A Study of the Competitiveness of Regions based on a Cluster Analysis: The Example of East Germany

Franz Kronthaler

July 2003

Halle Institute for Economic Research
P.O. Box 11 03 61
06017 Halle/Saale
Germany

Tel: 0049-(0)345-7753-773
Fax: 0049-(0)345/7753-820
Email: fkr@iwh-halle.de

Paper prepared for the 43rd European Congress of the Regional Science Association "*Peripheries, centres, and spatial development in the new Europe*", University of Jyväskylä, Jyväskylä, Finland, 27th-30th August 2003.

Abstract

The Eastern enlargement of the EU is likely to increase regional disparities within the EU. At the same time regional cohesion is a primary objective of EU economic policies. This raises the question of whether and when the regions of the accession countries will reach the average level of economic development of the EU. East German regions are considered as an instructive example of a transition country for the integration process of the new Central and Eastern European member states.

The objective of this study is to analyse whether some East German regions have already achieved the same economic *capability* as the regions in West Germany, so that they are on a competitive basis with the West German regions and are able to reach the same economic level in the long run. If this is not the case, it is important to know more about the reasons for the economic weakness of the East German regions twelve years after unification.

The study is based on a cluster analysis. Criteria for the cluster formation are several economic indicators, which provide information about the economic capability and their determinants. The choice of the indicators is based on a review of results of the theoretical and empirical literature on the new growth theory and new economic geography.

The results show that most of the East German regions have not yet reached the economic capability and competitiveness of their West German counterparts so that they - from the viewpoint of the new growth theory and the new economic geography - are not in the position to reach the same economic level. According to these theories economic disadvantages are most notably the consequences of less technical progress, a lack of entrepreneurship and fewer business concentration. Under these points it is especially noteworthy that young well educated people leave these East German regions so that human capital might will turn into a bottle-neck in the near future. Only a few regions in East Germany - those with important agglomerations - are comparable to West German regions that are characterised by average capability and competitiveness, but not to those with above average economic capability and competitiveness. Even those more advanced East German regions suffer from economic disadvantages such as slower technical progress and low business concentration.

There are important policy implications based on these results: regional policy in East Germany was not able to assist raising *all* regions to a sufficient level of competitiveness. It may be more effective to concentrate the regional policy efforts on a selection of important agglomerations. This has also strong implications for the EU regional policy assuming that the accession countries will have similar problems in catching up to the economic level of the EU as have the East German regions.

Keywords: Regional Disparities, Competitiveness, Cluster Analysis, East Germany, EU-Enlargement

JEL classification: R12; P52; O18

1 Introduction

Twelve years after German reunification the East German regions have not yet reached the economic level of their West German counterparts. The East German region with the highest per capita income only reaches about 80% of the average German per capita gross domestic product (GDP p. c.),[1] so it was just catching up to the West German regions with the lowest GDP p. c. In spite of considerable transfers to support the transformation process in East Germany, the East German regions as a whole reached only about 65% of the GDP p. c. of the West German regions average. During the period from 1991 to 1999 the governmental gross payments of the German State increased to an amount of 550 Billion Euro, which corresponds to 61 Billion Euro a year. About one third of this amount was used to develop the infrastructure and to support enterprises. During the same period the financial aid of the EU was about 3 Billion Euro per year (Ragnitz et al. 2000: 14 ff.).

It can not yet be anticipated whether and when the regions of East Germany will reach the economic level of West Germany. After positive convergence between East and West Germany up to the middle of the 1990`s a slight divergence can be seen since 1996 (DIW et al. 2002: 190 ff.).

This study therefore examines in how far some of the East German regions have achieved the same economic capability as the West German regions until today, so that they are on a competitive basis with the West German regions and are able to reach the same economic level in the long run. Above all an answer shall be found to the following questions: are there any East German regions today, which have the same economic capability as (successful) West German regions? Which regions could become independent of transfer payments in the near future? In which fields and regions do economic weakness still exist?

The study is organized as follows: section two contains the theoretical framework of the analysis, which is used to determine indicators of economic capability and competitiveness of regions. In Section three the methodical approach to compare the regions is first discussed and then the indicator system to be used is specified and described. Following in section four the empirical results are presented. Finally, chapter five closes with a brief summary and a discussion of some political implications.

2 Theoretical Basis of the Analysis

Regional disparities in Germany and the convergence process between West and East Germany are frequently observed in research projects. These studies come to different conclusions depending on the indicators they are using to measure the regional disparities.[2] Most of these studies use indicators referring to the level of economic development plus sometimes a few selected indicators, which allow to derive statements about the economic capability of regions (e. g. DIW et al. 2002; European Commission 2001; Härtel 2001; Eckey 2000). In contrast to these studies, the following research uses a larger set of indicators, in relation to the economic capability of regions.

The neoclassical growth theory can be seen as a initial point for the determination of these determinants. In the basic model (Solow 1956) capital is decisive for economic growth. Important assumptions in this model are constant returns to scale and diminishing marginal products in the production factors of labour and capital. Thus, economic growth p. c. – without technical progress – is only possible up to the steady state. Long term growth can only be explained by exogenous technical progress. A consequence of the diminishing marginal product of capital is that different regions will reach a similar per capita income over time (hypothesis of convergence). Regions with lower capital p. c. tend to have higher interest rates so that capital p. c. grows faster and so does GDP p. c. However this only describes conditional convergence, as output per capita depends on the saving rate, population growth and the position of the production function.

Within the new growth theory long term growth is explained endogenous by the abolishment of the diminishing marginal product of capital assumption in the regional/national economy. First ROMER takes up an approach by ARROW and explains long term growth on the basis of external effects of private investments (Romer 1986). Referring to this there are two explanations, one is “learning by doing”, the other “learning by investing”. With the inclusion of public investments a further explanation is given by BARRO. Accordingly public investments are able to create endogenous growth as the production function has – in case that private capital (K) and public capital (G) are increasing similarly –constant returns to scale in G and K (Barro 1990). Other models focus on the relevance of human capital. In the model developed by LUCAS the growth rate of a region depends on the endowment with human capital and the imbalance of physical und human capital. The growth rate increases with the extent of the imbalance if human capital is abundant, but decreases with the dimension of the imbalance if human capital is relatively scare (Lucas 1988). In another model by ROMER endogenous growth is created by research and development with the use of human capital (Romer 1990). In addition to that, further models have been developed e. g. by AGHION and

HOWITT, who emphasize the importance of research and development activities (Aghion; Howitt 1992). With respect to the convergence and divergence of regions, the models of the new growth theory remain ambiguous. Under certain conditions both divergence as well as convergence are possible. E. g. supposing that knowledge is a local public good divergence will follow. In contrary, if it is assumed that knowledge diffuses totally, or that in certain less developed regions a better allocation of resources will be reached than in other regions convergence can be expected.

The new economic geography provides another source of regional economic growth with the existence of agglomeration advantages (Fujita et al. 1999; Ottaviano; Puga 1998; Krugman 1991; Marshall 1920). Agglomeration advantages consist mainly of spill-over, synergy and labour market effects. A concentration of enterprises leads to the formation of a big workforce pool, enables technological spill-overs by way of transfer of technology and knowledge and can lead to a more intense networking of enterprises. This way spatial concentration may generate competitive advantages and increase the growth prospects of regional concentrated enterprises. However, the new economic geography is not really a growth theory. Rather it is an attempt to explain the spatial structure of the economy dependent on the agglomeration advantages and transport costs. In relation to convergence or divergence the new economic geography is also open to any result.

A comprehensive debate of these theories is not the purpose of this study. However, these theories supply arguments for relevant growth factors. The previous studies have not delivered conclusive empirical evidence on the quantitative impact of these growth factors on economic growth. But it seems to make sense to include all these factors by using suitable indicators for an empirical analysis of the different growth prospects of regions. The indicators included are: patent applications, expenses for research and development, employees with university degree, employees (including self employed), net migration of people between 18 to 25, number of entrepreneurs subject to turnover tax, net business registrations, industrial investments, regional accessibility, municipal investments and social assistance rate. They supply us with information about the innovation activity, the human capital, the private and public capital and the regional concentration (see 3.2).

The analysis is carried out at the level of specific German regions, so-called "Raumordnungsregionen (ROR)" (Böltken 1996). The definition of these regions mainly depends on commuter fluctuation of employees.[3] Thus the "Raumordnungsregionen" are the regions which determine the welfare of the regional population if nationwide migration is not considered. For those regional units, a number of economic indicators useful for our analysis are available.

3 Methodology and Data

3.1 Cluster Analysis

Cluster analysis can be described as a method to form homogenous groups of objects by their characteristics (Bacher 1996; Backhaus et al. 1996). Thus, it is possible to form several homogenous groups from a multitude of objects with similar characteristics and to compare them.

Basis of the cluster analysis in this study is a Rxv matrix with R as the observed regions and v as the characteristics. In the first step, it is necessary to check whether all regions can be used in the cluster analysis. Regions with incomplete data sets have to be excluded from the analysis. Thus 5 of 97 regions (Berlin, Bremerhaven, Bremen, Hamburg, Köln) can not be included in the cluster analysis. In a second step it is to check whether all variables can be used. For this the variables are tested on correlations. Variables which are highly correlated can dominate the cluster analysis and may distort the results (Backhaus et al. 1996: 313 f.). It is assumed, that variables with a correlation coefficient of $r>0.8$ (Schmidt 1995: 77) or even $r>0.9$ (Backhaus et al. 1996: 314) should be excluded. The calculation of the correlation coefficients (see appendix 1) shows that none of the variables are correlated to the mentioned extent. Another problem is the possibly different weighting of the variables by their scale unit. To avoid this the variables are standardized by z transformation (Bacher 1996: 173 ff.).

For the calculation of the clusters the Ward technique is used. This technique belongs to the hierarchical agglomerative methods. At this group of clustering methods every region is an individual cluster at the beginning of the algorithm. Then they are step by step joined together into groups. At first 92 clusters are available, then 91, 90, …, etc., until there remains one cluster. Fusion criterion of the Ward technique is – at the basis of the squared euclidean distance – the variance criterion. This means that this method minimizes the sum of squares of any two (hypothetical) clusters that can be formed at each step. Very homogenous clusters thereby are formed and in comparison to other fusion algorithm the real structure of cluster is identified correctly (Backhaus et al. 1996: 298).

As mentioned before the algorithm stops when there is only one cluster left. Therefore it is necessary to find out the optimal number of clusters. For this intention several critera are used: the agglomeration schedule, Mojena tests I and II, and the measure of homogeneity ETA^2.

At the agglomeration schedule increases of the distance levels are considered. Significant increases provide an indication for a possibly optimal number of clusters and can be seen

especially from cluster 29 to 28, 25 to 24, 18 to 17, 15 to 14, 10 to 9, 9 to 8 and 6 to 5 (see appendix 2). Therefore several cluster solutions are possible. To check these possible cluster solutions the Mojena test statistics I and II are calculated (Bacher 1996: 249 f.). Both provide different results. According to the test statistic I the permissible level of significance 0,997 has just exceeded from 29 to 28 clusters; according to test statistic II from 10 to 9 clusters. Considering these results the study continues with the 10-cluster-solution. This seems to be plausible also with regard to the measure of homogeneity ETA^2, which explains the variance within the cluster and outside the clusters. At the 10-cluster-solution ETA^2 is about 68% which means that 68% of the variance is outside the clusters and only 32% is within the clusters. All criterions considered, the 10-cluster-solutions seems to be a good model adjustment.

Finally the 10-cluster-solution is checked with the discriminant analysis (Backhaus et al. 1996: 90 ff.). The results show that 98,9% of the regions were classified correctly by cluster analysis. Only one region was taken to another cluster according to the results of the discriminant analysis. Additionally the regions which were excluded from the cluster analysis are assigned to the existing clusters using the discriminant analysis, whereby missing data are replaced by their population means.

The F-values, the T-values and the mean values of the variables are used to interpret the individual clusters. The F-value provides information about the homogeneity of the several groups. It is the quotient of the variance of a variable within the cluster and the variance of the variable in the population.

$$F_j^C = \frac{Var_j^C}{Var_j}.$$

The smaller this quotient the more homogenous is the cluster. F-values smaller than one indicate homogenous clusters (the variance of the variable j within the cluster is smaller than the variance of the variable j within the population). Values bigger than one in analogy identify heterogeneous cluster.
The T-value is used to characterise the clusters. It is calculated from the difference of the cluster mean value of the variable j and the mean value of the variable j of the population divided by the standard deviation.

$$T_j^C = \frac{\bar{X}_j^C - \bar{X}_j}{\sqrt{Var_j}}.$$

T-values smaller than zero indicate that the variable j is lower than the mean of the population. T-values bigger than zero analogous indicate that the variable j is higher than the mean of the population. In addition to the T-value the mean value of the variables is used in the interpretation as it provides information about the variable in their original scala.

3.2 Data

The possibility to describe the several growth factors by use of empirical variables is limited by the available statistic data. For some factors hardly any data exist, e. g. for the totality of public and private investments. With the following selection of data this is taken into account and it is attempted to approximate each growth factor. Further on, all variables are standardized by the number of inhabitants to guarantee regional comparability (see appendix 3 for the specific formation of the variables and their sources).

As indicators of the innovativeness patent applications and expenses for research und development are used. The number of patent applications can be seen as a measure of technical progress respectively of product and process innovations. As the number of the patent applications normally vary year by year a mean value is used for the period of time from 1995 to 2000. Data for the expenses for research and development are only available for the year 1999. These expenses can be interpreted as resources which are used for product and process innovations.

The supply of human capital is described by the employees with university degree, employees (including self employed), and by the net migration of people between the age of 18 to 25. The employees with university degree are on the one hand a indicator for the capacity of a region to generate knowledge, on the other hand they show the capacity of a region to adapt knowledge from other regions and to implement production improvements. The employees (including self employed) are used as an indicator of the accumulated regional knowledge gained at the production process. Each indicator is taken for one year (2001 resp. 2000). Yet not only the current available human capital of a region is of interest but also the gain or the loss of human capital. This factor is described by the migration movement balance of the 18 to 25 year olds. Here again the average of the years from 1995 to 2000 is considered.

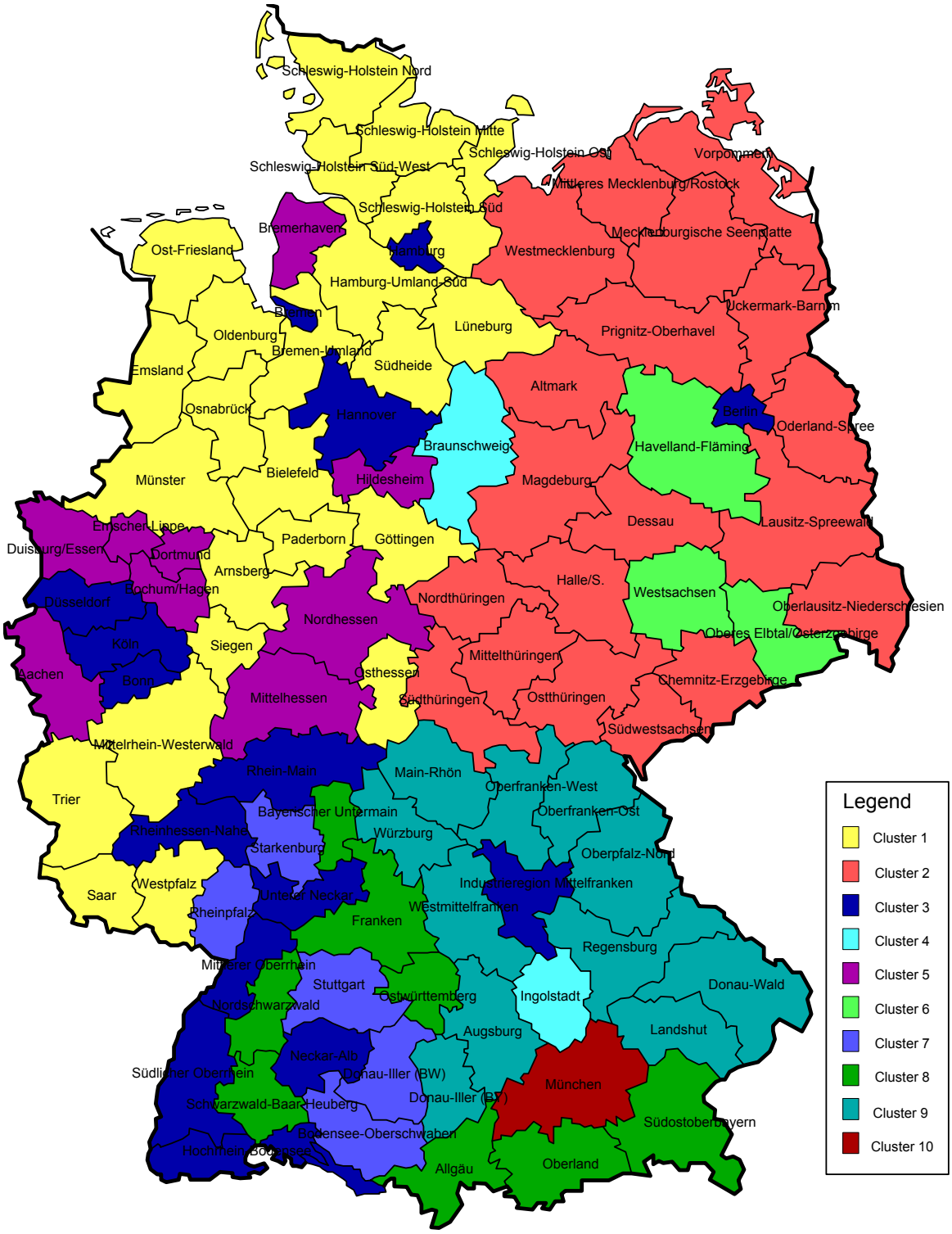
The spatial concentration is measured by the number of entrepreneurs subject to turnover tax. This indicator is used to measure the business density. It indicates the regional capability to produce technological spill-overs and to take advantage from horizontal and vertical enterprise linkages. The number of entrepreneurs subject to turnover tax are taken for the year 2000. In addition to that the number of net business registrations provide information about

the changes in the spatial business concentration and about the intensity of regional entrepreneurial initiative. For this indicator the average of the years from 1998 to 2001 is used.

The private capital is represented by the industrial investments. They are an approximation for the maintaining and the enlargement of the regional capital stock and of possible learning effects by investment activity. Further more they inform us about the regional attractiveness for private entrepreneurs. In order to take into account the fluctuation of the investment activity the average of the years from 1995 to 1999 is used.

The public capital is represented by the accessibility of regions, the municipal investments and the social assistance rate, as there are no regional data about the entire public investment activities. The regional accessibility may be considered as indicator for the endowment with nationwide traffic infrastructure. The municipal investments can be interpreted as a regional productive input factor for enterprises. The social assistance rate is used as a indicator that shows to which extent a region has the capability to invest in the municipal infrastructure.

4 Empirical Results



Map: Spatial Distribution of the Clusters

Source: Own calculations.

The cluster analysis identifies 10 clusters. The spatial distribution of the clusters can be seen on the map. A first result of the cluster analysis is that the regions of East Germany form two clusters (2 and 6) and the regions in West Germany eight clusters. This result point out that, even 12 years after unification, the economic capability in the East German regions is still different from the economic capability of the West German regions.[4]

The objective of this study is to find out whether some East German regions have reached the same economic capability as the regions in West Germany until today or in which fields there are still exist economic weaknesses. According to this objective only the East German clusters 2 and cluster 6 are presented more precisely and are compared with the other clusters.

The East German cluster 2 consists of 19 regions. Considering the characteristic profile of cluster 2 (see appendix 4) it can be seen that nearly all variables of cluster 2 have the lowest value in comparison to the other clusters. This applies to the patent applications, expenses for research and development, the employees (including self employment), the net migration of people between 18 to 25, the number of entrepreneurs subject to turnover tax, the net business registrations, the industrial investments and the regional accessibility. At first sight positive variables are the stock of employees with university degree, the municipal investment activity and the social assistance rate.[5] If it is taken into account that the positive fact of a high stock of human capital (employees with university degree) is negatively affected by a substantial loss of human capital (net migration of people between 18 to 25), and that an infrastructure gap between East and West Germany still exists, these indicators can not be interpreted as really positive with respect to the economic capability and competitiveness of cluster 2. In comparison to the other clusters only the West German clusters 1 and 5 have similar negative characteristic profiles but altogether better variable values.

Due to the results of the cluster analysis it can be assumed that the regions of cluster 2 have below average growth prospects. Thus these regions are probably not in the position to reach the economic level of the West German regions in the near future without transfers. Beside the low innovation activity, the lack of entrepreneurship, the low industrial investments and the loss of human capital are the most economic disadvantages. Our measure for entrepreneurial initiative shows that these regions are more likely not able to compensate the actual agglomerations disadvantages (business density) in the near future. Together with the low innovation activity it could be expected that the technical progress increases more slowly in these regions compared to the German average. The industrial investment activity is actually still supported by subsidies. If by 2007 the EU regional aid policy will probably get more restrictive in East Germany it can be expected that the investment activity diminishes if these regions do not increase their attractiveness. In this context human capital might play an

important role, but these regions are also confronted with a considerable loss of human capital.

The East German cluster 6 consists of three regions: the ROR Oberes Elbtal/Erzgebirge with Dresden, the ROR Westsachsen with Leipzig and the ROR Havelland-Fläming. The characteristic profile of cluster 6 is more positive than the characteristic profile of cluster 2. On the one hand there are still economic weaknesses, particularly in terms of patent applications, expenses for research and development, number of entrepreneurs subject to turnover tax and regional accessibility. On the other hand, however, there are very positive values particularly for employees with university degree, net business registrations and municipal investment activity. The municipal investment activity can be considered to be the same as in cluster 2. It can not be seen really positive because of the existing infrastructure gap between East and West Germany. In contrast to this is the high stock of human capital, an economic advantage which is not really affected by a substantial loss of human capital. In addition to that the positive value of net business registrations give reason to assume that the actual agglomeration disadvantage will disappear in the near future and might possibly become an agglomeration advantage in the long run. According to the results of the cluster analysis a real economic weakness exists only with respect to the low innovation activity.

In comparison to the West German clusters it can be seen that most of the variable values of cluster 6 are more positive than in cluster 1 and 5. Thus it can be assumed that the regions of cluster 6 have a higher economic capability than more than 50% of the German regions. As can be seen on the map the regions with a lower economic capability are mostly the northern peripheral regions, the old industrial regions in Germany and the other East German regions. Due to the results of the cluster analysis it can be expected that the regions of cluster 6 already have the potential to reach the economic level of several West German regions in the near future. However it can also be seen that, in comparison with cluster 6, the West German agglomerations and the South German regions (Cluster 3, 4, 7, 8, 9, 10) have economic advantages with respect to the economic capability, particularly in terms of their innovation activity.

5 Summary

The analysis has shown that most of the East German regions can not yet be compared with the West German regions in terms of their economic capability. Economic weaknesses can be seen particularly in the innovation activity, the business density, the entrepreneurial initiative, the industrial investments and the regional accessibility. Especially remarkable is the loss of human capital from which most of the East German regions suffer. The consequence of this loss could be that the still existing advantage in the endowment with human capital disappears and that human capital will become a bottle-neck in the near future. All in all most East German regions will probably not reach the same level of economic development as the West German regions in the near future. The analysis gives reason to fear that the gap increases even more. However it also has become clear, that some East German regions, especially those with agglomerations as Dresden and Leipzig, have a better economic capability than a lot of West German regions. So it can be expected that at least those regions might reach the same level of economic development as several West German regions in the near future. In these regions considerable weakness still exists only in the field of innovation activity. However the economic capability of these East German regions can not be compared with the economic capability of the successful West German regions. So probably they only catch up to the economically weak West German regions. This however could change if these regions succeeded in increasing their innovation activity.

All in all only a few East German regions have been successful in becoming economically competitive despite of the transfers. These are the regions with the important East German agglomerations. However these regions are probably not able to catch up to the successful West German regions. This raises the question if the German regional aid was successful by subsidising all East German regions in the same way. Maybe it would be more efficient to concentrate the regional aid more to the agglomerations. The results point into this direction but have to be proved by further studies.

[1] Excluding Berlin.

[2] Another possible reason is the use of varying definitions of regions.

[3] This definition is not used for all regions. Hamburg, Bremen and Berlin are defined by there administrative borders. It is important to keep this in mind when interpreting empirical results.

[4] An exception is Berlin, which belongs to cluster 3. But as mentioned before the result have to be interpreted cautiously because of the regional definition of Berlin by administrative borders.

[5] A negative T-value by regional accessibility and social assistance rate has to be interpreted positive. This can be seen by the mean in appendix 4.

References

Aghion, P.; Howitt, P. (1992): A Model of Growth trough Creative Distruction, in: *Econometrica*, Vol. 60, S. 323-351.

Bacher, J. (1996): Clusteranalyse, 2.Auflage, München.

Backhaus, K. et al. (1996): Multivariate Analysemethoden, 8.Auflage, Berlin.

Barro, R. J. (1990): Government Spending in a Simple Model of Endogenous Growth, in: *Journal of Political Economy*, Vol. 98, S. 103-125.

Bode, E. (1996): Ursachen regionaler Wachstumsunterschiede: Wachstumstheoretische Erklärungsansätze, Kieler Arbeitspapiere 740, Kiel.

Böltken, F. (1996): Neuabgrenzung von Raumordnungsregionen nach den Gebietsreformen in den neuen Bundesländern, Bundesforschungsanstalt für Landeskunde und Raumordnung, Bonn.

DIW; IAB; IfW; IWH; ZEW (2002): Fortschritte beim Aufbau Ost, in: *Wirtschaft Im Wandel*, H. 7-8, Jg.8, S. 182-234.

Eckey, H. F. (2000): Der wirtschaftliche Entwicklungsstand in den Regionen des Vereinigten Deutschlands, Kassel.

European Commission (2001): Unity, Solidarity, Diversity for Europe, its People and its Territory - Second Report on Economic and Social Cohesion, Brüssel.

Fujita, M.; Krugman, P.; Venables, A. J. (1999): Spatial Economy, MIT Press, Cambridge.

Härtel, H. H. (2001): Ostdeutschland im nationalen und europäischen Regionalvergleich, in: *Wirtschaftsdienst*, H. 5, Jg.81, S. 292-296.

Krugman, P. (1991): Geography and Trade, Leuven.

Lucas, R. E. (1988): On the Mechanics of Economic Development, in: *Journal of Monetary Economics*, Vol. 22, S. 3-42.

Marshall, A. (1920): Principles of Economics, London.

Ottaviano, G.; Puga, D. (1998): Agglomeration in the Global Economy - A Survey of the New Economic Geography, in: *The World Economy*, Vol. 21, S. 707-731.

Ragnitz, J. et al. (2000): Simulationsrechnungen zu den Auswirkungen einer Kürzung von Transferleistungen für die neuen Bundesländer, IWH-Sonderheft, Halle.

Romer, P. M. (1986): Increasing Returns and Long-Run Growth, in: *The Journal of Political Economy*, Vol. 94, S. 1002-1037.

Romer, P. M. (1990): Endogenous Technological Change, in: *Journal of Political Economy*, Vol. 98, S. 71-101.

Schmidt, B. (1995): Kreistypisierung und zentralörtliche Gliederung mit statistischen Verfahren, Regensburg.

Solow, R. M. (1956): A Contribution to the Theory of Economic Growth, in: *Quaterly Journal of Economics*, Vol.70, S. 65-94.

Sternberg, R. (2001): New Economic Geography und Neue regionale Wachstumstheorie aus wirtschaftsgeographischer Sicht, in: *Zeitschrift für Wirtschaftsgeographie*, H. 3-4, Jg. 45, S. 159-180.

Stough, R. R.; Johansson, B.; Karlsson, C. (2001): Theories of Endogenous Regional Growth. Lessons for Regional Policies, Berlin.

Appendix 1: Coefficient of Correlation (Bravais/Pearson)

	Patent applications	Expenses for research and development	Employees with university degree	Employees (including self employed)	Net migration of people between 18 to 25	Number of entrepreneurs subject to turnover tax	Net business registrations	Industrial investments	Regional accessibility	Municipal investments
Expenses for research and development	0,741**									
Employees with university degree	0,155	0,368**								
Employees (including self employed)	0,633**	0,485**	0,260*							
Net migration of people between 18 to 25	0,562**	0,425**	0,121	0,640**						
Number of entrepreneurs subject to turnover tax	0,654**	0,373**	0,039	0,557**	0,556**					
Net business registrations	0,231*	0,197	-0,113	0,309**	0,558**	0,414**				
Industrial investments	0,454**	0,556**	-0,015	0,304**	0,186	0,083	0,053			
Regional accessibility	-0,658**	-0,451**	-0,050	-0,396**	-0,698**	-0,586**	-0,350**	-0,334**		
Municipal investments	-0,072	-0,075	0,013	0,02	-0,391**	-0,016	-0,111	0,087	0,355**	
Social assistance rate	-0,452**	-0,224*	0,078	-0,316**	-0,031	-0,417**	-0,187	-0,344**	0,092	-0,600**

* The correlation coefficient is significant at the 5%-level; ** The correlation coefficient is significant at the 1%-level.

Source: Own calculations.

Appendix 2: Agglomeration Schedule (Ward-technique, last 30 steps)

Stage	Cluster Combined		Distance	Increase of distance	Mojena I	Significance	Mojena II	Significance
62	1	10	9,507	0,296	2,525	0,994	0,828	0,796
63	16	46	9,645	0,139	2,444	0,993	0,752	0,774
64	15	34	10,869	1,224	2,833	0,998	1,143	0,873
65	1	3	11,154	0,285	2,761	0,997	1,081	0,860
66	6	22	11,186	0,033	2,603	0,995	0,930	0,824
67	20	39	11,760	0,574	2,668	0,996	0,999	0,841
68	76	87	15,457	3,697	3,812	1,000	2,147	0,984
69	41	63	15,916	0,459	3,577	1,000	1,948	0,974
70	2	5	16,185	0,270	3,338	1,000	1,732	0,958
71	47	67	16,965	0,779	3,291	1,000	1,699	0,955
72	76	77	18,278	1,314	3,381	1,000	1,801	0,964
73	6	7	18,324	0,046	3,128	0,999	1,560	0,941
74	16	41	18,406	0,081	2,933	0,998	1,371	0,915
75	47	69	20,743	2,337	3,265	0,999	1,704	0,956
76	16	65	21,330	0,588	3,153	0,999	1,600	0,945
77	19	84	21,613	0,282	2,998	0,999	1,450	0,926
78	6	8	24,723	3,110	3,391	1,000	1,845	0,967
79	1	2	27,017	2,295	3,541	1,000	2,004	0,977
80	6	29	29,811	2,793	3,719	1,000	2,194	0,986
81	1	15	34,062	4,251	4,049	1,000	2,539	0,994
82	64	90	34,767	0,705	3,763	1,000	2,275	0,989
83	1	20	40,414	5,647	4,174	1,000	2,698	0,997
84	19	47	48,869	8,456	4,757	1,000	3,303	1,000
85	64	76	51,636	2,767	4,480	1,000	3,064	0,999
86	6	26	56,450	4,814	4,456	1,000	3,062	0,999
87	19	88	71,636	15,186	5,285	1,000	3,912	1,000
88	16	19	105,534	33,898	7,095	1,000	5,763	1,000
89	16	64	196,764	91,230	11,065	1,000	9,829	1,000
90	1	6	294,026	97,261	10,894	1,000	9,866	1,000
91	1	16	518,281	224,256	12,827	1,000	11,927	1,000

Source: Own calculations.

Appendix 3: Indicators used in the study

a) Patent applications

$$\frac{\textit{Average of the patent applications from 1995 to 2000}}{\textit{100 000 inhabitants}}$$

Source: Greif, S. (2002): Patentatlas Deutschland, Deutsches Patent- und Markenamt, München.

b) Expenses for research and development

$$\textit{Expenses for research and development of enterprises Euro 1999/inhabitant}$$

Source: Stifterverband für die deutsche Wissenschaft.

c) Employees with university degree

$$\frac{\textit{Employees subject to social insurance contribution with university or college degree 2001}}{\textit{1 000 Employee subject to social insurance}}$$

Source: German federal labour office.

d) Employees (including self employed)

$$\frac{\textit{Employees (including self employed) at their working place 2000}}{\textit{1 000 inhabitants}}$$

Source: Statistical offices of the German states.

e) Net migration of people between 18 to 25

$$\frac{\textit{Average net migration of people between 18 to 25 1995 to 2000}}{\textit{1 000 inhabitants between 18 to 25}}$$

Source: Statistical offices of the German federation and the German states.

f) Number of entrepreneurs subject to turnover tax

$$\frac{\textit{Number of entrepreneurs subject to turnover tax 2000}}{\textit{10 000 inhabitants}}$$

Source: Statistical offices of the German states.

g) Net business registrations

$$\frac{\textit{Average net business registrations1998 to 2001}}{\textit{100 000 inhabitants}}$$

Source: Statistical offices of the German federation and the German states.

h) Industrial Investments

Average investments in mining and manufacturing enterprises 1995 to 1999/inhabitant

Source: Statistical offices of the German federation and the German states.

i) Regional accessibility

Average driving time with car and plane to 41 european agglomerations 1998

Source: Federal office for building and regional planning.

j) Municipal investments

Average municipal expenses for fixed investments 1995 to 2000/inhabitant

Source: Statistical offices of the German federation and the German states.

k) Social assistance rate

$$\frac{\textit{Persons receiving social assistance 2001}}{\textit{1 000 inhabitants}}$$

Source: Statistical offices of the German states.

l) Gross domestic product per capita

Gross Domestic Product 2000 /inhabitant

Source: Statistical offices of the German federation and the German states.

m) Unemployment rate

$$\frac{\textit{Unemployed persons 2002}}{\textit{labor force}}$$

Source: German federal labour office.

Appendix 4: Characteristic profile of the clusters

		Patent appli-cations	Expenses for research and develop-ment	Employees with university degree	Employees (including self employed)	Net migration of people between 18 to 25	Number of entrepreneurs subject to turnover tax	Net business registrations	Industrial investments	Regional accessibility	Municipal investments	Social assistance rate	Reporting:	
													GDP p.c.	Unemploy-ment rate
ROR	Mean value	36,8	299,7	74,0	456,2	6,9	341,4	108,0	575,9	270,5	320,2	28,7	22554	10,5
Cluster 1	Mean value	24,7	115,5	54,1	441,3	8,0	336,2	127,9	482,9	275,8	252,0	33,4	21500	9,0
N=24	*T-value*	*-0,49*	*-0,47*	*-0,75*	*-0,30*	*0,06*	*-0,09*	*0,40*	*-0,37*	*0,17*	*-0,78*	*0,31*	*-0,19*	*-0,31*
	F-value	0,15	0,03	0,14	0,57	0,42	0,31	0,25	0,49	0,32	0,23	0,40	0,13	0,12
Cluster 2	Mean value	11,5	70,8	89,3	413,4	-18,3	287,1	48,6	470,2	311,9	392,3	27,3	15661	18,6
N=19	*T-value*	*-1,02*	*-0,59*	*0,57*	*-0,86*	*-1,44*	*-0,95*	*-1,19*	*-0,42*	*1,33*	*0,82*	*-0,10*	*-1,22*	*1,69*
	F-value	0,05	0,02	0,33	0,38	0,16	0,13	0,63	0,84	0,45	0,16	0,18	0,05	0,26
Cluster 3	Mean value	49,3	437,9	99,1	511,0	29,4	385,8	126,8	573,7	235,9	270,2	39,5	29323	8,5
N=15	*T-value*	*0,51*	*0,35*	*0,94*	*1,10*	*1,28*	*0,78*	*0,38*	*-0,01*	*-1,11*	*-0,57*	*0,73*	*1,20*	*-0,42*
	F-value	0,56	0,30	0,44	1,05	0,32	0,37	0,94	0,33	0,27	0,24	2,21	1,08	0,41
Cluster 4	Mean value	58,3	1578,9	66,9	462,6	14,1	294,0	141,0	1631,1	264,0	323,6	23,0	26423	8,1
N=2	*T-value*	*0,87*	*3,28*	*-0,26*	*0,13*	*0,41*	*-0,83*	*0,66*	*4,21*	*-0,21*	*0,04*	*-0,38*	*0,69*	*-0,50*
	F-value	0,02	0,30	0,08	0,08	0,20	1,36	0,74	1,58	0,90	2,39	1,86	0,03	0,49
Cluster 5	Mean value	29,8	174,1	66,8	418,0	7,8	300,1	84,7	483,7	257,8	219,3	43,9	21134	10,5
N=9	*T-value*	*-0,28*	*-0,32*	*-0,27*	*-0,77*	*0,05*	*-0,72*	*-0,47*	*-0,37*	*-0,41*	*-1,16*	*1,02*	*-0,25*	*0,00*
	F-value	0,22	0,11	0,21	0,30	0,39	0,23	0,22	0,23	0,49	0,43	0,34	0,15	0,12
Cluster 6	Mean value	21,1	214,0	127,6	459,6	6,1	310,5	148,0	547,9	286,3	420,7	29,4	17528	16,0
N=3	*T-value*	*-0,63*	*-0,22*	*2,01*	*0,07*	*-0,05*	*-0,54*	*0,80*	*-0,11*	*0,51*	*1,15*	*0,05*	*-0,89*	*1,13*
	F-value	0,28	0,18	0,85	0,30	0,05	0,02	0,16	0,86	0,04	0,07	0,15	0,02	0,10
Cluster 7	Mean value	85,1	1360,2	93,3	483,4	12,8	377,7	104,1	787,7	242,6	306,7	21,0	27930	6,0
N=5	*T-value*	*1,95*	*2,72*	*0,73*	*0,55*	*0,33*	*0,64*	*-0,08*	*0,85*	*-0,90*	*-0,15*	*-0,52*	*0,95*	*-0,93*
	F-value	0,56	0,59	0,58	0,72	0,26	0,17	0,34	0,25	0,73	1,19	0,12	0,30	0,03
Cluster 8	Mean value	69,9	303,2	53,0	470,4	5,9	422,8	90,8	689,9	252,5	381,3	13,1	24246	5,9
N=8	*T-value*	*1,34*	*0,01*	*-0,79*	*0,29*	*-0,06*	*1,43*	*-0,34*	*0,46*	*-0,58*	*0,70*	*-1,05*	*0,30*	*-0,95*
	F-value	0,42	0,08	0,02	0,18	0,12	1,28	1,47	0,37	0,39	0,18	0,05	0,11	0,01
Cluster 9	Mean value	43,6	194,1	48,2	480,6	10,4	341,0	150,3	673,2	271,6	426,9	11,1	23833	7,2
N=11	*T-value*	*0,27*	*-0,27*	*-0,97*	*0,49*	*0,20*	*-0,01*	*0,85*	*0,39*	*0,04*	*1,22*	*-1,19*	*0,23*	*-0,68*
	F-value	0,34	0,08	0,21	0,14	0,40	0,20	0,29	0,65	0,27	0,40	0,01	0,10	0,09
Cluster 10	Mean value	102,5	1655,7	157,5	605,4	46,3	561,8	190,2	610,5	226,0	347,0	13,1	42899	5,4
N=1	*T-value*	*2,65*	*3,48*	*3,13*	*2,99*	*2,25*	*3,87*	*1,64*	*0,14*	*-1,43*	*0,31*	*-1,06*	*3,60*	*-1,06*
	F-value	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Source: Own calculations.