Evaluation of Regional Disparities in Visegrad Four Countries, Germany and Austria using the Cluster Analysis

Ing. Eva Poledníková

Ing. Petra Lelková

VŠB – Technical University of Ostrava
Faculty of Economics
Department of European Integration

Katedra
evropské
integrace

ekf

VYSOKÁ ŠKOLA BÁŇSKÁ
TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

Content

1. **Introduction**

2. **Theoretical background of regional disparities in the EU**

   - cluster analysis as a selected method of regional disparities measurement

3. **Application of cluster analysis in the case of V4, Germany and Austria regional disparities**

4. **Conclusion**

Context of the Research Framework

„Disparities, Cohesion and Competitiveness in Countries and Regions of the Visegrad Four Group in the Context of Selected Advanced EU's Countries".

Project registration number:	SP 2012/153
Period of research:	1. 1. 2012 – 14. 12. 2012
Recipient:	VŠB–TU Ostrava, Faculty of Economics, Department of European Integration
Supervisor: **Team Leader:** **Team Members:**	prof. Ing. Karel Skokan, Ph.D. Ing. Lukáš Melecký doc. Ing. Jana Hančlová, CSc. Ing. Michaela Staníčková Ing. Eva Poledníková Ing. Petra Lelková

1. Introduction

- Elimination of disparities between less and more developed EU Member Countries and their regions is a long-term development goal of the EU.

- There are many opinions and methods of measurement and evaluation of regional disparities at national and European level.

- More sophisticated methods that can contribute to regional disparities measurement and evaluation represent multivariate methods. Cluster analysis identifies groups of objects that are very similar and assign them into clusters.

- *The main objective of paper is to classify the NUTS 2 regions of the Visegrad Four countries, Germany and Austria to the optimal number of the homogeneous clusters according to the similarity of the selected economic, social and territorial indicators. The paper aims to evaluate and compare the development of regional disparities in selected countries in the period 2000-2010.*

ekf VYSOKÁ ŠKOLA BÁŇSKÁ TECHNICKÁ UNIVERZITA OSTRAVA EKONOMICKÁ FAKULTA

2. Theoretical background of regional disparities in the EU

- In the frame of EU cohesion, three basic types of disparities are usually classified: economic, social and territorial.

- In the European concept, the level of disparities can be regarded as a measure of the level of cohesion and *the lower rate of disparities, the higher the level of cohesion territory achieves and vice versa* (Molle, 2007).

- There is no mainstream approach to measurement and evaluation of disparities.

- In the EU, the regional disparities are analysed mostly at the territorial level of NUTS 2 regions. The adequate indicators of regional disparities can be identified in the *Reports on Economic, Social and Territorial Cohesion (*2007, 2010). Another possible approach is the usage of *Structural indicators or indicator´s of Europe 2020*.

- Data base of convenient indicators for regional disparities evaluation is marked by significant differences in terms of indicators availability on the territorial levels within EU Member States and the reference period.

- The methods based on inter-regional comparison or mathematical and statistical methods are often used, namely:
  - variability level (e.g. standard deviation and variation coefficient); methods utilizing Geographical information system; multivariate statistical methods (e.g. method of main components and factor, cluster or discrimination analyses); simplistic models; method of real convergence; modified territorial Gini coefficient or method of artificial neuron nets.

- Multivariate statistical methods are an ever-expanding set of techniques for data analysis that encompasses a wide range of possible research situation.

ekf VYSOKÁ ŠKOLA BÁŇSKÁ
TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

- Cluster analysis is a group of multivariate method whose primary purpose is to group objects based on the characteristics they possess. The resulting cluster of objects should exhibit high internal homogeneity and high external heterogeneity.

- The most popular procedures represent the hierarchical methods and nonhiearchical methods.

- The hierarchical cluster analysis uses the dissimilarities such as distances between objects when forming the clusters. The distance is mostly defined as *Euclidean distances* or the *Squared Euclidean distance* .

$$E_{i,c,t} = \left(u_{i,\mathrm{r},t} - \hat{u}_{50i} \right)^2 \qquad (2)$$

Where:

|	
E	Euclidean distance ;
$u_{i,r,t}$	standardized value of *i-th* indicator for *r-th* region in time *t*;
$\hat{u}_{50i}$	median for *i-th* indicator ;
i	indicator of disparities;
r	region; r = {1, ……..83};
t	time; t = {2000, …, 2010}.

6

- After the determination of the distance measure, the clustering algorithm has to be selected. The most frequently used methods are:
  - *nearest neighbour*,
  - *furthest neighbour,*
  - *average linkage with (between) groups,*
  - *Ward´s method,*
  - *centroid method,*
  - *median method.*

- For cluster analysis, it was selected 15 indicators that represent the most frequently indicators used in *Cohesion Reports* (2007, 2010), some of them represent the *EU Structural indicators*.

Type of disparities	Indicator	Abbreviation
Economic disparities	GDP per head (Purchasing Power Standards per inhabitant)	GDP
	Gross fixed capital formation (Millions of euro)	GFCF
	Disposable income of households (Purchasing Power Standard per inhabitant)	DI
	Human Resources in Science and Technology (% of active population)	HRTS
	Patent applications to the European Patent Office (EPO) (Number per million of inhabitants)	EPO
	Employment in technology and knowledge-intensive sectors (% of active population)	ETKI
Social disparities	Employment rate (% of population aged 15-64)	ER15to64
	Employment rate of older workers (% of population aged 55-64)	ER55to64
	Employment rate of woman (% of woman population aged 15-64)	ERw15to64
	Unemployment rate (% of labour force aged 15-64)	UR15to64
	Unemployment rate of youth (% of labour force aged 15-24)	URy15to24
Territorial disparities	Collective tourist accommodation establishments (Number)	TE
	Tourism intensity (Number)	TI
	Crude death rate (Number per 100,000 inhabitants)	CD
	Victims in road accident (Number)	VRA

7

Source: European Commission, 2007, European Commission, 2010, Eurostat, 2012; own elaboration
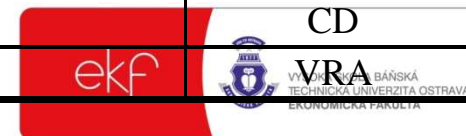
- Because of the *correlation* it was necessary to *remove* two indicators from the follow up analysis. The final input matrix for cluster analysis is created by 5 economic indicators, 4 social indicators and 4 territorial indicators in the years 2000 and 2010.

- Firstly, the attention should be paid to the *characteristic and quality* of *data file*. On the basis of *descriptive statistics*, the significant differences between regions of V4, Germany and Austria exist within occur within the economic indicators GDP, GFCF, EPO, ETKI and the territorial indicators TE, TI, VRA. The positive development is recorded in the social indicators.

ekf

VYSOKÁ ŠKOLA BÁŇSKÁ
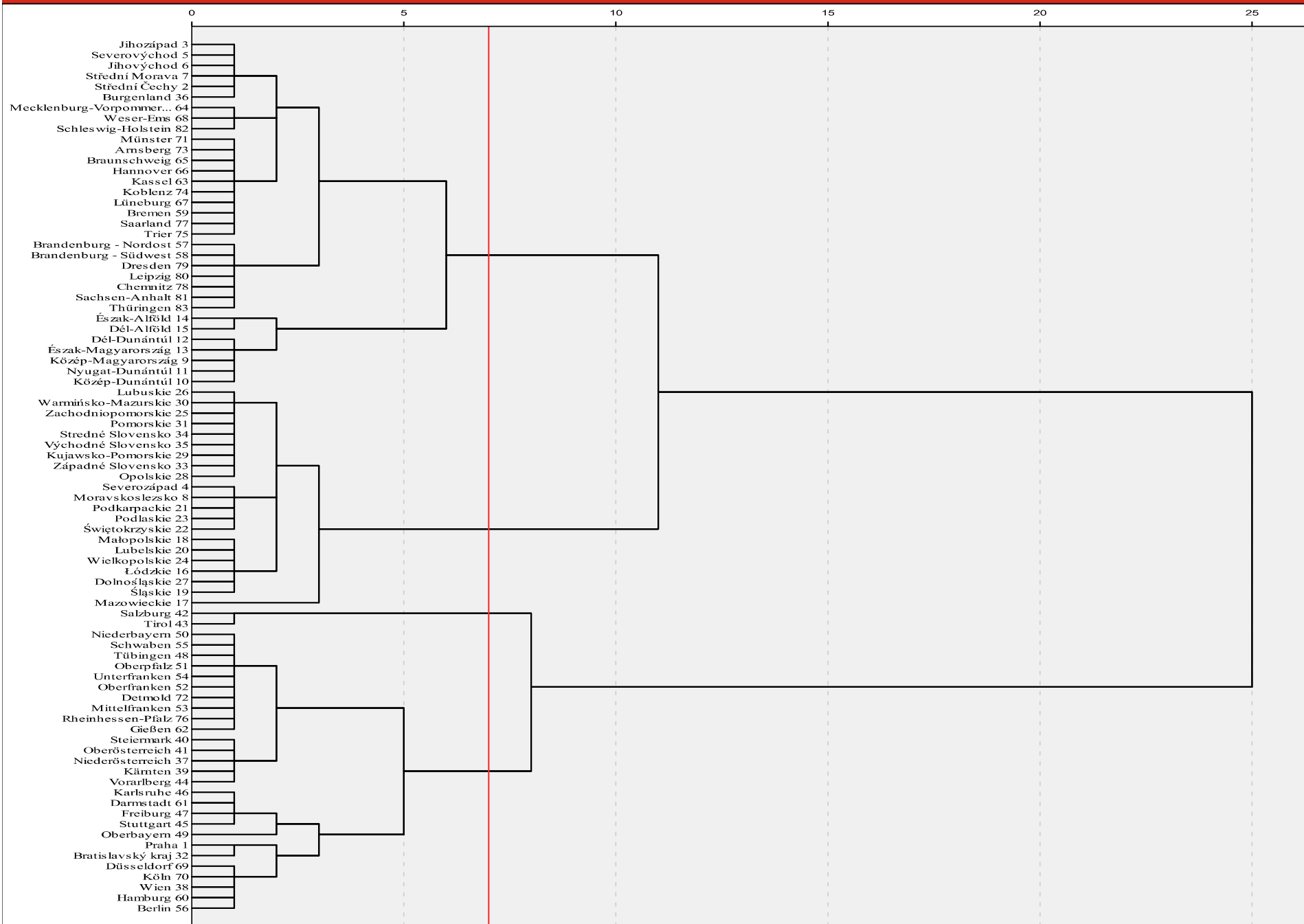TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

Results

- Proximity Matrix - Squared Euclidean Distance as a measure of dissimilarity:

- In 2000, the *highest* distance (a thus the highest rate of the dissimilarity) was recorded between region *Lubuskie* and region *Oberbayern* (138,816). The lowest distance (and the highest rate of the similarity) was between regions *Jihovýchod* and *Severovýchod* (0,810).

- In the 2010, the disparities between *Lubuskie* and *Oberbayern* were **reduced** when distance decreased on the value 113,254. *The highest* distance was recorded between Hungarian region *Észak-Magyarország* and Austrian region *Tirol* (151,667). The *lowest* distance was noticed between two Polish regions *Lubuskie* and *Opolskie* (0,342).

9

Results

- Agglomeration Schedule- provides the information about the hierarchical clustering process – especially includes the value of *Coefficient*s that help us to decide how many clusters are optimal for representation of the data.

- In the year 2000 and 2010 the optimal four-cluster solution was determined.

- However, the results of cluster analysis in the year 2010 differ in the terms of the steps at which cluster was joined to each other and in the structure of the cluster.

ekf

VYSOKÁ ŠKOLA BÁŇSKÁ
TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

Dendogram (2000)

Jihozápad 3
Severovýchod 5
Jihovýchod 6
Střední Morava 7
Střední Čechy 2
Burgenland 36
Mecklenburg-Vorpommer... 64
Weser-Ems 68
Schleswig-Holstein 82
Münster 71
Arnsberg 73
Braunschweig 65
Hannover 66
Kassel 63
Koblenz 74
Lüneburg 67
Bremen 59
Saarland 77
Trier 75
Brandenburg - Nordost 57
Brandenburg - Südwest 58
Dresden 79
Leipzig 80
Chemnitz 78
Sachsen-Anhalt 81
Thüringen 83
Észak-Alföld 14
Dél-Alföld 15
Dél-Dunántúl 12
Észak-Magyarország 13
Közép-Magyarország 9
Nyugat-Dunántúl 11
Közép-Dunántúl 10
Lubuskie 26
Warmińsko-Mazurskie 30
Zachodniopomorskie 25
Pomorskie 31
Stredné Slovensko 34
Východné Slovensko 35
Kujawsko-Pomorskie 29
Západné Slovensko 33
Opolskie 28
Severozápad 4
Moravskoslezsko 8
Podkarpackie 21
Podlaskie 23
Świętokrzyskie 22
Małopolskie 18
Lubelskie 20
Wielkopolskie 24
Łódzkie 16
Dolnośląskie 27
Śląskie 19
Mazowieckie 17
Salzburg 42
Tirol 43
Niederbayern 50
Schwaben 55
Tübingen 48
Oberpfalz 51
Unterfranken 54
Oberfranken 52
Detmold 72
Mittelfranken 53
Rheinhessen-Pfalz 76
Gießen 62
Steiermark 40
Oberösterreich 41
Niederösterreich 37
Kärnten 39
Vorarlberg 44
Karlsruhe 46
Darmstadt 61
Freiburg 47
Stuttgart 45
Oberbayern 49
Praha 1
Bratislavský kraj 32
Düsseldorf 69
Köln 70
Wien 38
Hamburg 60
Berlin 56

| Year	2000	2010	Year	2000	2010
Region	Cluster	Cluster	Region	Cluster	Cluster
Praha	1	1	Stuttgart	1	3
Střední Čechy	2	2	Karlsruhe	1	1
Jihozápad	2	2	Freiburg	1	3
Severozápad	3	2	Tübingen	1	3
Severovýchod	2	2	Oberbayern	1	1
Jihovýchod	2	2	Niederbayern	1	3
Střední Morava	2	2	Oberpfalz	1	3
Moravskoslezsko	3	2	Oberfranken	1	3
Közép-Magyarország	2	3	Mittelfranken	1	3
Közép-Dunántúl	2	2	Unterfranken	1	3
Nyugat-Dunántúl	2	2	Schwaben	1	3
Dél-Dunántúl	2	2	Berlin	1	1
Észak-Magyarország	2	2	Brandenburg - Nordost	2	3
Észak-Alföld	2	2	Brandenburg - Südwest	2	3
Dél-Alföld	2	2	Bremen	2	3
Łódzkie	3	2	Hamburg	1	1
Mazowieckie	3	2	Darmstadt	1	1
Małopolskie	3	2	Gießen	1	3
Śląskie	3	2	Kassel	2	3
Lubelskie	3	2	Mecklenburg-Vorpommern	2	3
Podkarpackie	3	2	Braunschweig	2	3
Świętokrzyskie	3	2	Hannover	2	3
Podlaskie	3	2	Lüneburg	2	3
Wielkopolskie	3	2	Weser-Ems	2	3
Zachodniopomorskie	3	2	Düsseldorf	1	1
Lubuskie	3	2	Köln	1	1
Dolnośląskie	3	2	Münster	2	3
Opolskie	3	2	Detmold	1	3
Kujawsko-Pomorskie	3	2	Arnsberg	2	3
Warmińsko-Mazurskie	3	2	Koblenz	2	3
Pomorskie	3	2	Trier	2	3
Bratislavský kraj	1	1	Rheinhessen-Pfalz	1	3
Západné Slovensko	3	2	Saarland	2	3
Stredné Slovensko	3	2	Chemnitz	2	3
Východné Slovensko	3	2	Dresden	2	3
Burgenland	2	3	Leipzig	2	3
Niederösterreich	1	3	Sachsen-Anhalt	2	2
Wien	1	1	Schleswig-Holstein	2	3
Kärnten	1	3	Thüringen	2	3
Steiermark	1	3			
Oberösterreich	1	3			
Salzburg	4	4			
Tirol	4	4			
Vorarlberg	1	3			

Cluster´s Profile (2000)
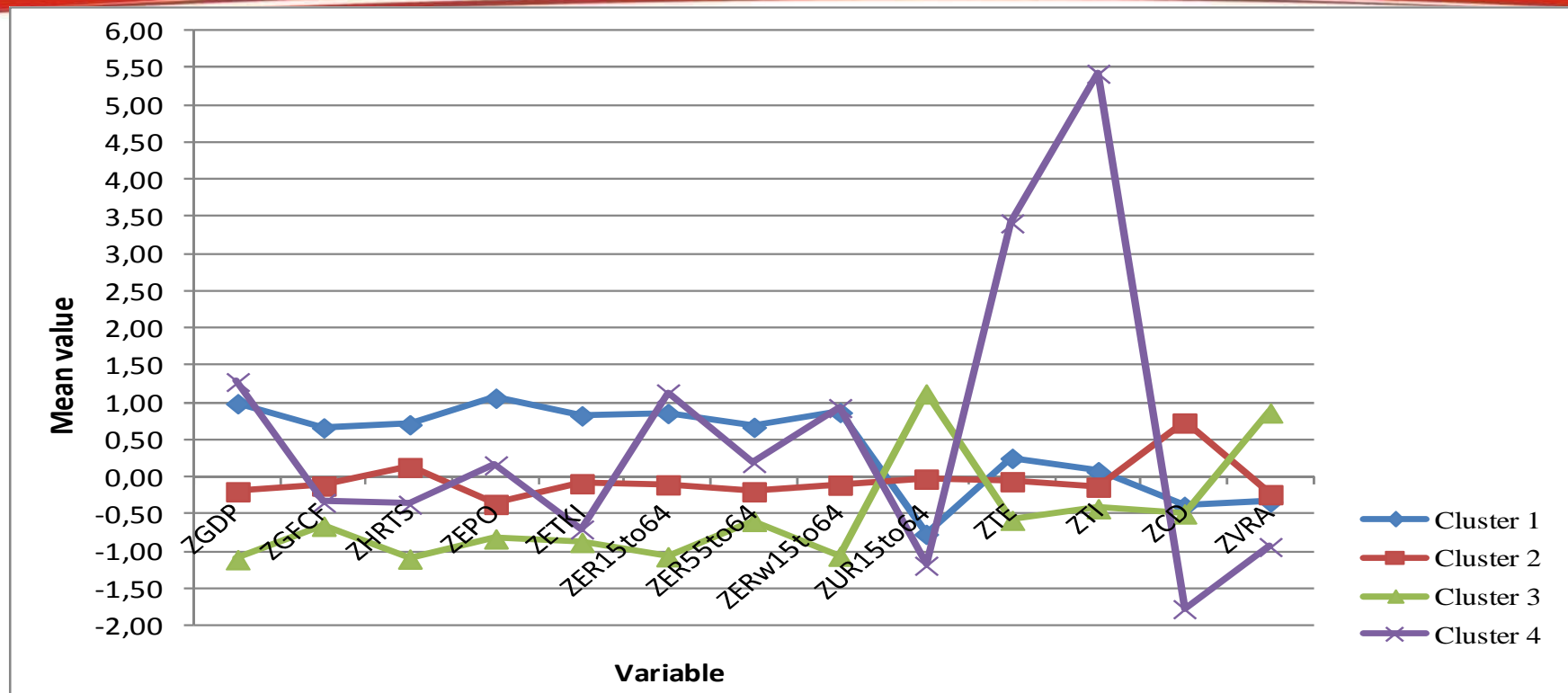
Cluster 1 achieves the highest average value of economic indicators with exception of indicator GDP and one of the best results in the labour market. The lowest average rate of unemployment and the best development of territorial indicators are reached by Cluster 4. Cluster 3 is characterized by the worst results in the economic, social and territorial segment.

Cluster´s Profile (2010)
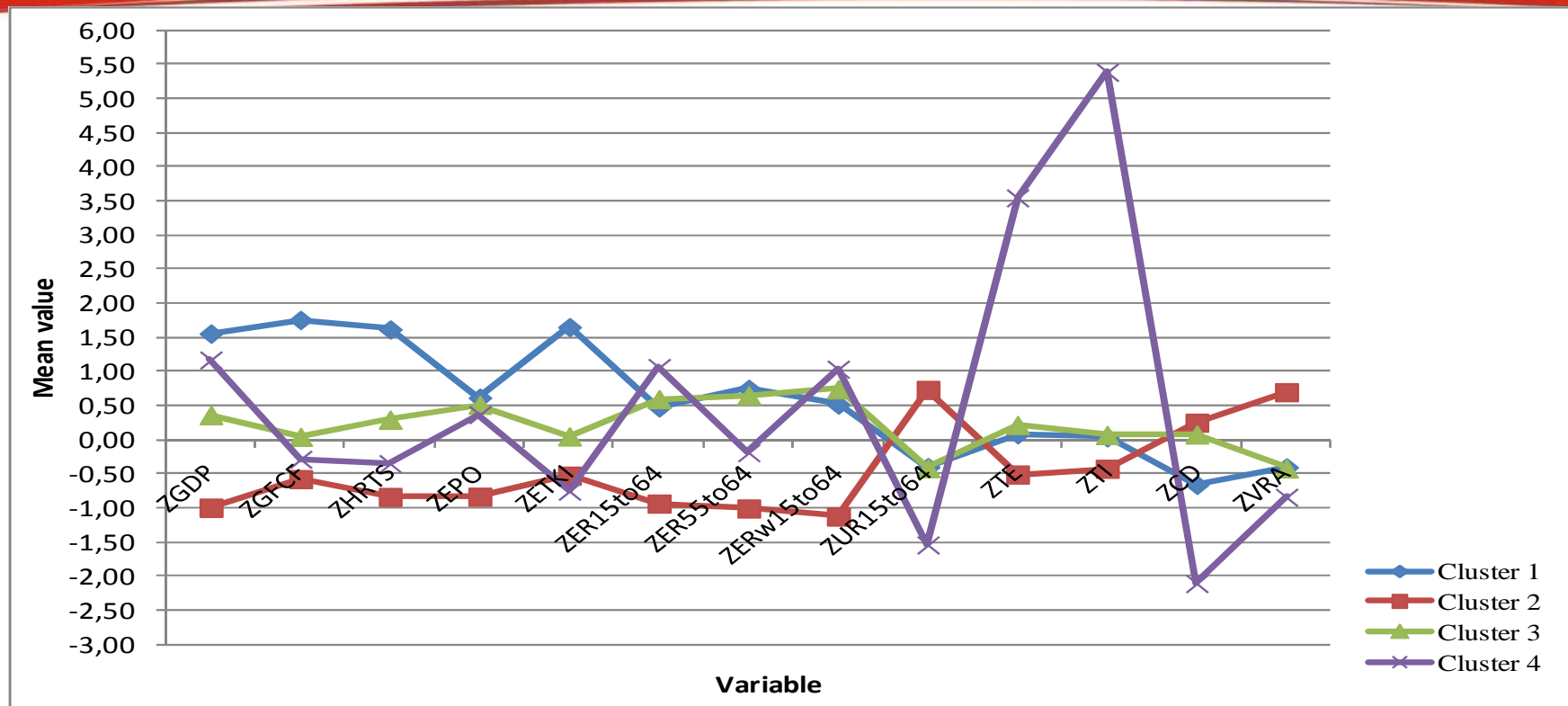
Cluster 1 achieves the highest average value of economic indicators and the highest average value of rate of employment of older workers. Cluster 2 is cluster with biggest troubles although the regions have recorded the positive development of the analyzed indicators. The modification of the Cluster 2 can be seen as a positive development in the increasing of the convergence between the less and more developed regions and reducing disparities. Cluster 4 achieves the lowest rate of unemployment and the best results in the territorial cohesion.

4. Conclusion

- On the basis of *Ward's method* applying the *Squared Euclidean Distance*, the optimal four-cluster solution has been determined in the study of V4, Germany and Austria regional disparities in the reference period 2000-2010.

- The results partially confirms the hypothesis that NUTS 2 regions with *capital cities* (*Praha*, *Bratislavský kraj*, *Wien* and *Berlin*) have had significant and different soci-economic position from the other regions in the year 2000 as well as in the year 2010. These regions have tended to be naturally grouped into one homogeneous cluster that has separated from the other clusters.

- In the year 2010, the modification of the cluster membership can be seen as the increasing of the similarity (reducing the disparities) of less and more developed regions within economic, social and territorial indicators.

- Despite of, the significant disparities has persisted between the regions in Hungary, Poland and Slovakia and the regions of capital cities and high developed regions in Austria, Germany and Czech Republic.

13

ekf

VYSOKÁ ŠKOLA BÁŇSKÁ
TECHNICKÁ UNIVERZITA OSTRAVA
EKONOMICKÁ FAKULTA

4. Conclusion

- Informative level of the results of cluster analysis depends on the researcher´s attitude to the selection of indicators (e.g. occurrence of outliers), technique of the clustering, criterion of the distance and algorithm (method) of the clustering.

- Multivariate methods eligible to concentrate the information and detect the relationship and coherence between them. Therefore these methods offer the great research´s potential.

- In the study of the regional disparities cluster analysis can provide useful information that can be used to facilitate the determination of the development measures and the decision about the subsidies allocation for the needy regions.

References

- Burns, R. B., Burns, R. A. (2008). Business Research Methods and Statistics Using SPSS. 2008. London: Sage, 2008. 544 p. ISBN 978-1-4129-4529-5.

- European Commission (2010). Fifth report on economic, social and territorial cohesion - Investing in Europe's future. Brussels: European Commission, 2010. 304 p. ISBN 978-92-79-17800-9.

- European Commission (2007). Growing Regions, Growing Europe. Forth Report on Economic and Social Cohesion. Luxembourg: Office for Official Publications of the European Communities, 2007. 222 p. ISBN 92-79-05704-5222.

- Eurostat (2012). Regions and cities. Regional Statistics [online]. 2012 [cit. 28.2.2012]. Available from WWW: http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/regional_statistics/data/database.

- Kutscherauer, A. a kol. (2010). Regionální disparity. Disparity v regionálním rozvoji země, jejich pojetí, identifikace a hodnocení. Series on Advanced Economic Issues, Vol. 3. Ostrava: VŠB-TU Ostrava, 2010. 235 p. ISBN 978- 80-248-2335-5.

- Meloun, M. (2005). Počítačová analýza vícerozměrných dat v příkladech. 1. vyd. Praha: Academia, 2005. 449 s. ISBN 80-200-1335-0.

- Mooi, E., Sarstedt, M. (2011). A Concise Guide to Market Research. The Process, Data, and Methods Using IBM SPSS Statistics. 1st Edition. Heidelberg: Springer, 2011. 307 p. ISBN 978-3-642-12540-9.

- Molle, W. (2007). European Cohesion Policy: Regions and Cities. London: Routledge, 2007. 347 p. ISBN -415-43811-X.

- Řezanková, H. (2009). Shluková analýza dat. 2. vyd. Praha: Professional Publishing, 2009. 218 s. ISBN 978-80-86946-81-8.

- Staníčková, M., Melecký L., Poledníková. E. (2011). Data Base Analysis for Exploration of EU Cohesion and Competitiveness. ACC Journal, vol. 3, iss. C, pp. 136-146.

- Tuleja, P. (2008). Možnosti měření regionálních disparit – nový pohled [online]. Šilheřovice, 2008 [cit. 1.4.2012]. Available from WWW: http://disparity.vsb.cz/regdis_2008/pdf/Sbornik%20regdis_2008.pdf.

Time for question and discussion……

Thank you for your attention.

eva.polednikova@vsb.cz
petra.lelkova.st@vsb.cz

VŠB – Technical University of Ostrava
Faculty of Economics
Department of European Integration