1 Uncovering Regional Disparities – the Use of Factor and Cluster Analysis*

Ivana Rašić Bakarić**

Abstract

The aim of this paper is to present a new methodology for the classification of spatial-economic entities in order to support a regional development policy. Such a classification may be obtained through the use of multivariate statistical methods – factor and cluster analysis, and is based on eleven economic and social indicators. The local government units of three Croatian counties are used as the working sample. Those are as follows: the County of Istria, the Lika-Senj County and the Primorje-Gorski Kotar County. Factor analysis leads to identification of a small number of socio-economic dimensions that summarise adequately the information contained in the original set of variables. Cluster analysis is used to look for groups of towns and municipalities with similar levels of socio-economic development. Multivariate techniques were successful in identifying the main axes of socio-economic characterisation and the regions of the observed counties with differing degrees of development. The new methodology for spatial-economic entities' classification enables a much more useful characterisation of the territory for policy-making purposes.

Keywords: factor analysis, cluster analysis, region, regional policy
JEL classification: C49, R12, R58

---

* *This paper was originally published in* Privredna kretanja i ekonomska politika *(Economic Trends and Economic Policy), No. 105, 2005, pp. 52-77.*

** *Ivana Rašić Bakarić, Research Assistant, The Institute of Economics, Zagreb.*

1 Introduction

In this paper we focus on three counties in the Republic of Croatia: the County of Istria, the Lika-Senj County, and the Primorje-Gorski Kotar County. The subject of analysis are the local government units of the above counties. County as a territorial unit is characterized by large differences between its individual regions, from the coastal regions of Istria and Kvarner with islands to inland Istria and the hilly and mountainous regions of Gorski Kotar and Lika. According to the initial proposal for defining the Nomenclature of Territorial Units for Statistics (NUTS), three counties make up one of the regions of the NUTS level II.[1] And it is NUTS level II that represents the main framework for implementation of regional policy and analysis of regional issues. Considering territorial differences, implementation of unique regional policy measures both on the territory of each of the three counties and on the territory of the regional unit comprising the three counties would not yield equal and desired results. As a precondition for formulating the goals and measures of the regional policy, it is necessary to conduct an analysis of social and economic characteristics of a region. This paper wishes to point to the use of factor and cluster analysis in recognizing the characteristics of individual territorial units and in grouping these units into homogenous groups.

In the observed year 2001, the selected counties comprised 86 local government units (JLS). Like the state territory of the Republic of Croatia, regional development issues of these local government units are governed by the Local and Regional Government Act (Official Gazette No. 33/01). There is also a wide range of legislations that cover specific territories of similar characteristics (Act on Areas of Special State Concern, Islands Acts, Act on Town of Vukovar and Act on Hilly and Mountainous Areas). Nonetheless, the selection of these territories was quite arbitrary and as such is not the result of economic analysis of a region's basic

[1] *According to the initial proposal for defining NUTS regions, adopted by the Croatian Parliament, Croatia is divided into five second-level NUTS regions. However, EUROSTAT, the statistical office of the European Union, rejected this proposal and adopted a new one which divides Croatia into four second-level NUTS regions: East Croatia (County of Virovitica and Podravina, County of Osijek and Baranja, County of Vukovar and Srijem, County of Slavonski Brod and Posavina, and County of Požega and Slavonia), Adriatic Croatia, Central Croatia (County of Krapina and Zagorje, County of Varaždin, County of Međimurje, County of Koprivnica and Križevci, County of Bjelovar and Bilogora, County of Karlovac) and the Zagreb Region (City of Zagreb and Zagreb County).*

characteristics. Without such analysis it is possible neither to define and implement regional policy measures nor trace their effects.

The main hypothesis of the paper is to show that by using the methods of factor and cluster analysis we may analytically single out those local government units that represent a more suitable subject matter of regional policy instruments and measures.

Results lead to the identification of the main dimensions of socio-economic development of the observed units (factor analysis), and the grouping of the units of similar dimensions into clusters (cluster analysis).

2 Local Government Units of Selected Counties

In 2001, the three observed counties comprised 28 towns and 58 municipalities – a total of 86 local government units.

The County of Istria comprises 10 towns (Buje, Buzet, Labin, Novigrad, Pazin, Poreč, Pula, Rovinj, Umag and Vodnjan) and 29 municipalities (Bale, Barban, Brtonigla, Cerovlje, Fažana, Gračišće, Grožnjan, Kanfanar, Karojba, Kaštelir – Labnici, Kršan, Lanišće, Ližnjan, Lupoglav, Marčana, Medulin, Motovun, Oprtalj, Pićan, Raša, Sveta Nedelja, Sveti Lovreč, Sveti Petar u Šumi, Svetvinčenat, Tinjan, Višnjan, Vižinada, Vrsar, Žminj). The county seat is Pazin.

The Lika-Senj County comprises 4 towns (Gospić, Novalja, Otočac, Senj) and 8 municipalities (Brinje, Donji Lapac, Karlobag, Lovinac, Perušić, Plitvička Jezera, Udbina and Vrhovine). The county seat is Gospić.

The Primorje-Gorski Kotar County comprises 14 towns (Bakar, Cres, Crikvenica, Čabar, Delnice, Kastav, Kraljevica, Krk, Mali Lošinj, Novi Vinodolski, Opatija, Rab, Rijeka, Vrbovsko) and 21 municipalities (Baška, Brod Moravice, Čavle, Dobrinj, Fužine, Jelenje, Klana, Kostrena, Lokve, Lovran, Malinska-Dubašnica,

Matulji, Mošćenička Draga, Mrkopalj, Omišalj, Punat, Ravna Gora, Skrad, Vinodol Municipality, Viškovo, Vrbnik). The county seat is Rijeka.

Like the state territory of the Republic of Croatia, regional development issues of these local government units are governed by the Local and Regional Government Act (Official Gazette No. 33/01). In the selected counties, the local government units situated either on territories of special government concern, on islands, or in hilly and mountainous regions are additionally regulated by special legislation (Act on Areas of Special State Concern, Official Gazette Nos. 44/96, 57/96, 124/97, 73/00, 87/00, Islands Act, Official Gazette No. 34/99, Act on Hilly and Mountainous Regions, Official Gazette Nos. 12/02, 32/02, 117/03).[2] The Act on Areas of Special State Concern additionally governs the following local government units: municipalities of Donji Lapac and Plitvička Jezera, which belong to the first category of special state concern areas; the town of Gospić and the municipalities of Lovinac and Udbina, which belong to the second category of special state concern areas; and the municipalities of Brinje, Brod Moravice, Grožnjan, Lanišće, Oprtalj, Perušić, which belong to the third category of areas of special state concern.

Covered by the measures stipulated by the Act on Hilly and Mountainous Areas are Buzet, Cerovlje, Gračišće, Lupoglav, Motovun, Senj, Karlobag, Čabar, Delnice, Vrbovsko, Čavle, Fužine, Jelenje, Lokve, Matulji, Mrkopalj, Ravna Gora, Skrad and the Vinodol Municipality.

The Islands Act regulates local government units that are spread across Istrian archipelagos (Poreč, Vrsar, Rovinj, Brijuni, Pula and Medulin archipelagos), and the local government units on the islands of Krk, Mali Lošinj and Rab, and Novalja on the island of Pag.

[2] *These acts, along with the Act on Reconstruction and Development of the Town of Vukovar and the Regional Development Fund Act, are the five acts that explicitely regulate the issues related to regional development. In addition to these, there is a range of legislations that indirectly regulate individual issues related to development of local government units.*

3 Data Used in the Analysis – Indicators of Socio-Economic Development

The analysis is aimed at grouping towns and municipalities into homogenous groups (clusters) with respect to the prevalence of specific features of socio-economic development. For this purpose, eleven variables – indicators of socio-economic development measured at the level of observed towns and municipalities - were selected as input variables for factor analysis. These are: own local government per capita revenue; persons employed in agriculture and fisheries as percentage of total employment; persons employed in manufacturing as percentage of total employment; persons employed in hotels and restaurants and real estate, renting and business activities as percentage of total employment; unemployment rate; ageing index; vital index; persons without completed primary education as percentage of total population; tourist nights; gross domestic product (GDP) per capita; and persons in employment as percentage of total population. All data refer to 2001.

Table 1 **List of Variables with Respective Codes**

Code	Description
X1	Own local government per capita revenue
X2	Persons employed in agriculture, hunting and fisheries as percentage of total employment
X3	Persons employed in manufacturing as percentage of total employment
X4	Persons employed in hotels and restaurants and real estate, renting and business activities as percentage of total employment
X5	Unemployment rate
X6	Ageing index
X7	Vital index
X8	Persons without completed primary education as percentage of total population
X9	Tourist nights
X10	Gross domestic product per capita
X11	Persons in employment as percentage of total population

Source: Author's research.

Own local government per capita revenues comprise own revenue from property and other non-tax revenue sources as well as own tax revenues. In most counties, towns

and municipalities these revenues are very modest and insufficient for funding local government activities. However, they are a much better indicator of economic strength of local governments than total local government revenues, because they exclude joint taxes and various central government grants. The source of data are local government budgets that are published by the Ministry of Finance.[3] Table 2 contains mean values of variables that have been subjected to factor analysis (arithmetic mean and median) and standard deviation as a measure of dispersion. By looking at the table, a large variance can be noticed in the value of this indicator within observed towns and municipalities. For example, the municipality of Gračišće has the lowest own per capita budget revenue of HRK 42.8, while the municipality of Kostrena has the largest budget of HRK 7,092.00. The own budget revenues of observed local government units in 2001 averaged HRK 1,189.20, while as many as half of the observed local governments generated own revenues of less than HRK 869.53.

As indicators of the economic structure of observed local government units the following were used: persons employed in agriculture, hunting and fisheries as percentage of total employment; persons employed in manufacturing and persons employed in hotels and restaurants, real estate, renting and business activities as percentage of total employment;[4] and the number of tourist nights. The source of data for the first three indicators is the "Census of Population, Households and Dwellings of 31st March 2001."

Persons employed in agriculture, hunting and fisheries as percentage of total employment. This variable indicates the share of primary sector of industry in the economic structure of the observed town/municipality. It is obtained by aggregating the number of persons employed in these two activities and dividing it by the total

[3] *Funding of local and regional self-government units is regulated by the Act on Funding Local Self-Government and Administration. This act sets out the sources of funding for the self-government activities of counties, towns and municipalities, and delegated government administration affairs under special legislation. Financial resources for funding these activities are provided in local budgets as own source (revenues from own property and other non-tax revenues, and revenues from county, municipal and town revenues respectively), as a share in joint taxes and state and county budget grants.*

[4] *The number of persons in employment by branches of industry is collected according to National Classification of Activities (NKD) based on European Classification of Economic Activities (NACE), Rev. 1, which is obligatory for all EU member states. NACE, Rev. 1 has been in use since 1995 in accordance with the European System of Accounts 1995 (ESA 1995).*

number of employed persons in the observed town/municipality. This variable, too, is characterized by large differences in value between observed local government units. In 2001, the lowest employment in these two sectors was registered in Bakar (0.7 percent), with the highest employment reported in Delnice (54 percent). During the same period, the average proportion of persons employed in these sectors in all of the observed local government units was 8.7 percent of total employment (Table 2).

The variable *persons employed in manufacturing as percentage of total employment* in the observed town/municipality indicates the share of secondary sector of industry in the economic structure of the observed local government unit. This variable is obtained in the same manner as the preceding variable, with the difference that the numerator is now the number of persons employed in manufacturing. The lowest proportion of employees in this sector was recorded in the municipality of Brod Moravice (3.2 percent in 2001), with the highest employment in manufacturing reported in Mrkopalj (47 percent). The average value of this indicator for the observed local government units in 2001 was 20.7 percent. In 50 percent of the observed local government units the value of this indicator is lower than 21.9 percent (Table 2).

Persons employed in hotels and restaurants and real estate, renting and business activities as percentage of total employment. This variable measures the share of tourism, as a tertiary-sector activity, in each of the observed local government units. The indicator value is calculated as the sum of persons employed in these activities divided by total employment in the observed local government unit. The value ranges from 3.4 percent (municipality of Vrhovine) to 47 percent persons employed in these two activities (municipality of Vrsar). In the observed towns and municipalities in all three counties an average of 17.0 percent of persons are employed in this sector (Table 2).

Tourist nights is another indicator of the share of tourism employment in total economy. The information on the number of tourist nights for the observed towns and municipalities is derived from the Central Bureau of Statistics (CBS) publication entitled Statistical Report Tourism 1176/2002 for 2001.

The following were used as indicators of the quality of living and working conditions: unemployment rate; persons without completed primary education as percentage of total population; persons in employment as percentage of total population; and gross domestic product per capita.

Unemployment Rate. Unemployment rate was calculated on the basis of standard definition of unemployment rate as the ratio of unemployed to total active population. Active population consists of persons in employment, persons who are actively engaged in some sort of occupation but are not in employment, and the unemployed. Information on active population and the number of unemployed in the observed local government units was obtained from the Population Census 2001. This indicator demonstrates quite a large scope of values (see Table 2). The lowest unemployment rate was measured in the municipality of Sveti Petar u Šumi (6.7 percent), and the largest unemployment rate of 70.0 percent was measured in the municipality of Donji Lapac. In fifty percent of the observed towns/municipalities the unemployment rate in 2001 was 16.4 percent or larger. The average unemployment rate in all of the observed local government units was 17.5 percent.

Persons without completed primary education as percentage of total population. This indicator shows the percentage of uneducated population. The indicator value is obtained as a ratio of the number of persons without completed primary education to total population. The lowest proportion of persons without completed primary education in 2001 was recorded in the municipality of Punat (0.3 percent), with the highest proportion of uneducated persons recorded in the municipality of Vrhovine (13 percent). The arithmetic mean for this indicator is 2.3 percent, the median 1.5 percent (see Table 2). The indicator values were calculated on the basis of data obtained from the "Census of Population, Households and Dwellings of 31st March 2001", CBS.

Persons in employment as percentage of total population. This indicator was calculated using the data on the proportion of persons in employment in total population of the observed town/municipality. As indicated in Table 2, the indicator value ranges between 11.2 percent of persons in employment in total population in the municipality of Donji Lapac and 45 percent of persons in employment in total

population in Buzet. In as many as fifty percent of towns and municipalities in selected counties this indicator was equal to 36.7 percent or higher. In 2001, the observed local government units had an average of 35.6 percent of employed in total population. The indicator was calculated on the basis of the Census of Population, Households and Dwellings 2001 data.

Gross domestic product per capita. In official statistics, information on gross domestic product is available only at national level. The value of GDP at the level of towns and municipalities was obtained on the basis of estimates assuming a constant labour productivity. The proportion of persons in employment in each of the 17 economic activities (according to national classification of activities) in a specific town/municipality is multiplied by gross added value (GAV) of the respective activity at national level. Then, estimated gross added values of all economic activities are added up and accordingly the gross domestic product for the respective town/municipality is calculated, i.e. estimated.[5] This indicator is put in relation to a town's/municipality's total population. If we look at Table 2, we can see that in 2001 the observed local government units generated an average of HRK 37,680.7 per capita. The lowest estimated per capita GDP was generated in the municipality of Vrhovine (HRK 8,968.50), with the highest GDP of HRK 62,207.0 generated in the municipality of Medulin. In 2001, fifty percent of observed towns/municipalities generated a per capita GDP of 39,052.80 or higher. The estimate was made using data from the 2001 Census and the Statistical Yearbook 2002.

Ageing index and vitality index. Out of indicators showing the demographic structure of population, ageing index and vitality index were selected. Ageing index is the proportion of persons aged 60 or older in total population.[6] When the proportion of persons aged "60 or older" reaches 12 percent, the population of the

[5] *Gross domestic product in market prices expresses the value of all manufactured goods and services of resident units, i.e. the sum of added values by economic activity, with items not allocated by activity included into calculation at national economy level: financial intermediation services indirectly measured (FISIM) and taxes on products less subsidies on products (Statistical Yearbook 2004).*

[6] *Formula for calculation of the ageing index:* $x_{(60 \text{ and older})} = \frac{P_{60+}}{P} \times 100$, *where* P_{60+} *is the number of persons aged 60 and older, and P total population.*

country or region is considered to have started ageing (Wertheimer-Baletić, 1999). Vital index[7] is also used to evaluate the age level and ageing of population. It represents a synthesis of indicators of natural trends in the population (mortality (m) and fertility (f) and population age structure (Wertheimer-Baletić, 1999). In 2001, the lowest aging index among the observed local government units of 12.8 percent was reported in the municipality of Omišalj, with the highest ageing index of as many as 51 percent recorded in the municipality of Lovinac (Table 2). In the same period, the highest vital index value of 209 percent was recorded in the municipality of Viškovo, with the municipality of Lanišće having the lowest vitality index of 6.3 percent. The conclusion is that all local government units have been affected by an ageing process, while some have a good demographic potential. Source of data: "Census of Population, Households and Dwellings of 31st March 2001."

Table 2 **Descriptive Statistics of Input Variables (Indicators of Socio-Economic Development of Local Government Units)**

Variable	Descriptive Statistics					
	n	Arithmetic Mean	Median	Minimum	Maximum	Standard Deviation
X1	86	1,189.2	869.5	42.8	7,092	1,223.5
X2	86	8.7	6.1	0.7	54	8.1
X3	86	20.7	21.9	3.2	47	11.5
X4	86	17.0	14.2	3.4	47	10.1
X5	86	17.5	16.4	6.7	70	8.0
X6	86	25.9	25.5	12.8	51	6.8
X7	86	66.7	62.7	6.3	209	35.0
X8	86	2.3	1.5	0.3	13	2.5
X9	86	314,526.0	8,151.5	0.0	4,896,520	709,260.8
X10	86	37,680.7	39,052.8	8,968.5	62,207	8,552.1
X11	86	35.6	36.7	11.2	45	5.6

Source: CBS data, Ministry of Finance, author's calculations.

[7] *Vital index is calculated using the following formula:* $V = \frac{f \times \frac{P_{(20-39)}}{P}}{m \times x_s} \times 100$. *Where* $P_{(20-39)}$ *is the number of persons aged between 20 and 39, P total population, and* x_s *ageing index.*

4 Identification of Principal Dimensions of Socio-Economic Development Using Factor Analysis

Factor analysis is used to identify a smaller number of dimensions of socio-economic development that adequately summarise the information contained in the original set of variables.

Since no prior hypothesis is made about the number and name of factors, explorative factor analysis is used. The extracted factors are independent linear combinations of original, correlated variables. Following selection of input variables and standardization of their values, justification of the use of analysis must be examined and a decision made on the selection of the method of factor analysis. Therefore, the correlation matrix of original variables has to be calculated and examined first. If correlations between manifest variables are high, the factor analysis is justified. The most frequently used approaches of factor analysis are the component analysis[8] and common factor analysis. The choice of approach depends on the objective of the analysis. Depending on how easy it is to interpret the obtained results, decision is made on rotation of factors. Since the results of this analysis will be used as clustering variables, factor scores must be calculated.[9]

4.1 Justification for Using Factor Analysis

Justification for using the factor analysis implies determining whether input, manifest variables[10] are significantly and sufficiently correlated. Only if manifest

[8] *The best-known and most widely used component analysis method is the principal component analysis.*

[9] *Component analysis is used when the objective is to summarise most of the original information (variance) with a minimum number of factors for forecasting purposes, and when factor analysis results are used as input variables for further analyses. Common factor analysis is used to identify supporting factors or dimensions that are not easy to discern.*

[10] *Original variables that are observed in factor analysis are called manifest variables.*

variables are correlated, can factors be identified as hypothetical components of a non-correlated variable, and sufficient for expressing manifest variables.[11]

Table 3 shows the manifest variable correlation matrix. Correlation matrix indicates that each variable has at least one correlation coefficient with an absolute value higher than 0.3, which is the minimum value proposed by Kinnear and Gray (1994) as a criterion for inclusion of variables into analysis. Therefore, all eleven variables have been included into analysis. Marked in Table 3 are those correlation coefficients that are significant with a significance level of 5 percent.

Table 3 **Correlation Matrix for Input Variables**

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
X1	1.00	**-0.26**	-0.07	**0.28**	-0.02	**-0.36**	**0.36**	**-0.23**	**0.29**	0.13	**0.25**
X2	**-0.26**	1.00	**-0.25**	**-0.22**	0.19	**0.61**	**-0.38**	**0.67**	-0.21	**-0.55**	**-0.50**
X3	-0.07	**-0.25**	1.00	**-0.63**	**-0.25**	-0.13	-0.05	**-0.29**	**-0.33**	**0.34**	**0.21**
X4	**0.28**	**-0.22**	**-0.63**	1.00	-0.12	**-0.31**	**0.27**	**-0.25**	**0.56**	0.06	**0.30**
X5	-0.02	0.19	**-0.25**	-0.12	1.00	**0.25**	**-0.23**	**0.46**	-0.03	**-0.55**	**-0.69**
X6	**-0.36**	**0.61**	-0.13	**-0.31**	**0.25**	1.00	**-0.58**	**0.68**	**-0.33**	**-0.69**	**-0.81**
X7	**0.36**	**-0.38**	-0.05	**0.27**	**-0.23**	**-0.58**	1.00	**-0.39**	**0.26**	**0.43**	**0.49**
X8	**-0.23**	**0.67**	**-0.29**	**-0.25**	**0.46**	**0.68**	**-0.39**	1.00	-0.19	**-0.73**	**-0.75**
X9	**0.29**	-0.21	**-0.33**	**0.56**	-0.03	**-0.33**	**0.26**	-0.19	1.00	0.14	**0.26**
X10	0.13	**-0.55**	**0.34**	0.06	**-0.55**	**-0.69**	**0.43**	**-0.73**	0.14	1.00	**0.86**
X11	**0.25**	**-0.50**	**0.21**	**0.30**	**-0.69**	**-0.81**	**0.49**	**-0.75**	**0.26**	**0.86**	1.00

Source: CBS data, Ministry of Finance, author's calculation.

4.2 Selection of Factor Analysis Method, Criteria for Deciding on the Number of Factors, Analysis Results

The selection of the factor analysis method is dependant on the goal of analysis. As mentioned above, there are two main factor analysis approaches: principal component analysis and common factor analysis. Since the results of factor analysis will be used as clustering variables (to which end factor scores must also be calculated), it is recommended to use the principal component analysis (Morrison,

[11] *Factor analysis and factor models are applicable to those variables that are interdependent, i.e. mutually correlated.*

1987). The principal component analysis allows direct identification of factor scores as opposed to the common factor analysis, where such scores are estimated. For selecting the number of factors the eigenvalue criterion was used according to which the amount of variation explained by each factor must be larger than 1. Unrotated factors obtained by analyzing the principal components of 11 variables in 86 municipalities and towns in the three observed counties are shown in Table 4.

Table 4 **Unrotated Factor Matrix**

Input Variable	Factors		
	1	2	3
X1	-0.39	-0.38	0.52
X2	**0.72**	-0.06	-0.34
X3	-0.20	**0.85**	0.27
X4	-0.38	**-0.79**	-0.28
X5	0.56	-0.30	**0.61**
X6	**0.87**	0.07	-0.18
X7	**-0.63**	-0.22	0.21
X8	**0.85**	-0.16	-0.00
X9	-0.37	**-0.66**	-0.06
X10	**-0.85**	0.31	-0.12
X11	**-0.92**	0.12	-0.24
Accounted Variance	4.75	2.20	1.06
Proportion of Accounted Variance	0.43	0.20	0.10

Source: CBS data, Ministry of Finance, author's calculations.

In the unrotated factor solution, factors are extracted according to their significance. The first factor is general and almost every variable has a high loading on this factor, with each following factor accounting for an ever smaller proportion of variance. In this solution, the first factor accounts for 43.2 percent of variable variance, the second factor accounts for 20 percent, and the third factor for as little as 10 percent of variance. It is therefore desirable to perform factor rotation, because it allows the variance to be redistributed from the factors that are first in order to those that come later. This results in a structure that is simpler and easier to interpret without changing the total variance. In addition, in cases when factor analysis solutions are used in further analyses, the theory recommends

orthogonal factor rotation.[12] In an orthogonal solution, the factors are extracted in such a way that the factor axes are maintained at 90 degrees meaning that each factor is independent of all other factors. And it is the *varimax* rotation[13] that yields solutions that are the easiest to interpret. The theory recommends *varimax* rotation also in cases when the obtained factors are used as the basis for calculating factor scores which serve as input variables for further analyses, in this case cluster analysis (Johnson and Wichern, 1992). The factor loading matrix obtained through *varimax* rotation is shown in Table 5.

Table 5 **Varimax Rotated Factor Matrix**

Input Variables	Factors		
	1	2	3
X1	**0.71**	0.21	-0.12
X2	**-0.69**	0.08	-0.37
X3	0.15	**-0.85**	0.29
X4	0.19	**0.88**	0.17
X5	0.15	-0.03	-0.87
X6	**-0.70**	-0.12	-0.53
X7	**0.60**	0.22	0.28
X8	-0.51	0.03	**-0.70**
X9	0.32	**0.68**	0.09
X10	0.39	-0.13	**0.81**
X11	0.39	0.11	**0.83**
Accounted Variance	2.61	2.10	3.31
Proportion of Accounted Variance	0.24	0.19	0.30

Source: CBS data, Ministry of Finance, author's calculations.

For the solution to be accepted, it is necessary to examine the significance of obtained factors that represent dimensions of socio-economic development of the observed towns and municipalities.

[12] *As an alternative to orthogonal rotation, oblique rotation of factors, i.e. factor axes is used. In oblique rotation, the factor axes do not form right angles, which means that factors may be intercorrelated.*

[13] *The principal methods of orthogonal rotation are quartimax (simplification of factor matrix rows) and varimax rotation (simplification of factor matrix columns).*

Three factors meet not only the eigenvalue criterion, but also the variance proportion criterion. In social sciences, the lowest limit of acceptability is 60 percent of variance accounted by obtained factors (Hair, Anderson and Tahtam, 1987). This solution accounts for 73 percent of total variance.

Three factors were obtained through *varimax* rotation of the initial solution yielded by the principal component analysis. The first factor has a high positive factor loading on variables: X1 (municipality's own per capita revenues), X7 (vital index). This means that it positively correlates to the respective characteristics of local government units. The first factor has a high negative factor loading on X2 (persons employed in agriculture and fisheries as percentage of total population) and X6 (ageing index). This factor is therefore called *"relatively young population; economically strong local governments; low share of primary sector"*.

The second factor has a high positive factor loading on the "proportion of persons employed in tourism" and "tourist nights", and a negative factor loading on the "proportion of persons employed in manufacturing". This factor is labelled *"developed tourism"*.

The following indicators have a high loading on Factor 3: unemployment rate (X5) and persons without completed primary education (X8) with a negative sign; gross domestic product per capita (X10) and proportion of persons in employment (X11) with a positive sign.

This factor is labelled *"relatively high employment; better educational attainment of population; and higher per capita GDP"*.

5 Classification of Local Government Units Using Cluster Analysis

For classifying local government units in the three observed counties into groups characterized by similar features of socio-economic development cluster analysis was used. In social sciences cluster analysis was recognized as the most suitable

method of classifying units into groups of similar characteristics. Serving as input variables for cluster analysis were factor scores. Since the factor analysis resulted in three factors, for each observed local government unit three factor scores will be calculated. The factor score indicates the extent to which each town/municipality has a high score on a group of characteristics that have a high loading on a relevant factor. This means that each local government unit that has a high score on variables with high factor loading on one of the three obtained factors also has a high factor score on this factor. The factor score indicates the extent to which an individual town/municipality has a characteristic indicated by the factor.

5.1 Selection of Cluster Analysis Method

For grouping objects into clusters, non-hierarchical clustering method, the *"k-means"* method[14] was used. The main argument in favour of this clustering method is that this method of grouping objects into clusters is more suitable when grouping units (objects) on which specific characteristics were measured, and not when grouping characteristics, i.e. variables (Johnson and Wichern, 1992). Decision on the number of clusters is based on the analysis of variance (ANOVA).

5.1.1 Decision on the Number of Clusters – Analysis of Variance

What is characteristic for this method is that the number of clusters is defined in advance and the significance of the obtained solution is tested. In step one, the significance test for the two proposed clusters was performed. In ANOVA, the significance test examines between-group variability with within-group variability when testing the hypothesis that means differ between groups.[15] At the theoretical significance level of 5 percent, the ANOVA results for the two proposed clusters are not significant. Since on factor 2 the empirical *p value* exceeds the theoretical

[14] *There are two principal approaches to cluster analysis: hierarchical cluster analysis and non-hierarchical cluster analysis.*

[15] *This is actually ANOVA in reverse. Analysis of variance implies testing of a hypothesis of equality of arithmetic means of a number of basic sets.*

value ($p = 0.07 > 0.05$), the hypothesis[16] H_1 must be rejected, which means that the means between groups do not differ significantly (see Table 6 ANOVA results for two clusters). The solution that groups the observed local government units into two clusters cannot be accepted.

Table 6 **ANOVA Results for Two Proposed Clusters**

Factor	Means between clusters	s.s.	Means within clusters	s.s.	F ratio	p-value
Factor 1	17.19	1	67.81	84	21.30	0.000014
Factor 2	3.39	1	81.62	84	3.48	0.065458
Factor 3	30.30	1	54.70	84	46.53	0.000000

Source: CBS data, Ministry of Finance, author's calculations.

We therefore proceed to the next step to perform a significance test in ANOVA for three clusters. At the given significance level of 5 percent and empirical significance level of 0.001627 for factor one, 0.0000 for factor two and 0.0000 for factor three, hypothesis H1 is accepted, i.e. we may say that the means between the three proposed clusters differ significantly (see Table 7 ANOVA results for three clusters).

Table 7 **ANOVA Results for Three Proposed Clusters**

Variable	Means between clusters	s.s.	Means within clusters	s.s.	Empirical F ratio	p–value
Factor 1	12.18	2	72.82	83	6.9445	0.001627
Factor 2	61.21	2	23.79	83	106.7699	0.000000
Factor 3	44.52	2	40.48	83	45.6396	0.000000

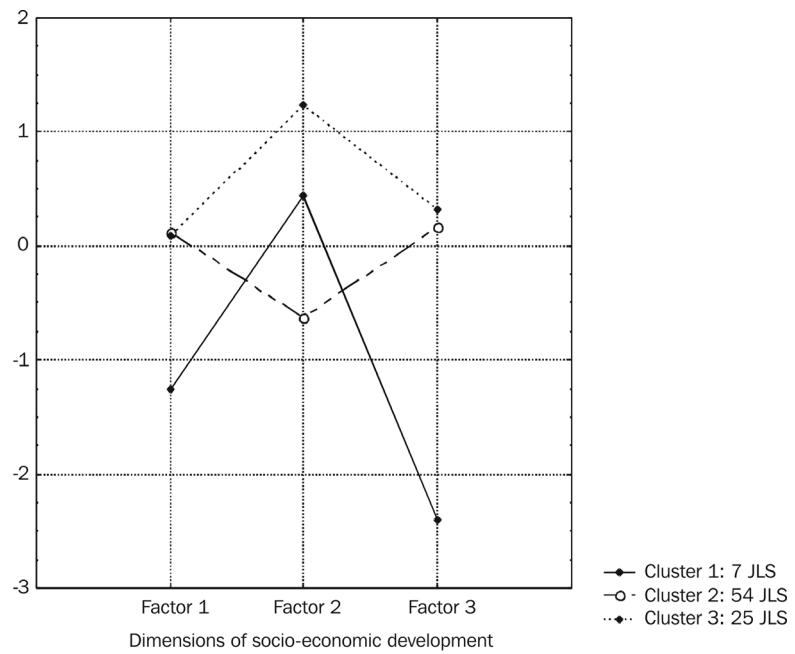
Source: CBS data, Ministry of Finance, author's calculations.

The results indicating grouping of local government units into three different clusters are significant. The goal to be aimed for is that means that each cluster has on an individual dimension differ significantly. As this is confirmed by the significance test in ANOVA, the same can be verified by looking at the graph of

[16] *H_1 hypothesis claims that means between clusters significantly differ on each dimension (factor). The alternative, null hypothesis (H_0) claims that cluster means are equal.*

means, i.e. by factor analysis identified dimensions of socio-economic development and means of an individual cluster (see Figure 1).

Figure 1 **Means of Each Cluster**



Source: CBS data, Ministry of Finance, author's calculations.

5.2 Identified Homogenous Spatial-Economic Entities

The above graph illustrates the differences between clusters, i.e. peculiarities of each cluster. Most conspicuous are the features of socio-economic development for cluster one, which differs from the other two clusters in all of the three dimensions being observed. As opposed to this, cluster two and cluster three are almost equal in regard to characteristics of socio-economic development that are presented by factors 1 and 3, but significantly differ in relation to features of socio-economic development presented by factor 2.

Cluster 1 is characterized by a strong negative relationship to factor score 3, which presents high employment, high per capita GDP. Also, there is a negative relationship to factor 1, which indicates a large share of primary sector and a predominance of older population, relatively uneducated. This group comprises only seven local government units, all located in the Lika-Senj County: town of Gospić and municipalities of Donji Lapac, Lovinac, Perušić, Plitvička jezera. Vrhovine and Udbina.[17] This cluster may be labelled: *group of local government units with poor economic potential.*

Cluster 2 comprises as many as 54 local government units. The predominant feature in this group of towns/municipalities is the one presented by factor 2: predominance of secondary sector in economic structure and undeveloped tourism.[18] Considering the weak, yet positive relationship to factor 3, to this group of municipalities we may also attribute the feature of non-existence of a significant employment problem compared to cluster 1, where this problem is most pronounced.[19] This cluster may be labelled as a group of local government units with *secondary sector predominating in economic structure.* In the County of Istria, this cluster comprises the following local government units: towns of Buje, Buzet, Labin, Pazin, Pula, Vodnjan and municipalities of Bale, Barban, Brtonigla, Cerovlje, Fažana, Gračišće, Grožnjan, Kanfanar, Karojba, Kršan, Lanišće, Ližnjan, Lupoglav, Marčana, Motovun, Oprtalj, Pićan, Raša, Sveta Nedjelja, Sveti Petar u Šumi, Svetvinčenat, Tinjan, Žminj. Towns and municipalities in the County of Primorje and Gorski Kotar belonging to this cluster are: Bakar, Čabar, Delnice, Kastav, Kraljevica, Rijeka, Vrbovsko, Brod Moravice, Čavle, Fužine, Jelenje, Klana, Kostrena, Lokve, Matulji, Mrkopalj, Omišalj, Ravna Gora, Skrad, Vinodol municipality, Viškovo, Vrbnik. Towns of Otočac and Senj and the municipality of Brinje are local government units in the County of Lika and Senj that also bear the features of this cluster.

[17] *The said municipalities belong to areas of special state concern.*

[18] *What is characteristic is a strong negative relationship to factor score 2, which presents the features of factor 2: developed tourism and low proportion of manufacturing activities. Since the relationship is negative, it is interpreted as larger share of secondary sector in economic structure and undeveloped tourism.*

[19] *It is important to note that the problem of unemployment is presented in relative terms, since as such it exists not only in the observed local government units, but in the country at large. Non-existence of a significant unemployment problem therefore refers to comparison with other clusters (especially cluster 1).*

Cluster 3 comprises 25 towns and municipalities. These are without exception local government units located in the coastal region. In the County of Istria: towns of Novigrad, Poreč, Rovinj, Umag and the municipalities of Kaštelir-Labinci, Medulin, Sveti Lovreč, Višnjan, Vižinada and Vrsar. In the Lika-Senj County: Karlobag and Novalja. In the Primorje-Gorski Kotar County, thirteen towns and municipalities belong to this cluster: Baška, Cres, Crikvenica, Dobrinj, Krk, Lovran, Mali Lošinj, Malinska-Dubašnica, Mošćenička Draga, Novi Vinodolski, Opatija, Punat and Rab. As can be assumed, in regard to the geographical position of these local government units, the most pronounced features in this cluster are those presented by factor 2. The predominant branch of industry in this cluster is tourism, and compared to other clusters it has the least pronounced problem of unemployment, i.e. the highest employment rate in total population. This cluster is therefore designated as the group of local government units with *markedly tourism-oriented economy and relatively high employment rate.*

6 Conclusion

Regional policy measures can lead to desired and equal effects in a specific region only if implemented in territorial entities that are homogenous in regard to dimensions of social and economic development. Since, by using factor and cluster analysis methods, we have been able to single out the local government units of similar characteristics, that are the subject matter of regional policy instruments and measures, the principal hypothesis of the paper has been confirmed.

The next conclusion is that the distance between individual territorial entities, in this case towns and municipalities, does not necessarily have to imply a "distance", or difference, in regard to characteristics of socio-economic development of such territorial entities. For example, the town of Pula and the municipalities of Ližnjan and Vodnjan, despite being located in the coastal region, as well as the majority of entities in cluster 3, do not belong to cluster 3, but to cluster 2. The same goes for the municipality of Vižinada, located in inland Istria, which is mostly surrounded by local government units that form cluster 2, while itself it belongs to cluster 3.

This paper allows consideration in the Croatia's practice of new approaches to classification of space into more homogeneous entities for the purpose of formulating regional policy measures. It presents the possibility of using two mathematical-statistical methods for singling out territorial entities of similar characteristics that will be the subject matter of such policy. Even though regional policy is in place in Croatia, its subject matter has not been clearly defined yet. Besides, it consists of a series of uncoordinated measures and initiatives. A number of legislations regulate territories of similar characteristics (Act on Areas of Special State Concern, Islands Act, Act on Town of Vukovar and Act on Hilly and Mountainous Areas). Nonetheless, the selection of these areas was quite arbitrary and as such is not the result of economic analyses that are based on principal features of a specific region. Without them it would be impossible to define and implement regional policy measures or monitor their effects. A major constraint to implementing such analyses is the lack of statistics at lower levels. Better statistical information would undoubtedly facilitate much better analyses. At this level, most data are collected through population censuses which are carried out about every ten years. For such analyses, that would serve as a basis for formulating regional policy goals and measures, research should be commissioned.

Literature

Čižmešija, M. and N. Kurnoga Živadinović, 2002, "Faktorska analiza rezultata konjukturnih testova Hrvatske", *Ekonomski pregled,* (7-8), pp. 684-704.

Ding, C. G. and F. M. Liou, 2002, "Subgrouping Small States Based on Socioeconomic Characteristics", *World Development,* 30(7), pp. 1289-1306.

Central Bureau of Statistics, 2002, *Statistical Yearbook 2002.*

Central Bureau of Statistics, 2002, *Statistical Report "Employment and wages" 1162/2002.*

Central Bureau of Statistics, 2002, *Statistical Report "Tourism" 1176/2002.*

Central Bureau of Statistics, 2002, Census of Population, Households and Dwellings*, 31March 2001,* http://www.dzs.hr/Popis%202001/popis20001.htm.

EUROSTAT, 2003, EUROSTAT European Regional Statistics - Reference Guide 2003, http://www.eu-datashop.de/download/DE/klassifi/ncronos/thema1/reg_guid.pdf.

Fulgosi. A., 1988, *Faktorska analiza,* Zagreb: Školska knjiga.

Gaur. S., 1997, "Adelman and Morris Factor Analysis of Developing Countries", *Journal of Policy Modeling,* 19 (4), pp. 407-415.

General provisions on the Structural Funds, http://europa.eu.int/scadplus/leg/en/lvb/l60014.htm.

Gilbert, A. and A. G. Churchill, 1995, *Marketing Research - Methodological Foundations*, Dryden Press: Fort Worth, TX.

Hair, F. J., E. R. Anderson and L. R. Tahtam, 1987, *Multivariate Data Analysis*, New York: Macmillan Publishing Company.

Johnson, A. R. and W. D. Wichern, 1992, *Applied Multivariate Statistical Analysis*, 3rd Ed., New York, Upper Saddle River: Prentice Hall.

Kinnear, P. and C. D. Gray, 1994, *SPSS for Windows Made Simple*, Hove, UK: Lawrence Erlbaum Associates Publishers.

Kronthaler, F., 2003, "A Study of Competitiveness of Regions based on a Cluster Analysis: The Example of East Germany", paper presented at the conference: "Peripheries, centres, and spatial development in the new Europe", organized by European Regional Science Association (ERSA), Jyväskylä, August.

Lipshitz, G. and A. Raveh, 1998, "Socio-economic Differences among Localities: A New Method of Multivariate Analysis", *Regional Studies Journal of the Regional Studies Association,* 32 (8), pp. 747-758.

Loedwijk, B. and D. Terweduwe, 1988, "The Classification of Countries by Cluster and by Factor Analysis", *World Development*, 16(12), pp. 1527-1545.

Marinović-Uzelac, A., 2001, *Prostorno planiranje*, Zagreb: Dom i svijet.

Ministry of Finance, 2001, *Proračuni jedinica lokalne samouprave Istarske županije za 2001. godinu*, http://www.mfin.hr/lokalna/2001/zup18.htm.

Ministry of Finance, 2001, *Proračuni jedinica lokalne samouprave Ličko-senjske županije za 2001. godinu*, http://www.mfin.hr/lokalna/2001/zup09.htm.

Ministry of Finance, 2001, *Proračuni jedinica lokalne samouprave Primorsko-goranske županije za 2001. godinu*, http://www.mfin.hr/lokalna/2001/zup08.htm.

Morrison, F. D., 1987, *Multivariate Statistical Methods*, New York: McGraw-Hill Book Co.

Official Gazette of the Republic of Croatia, different issues.

Punj, G. and D. W. Stewart, 1983, "Cluster analysis in marketing research; a review and suggestions for application", *Journal of Marketing Research,* 20, pp. 134-148.

Rašić, I., 2003, "Introduction of Nomenclature of Territorial Units for Statistics in Croatia", paper presented at the conference: "Peripheries, centres, and spatial development in the new Europe", organized by European Regional Science Association (ERSA), Jyväskylä, August.

Starc, N. ed., 2003, "Analiza gospodarstva u područjima posebne državne skrbi", Zagreb: The Institute of Economics, Zagreb, documentation.

STATISTICA, 1999, Program Overview, StatSoft.

Šošić, I. and V. Serdar, 1992, *Uvod u statistiku,* Zagreb: Školska knjiga.

Ward, J., 1963, "Hierarchical grouping to optimize an objective function", *Journal of American Statistical Association,* 58, pp. 236-244.

Wertheimer-Baletić, A., 1999, *Stanovništvo i razvoj,* Zagreb: MATE.