Oesterreichische Nationalbank

Stability and Security.

Workshops

Proceedings of OeNB Workshops

New Regional Economics in Central European Economies: The Future of CENTROPE

March 30 to 31, 2006

Eurosystem

No. 9

Regional Convergence within the EU-25: A Spatial Econometric Analysis

Martin Feldkircher

Institute for Advanced Studies

Abstract

This study investigates absolute convergence within the EU-25 for the time period 1995–2002. It is shown that growth performance and convergence depend crucially on the development of a region's surrounding. The detected spatial autocorrelation is of substantive form indicating that least squares estimation of the absolute convergence model yields biased results. A yearly convergence rate of 0.7% to 0.9% is estimated by using a spatial autoregressive model specification. Several robustness checks are carried out: First, it is examined whether the functional relationship of the convergence equation is stable over space. Secondly, the sensitivity of the estimation results on the specified weight matrix is investigated. Third, the paper identifies the source of spatial dependence.

1. Introduction

Since the 1990s considerable attention has been drawn to the question of regional income convergence in Europe. A lot of quantitative research has been conducted, and several new theoretical approaches have been proposed. A similarity of most studies is the neglected spatial interaction of the underlying observations. Now, there seems to be wide-spread agreement that spatial dependencies should be taken into account when analyzing growth. Recent studies suggest that geographic location does matter for a region's growth performance and consequently its pace of convergence. Spatial interactions such as technological spillovers[1] or factor mobility, both being important forces for the process of convergence, should not be neglected. There are two ways to deal with this phenomenon using standard econometric techniques: The data can be either nationally weighted or country dummy variables can be incorporated in the regression equation. Indeed, as this study will show, a great extent of spatial correlation is based on country-specific

[1] See López-Bazo et al. (2004) for a spatial econometric estimation of technological spillover effects.

effects. These approaches have been criticized[2] as being too restrictive for two reasons: First, spatial effects across national borders are excluded and secondly, the assumption that all regions of a country belong to the same national growth cluster seems not to be in line with reality. In addition, these approaches aim solely to eliminate possible spatial correlation in the regression's disturbance term and do not provide any further insights of the convergence process itself.

Spatial econometric regressions are thus more flexible in comparison to other approaches, and will pose the econometric rationale for this study. One focus of this paper is on the sensitivity of estimation results with respect to spatial proximity. Consequently all models and descriptive statistics are estimated using several weight matrices. Another issue discussed is the source of spatial dependence. Do spatial spillovers have a bigger influence on a region's growth performance than country effects?

The paper is organized as follows: Section 2 introduces the unconditional β-convergence model. Chapter 3 provides a description of the data. Section 4 examines the spatial structure of the underlying data by means of exploratory spatial data analysis. Section 5 consists of estimation results. Section 6 deals with several tests for robustness of the results, including estimations of the two-club-convergence model and section 7 concludes.

2. Convergence Based on the Neo-Classical Growth Theory

In the neo-classical growth theory, growth is solely determined by the rate of technology which is assumed to be exogenous. The main force that drives economies (homogenous countries, regions) to converge is the fact that returns to physical capital are diminishing. Localities with low initial income per capita have low ratios of capital to labor, and hence they also exhibit a higher marginal product of capital.[3] Thus there is a point at which per capita income growth converges to zero assuming that technology does not grow. This so-called steady state y^* can be assumed to vary (conditional β-convergence) or to be equal for all analyzed economies (unconditional β-convergence). The diminishing returns to physical capital imply that economies far away from y^* grow faster than those that are closer to y. In a regression context, absolute β-convergence can be estimated by regressing yearly average growth rates on a constant term and initial income.[4] Evidence of convergence is found whenever the β-coefficient is significantly different from zero and negative, thus implying that economies (regions) with low initial GDP per capita grow on average faster than others having a relatively high initial GDP. The underlying assumption here is that all economies are intrinsically

[2] Niebuhr (2001).
[3] Jones (2002).
[4] I will use GDP per capita in purchasing power parities as a proxy for income. Henceforth, the terms income and GDP per capita will be used both to denote GDP per capita (PPP).

the same (i.e. they share the same production function, savings rates, etc.), except their initial conditions making the concept of unconditional β-convergence applicable. A spatial regime switching model is estimated in section 6 devoting attention to the stability of the regression model over the data. Motivation for this model specification can not only stem from a spatial econometric point of view but also from economic theory. Here the identified regimes are called "convergence-clubs". Regions within these clubs are assumed to interact more with members of the club than with others from outside. The assumption of a single steady state for all regions belonging to the EU-25 is relaxed by allowing for club-specific steady states.

3. Data

The data used in this study is taken from the Eurostat-database "Regio"[5]. The explanatory variable is initial GDP per capita (purchasing power parities) in 1995 in logarithms; the dependent variable is the yearly average growth rate from 1995 to 2002. Although recent convergence studies[6] analyze data for a larger time horizon, this makes no sense for the purpose of this study for several reasons. Firstly, there is no reliable data available for the new Member States of the enlarged EU-25 before 1995. Secondly, even if available, interpretation and comparison of data on income with that for the old Member States could not be done in a meaningful way. This is due to the transition of the former CEE countries from a centrally planed to a market economic system.[7]

The data consists of 246 NUTS 2 regions for all the Member States of the EU-25 except Cyprus and some regions of France and Portugal. Those were dropped due to their isolated geographic position. The territorial classification "NUTS" (Nomenclature of Units for Territorial Statistics Classification) is proposed by Eurostat and does not deviate in most instances from administrative borders set by the specific countries. Hence, this NUTS classification is not based on functional, economically integrated units, which is the source of frequent criticism.[8]

4. Exploratory Spatial Data Analysis (ESDA)

According to Anselin (1988) one can distinguish between two spatial effects: Spatial dependence and spatial heterogeneity. Intuitively, observations from adjacent regions can on the one hand be correlated (*Spatial dependence / Spatial autocorrelation),* or on the other hand a functional relationship can vary across the regions (*Spatial Heterogeneity*).

[5] http://epp.eurostat.cec.eu.int.
[6] Mella-Marquez and Chasco-Yrigoyen (2004), Niebuhr (2001).
[7] Fischer and Stirböck (2004).
[8] Martin (2001).

The first effect – *Spatial autocorrelation* – can stem from aggregation of variables[9]. Because the underlying spatial scale of the variable is not correctly reflected within the aggregated variable, the result might be exposed to spatial autocorrelation. Although this kind of measurement error is likely to occur – and definitely is evident in the data underlying this study – it is not the main source of spatial dependence. Spatial autocorrelation derives to a large extent from the fact that localities interact with each other. The relationship of correlation and distance is in most instances a negative one. The second effect – *Spatial Heterogeneity* – can be dealt with by standard econometric methods. In many cases the assumption of a stable functional relationship across space might not hold. The following section introduces descriptive spatial statistics to assess whether the first spatial effect is present in the data.

4.1 Local Moran's I and Getis-Ord G_i*

The Local Moran's I statistic can be used to test whether the variables of the absolute convergence equation are clustered in space:

$$I_i = \frac{(x_i - \overline{x})}{\frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2} \sum_{j=1}^{n} w_{ij}(x_j - \overline{x}) \tag{1}$$

where x_i represents the underlying variable for region i, $\overline{x}$ the sample mean and w_{ij} the corresponding elements of a specified weight matrix W[10]. The null-hypothesis of the test statistic is the absence of spatial autocorrelation, implying that location does not matter. Inference is based on the z-transformed values of the statistic. The Local Moran's I decomposes the global spatial pattern and indicates to which extent a geographic locality is surrounded by similar / dissimilar values forming a geographical pattern. This implies that some structure is present in the data, which can be regarded as additional information. Most economic variables display positive spatial autocorrelation. Similar values are likely to cluster in space. Negative autocorrelation implies that contiguous areas are more likely characterized by dissimilar values than in a random pattern, which is a result not to expect intuitively, since it is the opposite of clustering. The four possible decomposition categories are:

Positive spatial correlation:

1) high-high
2) low-low

---

[9] Anselin (1988).

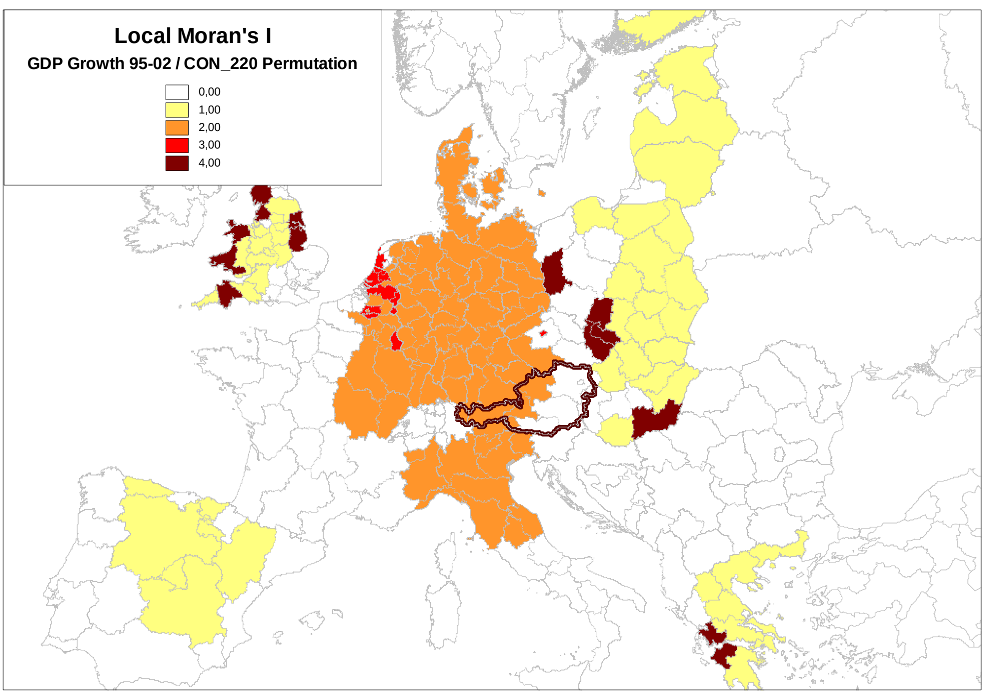
[10] For a description of the weight matrices consider the Appendix section.

Negative spatial correlation:
3) high-low
4) low-high

A region belonging to one of the two first categories is surrounded by observations that are characterized by similar values in magnitude. Spatial outliers (hot-spots) are found in categories 3) and 4).

Chart 1 shows the Local Moran's I significance map (at the 10% level[11]) for the yearly average growth rates 1995–2002 using the color-coding scheme from above. It was computed using a permutation approach, by empirically generating a reference distribution from which mean and variance are taken. This reference distribution is simulated under the null-hypothesis of no spatial dependencies. The permutation approach is then carried out by randomly reshuffling the observed values over all locations and by re-computing the *I* statistic for each sample.[12]

Chart 1: Local Moran's I – Yearly Average Growth Rates 1995–2002



Source: Author's calculations.

[11] Regions for that the test statistic did not reject the null-hypothesis are not assigned a color.
[12] For further description of Local Moran's I test statistic see Anselin (1992).

The chart reveals that Europe is divided into three growth zones: Clusters of fast growing regions in the East and West of Europe and in between a cluster of slow growing regions. Significant growth clusters indicate that regions located in a dynamic surrounding of high growing localities are more likely to show high growth rates than ones that are neighbors of "slow-growing" areas. This clustering phenomenon can be due to the existence of regional spillovers. A similar pattern with respect to the three clusters can be identified for per capita initial income in 1995 as well as in 2002. The overall structure with respect to the three zones remains the same but the low-low clusters are located in the East and West and the high-high cluster in between.

A second way to examine the spatial pattern of the data is by using the Getis and Ord Gi* distance statistic. It is used to identify the regimes of the spatial regime-switching model estimated in chapter 6. Unlike the Global Moran's I, which is a kind of correlation coefficient between observed values and locations, the Gi* statistic measures the concentration of a spatially distributed variable. It can be calculated as a global measurement or as a local indicator of spatial association. The local version of the distance statistic is defined as:

$$G_i^* = \left(\frac{\sum_{j=i}^{N} w_{ij}(\delta) x_j}{\sum_{j=i}^{N} (x_j)} \right) \qquad (2)$$

The w_{ij} elements correspond to a weight matrix (not standardized in rows) that is based on a threshold distance point δ. For every region *i*, the numerator of (2) gives the sum of the underlying variable for all regions lying within δ, including the observation *i* itself.[13] If large values of the variable examined are clustered close to region *i*, G_i* will be large as well. Inference is based on the z-transformed values of the statistic, and indicates to which extent an observation is surrounded by high or low values. This means that the G_i^* statistic shows solely positive spatial correlation, "high-high" clusters are indicated by positive z-values of the statistic, and "low-low" clusters by negative ones.

5. Estimation

As mentioned in section 2 the unconditional β-convergence model is given by:

[13] The G_i^* distances statistic includes also the values for the region under consideration *i* in the sum of the denominator, whereas G_i –not used in this study – does not.

$$\frac{1}{t}\log\left(\frac{y_{i,02}}{y_{i,95}}\right)=\alpha+\beta\log\left(y_{i,95}\right)+u_i$$
$$u_i \sim i.i.d.(0,\sigma^2{}_u) \tag{3}$$

with the disturbance term assumed to be *i.i.d*. Six dummy variables are added on the right-hand side of (3). Three of them ("Southern and Eastern Ireland", "Közép-Magyarország" and "Mazowieckie") were identified by examining the residuals of least squares estimation of (3). By including them into the regression equation, the Jarque-Bera test does not reject the null-hypothesis of non-normality of the error term. The remaining three dummy variables correspond to outlying regions identified by the Cook's Distance statistic. According to the statistic, the regions "Luxembourg", "Latvia" and "Inner London" were recognized to possibly have serious influence on the regression coefficient.

As outlined in chapter 4 there are two main sources of spatial correlation: The measurement error and the interaction of localities. In the terminology of Anselin (1988) he refers to the first one as a "by-product of measurement errors" (sometimes also called *nuisance dependence*). The latter one is due to "the existence of a variety of spatial interaction phenomena" which is in the literature referred to as *substantive form* of spatial autocorrelation. The former is more likely to occur and evident in most data sets of empirical cross-sectional studies. In case that the data exhibits spatial dependence of the nuisance form, spatial error models (henceforth SER) are a proper econometric model class to work with. They model the error term of equation (3) as a spatial moving average or spatial autoregressive process. In words this means that if spatial dependencies are present in the data, but to a rather "small" extent, modeling the error term is sufficient to get efficient estimates. In contrast, ignoring spatial correlation would yield still unbiased but *inefficient* OLS estimates. The SER model estimated in this paper is of the form:

$$\frac{1}{t}\log\left(\frac{y_{i,02}}{y_{i,95}}\right)=\alpha+\beta\log\left(y_{i,95}\right)+dummies+\varepsilon_i$$
$$\varepsilon_i=\lambda W\varepsilon_i+u_i \tag{4}$$
$$u_i \sim i.i.d.(0,\sigma^2{}_u)$$

with λ being a spatial parameter and W a specified weight matrix. In contrast to the former case, severe consequences occur whenever spatial dependence is of *substantive* form. In accordance to time series analysis, auto-correlated disturbances might point to an omitted lagged variable. Put differently, if the error term reveals a certain structure, it could be that not all of the information given by the data was properly taken into account. With respect to convergence, spatial autocorrelation of the substantive form means that regional spillovers do not only exist but are even determining a region's convergence process. The so-called spatial autoregressive model (henceforth SAR) – designed for this problem – explicitly adds a spatially lagged variable on the right-hand side of equation (3). In most, but not necessarily all instances, the added regression coefficient is a spatial lag of the dependent variable (therefore spatial "autoregressive" model).

In the context of convergence the spatial autoregressive model is given by:

$$\frac{1}{t}\log\left(\frac{y_{i,02}}{y_{i,95}}\right)=\alpha+\beta\log\left(y_{i,95}\right)+\rho W\left(\frac{1}{t}\log\left(\frac{y_{i,02}}{y_{i,95}}\right)\right)+dummies+u_i \qquad (5)$$

$$u_i \sim i.i.d.(0,\sigma^2_u)$$

where ρ is the autoregressive parameter and *W* the weight matrix. The estimation results are given in Table 1.

Table 1: Estimation of Convergence

	OLS Model		*SER Model*	*SAR Model*
α	0.181669	α	0.146387	0.077043
t-value	9.831427	z-value	6.110031	4.091827
Pr(>|t|)	0.000000	Pr(>|z|)	0.000000	0.000043
S.D.	0.018478	S.D.	0.023958	0.018829
β	–0.014192	β	–0.010546	–0.006743
t-value	–7.320362	z-value	–4.180739	–3.735143
Pr(>|t|)	0.000000	Pr(>|z|)	0.000029	0.000188
S.D.	0.001939	S.D.	0.002522	0.001805
	-	ρ / λ	0.729551	0.714431
	-	z-value	10.540210	10.676990
	-	Pr(>|z|)	0.000000	0.000000
	-	S.D.	0.069216	0.066913
Log.Lik.	747.107		777.804	782.201
AIC	1478.210		–1539.61	–1546.40
Obs.	246		246	246
Weight matrix	-		INV2_400	INV2_400

Source: Author's calculations.

All three models confirm the convergence hypothesis but the β-coefficient is varying in size. It is about two times larger than that of the SAR model. Compared to least squares estimation, both spatial models obtained a better fit indicated by the value of the maximized log likelihood function and a smaller AIC information criterion.[14] The significance of the two spatial coefficients ρ / λ indicates that the OLS model is not appropriate, which will be further explored in the next table consisting of selected specification diagnostics.

Table 2: Diagnostics for Spatial Dependence of the OLS Model

Test	*MI/DF*	*Value*	*Prob.*
Moran's I (error)	0.281891	10.288409	0.000000
RLMerr	1	2.700637	0.100308
RLMLag	1	10.572667	0.001148
Lagrange Multiplier (SARMA) Weight matrix INV2_400	2	105.706119	0.000000

Source: Author's calculations.

The Moran's *I* test (error) points to spatial dependence. Since this test is a measurement of global spatial dependence, it gives no conclusions about the source of spatial autocorrelation, which is the task of several Lagrange Multiplier tests. Even more so, they are the most important decision tools in spatial econometrics, clarifying whether spatial dependence is of substantive or nuisance form. There are robust versions of the LM-tests[15], which both take the possible specification of the respective other test into account. For example the "RLMerr" tests for spatially autocorrelated error terms, and also controls for the possible presence of a missing spatially lagged variable. The opposite is true for "RLMLag". Since the RLMLag rejects the null-hypothesis of no omitted spatial lag, inference goes in favor of the SAR model specification. The autoregressive parameter ρ indicates a positive relationship of the dependent variable and its spatial lag. With respect to convergence, this means that convergence speed is not solely determined by a region's initial income, but also by a high degree of its neighbourhood region's growth performance.

6. Robustness of the Results

The sensitivity of estimation results to the definition of spatial proximity is often criticized as a severe drawback in spatial econometrics. Hence it has to be assessed whether the estimated convergence speed is sensitive to the choice of the weight matrix. Chapter 5 outlined the economic implications of the spatial autoregressive model in contrast to that of the spatial error model. It would be unsatisfactory if the

[14] The standard R^2 is not appropriate to value the fit of a spatial model (Anselin 1988).
[15] See Anselin (1992) for a description of the test statistics.

model specifications as well as its implications are sensitive to spatial proximity, which is incorporated by the design of the weight matrices. Thus equation (3) including the six dummy variables is re-estimated using five different weight matrices. In every case the Lagrange multiplier tests come to the same conclusion: The detected spatial correlation is of substantive form. Table 3 gives a summary of the estimation results using different weight matrices.

Table 3: Summary Convergence Speed

Matrix	Model	β-Coefficient[16]	Convergence Speed (%)	AIC	HD[17]
-	OLS	–0.0142	1.49480	–1478.21	46.37
CON350	SAR	–0.0088	0.90778	–1520.84	76.36
CON220	SAR	–0.0075	0.76932	–1519.82	90.01
INV1_400	SAR	–0.0073	0.75290	–1532.26	92.06
INV2_400	SAR	–0.0067	0.69077	–1546.40	100.34
INV2_220	SAR	–0.0069	0.70578	–1544.32	98.21

Source: Author's calculations.

Since in the case of substantive spatial correlation the least squares estimator is biased, it is not surprising that the convergence rate also differs for the results based on the other matrices when compared to that of the ordinary least squares results. This is also reflected in the implied "half-distances" to steady state indicating how many years it takes the region to pass half of the distance to the common steady state. Table 3 reveals that the annual convergence rate falls into a certain range of 0.7% to 0.9%. Hence it is concluded that the SAR model specification holds for a range of matrices, and the specification of the matrix does not seem to be a source of non-robustness of the obtained results.

6.1 Spatial Heterogeneity

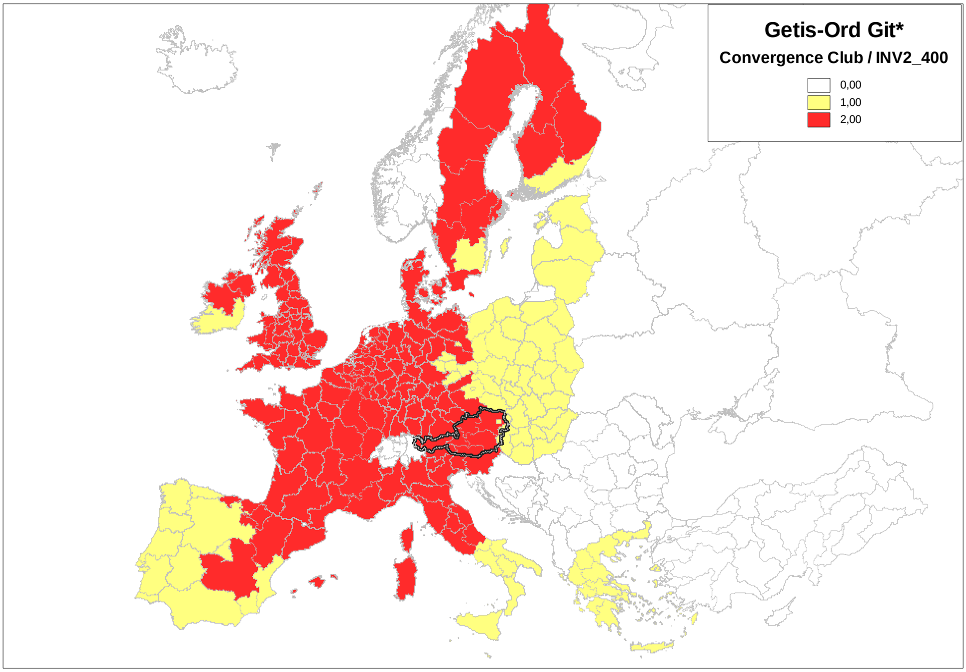
To check for spatial heterogeneity in the data a regime switching model is estimated. The previously calculated z-values of the Gi* statistic are used to identify the clubs, with every positive z-value belonging to club "A", and every negative z-value to club "B" [18].

[16] Computed as $cs = \left(-\log\left(1+\hat{\beta}t\right)/t\right)$.

[17] Computed as $\log(2)/cs$.

[18] Fischer and Stirböck (2004).

*Chart 2: Convergence Clubs Based on Git**



Source: Author's calculations.

The chart shows the two clubs identified by the Getis and Ord distance statistic based on the weight matrix "INV2_400". Slightly different clubs result for one of the other matrices. The classification seems to be quite reasonable: Regions with a relatively low income in 1995 are forming club "A", whereas mainly the old members of the EU-25 form club "B". It should be mentioned that identifying the clubs based on initial income, is only one way and maybe just the most obvious.[19] The diagnostics for spatial dependence of a least squares-estimation of the regime switching model are given in Table A.4 in the appendix section. Based on the "RMLag" test, again the SAR model specification is chosen. Table 4 consists of the estimation results:

[19] Niebuhr et al. (2005) focus on another approach that distinguishes between rural and urban regions based on population density.

Table 4: SAR Regime-Switching Model

Dependent Variable:	$\left(1/t\right)\log\left(y_{i,02}/y_{i,95}\right)$			
	Estimate	*Std. Error*	*t- value*	*Pr(>|t|)*
ρ	0.734773	0.064447	11.401133	0.000000
α_1	0.079586	0.027477	2.896462	0.003774
β_1	–0.007259	0.002808	–2.584266	0.009759
α_2	0.073150	0.022659	3.228424	0.001245
β_2	–0.006255	0.002259	–2.769184	0.005620
AIC / LOG.LIK	–1506.23 / 758.115			
TEST ON STRUCTURAL INSTABILITY FOR 2 REGIMES – CHOW TEST				
	DF	*Value*	*Prob.*	
Chow Test	2	4.531363	0.103759	
STABILITY OF INDIVIDUAL COEFFICIENTS				
	DF	*Value*	*Prob.*	
A_1	1	0.037881	0.845682	
B_2	1	0.083282	0.772898	

Source: Author's calculations.

The coefficient of the spatially lagged dependent variable is again positive and statistically different from zero. The β-coefficients for both clubs are negative pointing to a catching up process. They do not vary significantly from those of the former estimated SAR model based on the whole sample which indicates that we do not have club convergence in the EU-25. This is confirmed by running Chow tests. The tests for structural instability yield the conclusion that the regression as a whole and the individual coefficients do not vary significantly across the two regimes. Summing up I cannot detect a significant variance of the slope coefficient nor the functional relationship across the two regimes, while absolute convergence still holds for both convergence-clubs. In deviation to Fischer and Stirböck (2004) the chosen model specification is a spatial autoregressive regime-switching model with a homoskedastic error term. This difference might be caused by the different time period of analysis as well as by the smaller data set of this study[20]. The implied speed of convergence for club "A" is 0.7449% and for club "B" 0.6396% resulting into half-distances to steady states of approximately 93.05 and 108.38 years. The lack of significance concerning variation of relationship or variance indicates that the EU-25 regions are not characterized by two different clubs. Thus regions do not interact significantly more with a specific sub-group of the sample than with the rest of the EU-25.

[20] Fischer and Stirböck (2004) analyze regional convergence for the period of 1995-2000 including accession countries Bulgaria and Romania.

6.2 Growth Effects of Spillovers

The previous analyses showed that spatial dependencies are evident in the absolute convergence model. A possible conclusion could be the significant influence of regional spillovers on the convergence process. It seems reasonable to assume that spatial interaction of localities is highest within the regions of a country. To which extent does the detected spatial dependence stem from national factors and to which extent from regional spillover effects? National factors (or country effects) are considered as being the fact that regions forming a country share the same economic policies, legislation and institutions.[21] Quah (1996) draws attention to that question by analyzing income distributions. His conclusion is that regional spillovers matter more for the convergence process than national factors, which is in contrast to recent findings (based on a dummy variable approach) by Niebuhr et al. (2005).

For this purpose a special weight matrix "INV1_NAT" is constructed that displays within-country interaction. Here, regions are only allowed being neighbors of each other when they stem from the same country. I have re-estimated the convergence model including the 6 dummy variables starting again with the OLS specification. It is striking that this time all the Lagrange Multiplier specification tests point to the SER model as the specification fitting the data well. This means that, once controlled for national influences incorporated in the model by the specific weight matrix "INV1_NAT", the spatial dependence is of the nuisance form. It can be concluded that spillovers across regions are to a less extent influential to growth than national effects. Thus spatial dependence results only to a small part from spillovers. Table 5 summarizes the model.

Table 5: SER Model

Dependent Variable:	$\left(1/t\right)\log\left(y_{i,02}/y_{i,95}\right)$			
	Estimate	*Std. Error*	*z-value*	*Pr(>|z|)*
α	0.114821	0.023972	4.789819	0.000002
β	–0.006962	0.002535	–2.746861	0.006017
λ	0.694379	0.051955	13.365094	0.000000
Log.Lik.:	782.276000			
AIC:	–1548.550000			

Source: Author's calculations.

The coefficient of convergence speed does not deviate from the previous findings in section 5. Based on this model specification it can be concluded that rather country-specific-effects than spatial spillovers cause the spatial dependence of

[21] Niebuhr et al. (2005).

regional growth. This supports the findings of Niebuhr et al. (2005) and is thus in contrast with those of Quah (1996). The reason can be found in his definition of a neighbor. Quah considered only those regions as neighbors that are adjacent to each other, i.e. neighbors share a common border. Hence, in his sample of 78 regions, only 13 had neighbors belonging to another country. In this context this study differs considerably from Quah's research: The intra-country spatial correlation is compared with regional interaction, incorporated by weight matrices that allow for a multitude of neighbors.

7. Conclusions

This study analyzed absolute income convergence across EU-25 regions. The traditional OLS cross-sectional regression was the initial reference point. Exploratory spatial data analysis as well as several tests showed that spatial autocorrelation is present in the data. Depending on the specified weight matrix, in most instances spatial dependence turned out to be of substantive form pointing to biased OLS estimates. Hence, the already low "OLS-convergence rate" of 1.5% per year cannot be confirmed. In contrast, estimates based on spatial regressions lead to a lower annual rate ranging from 0.7% to 0.9%. Results are fairly robust to a wide range of possible misspecifications. In this study several weight matrices are used that allow for a wide range of spillovers. From an economic point of view the spatial autoregressive model bears important policy implications: It indicates a significant influence of regional spillover effects on convergence – a dynamic surrounding influences a region's growth performance. The framework of the two-club convergence model allows for examinations of distinct sample parts' behavior. The estimated pace of convergence for the two clubs lies again in the range of 0.7% to 0.9% per year. Since convergence rates of the two clubs differ only slightly, evidence for spatial heterogeneity is rather weak. The model showed no variance of the functional relationship across the two regimes. As before, a spatial autoregressive model is the final specification. Thus it can be concluded that the SAR model specification also holds for sub-samples of the data.

Besides, this study gives insights about the source of spatial autocorrelation. Estimating convergence with the intra-country weight matrix, spatial spillovers seem to be less effective. This means, once controlling for country effects, a large part of spatial autocorrelation vanishes. In line with Niebuhr et al. (2005) it might be concluded that most of the spatial autocorrelation is based on differences in national policies, legislation, tax-systems and other country-specific effects. These national factors play a more important role in determining growth than spillovers do.

References

Anselin, Luc (1988): "Spatial Econometrics: Methods and Models", Kluwer Academic Publishers.

Anselin, Luc (1992): "SpaceStat Tutorial – A Workbook using SpaceStat in the Analysis of Spatial Data", http://www.terraseer.com

Anselin, Luc (1995): "SpaceStat V.180 – User's Guide". http://www.terraseer.com

Anselin, Luc (2003): "An Introduction to Spatial Regression Analysis in R ". http://sal.agecon.uiuc.edu/stuff_main.php#tutorials

Anselin, Luc (2005): "Spatial Regression Analysis in R – A Workbook". http://sal.agecon.uiuc.edu/stuff_main.php#tutorials

Baltagi, Badi H. (2002): "Econometrics, 3rd Edition", Springer Verlag.

Barro, Robert J. and Sala-i-Martin, Xavier (1995): "Economic Growth", McGraw-Hill.

Barro, Robert J. (1997): "Determinants of Economic Growth – a Cross-Country Empirical Study", The MIT Press.

Brülhart, Marius, Crozet Matthieu and Koenig, Pamina (2004): "Enlargement and the EU Periphery: The Impact of Changing Market Potential", HWWA Discussion Paper 270.

Doppelhofer, Gernot, Miller, Ronald I. and Sala-i-Martin, Xavier (2004): "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", American Economic Review, 94(4), pp. 813–35.

Crespo-Cuaresma, Jesús, Ritzberger-Grünwald, Doris and Silgoner, Maria Antoinette (2002): "Growth, Convergence and EU Membership", Working Paper 62, Oesterreichische Nationalbank.

Faína, J. Andrés and López-Rodríguez, Jesús (2004): "The Economic Geography of EU Income: Evolution since the Eighties", JEL classification: F12, F14.

Fischer, Manfred and Stirböck, Claudia (2004): "Regional Income Convergence in the Enlarged Europe, 1995–2000: A Spatial Econometric Perspective", ZEW Discussion Paper Nr. 04–42, http://www.ftp.zew.de/pub/zew-docs/dp/dp0442.pdf.

Greene, William (2003): "Econometric Analysis", Prentice Hall.

Hanson, H. Gordon (2004): "Market Potential, Increasing Returns, and Geographic Concentration".

Krugman, Paul (1991): "Geography and Trade", Leuven University Press and the MIT Press.

LeSage, James P. (1999): "The Theory and Practice of Spatial Econometrics". Download: http://www.spatial-econometrics.com.

Levine, Ross and Renelt, David (1992). "A Sensitivity Analysis of Cross-Country Growth Regressions", American Economic Review 82, 4 (September), 942–963.

López-Bazo, Enrique, Vayá, Esther and Artís, Manuel (2004): „Regional Externalities and Growth: Evidence from European Regions“, Journal of Regional Science, Vol. 44, No. 1, pp. 43–73.

Kosfeld, Reinhold, Eckey, Hans-Friedrich and Dreger, Christian (2005): “Regional Convergence in Unified Germany: A Spatial Econometric Perspective”, in: C. Dreger, H. P. Galler (Eds.), Advances in Macroeconometric Modeling, Christian Dreger Baden Baden 2005, pp. 189–214.

Martin, Ron (2001): “EMU versus the Regions? Regional Convergence and Divergence in Euroland”, Journal of Economic Geography 1, pp.51–80.

Mayerhofer, Peter (2004): “Wien in einer Union der 25 – Ökonomische Effekte der EU-Osterweiterung auf die Wiener Stadtwirtschaft”, (thesis), (University of Linz)

Naude, W.A. and Krugell, W.F. (2004): “The Spatial Dimensions of Economic Growth in Africa: The Case of Sub-National Convergence and Divergence in South Africa”, Discussion Paper for the Convergence on Growth, Poverty Reduction and Human Development in Africa.

Niebuhr, Annekatrin (2001): “Convergence and the Effects of Spatial Interaction”, Jahrbuch für Regionalwissenschaft 21, pp.113–133.

Niebuhr, Annekatrin and Bräuninger, Michael (2005): „Convergence, Spatial Interaction and Agglomeration Effects in the EU“; HWWA Discussion Paper Nr. 322.

Portnov, Boris A. and Wellar, Bary (2004): “Development Similarity Based on Proximity: A Case Study of Urban Clusters in Canada”, Papers in Regional Science Vol. 83, pp. 443–465.

Quah, Danny T. (1993): “Empirical Cross-Section Dynamics in Economic Growth”, European Economic Review 37, pp. 426–434.

Quah, Danny T. (1996): “Empirics for Economic Growth and Convergence”, European Economic Review 40, pp. 1353–1375.

Sala-i-Martin, Xavier, (1997a):, “I Just Ran 2 Million Regressions”, American Economic Review.

Sala-i-Martin, Xavier, (1997b):, “I Just Ran Four Million Regressions”, NBER Working Paper 6252.

Software „R“ Development Core Team (2005): R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Software Regiograph 8: http://www.macon.de.

Software SpaceStat: http://www.terraseer.com.

Venables, William N. and Smith, David M. (2004): “An Introduction to R”, Download: http://www.R-project.org

Ying, Long Gen (2000): “Measuring the Spillover Effects: Some Chinese Evidence”, Papers in Regional Science Vol. 79, pp. 75–89.

Appendix

I have constructed two different types of weight matrices, binary contiguity matrices and inverse distance matrixes, both using a distance cut-off point δ. Every region j, with $i \neq j$ that lies within this distance is considered a neighbor of region i and gets assigned a nonzero weight. Distance is calculated using great circle distance based on longitude-latitude data for every NUTS 2 capital city of the EU-25 assuming that the capital reflects a region's centre of economic activity. Formally this is given by:

$$w_{ij} = \begin{cases} 1/d_{ij}^{\alpha} & \text{if } d_{ij} \leq \delta \text{ for } i \neq j \\ 0 & \text{if } i = j \\ 0 & \text{if } d_{ij} > \delta \text{ for } i \neq j \end{cases} \tag{6}$$

The binary contiguity matrices "CON350" and "CON220" use weights $w_{ij}=1/d_{ij}$ with a threshold point at a distance of δ=350 miles (ca. 563 km) and δ =220 miles (ca. 350 km) respectively. Weight matrices based on inverse distances are the matrices "INV1_400" "INV2_400" and "INV2_220". The first one assigns a weight to every region lying in a 400 miles (ca. 643 km) distance band according to the inverse distance $w_{ij}=1/d_{ij}^{\alpha}$ with α=1. The second one resembles the same matrix, only differing in α being 2. The last one, "INV2_220" uses the squared, inverse distances, i.e. $w_{ij}=1/d_{ij}^{\alpha}$ (α=1) for a distance band of 220 miles. The "INV1_NAT" matrix was designed aiming to get insight of intra-country spillovers. It reflects spatial interaction of regions within a country assigning weights $w_{ij}=1/d_{ij}$ for each region $i \neq j$ with i and j from the same country (otherwise the weight is zero).

Table A1:

REGRESSION DIAGNOSTICS for SER Model			
Diagnostics for heteroskedasticity			
Test	DF	Value	Prob.
Breusch-Pagan Test	7	3.451172	0.840370
Spatial B-P test	7	3.451194	0.840368
DIAGNOSTICS FOR SPATIAL DEPENDENCE			
Weights matrix INV_400			
Test	DF	Value	Prob.
Spatial Error dependence	1	61.393642	0.000000
TEST ON COMMON FACTOR HYPOTHESIS			
Likelihood Ratio Test	7	29.363158	0.000124
Wald Test	7	30.032716	0.000094
LAGRANGE MULTIPLIER ON SPATIAL LAG DEPENDENCE			
INV2_400	1	5.961694	0.014620

Source: Author's calculations.

Table A2:

REGRESSION DIAGNOSTICS for SAR Model			
Diagnostics for heteroskedasticity			
Test	DF	Value	Prob.
Breusch-Pagan Test	7	4.979513	0.662464
Spatial B-P test	7	4.979530	0.662461
DIAGNOSTICS FOR SPATIAL DEPENDENCE			
Weights matrix INV2_400			
Test	DF	Value	Prob.
Spatial Lag dependence	1	70.187049	0.000000
LAGRANGE MULTIPLIER ON SPATIAL ERROR DEPENDENCE			
INV2_400	1	0.032505	0.856925

Source: Author's calculations.

Table A3:

REGRESSION DIAGNOSTICS for OLS Model			
Test on normality of errors			
Test	DF	Value	Prob.
Jarque-Bera	2	4.331083	0.114688
Diagnostics for heteroskedasticity			
Test	DF	Value	Prob.
Breusch-Pagan Test	7	3.163053	0.869520

DIAGNOSTICS FOR SPATIAL DEPENDENCE Weights matrix INV2_400			
Test	MI/DF	Value	Prob.
Moran’s I (error)	0.281891	10.288409	0.000000
RLMerr	1	2.700637	0.100308
RLMLag	1	10.572667	0.001148
Lagrange Multiplier (SARMA)	2	105.706119	0.000000

Source: Author’s calculations.

Table A4:

Diagnostics for Spatial Dependence			
Test	DF	Value	Prob.
Robust LM (error)	1	0.097765	0.754529
Robust LM (lag)	1	13.589435	0.000227

Source: Author’s calculations.