

ZSÓFIA FÁBIÁN

The role of neighbourhood in the regional distribution of Europe

Introduction

The role of geographical space in social and economic research has gained more importance, resulting in the increasing significance of mathematical-statistical methods. In order to determine the cause and effect relationship more thoroughly, factors representing territories have also been taken into account. These numerical factors are called geographical parameters, and with their assistance the importance of proximity in the spatial division of society and economy is revealed.

It is important to mention that thematic maps are used to visually present the distribution, implying a connection between the geographical space and certain phenomenon. In comparison spatial inequality indicators are used to measure the level of and change in distribution, but neither of them provides information on the role of spatiality. In order to gain information on these factors we need more complex methods and geographical parameters.

The most common spatial variables are distance, neighbourhood, borders, and geographical latitude and longitude. From the above mentioned factors I am going to be concerned primarily with neighbourhood. The analysis of neighbouring relationships is one of the major issues today, as more and more cases have shown that the influence of neighbours should also be taken into consideration. Waldo Tobler's statement made in 1970 "everything is related to everything else, but closer things are more closely related" is considered to be the first law of geography (Tobler 1970).

Previous studies on the same subject also testify to the importance of the influence of neighbours on the differences in development in Europe (Ertur & Le Gallo 2000, Tóth 2003, Szabó 2006).

In my study I examine the extent to which neighbouring regions in Europe are similar to each other regarding their economic development; how the high and low values are distributed in a given region, and whether they are clustered or random. Both European and regional processes are examined, which allows the findings of the whole process to be interpreted more clearly.

Spatial autocorrelation

In order to define the role of neighbourhoods, different mathematical-statistical methods are applied (for example autocorrelation, autoregression), spatial autocorrelation being the most prevalent among these.

Spatial autocorrelation differs from basic autocorrelation, which seeks to find connections between two different variables regarding the “connection” itself, whereas the influence of the variable on the region including its surrounding areas is measured by spatial autocorrelation. Thus the possibility of finding a connection between the data of neighbours and the data of a region is investigated.

There are different methods for measuring spatial autocorrelation (Pearson’s correlation, Moran’s I, Geary’s C). The most frequently used method is Moran’s I. Moran’s I (Moran 1950) tests for global spatial autocorrelation of continuous data. The main reason for its frequent use is that it makes calculation easier and provides information about individual regions with the assistance of a local variation of Moran’s I, Local Moran I (Eff 2004).

The process of calculation is the following:

$$I = \frac{\frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij}}}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

where x_i and x_j are the data of regions, \bar{x} is the mean of data, $(x_i - \bar{x})$ is the value of the region subtracted from the average of the data, W_{ij} is the element of the weight matrix, and N is the number of regions (Nemes Nagy 2005).

In this method the effect of proximity is represented with a weight matrix. A weight matrix is a matrix which contains the values 0 and 1 in the basic case. The value is 0 if there is no neighbourhood between two regions, and 1 if there is neighbourhood. The relationships among regions can be described in many different dimensions. The basic case is that adjacent regions are considered as a neighbourhood, while regions which are only connected with their vertex are disregarded. The neighbourhood could also be defined by considering its distance from something (Nemes Nagy 2009).

Depending on which factors are to be focused on, there is a possibility of weighting based on the different features of the neighbours. Weighting the distance is one of the approaches most widely used in order to emphasise the greater importance of nearer neighbours.

In a weight matrix there is a possibility to take neighbours with higher weight into consideration. This means weighting with distance. This distance is primarily defined between the centres of regions based on the assumption that if the centres of regions are located nearer each other, this induces closer interaction. Furthermore, there is also a possibility to weight the neighbourhood according to the length of the borders.

The value of Moran’s I differs from the value of correlation (–1, +1) because it does not reach, only approaches these terminal values based on the number of the observation units. It could equal 0 if the number of regions was infinite.

$I > -1/(n-1)$, positive spatial autocorrelation,

$I = -1/(n-1)$, without spatial autocorrelation,

$I < -1/(n-1)$, negative spatial autocorrelation (Nemes Nagy 2005).

Positive spatial autocorrelation means that regions with similar values are more spatially clustered than could be caused by chance. Negative spatial autocorrelation is

when the neighbours have no similar values. Perfect negative spatial correlation is characterised by a chequerboard pattern of high and low values. If Moran's I approaches the value 0, this represents a lack of autocorrelation.

In this scenario, the distribution of data is random, which means that there is an equal likelihood of finding a territory with a lower or higher value next to a certain region.

Local Moran I enables accurate information to be obtained from the similarity or difference between regions and their neighbours.

The process of calculation is the following:

$$I_i = Z_i \sum_{j=1}^n W_{ij} Z_j,$$

where Z_i and Z_j are the standardised values of the observed areas (Nemes Nagy 2005).

By analysing Local Moran I it may be a problem that it does not have any definite values, and therefore the result cannot be compared to the terminal values. However, it can show where the neighbouring contacts are stronger and the similarity is more apparent as the higher the value is, the stronger the connection is. It should also be taken into consideration that Local Moran I values only provide information on the connection of adjacent geographical areas but not on real values. Therefore, it cannot be determined whether the similarity of neighbourhood is caused by low or high original values (Mészáros 2008).

Interpretation of the results in their original forms is not simple since it is complicated to reveal the coherence of the results. However, presenting them on a map contributes to an easier understanding as the map clearly shows both the homogeneous and the differing regions, while also allowing analysis based on their real location.

Furthermore, depicting the basic values on a map enables hot and cold spots to be distinguished. Hot spots mean geographical areas where strong neighbouring connections derive from high real values, while cold spots are areas where the strong neighbouring correlation is caused by low basic values (Mészáros 2008).

In addition, there is also a possibility to examine whether certain regions and their surroundings exceed average. To analyse this, a Moran scatter plot is used in order to compensate for the deficiencies of Local Moran I. The Moran scatter plot also allows the examination of whether the values clustering in space are high or low. Each point in the scatter plot represents a geographical area. The horizontal axis (x) shows the values of the observed areas, while on the vertical axis (y) there are the average values of the neighbouring areas observed; in other words the spatially lagged equivalent of the x variable.

Another significant requirement is that the scatter plot figure can only be created with a row-standardised matrix. The vertical axis should present the mean of the neighbours irrespective of the method of weighting geographical areas (Ertur & Le Gallo 2000).

According to the four quarters of the Moran scatter plot the following four groups can be determined, which describe the connection between the geographical areas and their surroundings.

- High–high (HH), the values of both the given geographical area and its neighbours are above average;

- High–low (HL), the value of the given geographical area is above average, while the value of its neighbours is below;
- Low–low (LL), the values of both the given geographical area and its neighbours are below average;
- Low–high (LH), the value of the given geographical area is below average, while the value of its neighbours is above.

The first and third quarters represent positive spatial autocorrelation, whereas the second and fourth quarters represent negative spatial autocorrelation.

The depiction of groups based on the Moran scatter plot is the easiest way of interpreting the phenomenon. The examination of the maps showing these groups helps determine the ways in which the geographical areas and their surroundings differ from average. On the other hand, they provide no information about the strength of similarity compared to the Local Moran I maps (Lim 2003).

Moreover, making a time-based comparison examines change, meaning whether the value of the given geographical areas and their surroundings differ from average. Accordingly, four types may be distinguished.

- Type 0: neither the given geographical area nor its neighbours change
- Type 1: only the given geographical area changes, its neighbours do not
- Type 2: the given geographical area does not change, only its neighbours do
- Type 3: both the given geographical area and its neighbours change

Therefore, the Moran's I variable is suitable to measure "global" spatial autocorrelation and Local Moran I shows the strength of spatial autocorrelation at a local level, while with the assistance of the Moran scatter plot the similarity of the given geographical areas and their neighbours can be examined and whether the similarity is caused by their low or high values.

The role of neighbourhood in the regional distribution of Europe

Previous research on economic development has shown the significance of neighbouring connections, whereas considering this at a European level the connection loosens (Ertur & Le Gallo 2000, Tóth 2003, Szabó 2006). Therefore, the question of how this has changed over recent years arises.

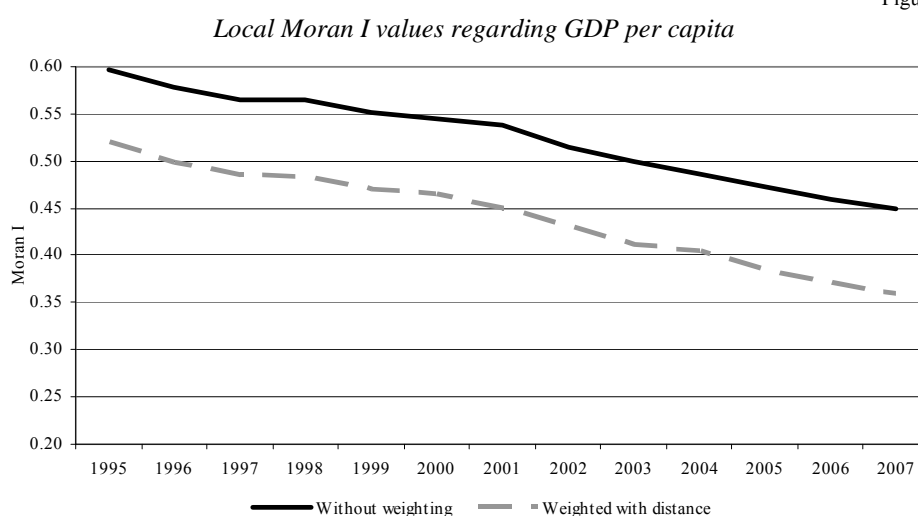
In this study, the spatial frame is Europe, and in particular the members of the European Union and certain other European countries – Norway, Switzerland and Croatia. The basic geographical areas are the NUTS 2 regions, which mean 277 regions in this case. The examined variable is GDP per capita by PPS. The period frame is 1995–2007. As regards the spatial frame, this is something new because the frame is broader than before. It includes all the member states after the last EU enlargement in 2007 as well as three other European countries which are significant for their neighbours. The change or the continuation of previously discovered processes can be examined with the help of the latest available data. The area analysed is henceforth referred to as Europe.

When establishing neighbouring connections, I considered geographical areas with any type of connection to be neighbours. To determine Moran's I values, I made calculations with a "basic" and distance-weighted matrix. For the Local Moran I

calculation, I used a row-standardised neighbourhood matrix in order to depict the Moran scatter plot.

After the calculations, I examined the results of Moran's I. The findings showed that the neighbouring connections are significant with regard to GDP per capita; neighbours have a stronger influence on each other and the values of the neighbouring geographical areas are more likely to be similar to each other than they are to geographical areas further away (Figure 1).

Figure 1



Source: author's calculations.

My second observation is that the strength of neighbouring connections constantly degraded as regards the examined variable. This tendency is due to the prominent development of the central large city regions as these regions have better facilities and opportunities (a larger highly qualified workforce, good infrastructure, favourable investment opportunities), which enable them to improve faster than their surroundings. In this way their similarity to their neighbourhood is decreasing, a tendency which can be perceived throughout Europe (EC 2007).

Moran's I values and changes concerning GDP per capita confirmed that this previously revealed process is continuing (the strength of neighbouring connections is decreasing in Europe). However, this is in harmony with the result of certain studies which assert that at a European level the decrease of spatial inequality is determinative (Novák & Papdi 2007, Szabó 2008) as the spatial differentiation can be observed in several countries (Szabó 2008).

Comparing the results of the calculations made by the weight matrix, the two curves are almost parallel, and the Moran's I values are lower when weighted by distance (Figure 1). This differs slightly from the previous hypothesis which stated that the values should be higher when weighted by distance, supposing that proximity encourages similarity. In the case of Europe it is not surprising that this does not apply because much longer distances are involved here than is the case when considering a single country.

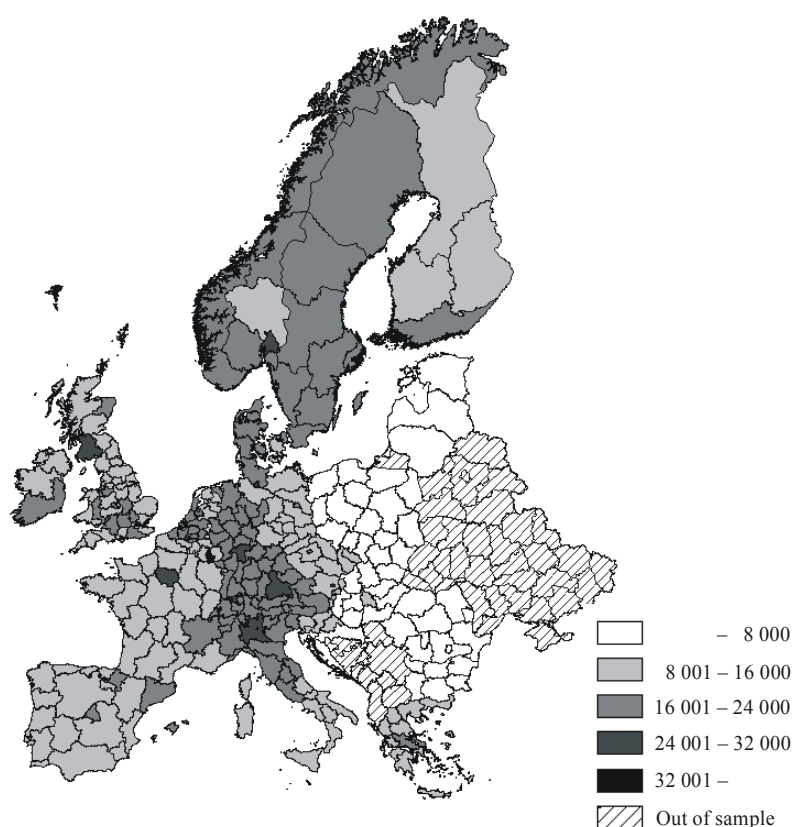
Due to the extent of the territory, its complexity and the variety of the relief, as well as representing the significance of the proximal regions on the basis of the distance between centres, it is possible for connections which are not important to be emphasised.

Turning to the analysis of Local Moran I, this can be best interpreted by depiction on a map, but in order to establish whether the similarity is caused by low or high basic values, we need to start with the depiction of GDP per capita values.

The regions located in the centre of Europe (the zone from Southern Germany to Northern Italy) had the most favourable values in 1995. In this respect, the regions of the countries which joined the EU in 2004 and 2007 were in the most unfavourable situation, except for Slovenia and the Czech Republic. In addition, it should be emphasised that the value of certain capital and urban regions exceeded that of their neighbours (Figure 2). Based on the Local Moran I values, the almost continuous area of Northern Italy, Switzerland and Southern Germany may be identified as a hot spot, and the zone extending from the Baltic countries to Bulgaria, as well as Portugal and Southern Spain may be considered as cold spots (Figure 3), meaning that between the regions situated in the centre of Europe there are strong neighbouring connections.

Figure 2

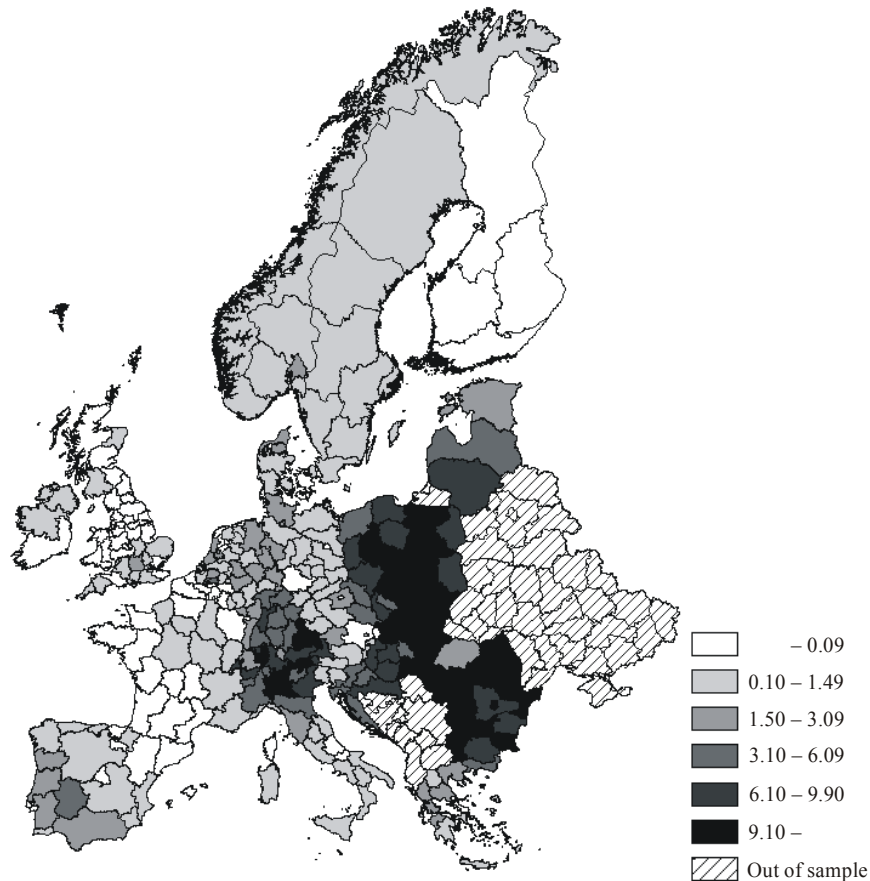
GDP per capita (PPS) in European regions in 1995



Source: author's calculations.

Figure 3

Local autocorrelation of GDP per capita in 1995 (Local Moran I values)

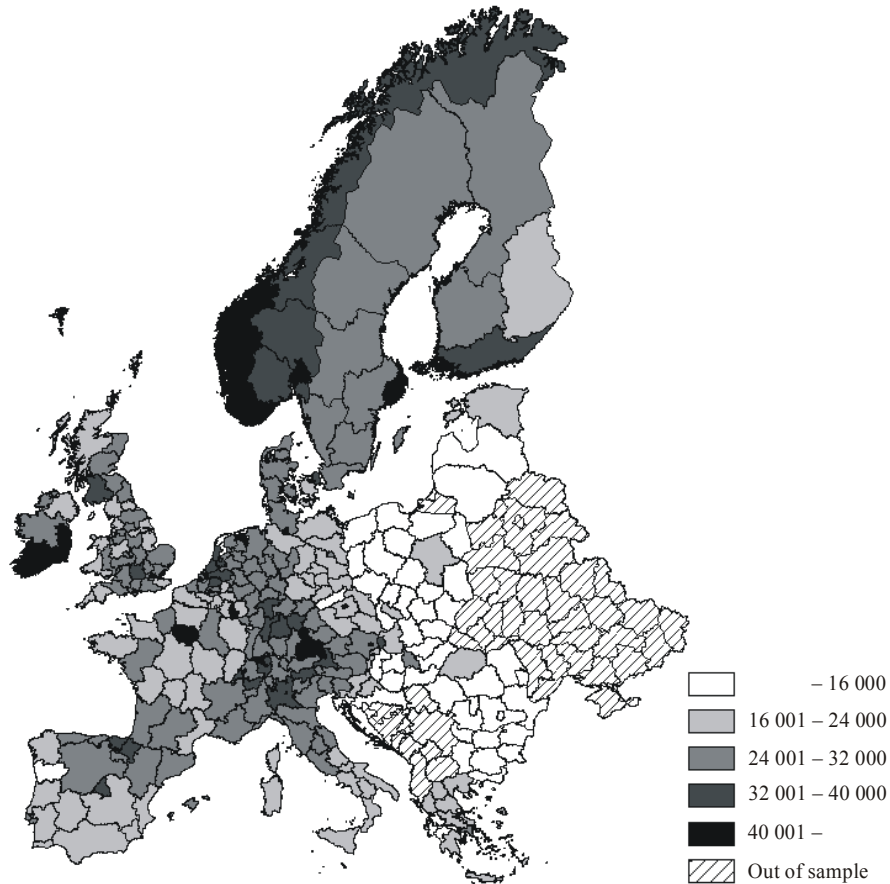


Source: author's calculations.

Similarly, the underdeveloped “East” represented strong neighbouring connections. Moreover, the zone extending from the Baltic countries to Bulgaria is the largest with the highest Local Moran I values along with the lowest GDP per capita values. The Portuguese and Southern Spanish spot is not so significant. In their case the neighbouring similarity is not so strong compared to the other two territories.

Turning to 2007, first it could be stated that polarised development took place compared to the initial year. According to GDP per capita in 2007, the Scandinavian values – among which primarily the Finnish region moved into a more favourable position – and the values of the Northern Spanish region were observed to have increased to a significant level. In a couple of regions in the newly acceded countries a remarkable development is noticeable. Divided development took place in the United Kingdom and France. In some regions of Greece there was a relapse (Figure 4).

Figure 4

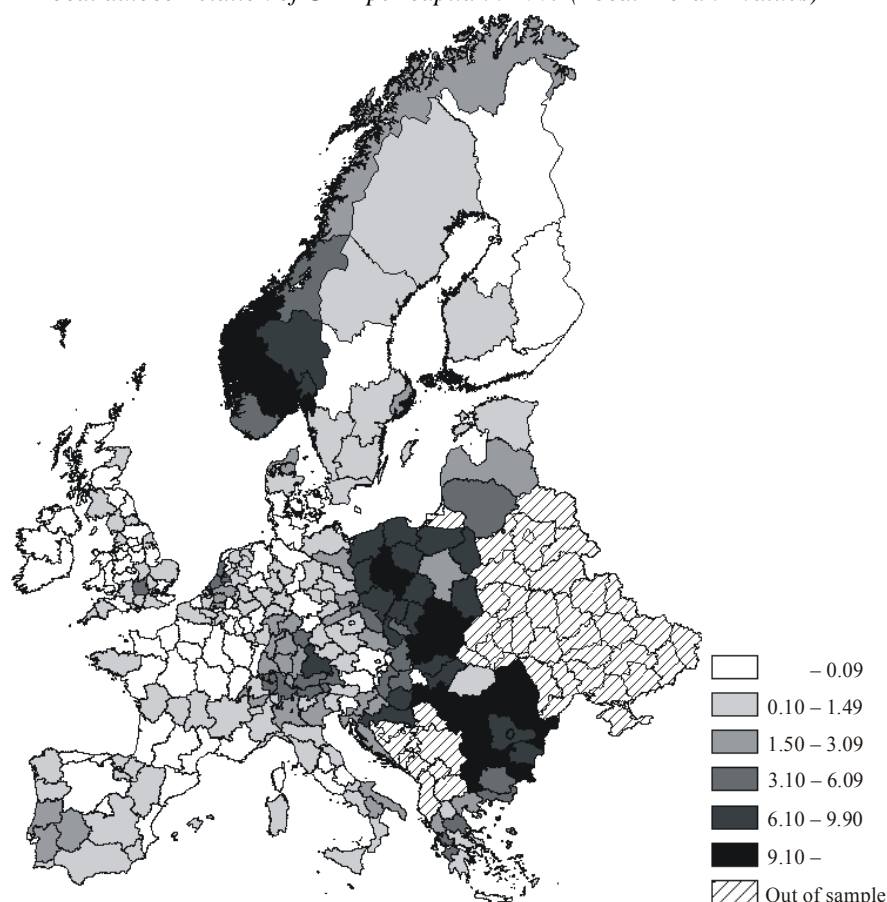
GDP per capita (PPS) in European regions in 2007

Source: author's calculations.

In terms of the Local Moran I, the picture has changed somewhat between 1995 and 2007 as Norway has appeared as a hot spot, while the German-Swiss-Italian spot has diminished. The Spanish-Portuguese cold spot faded, and similarities lost strength in the zone extending from the Baltic countries to Bulgaria (Figure 5). Due to the polarised development, the central regions of Europe proved to be less similar than in the early stages. In addition, the Southern Spanish and Portuguese zone cannot be referred to as a highly autocorrelated zone.

Figure 5

Local autocorrelation of GDP per capita in 2007 (Local Moran I values)



Source: author's calculations.

The decrease of spatial autocorrelation can be observed in the regions extending from the Baltic countries to Bulgaria. The main reason for this is the change of political system. With stable economic conditions, new ways emerged for the regions of Eastern and Central Europe to develop, which some regions were able to take advantage of.

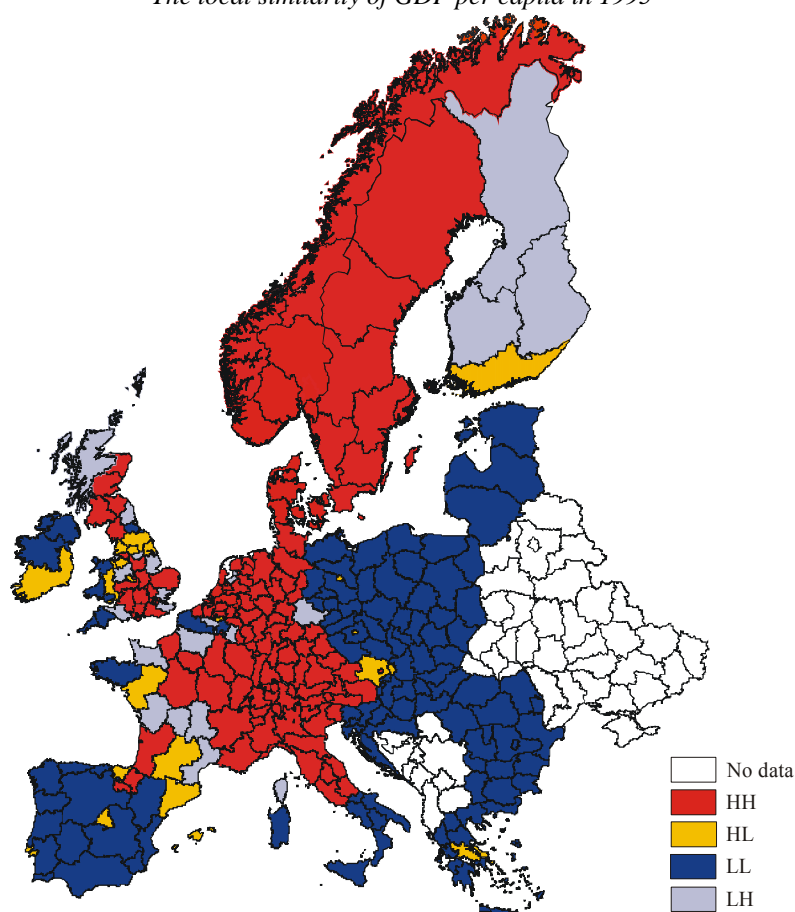
It is also worth mentioning that apart from one or two cases, the central regions had very low Local Moran I values since these territories exceed their neighbourhood in terms of GDP per capita value. Based on Local Moran I it is notable that in certain parts of Europe there were significant territorial differences in the mid 2000s.

After depicting the groups created by the Moran scatter plot, the following conclusions can be drawn. In 1995 the Southern European (Spanish, Portuguese, Southern Italian and Greek) regions were in an unfavourable situation (LL) in terms of GDP per capita as both the geographical area and its neighbours proved to be lower than average (Figure 6). Moreover, a broad continuous zone in an unfavourable position

extends through Eastern and Central Europe. Furthermore, this zone is far more extensive than is the case when examining the strength of neighbouring relations. This is because not only the regions extending from the Baltic countries to Bulgaria but also certain regions in Eastern Germany, the Czech Republic, Eastern Austria and Slovenia can be considered to have lower than average values. Areas where both the given geographical areas and their neighbours have above average values (HH) are Scandinavia (except for Finland), most of Germany, the Benelux countries, Northern Italy, Central and Eastern France, and large parts of Britain. All other areas have differing values in either direction (LH, HL).

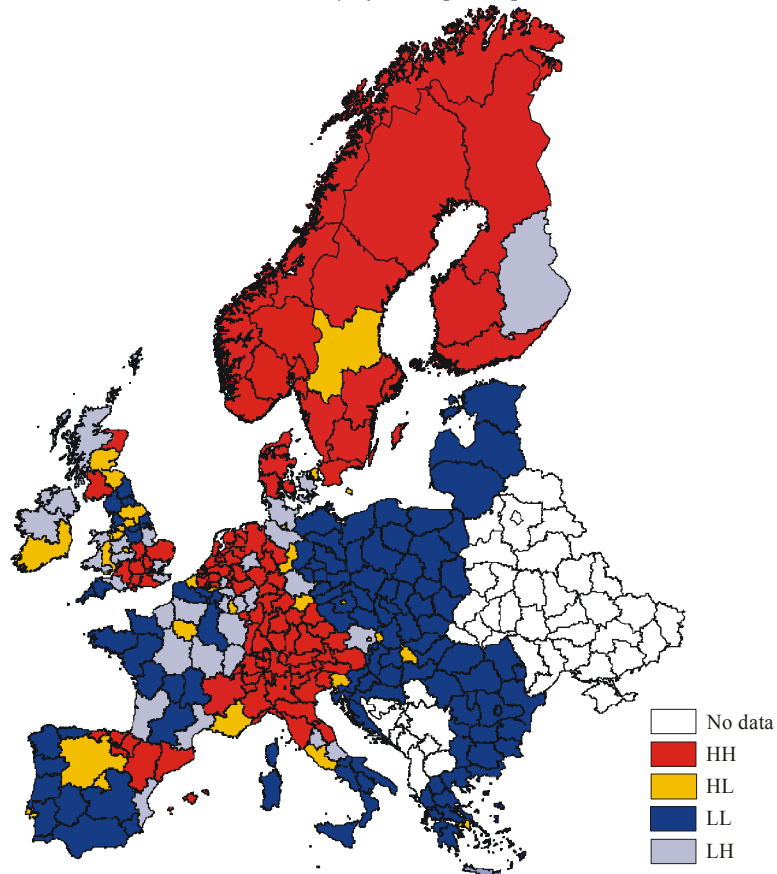
Figure 6

The local similarity of GDP per capita in 1995



Source: author's calculations.

Figure 7

The local similarity of GDP per capita in 2007

Source: author's calculations.

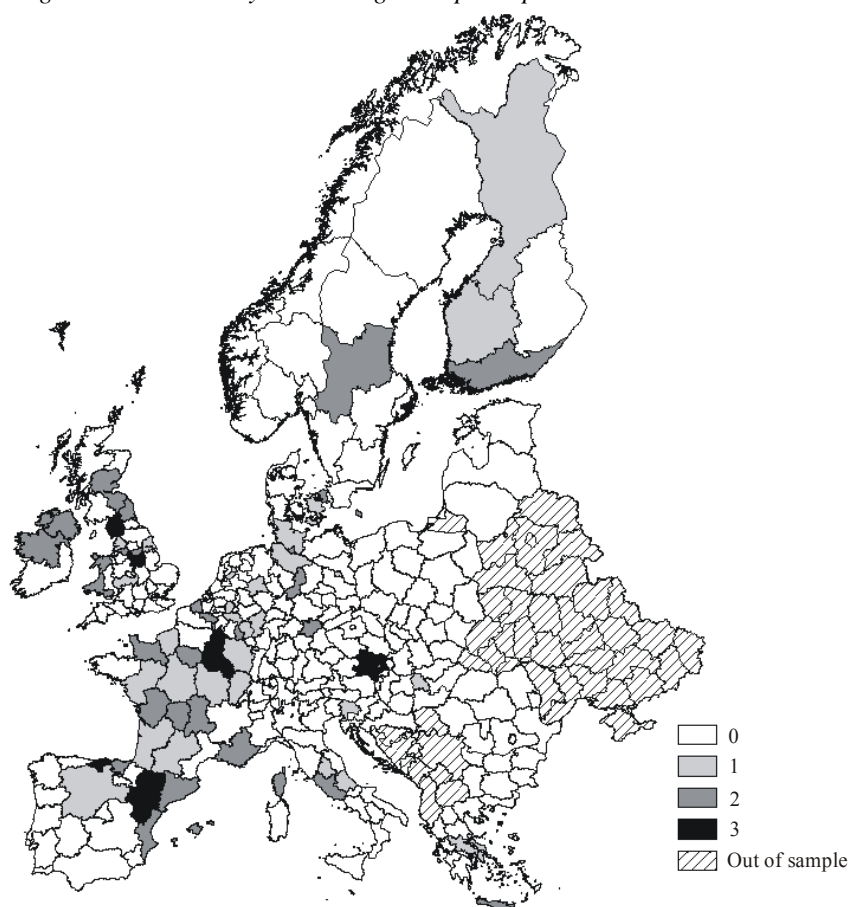
By 2007, the picture had become polarised (Figure 7) as far fewer continuous unicolour zones are represented. Only the unfavourable zone in Eastern and Central Europe remained at the same low level, while a decrease in continuous favourable and unfavourable regions is depicted everywhere else. Regarding the central regions of Europe, certain zones at a distance from the centre (Switzerland, Northern Italy) were exposed to a more unfavourable position and as a result now belong to the fourth group instead of the first. The value of the region is below average, while that of the neighbourhood exceeds average. A remarkable positive change can only be observed in the Northern Spanish and Finnish regions since they reached the level of their neighbourhood.

Comparing the results regarding time, a different classification method may be used to examine whether the values of the given geographical areas and their neighbours have changed compared to the average. Looking at the map depicting the change (Figure 8), it can be concluded that most of the regions included in the survey appear in the same

group of similarities as they belonged to in the initial year. In figures, only about one fourth of the 277 regions (67) were put into a different group. As regards spatial aspects, this change only characterises certain French, British, Irish, Northern Spanish, Finnish and German regions.

Figure 8

Change in local similarity concerning GDP per capita between 1995 and 2007



Source: author's calculations.

The result of categorisation based on the Moran scatter plot showed that the local similarity of GDP per capita decreased over the period examined.

Conclusions

Firstly, considering the process taking place in the whole continent, significant neighbouring connections can be observed according to GDP per capita based on the Moran scatter plot values. Between 1995 and 2007 spatial autocorrelation decreased

continuously. The reason for this was the remarkable development of the capital and urban regions.

Based on the results of the categorisation according to Local Moran I and the Moran scatter plot, in the period between the initial and final years of the survey it may be concluded that the extent of the great continuous areas and the interconnecting zones was reduced in the spatial structure of Europe. All in all, there was a shift towards a mosaic pattern.

BIBLIOGRAPHY

- EC 2007, *Growing regions, growing Europe. Fourth report on economic and social cohesion*, Office for Official Publications of the European Communities, Luxembourg.
- Eff, E. A. 2004, *Spatial, Cultural and Ecological Autocorrelation in U.S. Regional Data*, Department of Economics and Finance Working Paper Series, Middle Tennessee State University, Murfreesboro, USA pp. 1–54
- Ertur, C. & Le Gallo, J. 2000, *Exploratory Spatial Data Analysis of the distribution of regional per capita GDP in Europe, 1980-1995*, University of Burgundy, France. pp. 1–29
- Lim, U. 2003, 'A Spatial Analysis of Regional Income Convergence', *Planning Forum*, 9, pp. 66–80
- Mészáros, B. 2008, *Területi autokorreláció Magyarország példáján* [Spatial autocorrelation with the example of Hungary], Diplomamunka, ELTE Regionális tudományi tanszék, Budapest
- Nemes Nagy, J. (ed.) 2005, *Regionális elemzési módszerek* [Methods of regional analysis], Regionális Tudományi Tanulmányok, 11. ELTE Regionális Földrajzi Tanszék – MTA-ELTE Regionális Tudományi Kutatócsoport, Budapest.
- Nemes Nagy, J. 2009, *Terek, helyek, régiók – A regionális tudomány alapjai*, [Spaces, places, regions – The basis of regional science] Akadémiai kiadó, Budapest, pp. 266–274
- Novák, G. & Papdi, Á. 2007, 'Gazdasági egyenlőtlenségek a kibővült Európai Unióban' [Economic inequalities in the enlarged European Union], *Területi Statisztika*, 6, pp. 571–585
- Szabó, P. 2006, *A fejlettség makroregionális különbségeinek alakulása az Európai Unióban* [The changes of macroregional differences in development in the European Union], III. Magyar Földrajzi Konferencia, CD, MTA, Földrajztudományi Kutatóintézet, Budapest
- Szabó, P. 2008, 'A gazdasági fejlettség egyenlőtlensége az Európai Unió különböző területi szintjein' [The inequality of economic development at different territorial levels of the European Union], *Területi Statisztika*, 6, pp. 687–699
- Tóth, G. 2003, 'Regionális fejlettségi különbségek az egységesülő Európában' [Differences in regional development in the uniting Europe], *Területi Statisztika*, 3

Keywords: Neighbourhood, spatial autocorrelation, European spatial inequalities.