THE APPLICATION OF FACTORIAL AND CLUSTER ANALYSIS IN REGIONAL POLICY

Jasmina Osmanković¹, Rabija Somun-Kapetanović², Emina Resić³, Muamer Halilbašić³,

 ¹School of Economics and Business Sarajevo, Bosnia and Herzegovina, jasmina.osmankovic@efsa.unsa.ba
²School of Economics and Business Sarajevo, Bosnia and Herzegovina, rabija.somun@efsa.unsa.ba
³School of Economics and Business Sarajevo, Bosnia and Herzegovina, emina.resic efsa.unsa.ba
⁴School of Economics and Business Sarajevo, Bosnia and Herzegovina, muamer.halilbasic@efsa.unsa.ba

Abstract

The paper deals with the potential use of factorial and cluster analysis in defining a framework for regional economic policy, as well as a statistics framework acceptable to Eurostat. The spatial context of the study is Bosnia and Herzegovina. The paper uses factorial and cluster analysis and comparative analysis. The experience of Croatia in the application of these methods to the grouping of local entities by homogeneity is particularly relevant for the purposes of comparison. The topicality of the research is determined by the need to define an effective territorial or spatial framework for Bosnia and Herzegovina's regional economic policy, and to ensure that the framework will facilitate access to European funds, programmes and projects.

JEL classification: R11, R14, R58

Key words: factorial analysis, cluster analysis, regional policy, Bosnia and Herzegovina.

Introduction

The paper contain the results from scientific and research project Statistic (NUTS) Regionalization – the application of the factor and Klaster analysis on the example of Bosnia and Herzegovina financed by Fund for science. (Osmanković et al, 2009). Direct reason is in the Bosnia and Herzegovina obligation to select methods and manners of statistic, or NUTS regionalization in accordance to the goals, principles, the EUROSTAT criterions and experiences of the countries that defined NUTS regions of the second and the third level within the process of stabilization and association to EU. It understands the acceptance of the NUTS nomenclature, its principles, goals and

criterions. The purpose and essence of the statistic reorganization according to the EUROSTAT is in its use for the collection, improvement, harmonization of regional statistics in the countries EU members for the socio-economic analysis of the region, for defining regional policy of EU, for the accession to the funds, programs and projects of EU intended for the support of regional development. By signing the Agreement on stabilization and association all countries, including Bosnia and Herzegovina, are obliged to, besides others, within five years upon signing the Agreement, deliver to the Commission the indicators on GDP per capita for the level NUT 2. (Anon, 2007; Anon, 2007a) The similar importance was the interest for the application of relevant scientific, research, statistic methods within regional economic researches. The subject is the application of the factor and Klaster analysis in the procedure of the grouping of local units for the needs of the NUTS regionalization in Bosnia and Herzegovina.

The Methods

Certain statistic methods, especially factor and Klaster analysis, may be used in the procedure of grouping local units for the needs of the statistic (NUTS) regionalization in Bosnia and Herzegovina in order to define optimal regions for statistics, leading of regional policy and the access to the funds, programs and projects. It is the basic hypothesis being tested within this material. Its approval may contribute to the finding the solution for the salvation of actual situation which is "burdened" by the events from the first half of the nineties of the last century, by political tensions, by wrong or non systematic advises from local and international experts, by parallel existence of numerous regionalization for the sake of different social, administration, economic, political, social functions (Kronthaler, 2003; Somun-Kapetanović et al, 2007, Osmanković et al, 2008; Osmanković et al, 2008a).

Data contained in the publications of the Federal Institute for Statistics, Institute for Statistics of Republic Srpska, Federal Institute for programming, Central Bank of Bosnia and Herzegovina, tourist communities in Bosnia and Herzegovina, data from local communities, data from Institute for Employment of Republic Srpska, Institute for Employment of the Federation BiH, Institute for the employment in the cantons of the Federation BiH, data from the Ministry of Finance BiH, data from corresponding international organizations and institutions available in hard or/and e-form were used as the sources for the creation of relevant data base for factor and Klaster analysis (Kurnoga-Živadinović, 2007; Lovrinčević, et al, 2005; Marcou, 2005)

Factor analysis is a collection of mathematic-statistic proceedings providing in greater number of variables between which there is certain connection, to

determine smaller number of important variables which explain their connection. Prior to factor analysis, the analysis of coefficient of input correlation (input variables) is done. As input variables the following were used: area in square km, GDP per capita, employment rate in %, structure of employment – primary sector in %, structure of employment – secondary sector in %, structure of employment – tertial sector in %, total number of tourists, and total number of nights and coefficient of vitality. All variables were given for the year 2007. The analysis of the main components leads towards the main dimensions of development of treated areas. Upon conclusion of the factor analysis, its results are used in Klaster analysis in order to get homogenous regions (Klasters) composed from local communities of similar characteristics. (Rašić, 2005, 2). Klaster analysis is a technique of municipalities grouping into the group of municipalities or into the region. Municipalities within one region are similar and regions are different. Similarity between municipalities is identified by using different measures of distance, more often Euclid distance. The methods of the hierarchy Klaster analysis, k-means method, and alternative approach for the grouping of local units in Bosnia and Herzegovina were applied.

In the performing of research, the authors of the study found some difficulties. The first difficulty was related to relatively small number of the indicators of socio-economic development available on the level of municipalities in Bosnia and Herzegovina. The second closely connected problem was related to the fact that some indicators, such as GDP on the municipality level, were the results of the estimation from the Institute for programming of development in the Federation BiH.

NUTS regionalisation as framework for statistic and regional policy

The Decree number 1059/2003 from May 26 2003, users of the European statistics expressed the increasing need for the harmonization in order to get comparative data for the EU territory. Strategy of the Bosnia and Herzegovina integration to EU defined the following necessary measures under the items 3 and 5: to develop system of the official statistics in accordance to the EU and international standards and recommendations and as well to the local demands and assure continuous cooperation with Eurostat. Specific are the articles 88, 87 and 71 from the Agreement on Stabilization and Association between European Union, its members and Bosnia and Herzegovina. In accordance to this, the Commission is of the opinion that those conditions are fulfilled if the region, in accordance to the NUTS 2 level classification of geographic units, has gross domestic product (GDP) per capita, measured according to the parity of purchasing power, lower than 75% from the community average. Certain regions on the level NUTS 2, where GDP per capita overpasses 75% of the

Community average purely as a statistic effect of enlargement, are still, on the transition basis, considered adequate for such kind of help.

Further on, article 71, paragraph 7b of the Agreement says that by the end of the fifth year from the Agreement coming into effect, Bosnia and Herzegovina will deliver to the Commission the indicators on GDP per capita harmonized on the NUTS 2 level. Independent public body which will be formed meanwhile, and which will treat these issues, will together with the Commission make common evaluation on the feasibility of certain regions in BiH and the maximal intensity in this regard in order to complete the map on regional help on the basis of relevant instructions from the commission.

Results of the factor and klaster analysis

Results of the application of factor and Klaster analysis on the basis of socioeconomic characteristics in grouping local units for the needs of defining optimal statistic NUTS regions and the frameworks of corresponding economic regional politics on the example of Bosnia and Herzegovina and Croatia show that it is one of possible manners which is not used enough. Those analyses are mostly used in sociological, psychological and research marketing.

Starting from the criterion of homogeneity, on the territory of Bosnia and Herzegovina, more options are identified. Respect of the interpretation reduced their number to two: with four and six Klasters. Analysis confirmed that spatial distance between certain local units does not understand distance or difference with regard to their homogenous according to socio-economic characteristics.

For the grouping of the units of local self-governance a combination of factor and Klaster analyses is used. This methodology is common in a situation in which we have at disposal bigger number of variables. Using factor analysis, the information contained in greater number of analyzed variables are summarized; after this, Klaster analysis based on formed factors is performed.

| Factor | Klaster | | | | | |
|----------|----------|----------|----------|----------|--|--|
| | 1 | 2 | 3 | 4 | | |
| Factor 1 | 0,25832 | -0,42543 | 0,12723 | 6,66822 | | |
| Factor 2 | 1,05733 | -0,66823 | -0,12798 | -1,50965 | | |
| Factor 3 | -0,43768 | -0,53538 | 1,20994 | -0,89286 | | |

Table 1: Final centers of Klasters and influence of factors

Source: Author's calculation

Final centers of Klasters and influence of factors indicate positive correlation of the first Klaster and factor 1 and 2, third and fourth Klasters and factor 1.

| Table 2. Dist | Table 2. Distance between final centers of Masters | | | | | | | |
|---------------|--|-------|-------|-------|--|--|--|--|
| Klaster | 1 | 2 | 3 | 4 | | | | |
| 1 | | 1,859 | 2,034 | 6,920 | | | | |
| 2 | 1,859 | | 1,909 | 7,152 | | | | |
| 3 | 2,034 | 1,909 | | 7,008 | | | | |
| 4 | 6,920 | 7,152 | 7,008 | | | | | |

Table 2: Distance between final centers of Klasters

Source: Author's calculation

Basic socio-economic characteristics of local units grouped into homogenous Klasters are presented in the next table.

| Table 3: Basic characteristics of four K | lasters according to the results of factor |
|--|--|
| and Klaster analysis | |

| Variable | | Klaster 1 | Klaster 2 | Klaster 3 | Klaster 4 |
|--------------------------|--------------|-----------|-------------------|-----------|-----------|
| Number of municipalities | | 36 | 48 | 35 | 1 |
| Area | in square km | 11.822 | 12.357 | 23.099 | 148 |
| Number of population | | 1.226.369 | 882.680 1.259.323 | | 304.070 |
| Rate of employment | | 18,9 | 10,7 | 14,1 | 28,9 |
| GDP per capita KM | | 4.974 | 3.229 | 4.361 | 13.626 |
| GDP pc % EU ¹ | | 10 | 7 | 9 | 28 |
| Reprimary | | 2,5 | 4,7 | 10,4 | 0,7 |
| cture | Secondary | 50,5 | 25,4 | 30,0 | 19,8 |
| Stru | Tertial | 47,0 | 69,9 | 59,6 | 79,5 |
| Coefficient of vitality | | 18 | 11 | 12 | 15 |

Source: Author's calculation

¹ Average GDP pc EU for 2007. is 24.900 EUR according to <u>http://www.economic-growth.eu/Englis/updated_data/chart-EU-GDP-per-capita2007.htm 20. 12</u>. 2008

In the situation when the number of variables being at disposal is lower, there is no need for the creation of factors. In this manner it is tried to avoid the loss of important information. In this case, firstly, the corresponding method of hierarchy Klasterisation is performed. After this, results are used as input in k means Klaster method. Hierarchy method provides to determine a number of groups (Klasters) and centroide of groups. In the next step, k-method is used in order to improve the results of hierarchy method. Namely, the main lack of hierarchy method is that the allocation of units is final, with no possibility of regrouping into others (more corresponding) groups during the very procedure. K-means method, on the other side, is sensitive to proscribed initial values and may finish in a trap of the local optimum which is far from the global optimum. Empiric proves suggest that it is possible to come close to the global optimum if centroides from hierarchy method are used (Ferligoj, 1989) as the starting points for k-means method.

| Variable | Klaster | | | | | |
|---|---------|---------|---------|---------|---------|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| GDP per capita | 3993,11 | 6007,91 | 5571,00 | 3793,27 | 4940,00 | 13626,00 |
| Employment rate | 0,1510 | 0,1971 | 0,2457 | 0,1282 | 0,1991 | 0,2885 |
| Structure of employment - primary sector | 0,0489 | 0,0203 | 0,0200 | 0,0630 | 0,0011 | 0,0065 |
| Structure of employment - secondary sector | 0,3392 | 0,3587 | 0,3059 | 0,3480 | 0,1384 | 0,1976 |
| Structure of employment – tertial sector | 0,6119 | 0,6210 | 0,6742 | 0,5890 | 0,8605 | 0,7958 |
| Tourists total | 3019 | 12273 | 29018 | 412 | 55902 | 95644 |
| Nights total | 6385 | 25728 | 101615 | 688 | 184927 | 187202 |
| Vitality coefficient | 0,1604 | 0,1565 | 0,1655 | 0,1555 | 0,1183 | 0,1526 |

Table 4: Final centers of Klasters upon performed k-means method

Source: Author's calculation

Into the Klaster one 18 units of local self-governance is divided, in Klaster two 11, in Klaster three 4, while Klasters five and six contain only one unit each.

| | 1 | 18,000 | |
|---------|---|---------|--|
| ster | 2 | 11,000 | |
| | 3 | 4,000 | |
| Xla | 4 | 85,000 | |
| H | 5 | 1,000 | |
| | 6 | 1,000 | |
| Valid | | 120,000 | |
| Missing | | 10,000 | |

| Table 5: Number of | cases in each | Klaster according to non-l | hierarchy analysis |
|--------------------|---------------|----------------------------|--------------------|
| | | - | |

Source: Author's calculation

Basic socio-economic characteristics of local units which are grouped into six homogonous Klasters are given in the following table.

Table 6: Basic characteristics of six Klasters according to the results of k-means non-hierarchy analysis

| Vari | able | Klaster 1 | Klaster 2 | Klaster 3 | Klaster 4 | Klaster 5 | Klaster 6 |
|----------------------|--------------------|-----------|-----------|-----------|--------------|--------------|-----------|
| Num mun | ber of icipalities | 18 | 11 | 4 | 85 | 1 | 1 |
| Area in square km | | 7.627 | 6.478 | 2.904 | 30.045 | 225 | 148 |
| Num popu | ber of lation | 637.908 | 740.848 | 353.992 | 1.630.944 | 4.682 | 304.070 |
| Rate emp | of loyment | 15 | 20 | 25 | 13 | 20 | 29 |
| GDF KM | P per capita | 3.993 | 6.008 | 5.571 | 3.793 | 4.940 | 13.626 |
| GDF | $P pc \% EU^2$ | 8 | 12 | 11 | 8 | 10 | 28 |
| % | Primary | 5 | 2 | 2 | 6 | 0 | 1 |
| tructure | Secondary | 34 | 36 | 31 | 35 | 14 | 20 |
| | Tertial | 61 | 62 | 67 | 59 | 86 | 79 |
| S Coef vital | ficient of | 16 | 16 | 17 | 16 | 12 | 15 |

Source: Author's calculation

² Average GDP pc EU for 2007 is 24.900 EUR according to <u>http://www.economic-growth.eu/Englis/updated_data/chart-EU-GDP-per-capita2007.htm 20. 12</u>. 2008

Results of the performed factor and Klaster analysis may be resumed in the form of two options. The first option is a result of factor and Klaster analysis and the second is a result of the application of the alternative approaches such as k-means approach on the basis of result non-hierarchy analysis.



Map prepared by Mulaomerović, J.

The first option identifies four Klasters, and the other six. Within both options, the city of Sarajevo, according to its characteristics, significantly differentiates in relation to its narrow and wider surrounding. As it is about capital city, or a "real city", the city with the highest level of economic development, with the concentration of economic and the most qualitative demographic potentials, with the concentration of functions in education, research, health care, culture, financial sector, it is expectable. Still, demographic size, so as some other criterions, such as historic, institutional, competitive do not support the option according which the city of Sarajevo, as the city community of four municipalities, should be treated as an optimal statistic unit and an optimal framework of the regional policy.

Over two thirds of local units by k-means method are grouped according to the criterion of homogeneity into one Klaster. It directs to BiH as NUTS 2 level, the option which would parallel and in greatest measure satisfy all three criterions: demographic, access to funds and homogeneity. The following comments support this: it won't be a precedent in the formation of regions, there is a statistic system which may, with certain technical assistance, provide more quality statistic base, there are institutions for the preparation of applications, acceptance of funds, their implementation, monitoring and

evaluation, the approach which in the greatest extent decreased arbitration and subjectivism in reaching decisions.

REFERENCES

- 1. Anon (2007): Nomenclature of Territorial Units for Statistics, *Avaliable from* <u>http://www.answers.com/topic/nomenclature-of-territorial-units-for-statistics</u>, *Access* (15. 07. 2007)
- Anon (2007a): Nomenclature of territorial units for statistics NUTS Statistical Regions of Europe, *Avaliable from* <u>http://ec.Evropa.eu/comm/ eurostat/ramon/nuts/home_regions_en.html</u>, *Access* (15. 07. 2007)
- 3. Kronhaler, F. (2003). A Study of the Competitiveness of Regions based on a Cluster Analysis: The Example of East Germany, *Avaliable from* <u>http://www.iwh-halle.de</u>, *Access (*20. 11. 2008)
- Kurnoga-Živadinović, N. (2007). Multivarijantna klasifikacija županija Hrvatske. Zbornik radova Ekonomskog fakulteta Zagreb, Zagreb, broj 5/07, p., ISSN 1333-899
- Lovrinčević, Ž. et al (2005). Kako optimalno regionalizirati Hravtsku? Zbornik radova Ekonomska politika Hrvatske u 2006. HDE i Inžinjerski biro Zagreb, objavljen u skraćenoj i izmjenjenoj verziji u Ekonomski pregled, p 1109-1160, *Avaliable from <u>http://hde.hr/naslovnica.aspx</u>, <i>Access* (20. 02. 2009)
- Marcou, G. (2005). Regionalization for Development and Accession to the European Union: A Comparative Perspective Hungary, Lithuania, Poland, Slovakia, Ukrajina. *Avaliable from* <u>http://lgi.osi.hu/publications/2002/105/11-Intro.pdf</u> - Access (05. 5. 2007)
- Osmanković, J. et al. (2008). Analiza potencijalnih modela statističke regionalizacije u skladu sa NUTS kriterijima u BiH. Zbornik radova, broj 28/2008., Ekonomski fakultet Sarajevo p72-93, ISSN 0581-7439, *Avaliable from <u>http://www.epnet.com/titleLists/bt-jounrals.xls. Access</u> (10. 12. 2008)*
- Osmanković, J. et al. (2008a). The use of factor and cluster analysis to support regional development policy and adoption of NUTS classification in Bosnia and Herzegovina. Zbornik radova International Scinetific Conference on the Occasion of the 50th Anniversary of the IES CHALLENGES IN THE 21ST CENTURY. p75-83. Ekonomski institut Beograd, ISBN 978-86-80315-71-3, Beograd, 4 - 5. 12. 2008.
- Osmanković, J. et al. (2009). Statistička (NUS) regionalizacija, Ekonomski institut Sarajevo, ANU Odjeljenje ekonomskih nauka, ISBN 978-9958-687-06-8, Sarajevo

- 10. Rašić I. (2005). Primjena faktorske i klaster analize u grupiranju jedinica lokalne samouprave odabranih županija na temelju društvenogospodarskih pokazatelja, magistarski rad, Ekonomski fakultet Sveučilišta u Zagrebu, Zagreb
- Somun-Kapetanović, R. et al. (2007). Studija o procjeni uticaja regulative (RIA) u procesu evropskih integracija u Bosni i Hercegovini. DEI BiH, Sarajevo, *Avaliable from www.dei.ba*, *Access* (20. 09. 2008)