ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Bakaric, Ivana Rasic

Conference Paper

Methods of Multivariate Analysis to Uncover Socio-Economic Differences Among Spatial-Economics Entities

46th Congress of the European Regional Science Association: "Enlargement, Southern Europe and the Mediterranean", August 30th - September 3rd, 2006, Volos, Greece

Provided in Cooperation with:
European Regional Science Association (ERSA)

Suggested Citation: Bakaric, Ivana Rasic (2006) : Methods of Multivariate Analysis to Uncover Socio-Economic Differences Among Spatial-Economics Entities, 46th Congress of the European Regional Science Association: "Enlargement, Southern Europe and the Mediterranean", August 30th - September 3rd, 2006, Volos, Greece, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:
https://hdl.handle.net/10419/118180

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

METHODS OF MULTIVARIATE ANALYSIS TO UNCOVER SOCIO-ECONOMIC DIFFERENCES AMONG SPATIAL-ECONOMICS ENTITIES

Paper prepared for the
45th Congress of the European Regional Science Association
Volos, Greece, August 30th-Semptember 3rd, 2006

Ivana Rasic Bakaric
Institute of Economics Zagreb
Trg. J. F. Kennedya 7, 10000 Zagreb, Croatia
irasic@eizg.hr

Abstract

Main aim of regional policy is to reduce regional disparities. There is now doubt that action to tackle regional disparities is needed in Croatia. Main problem is the definition of region as a subject of regional policy. After discussion about definition of region and main criteria which should be considered during the process of establishment of regions, the paper will present one of the methodologies for spatial-economic entities' classification. The methodology will be used on the working sample of several Croatian counties. The classification will be obtained through use of multivariate statistical methods, and it is based on a wide number of demographic, economic, educations, employment indicators. Results leads to the identification of several axes of socio-economic characterisation, and the division of the observed territory into several homogenous clusters, regions with different degree of development. The main conclusion is that socio-economic reorients uncovered with this methodology provide useful characterisation and division of the territory, for policy making purpose.

JEL Classification: C49, R12, R58

Keywords: homogeneous regions, factor and cluster analysis, regional policy

Introduction

A precondition to formulating implementing and monitoring of regional policy measures is the existence of established regional boundaries. Division proposals depend on criteria used. Politicians and various professions use different criteria, in order to achieve the optimal territorial structure that could enable same or at least similar effects of certain policy measures within the same territorial unit, and could be easily monitored. Croatia is exceptionally heterogeneous country; regional differences in economic and social development are great and without clear defined regional policy will probably continue to increase in future (Rasic, 2003). The Croatian territory has been divided into 21 counties and 122 cities and 426 municipalities. The existing administrative division of the country's territory into 21 counties which does not take into consideration the natural and geographic factors, the existing economic structure, or some traditional divisions of Croatian territory, isn't the most appropriate one (Human Development Report Croatia 2002, to appear in 2003). After discussion about definition of region and main criteria which should be considered during the process of establishment of regions, the paper will present one of the methodologies for spatial-economic entities' classification. This approach assumes that the large geographic units are not all of one kind, and there are socio-economic differences among the localities that make up this units. Geographical proximity between localities does not necessarily mean socio-economic proximity. This approach gives great importance to classifying the localities in the different geographical regions of each country into "socio-economic" regions. This study takes into account smaller geographic units and broader spectrum of socio-economic indicators. The study deals with the 127 towns and municipalities into which the territory of four Croatian Counties has been divided. Those are coastal Counties: County of Zadar, County of Šibenik-Knin, County of Split-Dalmatia and County of Dubrovnik-Neretva. Table 1 (see Appendix) contains list of towns and municipalities in selected Counties. The classification is obtained through use of factor and cluster analysis, and is based on 12 socio-economic indicators. The first objective of this paper is to identify a smaller number of socio-economic dimensions that adequately summarise the information contained in a range of regional indicators (using factor analysis). The second objective is to obtain homogeneous regions in terms of socio-economic development (using cluster analysis).

The remaining part of this paper is organised in followed way. The definition of region and main criteria which should be considered during the process of establishment of regions has been presented in Section 1. In Section 2 details of the socio-economic data for each town/municipality have been presented. Section 3 deals with identification of the smaller number of socioeconomic dimensions (factor analysis). In Section 4 conducted cluster analysis and its' results have been presented. The main conclusions of this study are presented in Section 5.

1. Region, regionalisation

Region is mostly defined as area or division with or without definite boundaries and characteristics (Oxford-Advanced Learner's Dictionary of Current English, 1986). There are region within the country or cross borders regions. The two important aspects of regions within the country are delimitation of space on the basis of one or more criteria and administrative function of a region as an administrative level below that of nation state. The scope of a region is usually based on natural boundaries, historical boundaries and administrative boundaries (Rasic, 2003).

The regionalisation therefore depends on the criteria used. During the division of territory, different profession use different criteria - usually one criterion or a very narrow set of criteria. A criterion can be a single one, which is justified if the goal of the division is very clear and directly follows from the application of that particular criterion.

Country's territory needs to be divided in such a manner so that the established economic entities, i.e. regions, are the most suitable subject of the regional policy. Regions need to have suitable size, homogeneity, cognizable regional identity, mentality, customs, and social conveniences. The main point is to form regions in a manner so as to ensure that regional policy measures have the same effect in every part of a particular region and that the measures can be easily monitored which means that the regions are the most suitable subject of the regional policy. Therefore it is necessary to respect also homogeneity criteria of regions, i.e. regional statistical units, natural and geographic diversity, historical tradition and geopolitical situation, the structure of the economy and development of individual regional identities, as well as suitable size of territory. Homogeneous regions are those regions that share some uniform

characteristics (economic, social), i.e. territory whose parts show fewer differences among themselves then when compared with the units of the neighbouring territory (Uzelac, 2001, p. 290). This is the process of identifying the similarities within regional units of a territory.

Whether one talks of regions or smaller territorial units, this division requires multi-criteria judgment and decision-making. It is very important to take note of the characteristics of the regional geographic base, which includes economic and transport aspects, the direction of inhabitants' gravitation, antagonisms between the regions and central towns etc., with a goal of establishing logical and sustainable economic territorial entities (Rasic, 2003).

2. Main socioeconomic indicators

This study is based on 12 variables, socioeconomic indicators. Their selection was based on the fact that they are indicators of different town's/municipality's socioeconomic characteristics and on the availability of data at the level of towns/municipalities. Their code and description is presented in Table 1.

Table 1 Codes and descriptions of the variables

Code	Description
X1	Persons employed in *Agriculture, hunting and forestry* as percentage of total employment
X2	Persons employed in *Manufacturing* as percentage of total employment
X3	Persons employed in Tourism (*Hotels and Restaurants*) as percentage of total employment
X4	Unemployment rate
X5	Ageing index[1]
X6	Vital index[2]
X7	Persons without schooling as percentage of total population
X8	Persons without schooling as percentage of persons with at least basic school
X9	Tourist nights
X10	Persons in employment as percentage of total population
X11	Gross domestic product per capita
X12	Town's/municipality's budget revenues

[1] Ageing index is calculated as the number of persons 60 years old or over per 100 persons under age 15.
[2] The ratio of births to deaths within a population during a given time.

All variables report to the year 2001. With the exception of two variables (*Tourist nights* and *Town's/municipality's budget revenues*), main source of the rest of the data was the CENSUS 2001. During the Census 2001 there were 20 counties, the City of Zagreb with a special status, 122 towns and 423 municipalities in the Republic of Croatia. All data presented through CENSUS results refer to the territorial constitution of the Republic of Croatia with the situation as on 31st March 2001. The data concerning the variable "Tourist nights" were provided by Croatian Statistical Bureau Publication, "First Release", No. 4.4.2/11. The data on budget revenues were downloaded from the Ministry of finance web site[3]. Concerning other indicators, the available statistics doesn't offer many data on socioeconomic development at that level. In addition, there aren't any available data on health and culture. For example, the data of GDP and other data that are basis for estimation of other economic indicators have not been collected at the county's level and towns' and municipalities' level. GDP by towns and municipalities in 2001 was estimated using the constant average labour productivity assumption.

The table of the descriptive statistics shows huge differences in observed 127 towns and municipalities.

Table 2 Descriptive statistics

Variable	Valid N	Mean	Median	Minimum	Maximum	Std.Dev.
X1	127	12.10	6.00	0.00	62.00	13.60
X2	127	13.50	11.00	0.00	47.00	9.60
X3	127	14.10	11.00	0.00	64.00	10.70
X4	127	28.50	27.25	12.13	86.00	10.40
X5	127	26.80	25.70	12.30	76.00	9.30
X6	127	93.20	81.40	2.00	325.00	58.20
X7	127	6.80	3.82	0.51	34.00	7.00
X8	127	7.90	3.97	0.51	52.00	9.50
X9	127	117609.60	35681.00	0.00	1,084,375.00	192001.30
X10	127	27.60	27.72	3.85	44.00	6.60
X11	127	25876.90	25552.63	1720.35	43976.00	7944.30
X12	127	558.50	391.48	0.87	4745.00	668.60

Source: author's calculation

For example variable X1 ("Persons employed in *Agriculture, hunting and forestry* as percentage of total employment") has a range of 62.00 percentages. Comparing with its minimum value of 0.00 percentage of employed in that activity, its maximum value is 62 percentages (Municipality Ružić located in County of Šibenik-

[3] http://www.mfin.hr/str/11/

Knin). The variable "Tourist nights" has also a big range. The maximum value of the variable "Tourist nights" was reached in Dubrovnik, even 1,084,375 nights. The lowest GDP per capita (HRK 1,720.35) was produced in municipality Ervenik, located in County of Zadar. Municipality Dubrovačko Primorje produced the highest GDP per capita among observed towns and municipalities, HRK 43,976.00.

In order to eliminate the effects due to variables having been measured on different scales, further analysis will be carried out with standardised variables.
Table 1 (see Appendix) presents correlation matrix, which reveals the existence of strong relationships between some variables. Marked correlation are significant at p = 0.05.

3. Factor analysis – Common factors for socioeconomic indicators

3.1 Purposes of factor analysis

Factor analysis can be used to examine the underlying patterns or relationships for a large number of variables and determine if the information can be considered or summarised in a smaller set of factors and components (Hair et al., 1987). The purposes of factor analysis are two: data reduction and substantive interpretation. The first purpose is summarising the important information in a set of observed variables by a new, smaller set of variables expressing that which is common among the original variables. The second purpose concerns the identification of the constructs or dimensions that underlie the observed variables (Gilbert, Churchill, 1995). Therefore, factor analysis is applied as a data reduction or structure detection method. At the mathematical level, a factor is simply a linear combination of original, highly correlated variables. The various methods of factor analysis are differentiated in terms of how the weights used in forming the linear combinations are determined.

Basically, the factor model is motivated by the fact that variables within particular group are highly correlated among themselves but have relatively small correlations with variables in a different group. It is conceivable that each group of variables represents a single underlying dimension, or factor, that is responsible for the observed correlations.

3.2 Steps in application of factor analysis

Factor analysis should begin with deciding whether a factor analysis should be applied to a set of data. This could be done by evaluating the appropriateness of the factor model assessing whether the variables are significantly and sufficiently correlated with each other so that their number can be reduced by applying the factor analytic model. Factor analysis is concerned with homogeneity of items. This means that some of the items in the correlation matrix should be large, indicating they go together (Gilbert, Churchill, 1995). According to Kinnear and Gray (1994) all but two variables should have at least one correlation coefficient with an absolute value larger than 0.3. After evaluation of correlation matrix (Table 1, see Appendix) all variables are included into analysis.

In the next step it has to be decided which model of factor analysis will be used. There are different methods of extracting the factors from a set of data. The two most frequently employed factor analytic approaches used are component analysis (in this paper principal component analysis) and common factor analysis. Principal components analysis (PCA) is by far the most common form of factor analysis. It seeks a linear combination of variables such that the maximum variance is extracted from the variables. It then removes this variance and seeks a second linear combination which explains the maximum proportion of the remaining variance, and so on. Common factor analysis (CFA) is a form of factor analysis which seeks the least number of factors which can account for the common variance (correlation) of a set of variables, whereas the more common principal components analysis (PCA) in its full form seeks the set of factors which can account for all the common and unique (specific plus error) variance in a set of variables. CFA uses a PCA strategy but applies it to a correlation matrix in which the diagonal elements are not 1's, as in PCA, but iteratively-derived estimates of the communalities[4].

As PCA seeks the set of factors which can account for all the common and unique (specific plus error) variance in a set of variables, this method was used in our analysis.

The next issue that needs to be addressed is determining the number of factors that are necessary to account for the variation of data. The main mostly popular criteria

[4] The amount of variance an original variable shares with all other variables included in the analysis.

that can be used to decide on proper number of factors are the eigenvalue (latent root) criterion and the percentage of variance criterion. The eigenvalue (latent root) criterion holds that that amount of variation explained by each factor must be greater of one. Each factor should account for the variation in at least one variable. According to the percentage of variance criterion, in the social sciences, the cumulative percentage of variance extracted by successive factors should be at least 60 percent of total variance.

In addition to selecting the factor model, the analyst must specify how the factors are to be extracted. Unrotated factor solution is difficult to interpret. All of the most significant loadings (highlighted) are on the first factor. This is a common pattern. One way to obtain more interpretable results is to rotate your solution. The ultimate effect of rotating the factor matrix is to redistribute the variance from earlier factors to later ones to achieve a simpler, theoretically more meaningful, factor pattern. Two available options are orthogonal factors and oblique factors. Those factors can be obtained by two different rotation methods (orthogonal and oblique). 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. Three major orthogonal approaches are quartimax[5], varimax[6] and equimax[7] rotation. In an oblique factor solution, the extracted factors are correlated.

After evaluation of the correlation matrix of the 12 input variables, all the variables were included into the factor analysis. In the case when the factor analysis results are further used as input variables for cluster analysis (factor scores will have to be computed), using of PCA has been recommended (Morrison, 1987).

[5] Goal is to simplify the rows of the matrix.
[6] Goal is to simplify the columns of the matrix.
[7] Compromise between quatrimax and varimax rotation.

Table 3 Unrotated factor matrix

	Factor 1	Factor 2	Factor 3	Factor 4
X1	-0,271	0,474	0,470	0,391
X2	0,067	**-0,593**	0,421	**-0,553**
X3	0,505	0,488	-0,460	0,023
X4	**-0,682**	-0,254	**-0,520**	0,032
X5	**-0,718**	0,526	0,038	-0,221
X6	0,339	**-0,659**	-0,176	0,447
X7	**-0,809**	0,047	0,261	0,062
X8	**-0,797**	0,104	0,264	0,066
X9	0,531	0,327	-0,339	0,061
X10	**0,813**	0,195	0,462	0,166
X11	**0,798**	0,006	0,430	0,018
X12	0,430	0,309	-0,123	-0,545
Expl.Var	4,483	1,843	1,578	1,046
Prp.Totl	0,374	0,154	0,132	0,087

Source: author's calculation

The results of factor analysis are given in a form of a factor matrix. Each column of the matrix represents a separate factor. Number in each column is the factor loadings for each variable on each factor. Factor loading is the correlation between the original variable and the factor, and it is the key to understanding the nature of a particular factor. Factor loadings greater than 0.50 are considered very significant (Hair et al., 1987). Squared factor loadings indicate what percentage of the variance in original variable is explained by factor. (Hair et al., 1987). Table 3 presents unrotated factor solution obtained using principal component analysis. The four factors with eigenvalue greater than one are extracted, and they account for 74.6 percentage of total variance. All of the most significant loadings (highlighted) are on the first factor. This factor accounts for 37.4 percentages of total variance. The second factor is less important and accounts for 15.4 percentages of total variance. Than comes third factor (13.2 percentages of total variance) and least important fourth factor that accounts for 8.7 percentages of total variance. In order to obtain more interpretable results this solution had to be rotated. Rotation was done using varimax rotation. According to Hair varimax rotation should always been used when the resulting factor scores are to be analysed by other statistical procedures, as is the case in this study. The factor matrix obtained by varimax rotation is present in Table 4.

Table 4 Varimax rotated factor matrix

	Factor 1	Factor 2	Factor 3	Factor 4
X1	0,033	-0,334	0,200	**-0,721**
X2	**0,730**	0,060	0,216	0,506
X3	**-0,803**	-0,047	0,156	0,184
X4	0,101	0,055	**-0,887**	-0,035
X5	0,066	**-0,777**	-0,401	-0,269
X6	0,030	**0,882**	0,010	0,029
X7	0,464	-0,363	-0,409	-0,463
X8	0,426	-0,398	-0,388	-0,480
X9	**-0,650**	0,072	0,224	0,171
X10	-0,207	0,156	**0,933**	-0,036
X11	-0,071	0,213	**0,866**	0,144
X12	-0,311	-0,327	0,289	**0,552**
Expl.Var	**2,158**	**1,974**	**3,132**	**1,685**
Prp.Totl	**0,180**	**0,165**	**0,261**	**0,140**

Source: author's calculation

As aforementioned four extracted factors account for 74.6 percentage of total variance. First factor explains 18.0 percentage of the variance of each individual variable. Second factor accounts for 16.5 percentage of total variance. Third factor explains 26.1 percentages and forth factor explains 14.0 percentage of the variance of total variance. Marked are factor loadings greater than 0.55.

3.3 Naming the factors – dimensions of socioeconomic development

The first factor has positive high factor loadings on the variable "Persons employed in Manufacturing as percentage of total employment" and negative high factor loadings on the variables "Persons employed in Tourism (Hotels and Restaurants) as percentage of total employment" and "Tourist nights". This factor was labelled "*Undeveloped tourism*".

The second factor has high positive loading on the variable "Ageing index" and high negative loading on the variable "Vital index". This factor was named "*Younger demographic structure*". This factor should be considered very carefully. Average value of the ageing index for the Republic of Croatia is 26.8 percentages which means that Croatian population is generally old.

The third factor, labelled "*High employment and GDP per capita*" has positive high loadings on the variables "Persons in employment as percentage of total

population" and "Gross domestic product per capita" and negative high loading on the variable "Unemployment rate".

The forth factor has high negative loadings on the variables "Persons employed in Agriculture, hunting and forestry as percentage of total employment" and "Persons without schooling as percentage of persons with at least basic school". The negative high loading this factor has on the variable "Town's/municipality's budget revenues". The factor was named "*Underdeveloped primary sector, developed secondary sector, good educated population, financially stronger units*".

4. Grouping the municipalities

Cluster analysis (CA), a well-known statistical method, was used to subgroup towns and municipalities based on their socioeconomic characteristics. Cluster analysis is a technique for grouping individuals or objects into clusters so that objects in the same cluster are more similar than objects in other clusters (Hair et al., 1998). Cluster analysis seeks to identify a set of groups which both minimize within-group variation and maximize between-group variation

CA techniques may be hierarchical or non-hierarchical. Hierarchical cluster analysis can use either agglomerative or divisive clustering strategies. In agglomerative hierarchical clustering every case is initially considered a cluster, then the two cases with the lowest distance (or highest similarity) are combined into a cluster. The case with the lowest distance to either of the first two is considered next. If that third case is closer to a fourth case than it is to either of the first two, the third and fourth cases become the second two-case cluster; if not, the third case is added to the first cluster. The process is repeated, adding cases to existing clusters, creating new clusters, or combining clusters to get to the desired final number of clusters. There is also divisive clustering, which works in the opposite direction, starting with all cases in one large cluster. Others are non-hierarchical methods, for example k-means clustering. The first step in non-hierarchical clustering procedures is to select a cluster centre or seed, and all objects within prespecified threshold distance are included in the resulting cluster (Hair et al., 1987).

The first step in cluster analysis is establishment of the similarity or distance matrix. This matrix is a table in which both the rows and columns are the units of analysis and the cell entries are a measure of similarity or distance for any pair of cases.

Interobject similarity can be measured in a variety of ways. One way is to look at the distance or difference between the pairs of objects. Distance measures how far apart two observations are. Cases which are alike share a low distance. Euclidean distance is the most common distance measure. A given pair of cases is plotted on two variables, which form the x and y axes. The Euclidean distance is the square root of the sum of the square of the x difference plus the square of the y distance. It is common to use the square of Euclidean distance as squaring removes the sign. When two or more variables are used to define distance, the one with the larger magnitude will dominate, so to avoid this it is common to first standardize all variables.

4.1 K- means clustering method using the factor scores

Factors present the underlying dimensions (constructs) that summarize or account for the original set of 12 variables. The aim of this analysis is to obtain clusters of towns/municipalities with different socioeconomic development. As clustering variables were further used factor scores. So the entirely new set of a smaller number of original variables had to be computed. Factor scores are composite measure for each factor representing each object (Hair et al., 1987, pp. 259). After computation of factor scores each municipality/town was represented by 4 composite measure rather than 12 original measures. These four factor scores represent each of four extracted factors. As non-hierarchical clustering methods are designed to group items, rather than variables, k- means non-hierarchical clustering method was used. The second argument for using this method is the fact that non-hierarchical clustering methods can be applied to much larger data set than hierarchical techniques.

Table 5 Analysis of variance, 3 clusters, at p = 0.05

Factor scores	Between SS	df	Within SS	df	F	signif. p
Factor score 1	50.49	2	75.51	124	41.45	**0,000000**
Factor score 2	36.45	2	89.55	124	25.24	**0,000000**
Factor score 3	11.92	2	114.08	124	6.48	**0,002110**
Factor score 4	54.04	2	71.96	124	46.56	**0,000000**

Source: author's calculation

As result of conducted k-means cluster analysis three clusters were obtained. Computationally, the k-means method is the analysis of variance (ANOVA) "in

reverse." The analyst starts with k random clusters, and then move objects between those clusters with the goal to 1) minimize variability within clusters and 2) maximize variability between clusters. This is analogous to "ANOVA in reverse" in the sense that the significance test in ANOVA evaluates the between group variability against the within-group variability when computing the significance test for the hypothesis that the means in the groups are different from each other. In k-means clustering, the program tries to move objects (e.g., cases) in and out of groups (clusters) to get the most significant ANOVA results (STATSOFT) (see Table 5)

Figure 1 Three obtained clusters

Source: author's calculation

Figure 1 presents Graph of means of each cluster. Graph of means is a line graph of the means across clusters. This plot is very useful for visually summarizing the differences in means between clusters.

The first cluster contains of 78 cases (towns and municipalities) spreading over different counties. It is characterised by developed tourism, underdeveloped manufacturing and younger population. The factor score 3 three is not important for this cluster. Beside tourism, the second important sector is primary sector. This cluster was named "*tourist cluster*".

The second cluster is composed from 38 towns and municipalities. This group, contrary to cluster 1, is characterised by positive means on the factor score 1 (developed manufacturing and underdeveloped tourism), factor score 3 (high employment and GDP per capita and factor score 4 (underdeveloped agriculture, good educated population, financially strong units of local self government). This cluster was labelled "*developed towns and municipalities with high GDP per capita and high employment"*

The third cluster contains of 11 towns and municipalities. This cluster is characterised by higher proportion of workers involved in the agriculture, by older population, low GDP per capita and high unemployment. The towns and municipalities in this cluster are financially very week. This cluster was named "*less developed towns and municipalities with low GDP per capita and high unemployment rate"*.

5. Conclusions

This study has successfully characterised observed towns and municipalities of four Croatian counties using four factors extracted from 12 socioeconomic indicators. Using factor scores as input variables for cluster analysis these towns and municipalities were divided into three mutually heterogeneous but internally homogeneous groups. This allows for better investigation of their specific features and thus improved a better basis for regional policy formulation and implementation. According to obtained results it could be concluded that geographical proximity between localities does not necessarily mean socio-economic proximity. This approach gives great importance to classifying the localities in the different geographical regions of each country into "socio-economic" regions. The factor and cluster analysis as multivariate techniques were successful in identifying the main axes of socio-economic characterisation and in identifying regions of the observed territory (four Croatian counties) with different degrees of development.

The main conclusion of this paper is that socio-economic reorients uncovered with this methodology provide useful characterisation and division of the territory, for policy making purpose.

REFERENCES

Aaker, D. A., Kumar, V., Day, G.,S (1995), *Marketing Research*, John Wiley & Sons Inc., New York.

Armstrong, H., Taylor, J. (2000) *Regional Economics and Policy*, Third edition published by Blackwell Publishers.

Basic, F., Bogunovic, M., Husnjak, S. (2000), *Croatian Agricultural Regions and Subregions,* Institute for General Production of Plants and the Institute for Pedalogy, Zagreb.

County budget (2001) Ministry of finance, Web page: http://www.mfin.hr

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

First Release, No. 4.4.2/11. Web site http: www.dzs.hr

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

Gilbert, A. Churchill, A. G. (1995), *Marketing Research – Methodological Foundations*, Dryden Press, Fort Worth, TX.

Gupta, V., Hanges, P.J., Dorfamn, P. (2002) "Cultural clusters: methodology and findings", *Journal of World Business* 37, pp. 11-15.

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

Human development report Croatia 1999 (1999), UNDP, Sveucilisna tiskara d.o.o., Zagreb

Human development report Croatia 2002 (2003), UNDP, Sveucilisna tiskara d.o.o. Zagreb

Johnson, A. R., Wichern W. D. (1992) *Applied Multivariate Stastical Analysis*, 3rd Ed. Prentice Hall, Upper Saddle River, NY.

Kinnear, P., Gray, C.D. (1994) *SPSS for Window Made Simple*, Lawrence Erlbaum Associates Publishers, Hove.

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

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

Morrison, F. D. (1987). *Multivariate Statistical Methods*, Mc Graw - Hill Book Co. New York.

Oxford-Advanced Learner's Dictionary of Current English (1986). Oxford University Press, Oxford.

Population, household and apartment Census (2002). Croatian Bureau of Statistics. Web Page. http://www.dzs.hr/Popis%202001/popis20001.htm Accessed on 2002/09/16.

Rašić. I. (2003) "Introduction of Nomenclature of Territorial Units for Statistics in Croatia", European Congress of the Regional Science Association „Peripheries, centres, and spatial development in the new Europe", Jyväskilä, Finland.

STATISTICA -Program Overview (1999)., StatSoft.

Strategy of the Republic of Croatia (1997), Ministry of Physical Planning, Construction and Housing-Institute of Physical Planning, Zagreb.

Šošić, I., Sredar V. (1992). *Uvod u statistiku*, Školska knjiga, Zagreb.

Uzelac Marinovic, A. (2001). *Spatial Planning*, Dom i Svijet, Zagreb.

APPENDIX

Table 1 Correlation matrix

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1,00	**-0,28**	**-0,22**	-0,01	**0,38**	**-0,29**	**0,20**	**0,22**	**-0,20**	0,10	-0,05	-0,10
X2	**-0,28**	1,00	**-0,37**	-0,10	-0,17	0,10	0,00	-0,03	**-0,19**	0,05	**0,21**	-0,02
X3	**-0,22**	**-0,37**	1,00	**-0,29**	-0,09	0,02	**-0,38**	**-0,34**	**0,52**	**0,34**	**0,23**	**0,33**
X4	-0,01	-0,10	**-0,29**	1,00	**0,30**	0,02	**0,36**	**0,32**	**-0,24**	**-0,84**	**-0,69**	**-0,25**
X5	**0,38**	-0,17	-0,09	**0,30**	1,00	**-0,64**	**0,54**	**0,57**	**-0,23**	**-0,52**	**-0,53**	-0,12
X6	**-0,29**	0,10	0,02	0,02	**-0,64**	1,00	**-0,22**	**-0,24**	0,04	0,14	**0,22**	-0,09
X7	**0,20**	0,00	**-0,38**	**0,36**	**0,54**	**-0,22**	1,00	**0,99**	**-0,35**	**-0,46**	**-0,46**	**-0,34**
X8	**0,22**	-0,03	**-0,34**	**0,32**	**0,57**	**-0,24**	**0,99**	1,00	**-0,33**	**-0,44**	**-0,45**	**-0,32**
X9	**-0,20**	**-0,19**	**0,52**	**-0,24**	**-0,23**	0,04	**-0,35**	**-0,33**	1,00	**0,37**	**0,33**	**0,22**
X10	0,10	0,05	**0,34**	**-0,84**	**-0,52**	0,14	**-0,46**	**-0,44**	**0,37**	1,00	**0,83**	**0,25**
X11	-0,05	**0,21**	**0,23**	**-0,69**	**-0,53**	**0,22**	**-0,46**	**-0,45**	**0,33**	**0,83**	1,00	**0,29**
X12	-0,10	-0,02	**0,33**	**-0,25**	-0,12	-0,09	**-0,34**	**-0,32**	**0,22**	**0,25**	**0,29**	1,00

Source: author's calculation

Table 2 List of towns and municipalities by counties and cluster

Town	Municipality	Cluster	County
Dubrovnik		2	County of Dubrovnik-Neretva
Korčula		1	County of Dubrovnik-Neretva
Metković		1	County of Dubrovnik-Neretva
Opuzen		3	County of Dubrovnik-Neretva
Ploče		1	County of Dubrovnik-Neretva
	Blato	1	County of Dubrovnik-Neretva
	Dubrovačko primorje	1	County of Dubrovnik-Neretva
	Janjina	1	County of Dubrovnik-Neretva
	Konavle	1	County of Dubrovnik-Neretva
	Kula Norinska	1	County of Dubrovnik-Neretva
	Lastovo	1	County of Dubrovnik-Neretva
	Lumbarda	1	County of Dubrovnik-Neretva
	Mljet	1	County of Dubrovnik-Neretva
	Orebić	1	County of Dubrovnik-Neretva
	Pojezerje	1	County of Dubrovnik-Neretva
	Slivno	1	County of Dubrovnik-Neretva
	Smokvica	2	County of Dubrovnik-Neretva
	Ston	1	County of Dubrovnik-Neretva
	Trpanj	1	County of Dubrovnik-Neretva
	Vela Luka	2	County of Dubrovnik-Neretva
	Zažablje	2	County of Dubrovnik-Neretva
	Župa Dubrovačka	1	County of Dubrovnik-Neretva
	Unešić	1	County of Šibenik-Knin
Hvar		1	County of Split-Dalmatia
Imotski		1	County of Split-Dalmatia
Kaštela		1	County of Split-Dalmatia
Komiža		1	County of Split-Dalmatia
Makarska		2	County of Split-Dalmatia
Omiš		2	County of Split-Dalmatia
Sinj		2	County of Split-Dalmatia
Solin		1	County of Split-Dalmatia
Split		2	County of Split-Dalmatia
Stari grad		1	County of Split-Dalmatia
Supetar		1	County of Split-Dalmatia
Trilj		1	County of Split-Dalmatia
Trogir		2	County of Split-Dalmatia
Vis		2	County of Split-Dalmatia
Vrgorac		1	County of Split-Dalmatia
Vrlika		3	County of Split-Dalmatia
	Baška Voda	1	County of Split-Dalmatia
	Bol	1	County of Split-Dalmatia
	Brela	1	County of Split-Dalmatia
	Cista Provo	2	County of Split-Dalmatia
	Dicmo	2	County of Split-Dalmatia
	Dugi Rat	2	County of Split-Dalmatia

	Dugopolje	2	County of Split-Dalmatia
	Gradac	2	County of Split-Dalmatia
	Hrvace	2	County of Split-Dalmatia
	Jelsa	1	County of Split-Dalmatia
	Klis	1	County of Split-Dalmatia
	Lećevica	1	County of Split-Dalmatia
	Lokvičići	1	County of Split-Dalmatia
	Lovreć	2	County of Split-Dalmatia
	Marina	2	County of Split-Dalmatia
	Milna	1	County of Split-Dalmatia
	Muć	1	County of Split-Dalmatia
	Nerežišća	2	County of Split-Dalmatia
	Okrug	1	County of Split-Dalmatia
	Otok	2	County of Split-Dalmatia
	Podbablje	2	County of Split-Dalmatia
	Podgora	2	County of Split-Dalmatia
	Podstrana	2	County of Split-Dalmatia
	Positra	2	County of Split-Dalmatia
	Prgomet	1	County of Split-Dalmatia
	Primorski dolac	1	County of Split-Dalmatia
	Proložac	1	County of Split-Dalmatia
	Pučišća	1	County of Split-Dalmatia
	Runovići	2	County of Split-Dalmatia
	Seget	2	County of Split-Dalmatia
	Selca	2	County of Split-Dalmatia
	Sućuraj	2	County of Split-Dalmatia
	Sutivan	1	County of Split-Dalmatia
	Šestanovac	1	County of Split-Dalmatia
	Šolta	1	County of Split-Dalmatia
	Tučepi	3	County of Split-Dalmatia
	Zadvarje	2	County of Split-Dalmatia
	Zagvozd	2	County of Split-Dalmatia
	Zmijavci	1	County of Split-Dalmatia
Drniš		3	County of Šibenik-Knin
Knin		3	County of Šibenik-Knin
Skradin		2	County of Šibenik-Knin
Šibenik		3	County of Šibenik-Knin
Vodice		3	County of Šibenik-Knin
	Biskupija	3	County of Šibenik-Knin
	Civljane	1	County of Šibenik-Knin
	Ervenik	1	County of Šibenik-Knin
	Kijevo	1	County of Šibenik-Knin
	Kistanje	1	County of Šibenik-Knin
	Murter	3	County of Šibenik-Knin
	Pirovac	1	County of Šibenik-Knin
	Primošten	3	County of Šibenik-Knin
	Promina	2	County of Šibenik-Knin
	Rogoznica	1	County of Šibenik-Knin

	Ružić	1	County of Šibenik-Knin
	Tisno	3	County of Šibenik-Knin
Benkovac		1	County of Zadar
Biograd na moru		1	County of Zadar
Nin		1	County of Zadar
Obrovac		1	County of Zadar
Pag		1	County of Zadar
Zadar		1	County of Zadar
	Bibinje	1	County of Zadar
	Galovac	2	County of Zadar
	Gračac	2	County of Zadar
	Jasenice	1	County of Zadar
	Kali	1	County of Zadar
	Kukljica	1	County of Zadar
	Lišane Ostrovičke	1	County of Zadar
	Novigrad	1	County of Zadar
	Pakoštane	1	County of Zadar
	Pašman	1	County of Zadar
	Polača	2	County of Zadar
	Poličnik	1	County of Zadar
	Posedarje	1	County of Zadar
	Povljana	1	County of Zadar
	Preko	1	County of Zadar
	Privlaka	1	County of Zadar
	Ražanac	2	County of Zadar
	Sali	2	County of Zadar
	Stankovci	1	County of Zadar
	Starigrad	1	County of Zadar
	Sukošan	1	County of Zadar
	Sveti Filip i Jakov	1	County of Zadar
	Škabrnje	1	County of Zadar
	Tkon	1	County of Zadar
	Vir	1	County of Zadar
	Zemunik Donji	2	County of Zadar

Source: author's calculation