in Hungary

According to Hungarian micro-regional data, the median and mean value of the differences between men and women is around 8.5 years, which is somewhat higher than macro-level differences. The lowest difference, 6.35 years was observed in the Szentendre micro-region, where life expectancy at birth was 73.04 years (95%, KI: 73.44–73.64) for men and 79.39 years (95%, KI: 78.82–79.96) for women. In Budapest the difference was almost the same as in Szentendre (6.46 years). At the same time, especially considering that these are areas with the most equal values , it is worth mentioning that difference of more than six years is is greater than the difference of four years which is not rare in Western Europe. The largest regional gender difference was observed in the Abaúj–Hegyköz micro-region (12.16 years), where life expectancy of men was 63.54 years (KI: 62.11–4.97), while that of women was 75.7 (KI: 74.02–77.37) years.

Table 3

Descriptive statistics of differences in life expectancies of men and women

Denomination	Difference
Minimum	6.35
Maximum	12.16
Range	5.81
1st quartile	7.87
Median	8.52
3rd quartile	9.28
Interquartile range	1.41
Mean	8.59
N=	174

The map illustrating the differences (Figure 9) and the high level of the correlation coefficient mentioned earlier imply that genders are good predictors to each other. The differences are more moderate in those regions where life expectancies are higher, while absolute disadvantages mean relative disadvantages between genders as well. I applied the cluster analysis for the typisation of the regional features of gender differences.. Clusters were formed on the basis of only two vectors, the difference in life expectancy between genders and the life expectancy of women. Before clustering, both variables were standardized (z-score). Among the possible solutions of partitioning, I used agglomerative hierarchical and non-hierarchical methods (PAM, k-means). In the hierarchical method the identity of the single groups was examined by single and complete linkage clustering and also with the centroid, group mean and sum of squares methods (Ward). (For a detailed description of the methods see Kauffmann and Rousseeuw 2005, Füstös et al. 2004.) The most adequate method was the k-means cluster (R 'cclust' package). The distance matrix between observations was based on the most general Euclidean metric:

$$d(i,j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2} \ ,$$

Defining the optimum number of clusters is the most difficult task in clustering procedures. Numerous tools, which are not necessarily in line with each other, are available to carry out the most appropriate clustering (for evaluating the different cluster algorithms see first of all Gan et al. 2007, Borcard et al. 2011). In the present case, I applied the silhouette coefficient (SC) of Rousseeuw (Rousseeuw 2005). The indicator takes into account the dissimilarity of observations belonging to the cluster *a*, and examines the same in relation with cluster *b* being the closest to cluster *a.*

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \cdot$$

The highest s(i) mean coefficient value define the number of clusters. The value of the indicator is $-1 \leq s(i) \leq 1$. The definite advantage of the theory is that the silhouette value of each cluster member can be visualized, so clusters whose members do not fit properly to their groups can be seen well. The mean of the coefficients by cluster indicates the adequacy of clustering; in the case of unequal number of groups SC is not recommended.

Due to the coefficient two clusters had to be defined. The value of the average silhouette coefficient was 0.53, which showed a stable cluster scheme. In case of partitioning with a higher number of clusters, the value of SC was always below 0.4 (the value of SC was 0.36 in case of three and 0.23 in case of four clusters), indicating the weakness of the structure, i.e. the artificial nature of partitioning. For validating clustersthe variance analysis (ANOVA) was used separately for both variables. The normal distribution of residuals was verified by the Shapiro–Wilk-test, while the homogeneity of variances by the Bartlett-test. Both tests met the requirements specified for the variance analysis. The analysis confirmed that life expectancies at birth significantly differ in the function of the defined factor (cluster members).

Figure 9

Gender differences in life expectancies at birth

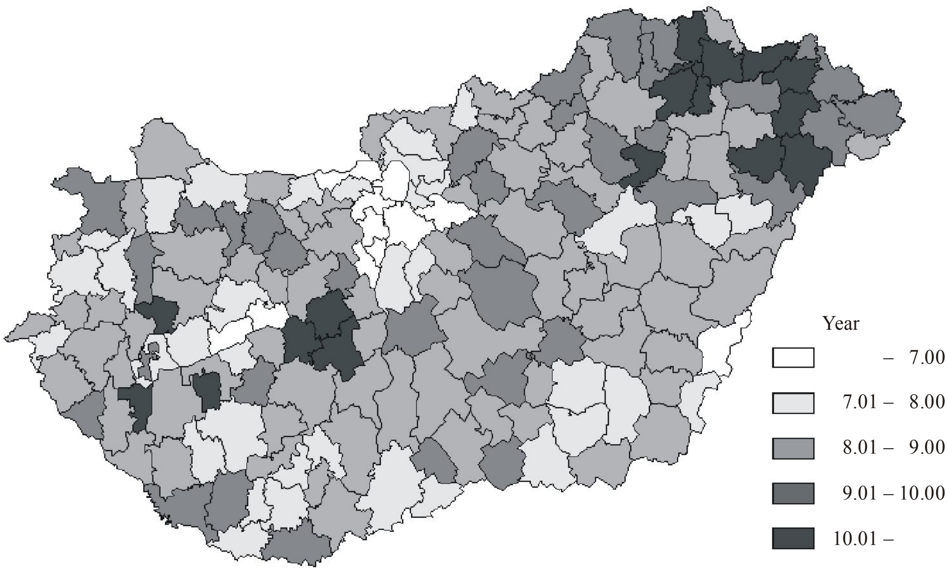
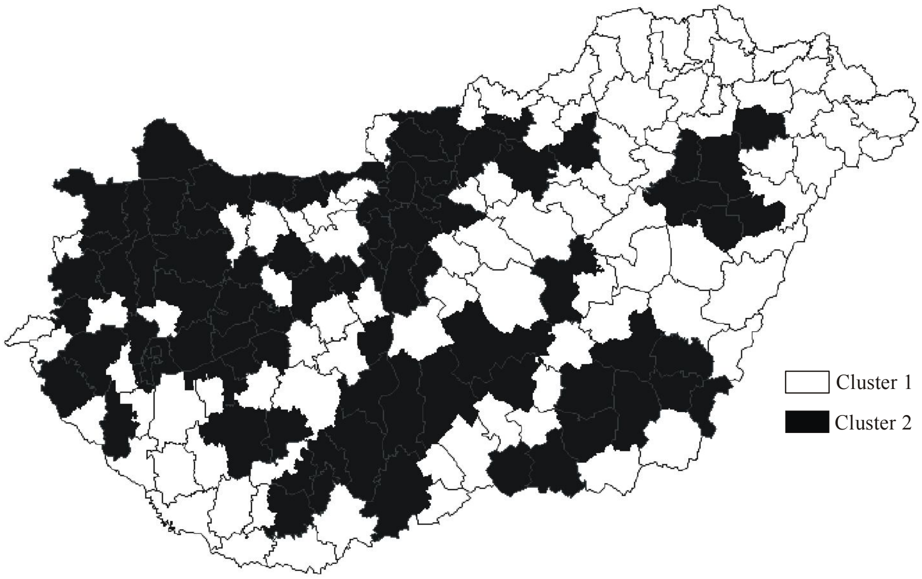


Figure 10

Result of k-mean clustering



According to the results the country can be divided into two types on the basis of regional gender differences, which are of course differences of degree. According to present data, the expected gender features are consistent in space. Where women can expect longer lifetime on the basis of the present level of mortality, men have similar life chances as well. So, my most important finding is that according to present data, inconsistent clusters cannot be demonstrated in the spatial structure in Hungary with the applied method. Life expectancies by gender strongly correlate with each other.

Summing up briefly the results of the cluster analysis, it can be stated that the first group is made up of micro-regions (n=93) with lower life expectancies for both genders. The mean lifetime of men was 67.3, while that of women was 76.4 years, and the mean difference amounted to 9.1 years. In the second cluster (n=81), the mean life expectancies were 70.3 and 78.3 years, respectively, and the difference was somewhat smaller. The spatial distribution does not correspond either to the distribution by east–west or to the one by centre–periphery. In the western part of the country those areas are in majority where differences are somewhat smaller on a national scale, and life chances are better for both genders. Such areas can be found in the regions Southern Great Plain, Northern Great Plain and Southern Transdanubia; moreover, they covers larger territories. The second-type cluster is dominant mainly in Southern Great Plain. The concentrated occurrence of observations belonging to the first cluster can be seen also in Southern Transdanubia, Northern Great Plain and mainly in Northern Hungary, however, they occur more sporadically in any other Hungarian region as well.

Conclusions

In the study I examined the spatial pattern of life expectancies at birth. The life expectancies were estimated with own computations on the basis of the method of Chiang (1984). Of the data analysis techniques the global and the local clustering provided a proper framework for describing spatial features and revealing local patterns. Local clusters were defined with the binary weight matrix and by assuming normality in case of both genders. The location of clusters proved to be similar by gender as well, while smaller differences occurred in the extent of the cluster.

Due to problems deriving from spatial delimitation and the different size of the observed sub-regions, regional inequalities in Hungary cannot be unambiguously compared to the similar data of other countries; the differences reveal significant regional disparity. Compared to the results of Western European researches, the differences in Hungary proved to be definitely more considerable. In the entire analysis presentation of the features by genders played a determinant role. It was not surprising that regional differences were more robust in the case of men. Gender differences were more than six years even in the most balanced micro-regions, which only just approach the mean in Western Europe, while areas with larger differences show similarities with the Eastern European, former Soviet region. We could observe a very strong correlation between life expectancies of men and women. The cluster analysis revealed only differences of degree, i.e. we cannot report any cluster or region that would refer to behaviours inconsistent by gender. In Hungary life expectancies by gender are clearly effective macro-level predictors to each other.

REFERENCES

Aldstadt, J. (2010): Spatial Clustering. In: Fischer, M. M. – Getis, A. (eds.): *Handbook of Applied Spatial Analysis*. Software Tools, Methods and Applications, Springer-Verlag, Berlin–Heidelberg.

Annandale, E. (2010): Health Status and Gender. In: Cockerham W. C. (ed.): *The New Blackwell Companion to Medical Sociology*. Wiley–Blackwell.

Anselin, L. (1995): Local Indicators of Spatial Association – Lisa. *Geographical Analysis* 27(2): 93–115.

Backhans, Mona C. – Lundberg, M.l – Mansdotter, A. (2007): Does increased gender equality lead to convergence of health outcomes for men and women? A study of Swedish municipalities. *Social Science and Medicine* 64 (9): 1892–903.

Bivand, R. (2009): „The Problem of Spatial Autocorrelation" forty years on. http://cran.r-project.org/web/packages/ spdep/vignettes/CO69.pdf (downloaded: May 2012).

Borcard, D. – Gillet, F. – Legendre, P. (2011): *Numerical Ecology with R. USE R!* Springer New York, Dordrecht, London, Heidelberg.

Chiang, C. L. (1984): *The Life Table and its Applications*. Robert E. Kriger Publishing Company Malaber, Florida.

Cockerham, W. C. (1999): *Health and Social Change in Russia and Eastern Europe*. Routledge, New York and London.

Cockerham, W. C. (2004): *Medical Sociology*. Ninth Edition, Pearson Prentice Hall, Upper Saddle River, New Jersey.

Cutler, D. – Deaton, A. – Lleras-Muney, A. (2006): The determinants of Mortality. *Journal of Economic Perspectives* 20(3): 97–120.

Daróczi Etelka – Hablicsek László (2008): A halandóság területi és időbeli különbségei az életkor függvényében. *Demográfia* 51 (1): 7–50.

Daróczi Etelka (1997): *A halandóság területi eltérései Magyarországon 1959/60–1992*. KSH Népességtudományi Intézet Kutatási Jelentések, 60., Budapest.

Daróczi Etelka (2004): Területi és társadalmi különbségek a középkorú férfiak és nők halandóságában Magyarországon 2001 körül. In: Daróczi Etelka – Kovács Katalin (szerk.): *Halálozási viszonyok az ezredfordulón: társadalmi és földrajzi választóvonalak*. KSH, Népességtudományi Kutatóintézet. Kutatási Jelentések, 77., Budapest.

Dasgupta, P. (2000): Valuation and Evaluation: Measuring the Quality of Life and Evaluating Policy. Reseurces for the Future, Discussion Paper 00–24. http://www.rff.org.

Emslie, C. – Mitchell, R. (2009): Are there gender differences in the geography of alcohol-related mortality in Scotland? An ecological study. BMC Public Health, 9 (58).

Fortin, M. J. – Dale, M. R. T. (2005): *Spatial Analysis – A Guide for Ecologists*. Cambridge University Press, Cambridge.

Füstös László – Kovács Erzsébet – Meszéna György – Simonné Mosolygó Nóra (2004): *Alakfelismerés (Sokváltozós statisztikai módszerek)*. Új Mandátum Könyvkiadó, Budapest.

Gan, G. – Ma, C. – Wu, J. (2007): *Data Clustering: Theory, Algorithms, and Applications*. ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA.

Geary, R. C. (1954): The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician* 5. (3): 115–127+129–146.

Getis, A. – Ord, J. K. (1992): The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24 (3): 189–206.

Getis, A. (1995): Spatial Filtering in a Regression Framework: Experiments on Regional Inequality, Government Expenditures, and Urban Crime. In: Anselin, L. – Florax, R. (eds.): *New Directions in Spatial Econometrics* Springer, Berlin.

Getis, A. (2010): Spatial Autocorrelation. In: Fischer, M. M. – Getis, A. (eds.): *Handbook of Applied Spatial Analysis*. *Software Tools. Methods and Applications* Springer–Verlag, Berlin–Heidelberg.

Groenewegen, P. P. – Westert, G. P. – Boshuizen, H. C. (2003): Regional differences in healthy life expectancy in the Netherlands. *Public Health* 117 (6): 424–429.

Hablicsek László (2003): Térségi halandósági különbségek a középkorúak körében. In: Daróczi Etelka (szerk.) (2003): *Kettős szorításban. A középgenerációk élete és egészsége*. KSH Hivatal Népességtudományi Kutatóintézet, Kutatási Jelentések, 74., Budapest .

HFA-DB: European Health for All Database, http://data.euro.who.int/hfadb/.

HMD: Human Mortality Database, http://www.mortality.org/.

Józan, P. – Forster, D. P. (1999): Social inequalities and health: an ecological study of mortality in Budapest, 1980–83 and 1990–93. *British Medical Journal* 318 (7188): 914–915.

Kauffmann, L. – Rousseeuw, P. J. (2005): *Finding Groups in Data. An Introduction to Cluster Analysis*. John Wiley & Sons, Hoboken, New Jersey.

Klinger András (2003): A kistérségek halandósági különbségei. *Demográfia* 46 (1): 9.

Klinger András (2006a): Újabb adatok a vidéki kistérségek és a budapesti kerületek halandósági különbségeiről (I.). *Demográfia* 49 (2–3): 197–231.

Klinger András (2006b): Újabb adatok a vidéki kistérségek és a budapesti kerületek halandósági különbségeiről (II.). *Demográfia* 49 (4): 342–365.

LLoyd, C. D. (2011): *Local Models for Spatial Analysis*. 2nd. Ed. CRC Press, Taylor & Francis Group, Boca Raton, London, Routledge.

Malczewski, J. (2009): Exploring spatial autocorrelation of life expectancy in Poland with global and local statistics. *Geojournal* 75(1): 79–92.

Mazumdar, K. (2001): Improvements in life expectancy: 1960–1995 an exploratory analysis. *Social Indicators Research* 55 (3): 303–328.

Moran, P. A. P. (1950): Notes on continuous stochastic phenomena. *Biometrika* 37 (1-2): 17–23.

Nathanson, C. A. (1984): Sex differentials in Mortality. *Annual Review of Sociology* 10: 191–213.

Oden, N. (1995): Adjusting Moran's I for population density. *Statistics in Medicine* 14(1):17–26.

Ord, J. K. – Getis, Arthur (1995): Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. *Geographical Analysis* 27 (4): 286–306.

Preston, Samual H. (1976): *Mortality Patterns in National Populations with Special Reference to Recorded Causes of Death*. Academic, New York.

Read, Jen'nan Ghazal – Borman, Bridget K. (2011): Gender and Health Revisited. In: Pescolido, B. A. –Martin, J. K. –McLeod, J. D. –Rogers, A. (eds): *Handbook of the Sociology of Health, Illness, and Healing*. A Blueprint for the 21st Century. Springer, New York, Dordrecht, Heidelberg, London.

Relaigh, Veena Soni – Kiri, Victor A. (1997): Life expectancy in England: variations and trends by gender, health authority, and level of deprivation. *Journal of Epidemiology and Community Health* 51(6): 649–658.

Sen, Amartya (1998): Mortality as an Indicator of Economic Success and Failure. *The Economic Journal* 108 (446): 1–25.

Waller, Lance A. – Gotway, Carol A. (2004): *Applied Spatial Statistics for Public Health Data*. John Wiley & Sons, New York.

Wilkinson, R. G. (1992): Income distribution and life expectancy. *British Medical Journal,* 304 (6820): 165–168.

Keywords: life expectancy at micro-regional level, global autocorrelation, local autocorrelation, k-means cluster, spatial demography.

Abstract

This article examines the spatial gender differences of life expectancy at birth in Hungary using the most recent data available. The estimation of life expectancies is based on modified version of Chiang's method which was calculated by the author. We experienced much larger differences amongst men than women. Our empirical result corresponds to the conventional experience. The range regarding men was 9.5 years and 6 years for women. Both values refer to enormous large spatial disparities. Nonetheless the strength of relationship between genders, based on a simple correlation coefficient was very high (r=0,83). It implies that genders are good predictors to one another. The exploration of spatial pattern of life expectancy is based on well-known global and local spatial autocorrelation tests. It has been found that the global spatial autocorrelations are somewhat larger for men than for women. However the differences are not remarkable. Not surprisingly the local patterns were very similar as well. Minor differences were found regarding the extent of local clusters. Areas with high life expectancies can be found in the central region, especially in the capital and its sorrounding zones, and Western Transdanubia region. The areas with the lowest life expectancy lie in South Western and North-Eastern regions of the country.

Finally we attempted to partition areas according to women's life expectancy and gender differences. We used k-means cluster method and silhouette coefficient (SC) to classify proper cluster structure. Because of the above mentioned strong relationship we could define only two clusters. SC value was more than 0.5. We experienced that, where the life expectancy was high or relatively high the gender differences were significantly smaller.

REGIONAL STATISTICS, 2012, VOL. 2: 129–148 DOI: 10.15196/RS02109

VIKTÓRIA SZIRMAI

Urban Sprawl in Europe*

The alienated atmosphere, the shallow relationships of the residential suburbs inhabited by the American middle classes and the spatial manifestations of the failure of the American dream and family model are spectacularly presented in Edward Albee's dramas written in the 1960s *Who's Afraid of Virginia Woolf?*, *Everything in the Garden*, or *A Delicate Balance*. The problems of some types of European, including French, suburbs such as suburban crime, violence or fear are best illustrated in the context of art cinema by films such as Mathieu Kassovitz's *La Haine* (Hate) or Luc Besson's *Banlieue 13* (District B13).

Naturally, for us examples from the context of art are not as important as scientific studies indicating the main causes of today's social problems, which are basically the lifestyle of the middle class, its self-fulfilling confinement and individualism in the American suburbs, and the concentrations of socially disadvantaged classes and social exclusion in Europe. Behind the differences, behind the characteristics of the American and European societies, and the traits of their history, economy and development, the different social structural features of cities are evident. The American middle classes live mostly in suburbs, and more disadvantaged groups are spatially concentrated in inner cities. The European middle classes are located predominantly in the inner city or better-off suburban areas, while the disadvantaged classes are increasingly concentrated in suburban areas.

Although European suburbs are different from their American counterparts, there are visible signs of increasing convergence. Now European cities also have to face tensions which hitherto characterised mainly American urban societies. European cities are also typified by fragmentation, the rise of individualism, and the disintegration of community sense and social cohesion (Cattan p. 1). Slums, which in the past were mainly typical of American cities, are now spreading in European cities as well and segregation is also increasing (Haussermann–Haila 2010, p. 60). The hereto positive features specific to Europe, namely the European welfare state which has played an outstanding role in comparison with America, the relatively limited enforcement of market impacts and public efforts to manage social tensions, have been damaged and needs to be limited (Kazepov 2010, Cities of Europe). It is a fact that the European welfare states are not always able to restrain the various tensions, particularly due to different social structural and economic difficulties and crises, and to the strengthening of European urban sprawl and its sociological impacts. Besides American suburbanisation, which fundamentally has historical relevance to America, European suburbanisation is also gradually spreading (Kazepov 2010, p. 13). Because of global economic interests, metropolitan

* This article is the introductory part of the book entitled *Urban Sprawl in Europe* published by Aula Kiadó (Budapest, 2011), edited by Viktória Szirmai.

areas are functionally transforming; the significance of the role of central business districts, so prominent in the United States, is growing in European inner cities. As a result there is an increasing 'tendency to shift consumption from the CBD to suburban shopping malls' (Haussermann–Haila 2010, p. 61).

Urban sprawl is strongly criticised in international literature, including the case studies of this book. Criticisms highlight the problems of urbanised but uncontrollably expanding areas, the decline of rural areas, the decreasing territory of agricultural areas and forests, and the depletion of green spaces (Görgl et al. 2011, Cattan 2011). Nearly all the criticisms point out the environmental threats of powerful motorisation, the effects on health of time-consuming commuting between home and work, the radical lifestyle shift related to residence change, and the negative consequences of car-dependent suburban lifestyle (Frumkin 2002, Reeh–Zerlag 2011, p. 14).

The issues of adverse economic effects and the costs of meeting the development needs of infrastructural investments, including road network development which slows the dynamism of the economy, are also raised (Williamson–Imbroscio–Alperovitz 2005). Concerns include citizens escaping to suburban zones due to environmental hazards and the visible signs of poverty in inner cities, the radically decreasing urban population and the disappearance of traditional compact cities (Munoz 2003).

The new structural characteristics of urban-regional societies transforming in sprawl, the segregation of residential communities populated by the upper and middle classes together with their negative urban effects (Le Goix 2004), and peri-urban/social exclusion are strongly criticised. According to the French study in this book, urban sprawl is perceived in France and other parts of Europe as the end of the social model, supporting a trend which may impair the integration of sections of society. Urban sprawl here appears to replicate all the faults of the models, concerning North American cities, that is to say increasing individualism, strengthening social segregation and the increased use of motor vehicles among other things (Cattan 2011, p. 11).

In Europe, instead of or in addition to the urban-rural conflict, problems are caused by the formation of a new kind of spatial social dichotomy, the disparities between the core and the periphery, the city and its urban neighbourhood (Vieillard-Baron 2008). The most extreme examples of social exclusion are generated by the incredible scale of urban growth in Third World countries, which is totally uncontrolled. In Africa, South America, Brazil, India and even China the poorest of the poor live on the edge of cities in peripheral slums under inhuman circumstances (Davis 2007).

Not everyone opposes urban growth, mainly because they acknowledge the social demand for a suburban environment, namely that more and more people prefer suburban zones when choosing a place of residence. Many people also believe that urban sprawl is a development tool of urban neighbourhood which enables the urbanisation of rural areas (Fishman 1990, Le-Goaziou–Rojzman 2006, p. 10).

The supporters of both viewpoints offer solutions to the problems. The opponents recommend the strengthening of centralised urban planning models, developments in inner city areas, and higher and denser buildings. The supporters of urban sprawl urge the dynamic development of public transport and powerful planning interventions.

Neither the arguments nor the proposed solutions are new. Of the numerous historical examples, perhaps Howard's concept of the garden city, the planning intervention of the

19th century, is the most significant, which many people believe is still valid. Garden cities built around but separated from major cities – as compact settlements – offer work and living space with full leisure facilities, and fully integrate the inhabitants, absorbing people wishing to live in big cities, thus preventing the unlimited growth of town centres (Howard 1902). In contrast, the so-called 'Good City' idea offered a solution for a capitalist town of the 19th century, especially for problems faced by workers (Haussermann–Haila 2010, p. 54), although in effect it envisaged a more middle-class (Kesteloot 2010, p. 126) than working-class suburban model of development. As regards the latter, the Paris suburban pavillonaires model built for workers proved to be more successful[1] (Haumont 2005).

Another example are the groups of avant-garde Russian architects in the 1920s and 1930s, the urbanists and deurbanists, who represented two different concepts, and their proposals for influencing urban sprawl. The supporters of the first group envisioned a model of concentrated urban development, while the other group planned to alleviate the perceivable urban problems by a decentralised, dispersed urban development model (Kopp 1979).

Although the suburbs born in the spirit of the garden city idea in several European cities tried to prevent the excessive expansion of the population by diverting the settlement of the population in specific directions, they were never really successful. The relevance of the garden city idea and the realisation of plans, i.e. the formation and long-term sustainability of communities propagating the unity of home and work in Europe, were supplanted in the meantime by processes transforming the economy: the development of the urban service sector attracting masses of people, which was not yet very noticeable in the early 1900s but later became dynamic, and the changing demands and forms of lifestyle as well as the weakening of local character, the trends of globalisation (Castells 1972, Merlin 1972, p. 62). The ideas of 'Fingerplan' in 1945 presented in the Danish case study, the positive effects elaborated and expected at that time on the unity of economic and social welfare, and the ambitions of the plan built on the aspects of reasonability and careful planning were questioned by these processes as well (Reeh–Zerlang 2011, p. 13).

Since then, the problems raised by Howard or even the avant-garde architects and their proposed solutions, and the processes of urbanisation and their impacts have changed radically because of the acceleration of the processes of globalisation and, as Enyedi suggests in his recent study, the enforcement of the new phase of urbanisation (Enyedi 2011). Originally the concepts attempted to respond to the problems of the first stages of urbanisation at that time, addressing the problems of urban growth, the accelerated flow to cities and the emergent problems of suburbanisation. Of course, the spatial problems of the early periods were also severe but differed from those of the present process and the effects of the advance of globalisation, behind which there are great mechanisms that today seem to be scientific banalities. One the one hand, in the developed Western European countries (and the United States and Japan) a strong economic and social centralisation of the service sector can be observed: the metropolitan

1 In the 1970s 70,000 state-subsidised, private family houses were built in the outskirts of Paris for families with modest means by usinga uniform house manufacturing technology and architectural style.

concentration of the service sector and skilled labour, the advance of multinational and transnational corporations, and as a result the strong development of large cities and their periphery (Veltz 1996, p. 33). On the other hand, the central role of the metropolitan regions in the world economy is evident, as is their aspiration for power for these reasons. The increasing concentration of the economy and population was seen in the growing territorial division of homes and workplaces and the fact that homes expanded outwards faster than workplaces. This determined the spatial movement of capital investment from the centre outwards, infrastructure development projects, and the establishment of trade and other services (Hall 1996). The result is the space-consuming expansion of urban agglomerations, the increased level of short- and long-term commuting, growing demand for transport and greater capacity, rising environmental damage, the transformation of spatial and social structure, i.e. urban sprawl, and its effects. Several factors lie behind the contemporary space consumption phenomena of metropolitan areas, including the expansion of the values of consumer society, the steadily increasing number of people living well and the changed residential demands. In addition the world system of power provides a regional development system which pushes social aspects into the background and enforces economic interests. The Danish case study shows that the global transformation, the evolving urban network systems between contemporary modern cities, and the goals of the project-oriented political intervention caused 'the shift from the welfare city to the growth city' (Reeh–Zerlang 2011, p. 13).

The evolution of urban sprawl correlates with the characteristic features of the roles of states in regional development and with the system of state subsidisation. In the United States of America, suburban development was initiated by the New Deal, the government's new economic, social and credit policy introduced in the 1930s, which was boosted by housing subsidies and long-term, 20-to-25-year loans in the 1950s. The marketing of suburban lifestyle promoting the ideal new American way of life also influenced the process as it met the demands of the middle classes tired of metropolitan life by offering a new way of living with the promise of less noise, less traffic and less crime. It even suggested that local community cohesion would afford protection from the poor (Williamson–Imbroscio–Alperovitz 2005). In the United States the importance of state initiatives has now decreased considerably; in addition to various government incentives and support systems, market actors have an increasing influence on regional trends.[2]

The case studies presented in this book show that, even though strong state intervention and modest market intervention was typical in Europe historically, this situation has now changed, and the convergence is becoming more visible; the role of the state has decreased and that of market actors has increased. This was partly because of the crisis of the welfare states and partly because of the increasing need of local governments to involve the market actors in planning thanks to the support of local

2 The role of civil society actors is also increasing. A number of initiatives have been launched not to support urban sprawl but to stop it, for example the New Urbanism movement or the preparation of Smart Growth guidelines aimed at building viable, multi-functional, compact settlements.

residential civil society and especially to the pressure of local and global economic lobby groups.

The case study of the Danish capital presented in this book shows how the state consciously managed the development of the Greater Copenhagen Region and the decentralisation processes, including moving the middle classes out of the city, by the construction of rented family homes and by creating the necessary transport and other infrastructure. However, planning possibilities were limited, if not supplanted, by urbanisation, globalisation mechanisms and the global interests of the market. An urban neighbourhood has developed as a result of state-backed bank loans, and the co-operation of the national and local authorities, and private land developers and investors (Zerlag–Reeh 2011).

In the case of the Paris Region the management of development processes was also a conscious public policy goal, especially in the 1960s and 1970s. Among them, the new urban development programme defined the directions of peri-urban development by designating the location of new towns as well as the main target areas for the accommodation of the middle classes (dissatisfied with housing estates) in the suburbs, which had outstanding significance (Szirmai 1991, Haumont et al. 1999). The French case study highlights the continuing importance of state intervention. As a result of an act passed in 2000, the population of Paris, which had been decreasing for a long time, began to grow again. The evolution of peri-urban areas today is mainly regulated by market forces and bank loans based on social and residential needs. As regards the latter, the high property prices in Paris are of particular importance. An increasing number of low social status and younger, career starter intellectual groups seek to buy cheaper land around the city (Cattan 2011).

The Austrian case study also shows that, although market forces are important, the intervention efforts of the state or even regional-level authorities are still not negligible. By using various planning tools, monitoring the processes and employing new development strategies reflecting today's conditions, these endeavours seek to mitigate the adverse effects of suburbanisation patterns (Görgl et al. 2011).

The growth of the suburban zone of Budapest in the socialist system was naturally different from the above phenomena. However, the benefit of hindsight reveals similarities, albeit historically belated. Let us see why.

In Hungary, including Budapest, especially in the first phase of the state socialist system, spatial processes were fully regulated by the intervention of the state and followed its political and ideological interests. There were no market mechanisms (following the economic reform in 1968 they operated only in a limited and controlled manner[3]), and consequently there was neither private capital nor private property or land; the means of production and companies were all state-owned. There were no market actors and there were no local governments in a real sense. In the so-called council

3 The so-called new economic mechanism was introduced in Hungary on 1 January 1968. In order to renew the half-depleted economy, to accelerate growth and to strengthen social welfare, three key reforms were introduced by the government: the central planning role of the state decreased and corporate self-reliance increased; prices were liberalised, i.e. in addition to officially fixed prices the prices of some products were defined in accordance with market demand, and finally the centrally determined wage system was replaced by a more flexible one, within certain limits set by corporate control. These reforms were unique in the countries of the socialist bloc.

system ‘local’ powers had neither legal nor financial or socially legitimate (political) means to manage their regional processes. Due to the characteristic features of the system, right up until the change of the political system in 1990, the growth of the peri-urban population could not be described as suburbanisation. It was rather the peri-urban concentration of the unemployed rural agricultural population excluded from the mechanisms of the system, especially from the economic and other benefits of regional development. These people flocked to urban areas, particularly around the capital city, in search of job opportunities. This involved mass commuting (mainly public transport and workers’ bus services), which consequently caused serious pollution. However, due to the interpretation of environmental policy at that time, this did not generate social problems[4].

In order to prevent the expansion of large cities, partly because of the hostility towards Budapest in particular and big cities in general, state-initiated new town development concepts also prevailed. Unlike the Western European model, these developments (one of them in the Budapest region) were mainly ideological goals, representing working class interests. However, in fact they provided a chance to exploit the raw materials necessary for accelerating the development of heavy industry. At the same time, through their attractivity they also contributed to the disintegration of the old urban bourgeois society (Szirmai 1991, 1998). The regional development control of new towns until the regime change can be considered successful. The population of such towns steadily increased until 1990. This means that they were able to absorb the immigrating rural population due to the continuous substantial public and private house constructions. Behind the dynamics of flat constructions it is also apparent that, despite the looming crisis, local large companies (including workplaces), which were also in a dire predicament, received significant support from the local government in order to prevent the outbreak of social conflicts.

The failure of the state socialist regional development policy was not caused by the rise of market forces, but rather by the consequences of the crises of the regime which became harder and harder to conceal (the foreign debt of the state, the absence of large companies, the uncertain market conditions, the consequences of the shortage economy). In addition, in large cities with county rank, in particular Budapest, local land development needed autonomy. These were the consequences of the so-called softness of the Hungarian political system, the results of the economic opportunities offered by the Hungarian model. In 1968, the introduction of the so-called new economic mechanism, the state-controlled quasi-market conditions, and the appearance of the second economy reinforced local forces, especially cities with county rank and Budapest. These became

4 The state socialist centralised social governance and political power structure eliminated environmental factors and all the social forces that were interested in protecting the environment. The socialist planned economy did not consider environment as an internal element of the economy. Therefore economic development concepts did not take ecological consequences into account and did not attribute objective value to natural resources. Thus the centralised economic management model operating without market and price mechanisms utilised natural elements free of charge and did not consider the consequent social costs. In this way it could more easily enforce its non-economic power, military and other interests, and it could better conceal the performance limits of the economic system as well as the problems of efficiency. Society was ignorant of both damage to the environment and its adverse effects on health. Social conflicts arising from environmental damages emerged only in the 1970s and 1980s (Szirmai 1993, 1996).

redistributive centres as well as regional, settlement development and resource distribution centres able to assert their own interests.

In the political and social transformation of the 1990s, the state further withdrew from the development of regional processes due to the legacy of the socialist period, the crisis that emerged and a lack of capital. The new local governments were given greater powers. By virtue of the new Local Government Act of 1990, local authorities became responsible for town planning, development and decisions related to land use. The new local authorities, funded by locally generated revenue, shared taxes and the central government's normative contribution, were self-managing. However, the difficulties arising in the development of a market economy and the deficits of the State Treasury saw the emergence of re-centralisation tendencies already in the 1990s. Not only the re-centralisation of development resources, but also the lack of local capital resources restricted the planning and development opportunities of local governments. As a result, local development decisions were reached to serve land owners' and real estate developers' interests throughout the whole country, including Budapest. This was partly because of the underfunding of local authorities and partly because of the interest in realising additional revenues from land sales and the re-classification of territories. Urban policy-makers in Budapest fully ceded the development formation and investment opportunities to market forces and economic lobbies interested in the development of the city's outer urban neighbourhood. It should be noted that not only large urban centres but also peri-urban areas attracted foreign direct investments and various developments during the change of the Hungarian political system.

Among urban citizens who were discontented with urban rehabilitation projects in central Budapest implemented in isolation due to signs of crisis in the inner city area (traffic anomalies, pollution, the increasingly visible signs of urban poverty), the long-suppressed unsatisfied needs for suburban individual residential houses gradually intensified. This was in part because of the contemporary housing policy and the characteristic features of the housing market. Their needs boosted the demand for moving out of cities and gave rise to suburbanisation processes.

Almost all the professional groups concerned and relevant professional documents warned of the foreseeable adverse consequences of the unrestrained territorial expansion. Nevertheless, no regulations with a powerful set of tools were made with the intention of controlling or the ability to influence this 'spontaneous' process. The suburban development policy of Budapest was clearly contradictory. These contradictions are also explained by the fact that, although ideas of limiting expansion appeared in plans at the level of the capital city (as well as state or regional level), the provision of resources and the assignment of legal and incentive instruments essentially failed.

The regional development policies and programmes approved in this period not only showed the lack of adequate state regulation but also demonstrated that there were no incentive mechanisms to encourage efficient land use. The regional level development approach did not develop either. This process resulted in the unstructured and uncoordinated land use practice that developed in the Budapest agglomeration area (especially in its outskirts) which was characterised by the uncontrolled expansion of urban areas, along with the deterioration of territorial potentials and by the oversupply of potential development areas.

Public housing and credit policies did not assign development priorities either. They could not significantly influence suburban development in which private equity established in part as a result of land and housing privatisation, the site-selection strategy of global companies and new demands for moving into the suburban zone were dominant.

The legislation that restricted the unlimited expansion of new 'areas to be built in' was delayed by more than a decade. In 2005 Parliament passed the Act on Spatial Planning in the Agglomeration of Budapest (Act LXIV of 2005), which, intending to coordinate regional land use, set the possibility of further increasing the ratio of built-in areas in the administrative area of municipalities at 3%. However, these overdue restrictions were ineffective because the local authorities concerned reclassified large areas either before the legal limitation or on hearing the news of its possibility, and in these areas developments and investments started later at random.[5]

The amendment of LXIV Act of 2005 on Spatial Planning in the Agglomeration of Budapest[6] further tightened the eligibility for including new areas and reduced the possibility of development to 2%. The amendment contains several new elements encouraging local authorities in the region to take greater care of the landscape, and natural and cultural values, and urging more thoughtful and efficient land management. However, because the Act does not revoke the previous ill-considered local decisions, its efficiency – and impact on real processes – would only be appropriate if the instruments of land use control were extended by development, support and incentive elements which served sustainable land management purposes and at the same time redefined interest relations.

The state, including the capital city's official actors, essentially ceded the suburban residential (including condominium) and other economic development-related areas to market players on the one hand, and on the other hand to the middle classes who were unsatisfied with the city centre and wanted to move out, and to the lower classes forced to leave the large city due to gradual impoverishment. In this way the capital city's social problems and certain tensions of urban poverty were relocated to the outskirts (forming the subject of our research) and less developed parts of the city.

Unlike the French case, social problems concentrated in the surrounding zone of Budapest are more the cumulative tensions of residential disadvantage and social backwardness, thus they are rather the problems of poverty[7] and not of immigration. Spatial conflicts arising from the immigration of foreigners in Hungary are not or hardly visible, which is partly due to the lower ratio of foreign immigrants in comparison to Europe (6.5% of the EU population are foreigners, while the Hungarian average is 2.0%

5 The size of reserve areas can be characterised by the fact that on the basis of land use surveys carried out in 2010 the total size of 'unused' areas designated for residential and municipal purposes in local land use plans was 6,740 hectares, which is 9.2% of the total area designated for this purpose. Of the area designated for economic use (7,450 hectares), about one third (35%) of the area designated for this purpose was still 'unused'. According to the investigations of PESTTERV Kft., of the municipal area of land designated, 15%, a total of about 14,000 hectares, was regarded as territorial reserves during this period (Schuchmann, 2011).

6 See Act LXXXVIII of 2011 on the amendment of Act LXIV of 2005 on Spatial Planning in the Agglomeration of Budapest.

7 For the sake of completeness, it should be noted that regional poverty is more prevalent in eastern Hungary, especially in rural areas and to a lesser extent in urban areas, and it is less of a regional phenomenon in Budapest.

(Vasileva, 2011)). The study on gentrification indicates a ratio of 5% in Budapest, but this is partly because segregated zones of immigrants have not been established in the Hungarian capital region as they have in Vienna or even Copenhagen.

Suburbanisation and gentrification

The social aspects related to the process of urbanisation are the expression of Fordist economic development, and of the spatial needs and interests of the resultant consumer society; the well-off and expanding middle class wanting to live in family-owned houses with a garden in a clean natural environment.

In the 1950s in the United States 35 million people, nearly a quarter of the population, were suburban residents. This ratio was over 30% at the end of the decade (Hobbes–Stoops 2002[8]) and in the 1970s and 1980s this figure continued to increase. Statistics show that this trend has continued and today about half of American society lives in suburbs.

The ratio of urban population further decreased and the ratio of peri-urban population continued to grow in big American metropolises between 1990 and 2000 (from 46% to 50%). Suburban population ratios are much higher than the rates of urban population and the vast majority of the population of urbanised regions lives in suburbs (See Table 1).

Table 1

USA – Change in the rate of urban and suburban population in 1990 and 2000

(Per cent)

Metropolitan region	1990		2000	
	urban rate	suburban rate	urban rate	suburban rate
New York-Northern New Jersey-Long Island	37	63	38	62
Los Angeles-Riverside-Orange County	24	76	23	77
Chicago-Gary-Kenosha	34	66	32	68
Houston-Galveston-Brazoria	44	56	42	58
Philadelphia-Wilmington-Atlantic City	27	73	25	75
San Diego	45	55	43	57
Detroit-Ann Arbor-Flint	20	80	17	83
Dallas-Fort Worth	36	64	33	67
Phoenix-Mesa	44	56	41	59
New York-Northern New Jersey-Long Island	71	29	72	28

Source: Based on http://www.demographia.com/db-2000metrocoreshare.htm and http://eh.net/encyclopedia/article/Smith. Urban. Decline.doc, edited by the authors.

However, trends in Europe are different. In major European metropolises the ratio of urban population in many cases is well above suburban population ratios. (This is particularly true of the capitals of former socialist countries like Prague, Warsaw and Budapest, but also of other capitals such as Berlin, Vienna, Brussels, London and Rome.) There are exceptions, such as Copenhagen and especially Paris, where the proportion of the suburban population is much higher.

8 http://www.census.gov/prod/2002pubs/censr-4.pdf

In the last few years there have been spectacular changes in the share of the population of the metropolitan regions of several European cities. For example, the proportion of the suburban population rose by a few per cent in Madrid, Berlin, Rome, Vienna, Warsaw, Prague and Ljubljana between 1996 and 2004. The biggest increase took place around Budapest (see Table 2).

Table 2

European Union – Change in the rate of urban and suburban population in 1996 and 2004

(Per cent)

Metropolitan region	1996		2004	
	urban rate	suburban rate	urban rate	suburban rate
London	61	39	62	38
Paris[a]	19	81	19	81
Brussels	55	45	56	44
Copenhagen	27	73	28	72
Madrid	56	44	53	47
Berlin	70	30	68	32
Rome[a]	75	25	74	26
Vienna	75	25	73	27
Warsaw	65	35	64	36
Prague	62	38	60	40
Ljubljana	55	45	54	46
Budapest	76	24	71	29
Stockholm	45	55	41	59
Dublin	35	65	30	70
Bratislava	73	27	71	29
Tallinn	77	23	75	25
Riga	76	24	73	27
Vilnius	79	21	78	22
Helsinki	47	53	46	54
Sofia[b]	96	4	86	14
Athens	20	80	20	80
Lisbon	27	73	22	78
Bucharest	91	9	90	10

a) Data of the 2001 Urban Audit.
b) Data of the 1991 and 2001 Urban Audit.

After peri-urban population growth following suburbanisation (i.e. the formation of the first suburbs, the settlements close to the urban centre) and later peri-urbanisation (i.e. the formation of settlement rings located further away), a new trend can be seen in almost all the big European cities of our study. For example, in Vienna, particularly in some inner city districts, the phenomenon of re-urbanisation has emerged in the last few years (Fasmann 2004, Görgl et al. 2011, p. 10). The 2007 statistics for Budapest show that the outflow from the capital also seems to have slackened and the population in the city has started to grow (Schuchmann 2011; Szirmai, et al. 2011). The Paris case study also reveals that 'the last ten years have shown a reversal in the trend. The return to the city centre has been particularly noticeable, demonstrated by an increase of 3% in the population of Paris and of 6% in its inner core' (Cattan 2011, p. 5).

However, the sources of population growth are often uncertain. In Budapest, for example, the absence of recent statistical data makes it very difficult to judge whether the reason for the increase is moving back from the urban neighbourhood or the concentration of population in metropolitan areas, which is experienced nationwide.

As regards Vienna, the explanation is clear as 'the population growth is generally based exclusively on migration processes … the growth is the result of immigration' (Görgl et al. 2011, p. 10).

In any case, it can be stated that urban population growth is obviously generated by gentrification processes as well. In European cities, including those we analysed and presented, these are increasingly characterised by burgeoning suburbanisation processes, which occur either in parallel or subsequently. In Kesteloot's opinion this is because 'the socio-spatial history of European cities has been characterised by tension between centrifugal and centripetal forces that, for the sake of simplicity, can be termed suburbanisation and gentrification' (Kesteloot 2010, p. 126). Essentially the same mechanisms are at work behind the process which formulates urban sprawl as well and thus the declining role of the state and the rise of market forces. However, these are not the aspects of groups interested in urban sprawl but of groups interested in urban renewal. Franz points out this process in the presentation of the gentrification process in Vienna: 'The Austrian welfare state system has changed over recent years, and now tendencies that are much more market-driven are being observed both in Vienna and elsewhere' (Franz 2011, p. 2). The factors originating from the rejection of suburban lifestyle, and from the reorganisation of cultural and consumer demands also form part of the dynamics. The study on Hungarian gentrification remarks that the better educated middle class consumers have demands that cannot be satisfied by the ordinary hypermarkets of the suburbs but only in the inner city (Csanádi, Csizmady and Olt 2011, p. 4).

The renaissance of historic city centres is taking place in a highly differentiated manner from country to country. The literature on this topic shows that the differences between European and North American cities are particularly significant. The renovation of the inner districts is fostered by various social groups and not just yuppies. Instances of this are global city centres such as New York, London and Tokyo (Sassen 1991). Different groups renew the inner districts using various intervention tools, which creates a differentiated built environment with different consumer opportunities and different social content (van Criekingen 2003, p. 96).

The urban renewal examples of Vienna and Budapest, which are described in the book, demonstrate local characteristics which originate from the historical conditions of urban development and from the belated nature of urban development in Hungary. The Austrian study deals with the post-renovation conditions, the purified phenomena of the physical and social fabrics, while the Hungarian case study records the trends of transformation and the currently emerging processes, and also demonstrates the rules of convergence. The expansion of the competitiveness of the two cities after a complete renovation and the rather segregated phenomena of the social composition of the residential neighbourhood are the reasons why the authors investigating the renewal of city centres, view the phenomenon of gentrification as a more positive process than the heavy outflow scenario, which strengthens social segregation as higher social classes move out from the city centre to suburban gated communities, leaving the poor behind.

The evaluation of the segregation effects of gentrification and suburbanisation is far from clear-cut: there is much research to show that urban renewal is not only followed by the continued residence or return flow of the higher social classes but also causes real estate prices to rise, displacing the poorer social groups from inner parts mostly to low-status suburbs or peri-urban settlements.

For example, real estate prices in the inner districts of Paris increased dramatically, tripling in the ten years after 2000. By the end of this period housing prices reached an average of 7,000 euro per square metre in 'Beaux Quartiers', that is the best neighbourhoods (Bronner, 2010, p. 9). This reinforces the existing socio-spatial hierarchy: binding the middle classes to metropolises, increasing the outward movement of poorer, lower social status groups (not just the uneducated, but rather people with a lower income), including school leavers and lower-middle class families, and accelerating the exclusion affecting the poor classes.

Suburban societies

The theatre and film examples mentioned briefly at the beginning of the study drew attention to two different suburban social formations. Both are types of mostly middle-class suburban societies, one with a lower and the other with a higher social status. Of course, these two types do not cover the spatial-social polarisation of the metropolitan societies of our time, which is much more differentiated today. For example, in Fassmann's and Hatz's opinion the city of Vienna 'is characterised by three separate social milieus, which are segregated from each other … The first sector is the city of the rich and educated, who benefit from the globalised economy … the second city refers to the city of the marginalised groups, consisting of the unemployed, an underclass dealing with multiple social problems like poverty, homelessness and drug abuse, and, finally, specific groups of immigrants … Last but not least, the third city has to be mentioned, the "normal" city providing work, supplies and housing. Its population consists of "ordinary" people, neither particularly affluent and well-educated nor extremely poor or marginalised. It is the city of the middle class … it comprises the "working-class" districts of the city itself' (Fassmann–Hatz 2007, pp. 75–76).

The problems of disadvantaged suburbs are topical due to the increasingly polarised urban areas and the concentrations of social tensions. In French society in recent years special political and scientific attention has been paid to this kind of urban problem. The reasons for this include the excessively high peri-urban population rate and the brutal 2005 Paris suburban riots and demonstrations, as well as the fact that the majority of French people regard the suburbs, including the suburban housing estates, as embodiments of social deviance[9] (Vieillard-Baron 2008).

9 According to French researchers, the influence of the media is particularly important in the development of a negative image, but it is also true that some professional groups such as architects, researchers and politicians have given the public an unfavourable or rather one-sided picture firstly of residential areas and then of suburbs.

The starting point of the housing-related social stigma was the so-called 'Sarcessism', i.e. the criticisms related to Sarcesses–Lochére, a new quarter built 15 kilometres from Paris in the 1960s, which over-emphasised isolation, the gigantic dull built environment and urban crime, which concentrated there. Contemporary studies, however, suggest that a significant part of the population in housing estates proved to be pleased with their place of residence, except where crime was significant (Vieillard-Baron, 2008, p. 28). Meanwhile actual ongoing processes also explained the negative attitude towards housing estates (Le-Goaziou–Rojzman 2006, p. 36). Discontent mainly among the middle class had been caused by post-war housing shortages and then the housing estates built as a response to the needs of the demographic explosion, the baby boom, because the necessary residential infrastructure was poorly provided and the monotonous built environment was disliked. For this reason higher-status social classes started to leave the new quarters in the 1970s and 1980s and moved back to urban centres that had been renovated in the meanwhile, neighbouring villages in a better ecological and social position, better suburban areas or new cities. From the mid 1960s an increasing number of foreign immigrants moved into housing estates, in part at the behest of the authorities and in part from the 'bidonville', that is the slums, created spontaneously by North and West African immigrants.

In the 1960s large companies (especially car factories) built suburban districts in Paris in part for French employees, but mostly for foreign immigrant, unskilled, uneducated workers arriving from North Africa, especially Algeria (Le Goaziou–Rojzman 2006, p. 21). Now the second and third generations of former migrants live there who, similarly to the first generation, are unskilled, uneducated and mostly unemployed.

About 5 million people in France live in the so-called 'Sensitive Urban Areas' (Zones Urbaines Sensibles, ZUS), in socially problematic peri-urban areas. The proportion of migrants concentrated in these areas is above average, nearly twice the average for French cities. Almost 40% of the residents of the sensitive areas did not complete their schooling and they account for a significant part of the unemployed today. In 2009 the unemployment rate there was nearly 18.6%, double the national average of 9.5%. They are mostly young people, mainly men (Le Goaziou, Rojzman 2006, p. 24, Bronner 2010, p. 9).

The uncertainties of the future, the problems of everyday living and residential social tensions are the sources of a number of specific conflicts. There are conflicts between the older and younger generations, which stem from the older generation's fear of the rowdy gangs of young people 'living' in the streets. (To illustrate that this phenomenon is general and not exclusively French, it is worth mentioning the 2009 British film *Harry Brown* directed by Daniel Barber. In this rather shocking thriller youth gangs terrorise the mostly retired residents of run-down housing estates on the outskirts of London.) Confrontations, sometimes quiet but sometimes loud, develop between the active and the unemployed, which often escalate into ethnic problems. The residential conflicts are essentially of a social, structural nature, sometimes between the immigrants or their descendants and the 'indigenous' French, but more often between minority suburban lower middle classes, middle classes and the majority underclass, the excluded.

Almost all the case studies presented here describe the problems of higher status, mostly middle-class suburbs. This kind of suburban model in Europe and North America was created as a result of different housing policy support schemes mostly for the middle-class and in many cases the upper-middle-class population (Jaumain–Lemarchand 2008). The developments intended to take into account the requirements of metropolitan expansion, the crisis of urban centres, the functional transformation of the suburbanisation of the middle class, including their aspirations for differentiation, car traffic and the needs of a family-based lifestyle (Harter 2008). This model is a result of social and structural inequalities, and the spatial separation of the upper and the lower social classes. In America, the model was shaped by ethnic contents as well due to the contradictions between the white upper and upper-middle class populated suburbs, the American 'mainstream', and the low social status (mostly inner) urban districts inhabited by problematic, often dark-skinned Americans, the so-called 'others' (Imbroscio, Williamson, Alperovitz 2005, p. 318).

The suburbs of the Budapest region can be divided into two types, which by using the case study, the statistical analyses and other empirical methods can also be defined as 'developed' and 'underdeveloped' peri-urban settlements. This division is based on social structural characteristics and other factors, such as infrastructure, housing and institutional coverage. The two types clearly differ in terms of the composition of local societies and the people moving in, the characteristic features of employment and commuting, and the economic processes of settlements. People moving to advanced villages are mainly members of higher status social groups (secondary and tertiary graduates, knowledge workers with high incomes), while underdeveloped rural areas typically attract newcomers who are blue-collar workers on low incomes with primary school education.

Ageing is one of the current problems of high status suburbs, which is not only a European but also a global phenomenon. This is proved, among other things, by a Canadian/French comparative analysis. In Canada in the 1950s and 1960s, and in France in the 1970s mainly young people, who were on average 25 years old and had a family, moved to suburban areas. They were the baby boomers, the generation of people who mostly entered the housing market around 1970 (Gill 2008, Luxembourg 2008). This group included people who no longer wanted to live with their parents, wished for their own home and desired to be independent.

However, the first half of the 2000s brought significant changes. By 2006 the baby boomers, typically graduates with high incomes, entered their sixties. In the meantime their children had left home and moved back to the city centres, partly due to the renewal of inner cities. This radically transformed the structure of the suburban households, significantly reducing the proportion of parents living together with children and increasing the ratio of single or single-parent households, which in part was caused by divorce. As a result of gentrification, the social structure also changed. For example, two thirds of active graduates, especially women aged between 25 and 54 years living alone in the Montréal metropolitan region, chose an urban form of life, living in the city as the urban environment was better suited to their contemporary lifestyle (Gill 2008, p. 61).

This is also the reason why the author of the Canadian case study asks how suburbs which were built mainly for the family life of the highly skilled middle classes with high incomes will be sustainable. The newly arrived young families have lower incomes and are less able to afford the suburban housing market, and it is unlikely that there will be a repeat of the higher income classes moving out from the city centres.

In the Vienna suburban area the problem of ageing mainly affects ‘the first generation of suburbanised settlements’. In inner city areas the concentration of people over the age of 60 is greater, while the proportion of younger age groups in the suburbs is significantly higher. In the first generation of suburbanised settlements parents who were young when they moved out in the 1970s and 1980s will slowly reach retirement age. This is another reason why the authors ask: ‘The question is how these communities will deal with the “collective ageing” of the original suburbanites. What will become of the many single-family homes when the children have moved away or when the “founding fathers and mothers” themselves are no longer alive? Likewise, how will the infrastructure adapt to the ageing society and who will assume the financial costs thereof? In particular the questions of caring for the elderly as well as retooling public institutions to cater for less mobile inhabitants will have to be answered satisfactorily’ (Görgl et al. 2011, pp. 16–17).

The Hungarian case study reveals the different trend of ageing in the developed and the undeveloped urban neighbourhoods. It points out that, apart from Budapest where ageing is higher than the national average, the phenomenon of ageing mainly affects settlements with less favourable environmental and infrastructural conditions. However, the population of the developed urban areas is younger as there is an inflow of young, active, well-qualified families. This may also be due to the belated, ‘catching up’ character of Hungarian suburbanisation processes. Possibly this is because the processes which formed the younger demographic structure of the first generation suburbs in developed European and Canadian metropolises are occurring in Hungary right now.

Urban and suburban dichotomies

Among the early and modern criticisms of urban sprawl several authors have warned of the threats of the disappearing compact city and the evolution of the fragmented city. In my view this problem is less relevant in the context of the modern European metropolitan spatial structure. This is because we can no longer speak of exclusive compact cities. In modern regional systems compact and disperse settlements and settlement groups exist simultaneously. The former consist of communities developed in a more concentrated manner and characterised by a densely built environment in a social sense which have transparent if not integrated community systems. The latter, on the other hand, can be characterised by developing in a more decentralised, horizontal way, with a family house environment and low population density, consisting of sub-groups in a social sense with more fragmented communities.

The Hungarian case study investigated the lack of clear models, which it confirmed. Different models exist side by side, complementing each other and showing how many other social and business relationships exist within the regional system examined. Furthermore, these relationships are formed within the structures of space ‘consumed by’ the everyday activities of social status groups of different spatial position (Szirmai et al. 2011). The Danish case study provided other kinds of evidence by presenting how spatial relationships are formed between the city and the agglomeration as well as spatial relationships that go beyond these (often crossing national borders), the cooperation systems established with various cities, and the big spatial axes of today’s modern lifestyle (Reeh, Zerlag 2011).

Compared to the compact or dispersed urban dichotomy, however, I believe that the modern hierarchical order of the societies of metropolitan areas is more important and socially more problematic. Aspects of this are the re-organised and even strengthening spatial social dichotomy, which is a negative side effect of urban sprawl, and the antagonisms between urban centres and outskirt areas.

Traditional and new segregation zones (enclaves) can also be seen in the Paris region. On the western edge the concentrations of suburban settlements are inhabited by high income, managerial classes, while the south-eastern area is populated by lower status settlement groups (Cattan 2011, p. 13).

The Vienna case study also emphasises the growth of the social separation process and its threats. Among the reasons for this is that for residents choosing the suburbs of Vienna, mainly to the south of the city, not only is easy accessibility essential, i.e. that the village should be close to the city centre, but so are land prices, that is to say the prices should be high enough to ensure a neighbourhood with people of similar social status and income (Görgl et al., p. 6).

A typical realisation of sociological rules is illustrated by the similar hierarchical arrangement of the different educational groups in the Hungarian and French capital cities. In the Paris Metropolitan Area (see Figure 6), the farther we go out from the regional centre (from the municipality of Paris) towards outer zones, the lower the ratio of people with tertiary education, and the higher the ratio of classes with average and poor education. The social structure of Budapest (and even of other Hungarian cities) also shows this hierarchical arrangement. In central areas the presence of higher social status is stronger (better educated, higher income groups with higher positions), while in outer districts and outskirt areas the ratio of lower social status (lower educated, lower income groups with lower positions) is more characteristic.

Conclusions: differences and/or similarities?

The case studies and their related literature state that the phenomena of European metropolitan areas are formed by urbanisation processes resulting from historical contexts and global trends. By this they prove that the historical roads or even the stages of urbanisation cannot really be bypassed, and less developed countries follow the paths designated or taken by those ahead of them due to the enforcement of the rules of global urbanisation. Moreover, these rules can sometimes override the objectives and the effects of conscious state-initiated regional and local development interventions. In the absence of this aspect, it would be hard to explain the European and Hungarian suburban social similarities, and the dichotomies between high and low social status settlements, which developed in different social and historical contexts in different economic conditions and by the activities of actors shaping different processes.

Global trends can also be explained by other processes causing convergence, namely the decreasing role of the state and authorities in regional development across Europe, the growing influence of market forces and the interest groups involved in market mechanisms. This convergence (on the basis of a relatively strong abstraction) was perceptible even in the state-socialist period as the state's decreasing role in regional development correlates with the crisis of the socialist type of welfare state, with the strengthening of local economic and political forces, and with the emergence of limited, state-controlled yet rising market forces.

The phenomenon of urban sprawl has been detected in all the investigated metropolitan areas in the stage of relative deconcentration. This has been explained by Enyedi (according to the Dutch school) as a process of suburbanisation, yet it is also clear that there is more to this case, since the process of re-urbanisation can also be felt here. The fact is that the suburbanisation process is accompanied by the phenomena of gentrification and they are complementary, parallel great mechanisms. Due to the absence of comprehensive studies, I cannot comment on the big questions raised by Enyedi, namely whether the return to city centres is merely a rearrangement of the population in urban agglomerations or whether it is a new stage of urban growth? The case studies presented here, however, clearly show that the population in urban centres (particularly in some inner city neighbourhoods) has tended to grow in recent years. There were several reasons for this: moving back, other regional social processes of concentration, and the increasing sense of well-being of the metropolitan middle classes in the renewed inner district, hence their disinclination to leave. Another contributing factor is certainly the fact that quite a few suburban residents are disappointed with the suburban way of life.

This disappointment can be explained by several reasons, including the adverse environmental and social impacts of urban sprawl, and the dependencies and inconveniencies of a commuting lifestyle. The development of segregation patterns is particularly relevant because the reasons mentioned indicate a problem, but not necessarily from a scientific viewpoint. Although segregation is usually regarded as harmful, separation from other positioned groups and isolation from social problems is a common expectation of the suburban middle classes. Naturally, they do not always achieve this despite the well-structured and spreading gated communities, as the urban poor are also located in the periphery of the metropolitan areas. These different residential groups, which have historically competed with each other for ‘elbow room’ in different ways and to varying degrees, supersede those in different situations or attract those in a similar situation.

Other reasons for the disappointment are related to market price mechanisms and to the fact that decisions on market development serve several people’s interests but not the local interests. A certain part of the population, which in post-socialist countries is the majority, cannot always afford and do not always want to pay for the costs of the infrastructure accompanying suburban development, for instance those of the road network or public transport, which neither the state nor the local government can afford or intends to pay.

However, the social significance of the suburban way of life cannot be questioned and the demand for individual family housing will continue to be significant (perhaps this is most apparent in the Danish case). Today there are still quite a few people intending, or at least desiring, to move out of the cities, undoubtedly because of the existing social and environmental values of living there and because suburban life is still attractive for them despite the fact that it shows symptoms of crisis. The current economic crisis (for example, the home loan debt in Hungary) will seriously restrict demand for suburbanisation, and the case of Budapest shows that relocation efforts have considerably decreased over the last few years.

Predicting the future processes of suburbanisation is not easy, not only because of the simultaneity of conflicting trends, that is to say that the effects of re-urbanisation are emerging during suburbanisation, but also because of the appearance of other major European processes. It must be remembered that the European (including the Hungarian) population is not only decreasing but at the same time European society is ageing, and the ageing phenomenon (especially of those in disadvantaged social positions) is a further factor limiting spatial migrations. Thus, today there is more uncertainty than clarity. But this might not matter, because if everything could be taken for sure, who would need science?

REFERENCES

Bronner, L. (2010): La politique de la villeimpuissante á résoudre les difficultés des quartiers. *Le Monde* 16 Octobre.

Castells, M. (1972): *La question urbaine*. François Maspero, Paris.

Cattan, N. (2011): Urban Sprawl in the Paris Metropolitan Area. In: Szirmai Viktória (ed): *Urban sprawl in Europe*. Aula Kiadó, Budapest.

Csanádi Gábor – Csizmady Adrienne – Olt Gergely (2011): Urban Renewal and Gentrification in Budapest City Centre. In: Szirmai Viktória (ed.): *Urban sprawl in Europe,* Aula Kiadó, Budapest.

Davis, M. (2007): *Le pire des mondespossbiles. De l'explosionurbaine au bidonville global.* La Découverte/Poche, Paris.

Enyedi György (ed.) (1998): *Social Change and Urban Restructuring in Central Europe*. Akadémia Kiadó, Budapest.

Fassmann, H. (2004): *Stadtgeographie 1. Allgemeine Stadtgeographie*. Westermann, Braunschweig.

Fasmann, H. – Hatz, G. (2007): The Austrian case study: Social Inequalities in the Vienna Metropolitan Region. In: Szirmai Viktória (ed.): *Social Inequalities in Urban Areas and Globalization. The Case of Central Europe*. Based on the results of the project Urban Areas, Socio-spatial Inequalities and Conflicts – The Socio-spatial Factors of European Competitiveness. pp. 51–76, Discussion Papers Special, Pécs.

Fishman, R. (1990): Megalopolis Unbound. *The Wilson Quarterly* 14 (1): pp. 25–45.

Franz, Y. (2011): Gentrification Trends in Vienna. In: Szirmai Viktória (ed.): *Urban sprawl in Europe*. Budapest, Aula Kiadó.

Gill, D. (2008): Modes de vie et modes d'habiter. La banlieue québécoise, un modele de moins en moins adapté mais toujours en expansion. In: Jaumain, S., Lemarchand N. (eds.): *Vivre en banlieue: unecomparaison France/Canada*. vol. 14, pp. 53–65. Étudescanadiennes-Canadian Studies.

Görgl, P. – Helbich, M. – Matznetter, W. – Fassmann, H. (2011): Spatial and Social Development Trends of Metropolitan Vienna: An Overview. In: Szirmai Vikória. (ed.): *Urban sprawl in Europe*. Budapest, Aula Kiadó.

Hall, P. (1996): La villeplanétaire. In: *Revue international des sciences sociales*, pp. 19–31. Mars, Nr. 147.

Harter, H. (2008): La suburbanisation á Ottawa: uneapprochehistorique. In: Jaumain, S., Lemarchand N. (eds.): *Vivre en banlieue: unecomparaison France/Canada.* pp. 95-109. Étudescanadiennes-Canadian Studies, vol. 14.

Haumont, N. (ed) (2005): *Les pavillonnaires. Etude psychologique d'un mode d'habitat.*L'Harmattan, Collection, Habitat et sociétés.

Haumont, N. – Jalowiecki, B. – Munro, M. – Szirmai, V. (1999): *Villesnouvellesetvillestraditionelles. Unecomparaison internationale.*Edition l?Harmattan, Paris.

Häussermann, H. – Haila, A. (2010): 'The European City: A Conceptual Framework and Normative Project.' In: Kazepov, Y. (ed.): *Cities of Europe 2010*. pp. 43–65. Blackwell Publishing, Oxford.

Howard, E. (1902): *Garden Cities of To-Morrow.* London.

Jaumain, S. – Lemarchand, N. (2008): *Vivre en banlieue: Unecomparaison France/Canada.* Peter Lang Pub.

Kazepov, Y. (ed.) (2005): *Cities of* Europe*: Changing Contexts, Local Arrangements, and the Challenge of Urban Cohesion.* Blackwell, Oxford.

Kesteloot, Ch. (2010): Urban Socio-Spatial Configuration and the Future of European Cities 2010. In: Kazepov, Y. (ed.): *Cities of Europe 2010*. pp. 123–149. Blackwell Publishing, Oxford.

Kopp, A. (1979): *Architecture et mode de vie, textes de annéesvingt en U.R.S.S.* Presses Universitaires de Grenoble.

Le Goaziou, V. – Rojzman, C. (2006): *Les banlieues.* Le Cavalier Bleu, Paris.

Le Goix, R. (2004): *Gated communities: Sprawl and social segregation in southern California*. http://hal.archives-ouvertes.fr/docs/00/03/58/67/PDF/legoix20041012.pdf

Luxembourg, N. (2008): La mobilité des personnes ágées en banlieue pavillonnaire. Regard croisés France-Québec. In :Jaumain, S. – Lemarchand, N.: *Vivre en banlieue: Une comparaison France/Canada.* pp. 65–79. Peter Lang Pub.

Merlin, P. (1972): *Les villes nouvelles: urbanisme régional et aménagement.*

Munoz, F. (2003): *Urban Sprawl in Mediterranean Cities.* Cities 20(6): 381–385. Presses Universitaires de France, Paris.

Reeh, H. – Zerlang, M. (2011): Fairytales from the Fingerplan: Representing Strategies for Suburban Copenhagen. In: Szirmai Viktória (ed.): *Urban sprawl in Europe.* Budapest, Aula Kiadó.

Rojzman, C. – Le Goaziou, V. (2006): *Les Banlieues.* Cavalier Bleu.

Sassen, S. (1991): *Global City: New York, London, Tokyo*. Princeton University Press.

Schuchmann Péter (2011): *A Budapesti Agglomeráció Területrendezési Tervéről szóló 2005. évi LXIV. törvény módosításának tervezete* [Draft to the Amendment of the Act on Spatial Planning in the Agglomeration of Budapest]. PESTTERV Kft., Budapest.

Szirmai Viktória (1991): New Towns in Hungary. (The future of the "Socialist" towns) In: P. Merlin – M. Sudarskis (ed.): *From garden city to urban reconstruction: New Towns in perspective.* pp. 101–119. INTA Press.

Szirmai Viktória (1993): The Structural Mechanisms of the Organisation of Ecological-social Movements in Hungary. In: Vári, A. – Tamás, P. (ed.): *Environment and democratic transition, Policy and Politics in Central and Eastern Europe.* pp. 146–157. Kluver Academic Publishers, Dordecht, Boston, London.

Szirmai Viktória (1996): Protection of the Environment and the Position of Green Movements in Hungary. In: Pickvance Láng, K.-Manning, N.-Szirmai, V.-Pickvance, Ch. (ed.): *Environmental and Housing Movements. Grassroots Experience in Hungary, Russia and Estonia.* pp. 23–88. Avebury.

Szirmai Viktória (1998): "Socialist" Cities (New Towns) in a Post-Socialist Era. In: Enyedi György (ed.): *Transition of Post Socialist Cities in Central Europe.* CEU Press, Pécs.

Szirmai Viktória (2007): *Social Inequalities in Urban Areas and Globalization. The Case of Central Europe*. Based on the results of the project Urban Areas, Socio-spatial Inequalities and Conflicts – The Socio-spatial Factors of European Competitiveness. Discussion Papers Special, edited by Szirmai Viktória, Pécs.

van Criekingen, M. (2003): La villerevit! Formes, politiques et impacts de la revitalisation résidentielle a Bruxelles. In: Bidou-Zacharisen, C. (ed.): *Retour en ville - des processus de "gentrification" urbaine aux politiques de "revitalisation" des centres*. pp. 73–107. Descartes et Cie, Paris.

Vasileva, K. (2011): *Population and social conditions. Statistics in focus.* 34/2011. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-SF-11-034/EN/KS-SF-11-034-EN.PDF

Veltz, P. (1996): *Mondialisation, Villesetterritoires* .L'économied'archipel, Presses Universitaires de France, Paris.

Vieillard-Baron, H. (2008): La constuction de l'imaginairebanlieusardfrancais: entre légende noire etlégendedorée. In: Jaumain, S. – Lemarchand N. (eds.): *Vivre en banlieue: unecomparaison France/Canada.* pp. 17-32. Études canadiennes – Canadian Studies, vol. 14.

Williamson, T. –Imbroscio, D. – Alperovitz, G. (2002): *Making a Place for Community: Local Democracy in a Global Era.* Routledge.

Keywords: urban sprawl, metropolitan regions, suburbanisation, middle class, urban planning, Budapest Agglomeration, decentralisation.

Abstract

During the 20th century, profound socio-economic changes typified the metropolitan regions of the World. One of the most outstanding alterations was the relative decentralisation of urban space, that's visible sign is urban sprawl. The following paper intends to introduce the most significant characteristics of this specific phenomenon.

Current urban processes are presented throughout both Northern American and European metropolitan regions, with special emphasis on the similarities and differences, by the help of adequate European examples (Paris, Vienna, Copenhagen, Budapest).

The research is aiming to give multilateral information on suburbanisation and gentrification, moreover the recentralisation or/and decentralisation of metropolitan areas.

A vaste introduction can be found in connection with the Hungarian urban planning system and procedures, especially during the Socialist Era and the Post-Socialist Period. The results of a deep research on the suburbia, suburban societies, the economic welfare of agglomeration settlements are presented concerning the Budapest Metropolitan Region, furthermore the recent effects of the world financial crisis that substantially determined.