Classifying Hungarian Sub-regions by their Competitiveness

Péter Kovács¹ – Miklós Lukovics²

During the beginning of the 1990s after the transformation of the political system in Hungary the regional differences sharpened remarkably. This process hasn’t stopped by the beginning of the 21st century. Parallel to the catching-up process of Hungary at the national level to the EU, there is another discernible process at the regional level: regional disparities are widening because the growth of the most developed spatial units is increasing while the less favoured ones are lagging behind. One of the possible tools to prove the above mentioned statements can be the notion of regional competitiveness. After 2004, the year of the enlargement of the European Union to 25 member states, the concept of economic and social cohesion furthermore the notion of regional competitiveness have been becoming key factors. Regarding the fact that development or underdevelopment does not spread evenly in space, we chose to base our empirical research on the smallest spatial unit for which statistical data are still available. Furthermore sub-regions are worth analyzing because nowadays local economic development is attached to local areas and commuting zones that are almost equivalent to sub-regions.

Present paper is aiming to measure and demonstrate inequality among the Hungarian sub-regions (local level) with the help of multi-variable data analyzing methods based on a determined system of viewpoints, and correctly chosen theoretical model (the pyramidal model of regional competitiveness) and statistical data. When weighting the indicators, we used a weighting system that was used for this reason first ever, following the logic of the correctly chosen theoretical models. In the course of our work, using cluster analysis, principal component analysis etc. the 168 Hungarian sub-regions will be classified into development phases. Parallely the rate of the useful data in our model will be measured by the PETRES’s RED.

1. Regional disparities and regional competitiveness

One of the most important effect of the global economy in connection with spatial trends is that regional disparities are increasingly growing. The larger aggregation level is proved, the larger is the possibility of significant regional disparities in the surveyed territorial unit. Parallel to the catching-up process of Hungary at the national level, there is another discernible process at the regional level: regional disparities are widening because the growth of the most developed regions and counties is increasing while the less favoured regions and counties are lagging behind (Lengyel 2004, Lukovics 2004). During the beginning of the 1990s after the transformation of the political system in Hungary the regional differences sharpened remarkably. This process hasn’t stopped by the beginning of the 21st century. Since the process of economic development comes along with the growth of regional disparities for a while.

One of the possible tools to prove the above mentioned statements can be the notion of regional competitiveness. In the global economy, competitiveness can be characterized as a central issue of regional economic development policies (Camagni 2002). The main target of regional economic development is to increase standard of living in the local area (Malizia-Feser 1999), its main tool is to strengthen competitiveness (Lengyel 2003). After 2004, the year of the enlargement of the European Union to 25 member states, the concept of economic

¹ Péter Kovács lecturer, University of Szeged, Faculty of Economics and Business Administration, Department of Statistics and Demography. e-mail: pepe@eco.u-szeged.hu
² Miklós Lukovics lecturer, University of Szeged, Faculty of Economics and Business Administration, Institute of Economics and Economic Development (Szeged). e-mail: miki@eco.u-szeged.hu
and social cohesion furthermore the notion of regional competitiveness have been becoming key factors. “Competitiveness and cohesion reinforce each other.” (EC 2004, pp. 4.).

There are several, well known definitions of regional competitiveness, which interpret the approach of competitiveness on territorial units variously. That’s why the correctly chosen definition determines the whole methodology of our research. Perhaps, the approach of regional competitiveness, published in the Sixth Periodic Report of the EU is based on the widest consensus: “The ability of companies, industries, regions, nations and supra-national regions to generate, while being exposed to international competition, relatively high income and employment levels” (EC 1999, pp. 75.). In our research we depend on this standard definition of competitiveness, which is increasingly used in the regional policy of the European Union (Lengyel-Rechnitzer 2000, EC 2004).

To carry out an analysis of competitiveness, there are more and more clear-out models, which can serve as the basis of an empirical research. The above mentioned standard definition and the resulting economic indicators enable us to measure competitiveness fairly precisely. The pyramidal model of regional competitiveness seeks to provide a systematic account to describe the basic aspects of improved competitiveness (Lengyel 2004). The development (programming) factors and success determinants placed in the model reinforce significantly regional disparities (EC 1999). Because of the logical framework (figure 1), and transparency of the pyramidal model based on wide professional consensus, it is serving as the basis of our empirical research. The model is internationally highly respected, it is more and more used as a theoretical basis of several competitiveness reports, spatial documents, decision preparation papers etc (Gadiner – Martin – Tyler 2004, Garlick 2003, GHK 2005). The pyramidal model, with its original logic and figure has been utilized in a governmental document of the United Kingdom (Pike et al 2006), however, the basic model – published in 2000 – has been rethought and developed by several authors during their research (LDY 2006, Parkinson et al 2006).

Figure 1. The pyramidal model of regional competitiveness

Based on the theoretical background and inheritance of the original model (figure 1), general and complex empirical researches have been carried out to identify and to compare the competitiveness of Hungarian counties (NUTS-3 level) (Lukovics 2006). In this paper, first of all we will examine the relevancy of the firstly selected indicators, representing the basic categories, development factors and success determinants of the pyramidal model. Than we will demonstrate an attempt to classify the 168 Hungarian sub-regions (NUTS-4 level) on their competitiveness, using our own weighting system. Finally, the rate of the useful data in our model will be measured. Before demonstrating our empirical research, some typologies of regions will be presented. The results of the following typologies will be highly considered in our research.

2. Possible typology of regions

In connection with the notion of competitiveness there are several well known typologies of regions available. In this paper we will emphasize three of them, which are the most important from the aspect of our research:

I. the stages of economic development by Porter,
II. the region types arising from the spatial aspects of the fordist and the post-fordist cycle,
III. the typology of the University of Cambridge about the European regions.

As early as 1990, Porter claimed that instead of the theory of comparative advantages, the analysis of countries’ international specialization must rely on that of competitive advantages (Porter 1990). Comparative advantages cannot explain the improvement of the economies of scale, companies’ networks and strategic alliances, the flow of production factors among countries, the growing significance of technology transfer, and so on. Based on competitive advantages, he classified countries in three groups (Porter 1990, 2003b):

1. Factor-driven economy (low-income countries/regions): globally competing companies compete with cost advantages (available labour force with low wages) based on cheap input costs (natural resources, agriculture). The quality of technology is low and these economies purchase it from other countries (it derives from imports) instead of developing it.

2. Investment-driven economy (medium-income countries/regions): modern technology is present also through foreign active capital. Advantages deriving from the economies of scale increase and productivity improve radically, however, the competitive advantages of global companies mainly come from improving the effectiveness of mass production.

3. Innovation-driven economy (high-income countries/regions): companies not only purchase technologies but also develop new ones that mainly tied to producing innovative products and services. The success of companies undoubtedly depends on innovation and companies can only produce high wages with the help of innovative activities.

Until the 1970’s territorial processes could be described by fordist features, while subsequently the regional processes of developed capitalist countries entered the postfordist stage constituting a new development phase. The extension of the fordist-postfordist cycles result in such region types that are based on today’s predominant global tendencies, the upvaluation of invisible property elements, the growing importance of knowledge, and so
forth. Based on their role in knowledge-based economies and competitive advantages three region types can be distinguished (Lengyel 2003, 2005):

1. In *neofordist regions* (relatively underdeveloped region type): region with low income, the traded companies operating in the region compete with cost advantages (e.g. cheap workforce, tax discounts). The region’s companies utilize purchased technology, usually with significant delays.

2. In *knowledge transfer regions* (medium developed region type): region with medium income, technology transfer assumes a significant role, the region’s traded companies buy developed technologies, but do not yet elaborate innovations. They usually create industrial parks.

3. In *knowledge creation regions* (relatively developed region type): region with high income, the source of companies’ competitive advantages lies in the creation of innovation results, they predominantly apply technologies developed by them. Companies have successful collaborations with universities and finance complex research programs. In many places high tech companies are concentrated in scientific parks.

The University of Cambridge analyzed the factors underlying differences in regional competitiveness, which will be of direct use in ensuring the appropriate formulation of the EU cohesion policy 2007 to 2013. Their research separated three different types of regions along two dimensions: GDP growth/capita and population density (Martin 2003, 6-23. o.):

1. *Space regions*: according this typology these are the least developed regions with low GDP per capita, population density and attractiveness of FDI.

2. *Regions as productions sites*: regions with lower to medium income levels, which derive their productivity above all from cheap inputs. Determinants of competitiveness often lie in the field of basic infrastructure and accessibility, such as low-cost sites and availability of human resources at reasonable costs.

3. a) *Regions as sources of increasing returns*: high growth regions with an average population density. A selected number of industries is an important source of wealth, this provides high and sustainable incomes for these regions.

   b) *Regions as hubs of knowledge*: regions with a higher population density and high and sustained GDP growth. These regions are open to international activities, they offer the best career opportunities that attract talented workers, they bring about the best matches between labour demand and supply, and are characterized by high levels of R&D, entrepreneurship, new firm formation and patent activity.

As shown above, the most important typologies of regions segregate three types of regions (figure 2). This fact encourages us to classify the Hungarian sub-regions into three relatively homogenous clusters.
3. **An attempt to measure competitiveness**

Regional competitiveness is a very complex notion, which can’t be described with one or two indicators alone. Measuring of competitiveness can be achieved by using indicator-systems. The key issue is to select the relevant and adequate indicators. In our survey, selection of the indicators is based on the logical framework of the pyramidal model, which seeks to provide a systematic account of the measuring and to describe the basic aspects of improved competitiveness (Lengyel 2004). The measurement of the regional competitiveness in the European Union is derived from GDP/capita, which can be factored into well known economic categories:

\[ \text{income generated in the region} = \text{labour productivity} \times \text{employment rate} \]

Measuring regional competitiveness has been traced back to three related economic categories: income generated in the region, labour productivity, and employment rate. The three coefficients of equation 1 correspond with the basic categories at the top of the pyramidal model, which fact underpins their accentuated role in describing regional competitiveness. According to the build-up of the chosen logical model, in our research we would like to characterize the basic categories (export), the development factors (improve competitiveness in short term directly) and success determinants (have indirect, long term impact on basic categories and development factors) in the first round with at least of three indicators. The model contains the three named indicators of the standard competitiveness definition (GDP per capita, labour productivity, employment rate) and other chosen indicators representing the basic categories, development factors and success determinants of the pyramid model. The first selection of the 64 indicators has been driven by economic theories and the principles of competitiveness. In the following chapters we would like to design a complex competitiveness picture about the Hungarian sub-regions with multivariate data analysis about the correctly chosen indicators.

4. **The second selection of the variables**

We attempted to classify the 168 Hungarian sub-regions to three homogenous groups similarly to the important typologies of the regions, sub-regions. The classifications is based on the competitiveness of the sub-regions. First of all, we study the right share of the 64 variables in the model.
So, we examine the amount of the information of the variables in each basic categories, development factors and success determinants. Using the Principal Component analysis for each basic category, development factors and success determinants, we left those variables, which had a bad goodness of fit in the representing Principal component(s). Naturally we used standardized variables because of the variety of the unit of the measure. We used principal component analysis because on the one hand the first selection of the 64 indicators has been driven by economic theories and the principles of competitiveness and the basic categories, development factors and success determinants of the pyramid. On the other hand a perspective aim of the examination is clustering the sub-regions directly by the basic categories, development factors and success determinants of the pyramid. So, we substitute each basic category, development factor and success determinant with one (ore more) principal component. The coordinates of these principal components that are the factor scores mark the sub-regions by the categories.

The main aspect of the 2nd selection of the variables was marking each basic categories, development factors and success determinants with one principal component, which has at least 75-80% amount of information. In each basic category, development factor and success determinant the numbers of the principal components were determined by the eigenvalues of the correlation matrix of the marking variables, which are greater then 1. If the result of Principal component analysis was one principal component we would attempt growing the amount of information of that by leaving those variables, which has low communality. Namely the low communality means that the principal component less interpret the variance of the variable. So the principal component less keeps the amount of information of the variable.

Naturally, there are such development factors and success determinants, for example the infrastructure and the human capital, which can’t be marked with one ‘good’ principal component. We analyzed the connection between the variables and the principal components by the loading variables. If the researcher couldn’t determine the means of the principal components, there isn’t right the application of the principal components method. If we could that, we determined the means of the components by separating the variables, thus each development factors and success determinants. If we couldn’t that, we attempted with the selection of the variables. Thus each development factors and success determinants were marked with right numbers of principal components according to theirs amount of information.

54 variables were entered for the model by the results of the Principal Component Analysis:

**Income level**
- Gross Value Added (GVA) per capita
- Volume of taxable incomes per one tax-payer
- Gross income serving as the basis of the personal income tax, per permanent population
- Earnings from main activity/number of tax-payers

**Labour productivity**
- Gross Value Added per employer
- Gross income serving as the basis of the personal income tax, per tax-payers
- Profit before taxes per employer

**Employment**
- Employment rate

---

3 The loading variables quantify correlation between the variables and the principal components.
- Unemployment rate
- Number of tax-payers per 1000 inhabitants

**Global integration**
- Income from export per inhabitants
- Integration of the trade ($\frac{\text{Export}}{\text{GDP}}$)

**Research and technological development**
- Number of members of public body of Hungarian Academy of Sciences
- Total number of patents between 2000 and 2004.

**Small and medium-sized enterprises (SME-s)**
- Number of active companies and partnerships per 1000 inhabitants
- Number of active small businesses (10-49 employers) per 1000 inhabitants
- Number of active corporations with legal entity per 1000 inhabitants
- Number of active small businesses (10-49 employers) with legal entity per 1000 inhabitants
- Proportion of partnerships from the active enterprises
- The sub-regional proportion of active companies with at least 250 employees from the Hungarian data.

**Foreign Direct Investment**
- FDI from the owner’s equity per capita

**Infrastructure and human capital**
- Number of university or college graduate unemployed per 1000 inhabitants
- Students in part time education in secondary schools per 1000 inhabitants
- Number of vocational and specialized vocational schools per 1000 inhabitants
- Number of computers in the educational institutions per 1000 inhabitants
- Number of task units with internet in the educational institutions
- ISDN-lines per 1000 inhabitants
- Bed-places in hotel per 1000 inhabitants
- Telephone main lines per 1000 inhabitants
- Number of dwellings connected to the cable television network per 1000 inhabitants

**Institutions and social capital**
- Active non-profit institutions per 1000 inhabitants
- Full-time students of higher educational institutions per 1000 inhabitants

**Economic structure**
- Proportion of active agricultural companies from all active companies
- Proportion of active industrial companies from all active companies
- Proportion of active economic service companies from all active companies
- Proportion of employees in agriculture from all employees
- Proportion of employees in industry from all employees
- Proportion of employees in services from all employees

**Skills of work force**
- Proportion of university or college graduate employees from all employees
- 25-x years old with diploma as a percentage of the same age group

**Regional accessibility**
- Complex regional accessibility indicator

**Social structure**
- Population aged 60 and over as percentage of permanent population
- Population aged 0-18 as percentage of permanent population
• Live births/deaths
• Share of inhabitants living in settlements with population density over 120
• Proportion of central settlement’s inhabitants from the sub-region’s inhabitants

**Decision centres**
• Number of active corporations with legal entity with at least 250 employees
• The sum of the company’s owner’s equity in the sub-region

**Environment**
• Number of discovered publicly indicted crimes
• Number of economy related discovered publicly indicted crimes
• Number of general practitioners per 1000 inhabitants
• Number of places of clubs for the aged providing day-time care per 1000 inhabitants aged 60 and over

**Regional identity**
• Arrivals per 1000 inhabitants
• Departures per 1000 inhabitants

5. Cluster analysis, weighting of variables

To fulfill classification, a multi-variable data analysis method, the cluster analysis will be first performed. This is an exploratory data analysis tool which aims to sort different objects into groups in a way that the degree of association between two objects is maximal if they belong to the same group and minimal otherwise. The theoretical background (the notion of competitiveness and the pyramid model) of our empirical research requires weighting the indicators according to their relevancy on the competitiveness. In case of the method, the Euclidean distance between pairs of objects was taken as basis in measuring the similarity of the objects. Version 13.0 of SPSS was used to carry out the analysis.

As mentioned above, the database defined by the variables mainly consists of variables with different units; the potential problems arising from this has been solved with the help of *standardisation*: the expected value of variables was 0 with their expected variance being 1. Identical variance practically means that all the variables have equal weight in the model. However, the logic of the pyramid model implicitly requires that the variables affecting the region’s competitiveness in different ways and with different relevance should be included in the model with different weight.

Accordingly we determined the weights of the 54 selected variables. The base of this process was a weighting method, which was published by Porter in the Global Competitiveness Report, one of the most highlighted publications on competitiveness. Porter (2003) constructed two sub-indexes. The weights were determined from the coefficients of a multiple regression of the sub-indexes on GDP per capita. The pyramidal model marks the competitiveness by an indicator system, thus we used a complex model. We also defined principal components, as indicators and we attempted for defining the objective weights of those in the given model. Our weight system and examination could be an advance in the effort of making commensurable the competitiveness. Opposite of Porters’ GDP/capita the pyramid model hasn’t metric dependent variable, thus we didn’t analyze the causality. We analyzed the state of the sub-regions.

---

4 Index of the national business environment, and the index of the company operations & strategy was defined by Porter, which based on 16 and 31 variables.
According to the variables selection method we used the principal component analysis to make an objective weight system in the given model. The determination of the weights is base on the following train of thought. If we substitute the standardized variables with principal components, the principal components represent the model in reduced dimension. One result of the principal component analysis the values of the communalities. As the communalities practically are coefficients of multiple determinations in a linear regression model, where the dependent variable is the given variable, and the independents are the principal components, the square roots of those are coefficients of multiple correlations. In general the coefficient of multiple correlation quantify the correlation between the effective (empirical) and the estimated values of the dependent variable. Thus it also quantifies the correlation between the dependent variable and the set of independent variables. Especially the coefficient of the multiple correlation means the correlation between the given standardized variable and the set of principal components, which represent the pyramid model. Thus, the coefficients mean the correlation between the variables and the model, namely the weight of the variables. First time we applied this method for the selected variables, and the second for the categories.

A subsequent research could be the examination the causality, and the delays of effects of the categories for the competitiveness and the change of it on the various levels in the pyramid, and the weight of the levels in the changes, which are measured at the various distances of time.

For the first time we determined the weights of the selected 54 standardized variables in the pyramidal model. The variables we could substitute with 11 principal components. Thus the model and namely the competitiveness were marked by 11 components. The weight of a variable is the square root of the communalities, which means the correlation between the principal components and the variable, thus the weight of the variable in the pyramid model.

For the second time we determined the weights of the categories. The applied method was the same.

Since the number of clusters should be generated has been former defined and proved (three theoretical types of regions), K-means clustering method has to be used. Going step by step, this method calculates cluster core-points and related objects until cluster centres do not change in one step any more. In the following, the 168 Hungarian sub-regions will be classified in two different ways: it will be based firstly on the 54 selected, weighted, standardized variables, secondly on the weighted principal components. Using the cluster analysis, the realization of the weighting means that we multiply the coordinates of each variable with the weight of it before.

According to the classification emerging from the first method\(^5\), Budapest, by itself, consists the most developed cluster, where values based on each variable are especially high. On the contrary, in the cluster of the 120 relatively underdeveloped sub-regions the majority of indicators show unfavourable values. In the case of the 47 sub-regions classified in the medium developed cluster the values of most indicators fall between the former two extreme values. The homogeneity of each cluster is defined by the distance of cluster members from the final cluster centre. The more classified objects are grouped, concentrated and „clustered” around the cluster centre, the more homogeneous the created cluster may be regarded.

\(^5\) The results of the two cluster analysis were close to each other. The situations of only 3 sub-regions from 168 were changed. It also means that, the principal components maintain at least 75-80% of the amount of information of the variables. We note that, we clustered the sub-regions by the 64 original variables. The result of it was close to the results of our methods.
Naturally, consisting of only one object, the relatively developed cluster of Budapest is the most homogeneous, while objects falling in the other two clusters have greater diversity.

As for the spatial concentration of relative development (competitiveness core) it can be stated that the only relatively developed sub-region (the capital) is circularly embraced by medium developed sub-regions. Furthermore, medium developed sub-regions are also the sub-regions of county towns themselves or the sub-regions of large towns. Medium developed sub-regions are concentrated in the vicinity of developed Western centres and motorways. Furthermore, it can be said that medium developed sub-regions are predominantly concentrated in the North-Western and Central part of the country, while relatively underdeveloped sub-regions are situated in the Northern and Eastern border zone (figure 3).

Figure 3: The situation of the theoretical sub-regions

Source: Our construction

We calculated the value of the Red indicator in our study. This indicator is based on the eigenvalues of the correlation matrix. This indicator quantifies the percentage of collinearity and the proportion of data with a useful content compared to the database of the given size and with minimum redundancy (Kovács-Petres-Tóth 2005). The Red quantifies the average correlation of the data of the database, and which can be regarded as the synthetic and normalized indicator.

In the case of the absence of redundancy the value of the above indicator is zero or zero percent, while in the case of maximum redundancy it is one or one hundred percent. The Red indicator measures the redundancy of the studied database of the given size. Kovacs proved the value of the Red indicator is the quadratic mean of the elements outside the main diagonal of the correlation matrix. The value of the Red indicator approximately 0.42 in all ways (clustering by the factors, clustering by 54 variables, and clustering by 64 variables). So our methods don’t draw down significant information loss.

One possibility generalization of the Red indicator is applying it for loading variables. In this case the value of the square of Red indicator is the same as the value of the redundancy
6. Summary

In Hungary, there are significant disparities in regional competitiveness among the Hungarian sub-regions. These disparities stem from structural deficiencies in key factors of competitiveness, like inadequate endowment of physical and human capital (of infrastructure and workforce skills), a lack of innovative capacity, of effective business support etc.

Present paper performed the classification of Hungarian sub-regions based on the pyramidal model of competitiveness. The use of various methods with different logics lead to similar results, therefore, it is likely that we managed to map the competitiveness of sub-regions realistically. Based on this, we believe that the pyramidal model and the methodology based on it are suitable to make regional competitiveness measurable and outline the possibilities of economic development. The results introduced in the present paper constitute the one of the first steps of our research, in the following we would like to test the statistical methodology, with special emphasis on checking the weight of indicators. Furthermore, we would like to define types of competitiveness that may also serve as the basis of realistic economic development strategies.

According to the logical framework of regional development strategies, in Hungary standard of living and welfare would be damaged without developing the core factors of regional competitiveness. Sources have to be concentrated in investments in order to enlarge human and physical capital to influence the growth of competitiveness.

Consequently, economic development should not be executed homogeneously, one should take into consideration the attributes and starting conditions of that certain territorial unit. The variety of starting conditions requires different interventions and strategies of economic development from sub-region to sub-region. Spatial units with different level of competitiveness should take variant steps on the road of economic development in order to achieve competitiveness in the global world.

References