

CONVERGENCE IN PER-CAPITA GDP ACROSS EU-NUTS2 REGIONS USING PANEL DATA MODELS EXTENDED TO SPATIAL AUTOCORRELATIONS EFFECTS

G. Piras, G. Arbia

1. INTRODUCTION

The popular regression approach to study regional convergence dynamics of per-capita income is mostly stemming from the neo-classical Solow-Swan (Solow, 1956; Swan, 1956) model of long run growth and from the framework developed by Mankiw *et al.* (1992) and Barro and Sala-i-Martin (1992; 1995). This theoretical framework led to the widely used β -convergence, approach, an empirically testable model that evaluates convergence by examining the inverse relationship between the growth in per-capita income over a definite time span and the income level measured at the beginning of the period.

The β -convergence model, therefore, is not properly a dynamic one, but it is rather based on a static comparison. This is indeed a major drawback under both the theoretical and the applied point of view. In fact the main interest is usually in studying the full dynamics of the convergence process, that is the path followed by per-capita incomes in the various regions over the whole period considered. Nonetheless, very different behaviours may lead to the same evidence in terms of the speed of convergence, this causing problems in the interpretation of the results and leading to unrealistic political decisions and irrational distribution of targeting resources (see Arbia, 2004). These are only some of the reasons why Islam (1995; 2003) suggests to move towards the application of panel data models, and other authors empirically test a continuous time specification for describing regional competitive performances (Arbia and Paelinck, 2003).

In addition, some recent literature (Rey and Montouri, 1999; Ertur and Koch, 2007, Fingleton, 2003; LeGallo *et al.*, 2003; Lopez-Bazo *et al.*, 1999, 2004; Rey and Janikas, 2005; Piras *et al.*, 2006; Abreu *et al.*, 2005) has highlighted the importance of considering spatial effects in the study of regional economic convergence dynamics. Furthermore, the New Economic Geography models emphasize the importance of spatial spillovers between economies, which should be formally integrated in convergence models (Fujita *et al.*, 1999). Thus, the spatial econometric application to the study of regional economic convergence has limited its atten-

tion to the classical specification of the cross-sectional spatial lag and error models. In other words, even though it has been proved that space matters in the explanation of competitive regional economic behaviours, again no emphasis has been placed on the specification of a truly dynamic model.

The bulk of the present paper is very much in line with these considerations. Indeed we do not pretend either to provide further insight to the development of a proper economic model to tackle regional convergence related problems or to add more evidence on the economic and policy related aspects of the unequal regional income distribution. Instead, our main interest is to enrich the methodological debate. In fact, we propose the estimation of convergence in per-capita GDP across European regions by making use of spatial panel data models both including a spatially lagged dependent variable and a spatial error specification (Anselin, 1988; Elhorst, 2001; 2003; 2005). The main idea developed is the advantage produced by the consideration of spatial dependence within a fixed-effect approach. Indeed, the control for fixed-effects allows us to be more confident that the spatial dependence coefficient may only capture regional interaction effects and not those due to omitted variables problems. The innovative aspect concerns the fact that spatial dependence has been formally considered in a panel data context only recently (Elhorst 2001; 2003) and that such a framework has never been employed to model regional convergence dynamics. In the present paper spatial effects are introduced both by adding a spatial lag of the dependent variable among the explanatory variables and by modelling the error term with a particular spatial structure. In other words, we want to prove that a spatial panel data model provides a suitable choice for the estimation of a regional convergence model for at least two reasons. First of all, because it explicitly accounts for the effect of space, which is a necessary condition to prevent biased estimates of the convergence coefficient as it has been recently addressed in Elhorst *et al.* (2006). Secondly, the inclusion of regional specific fixed effects in the model reflects the possible presence of omitted variables with a spatial dimension, which reflects differences in initial conditions (*i.e.* technology level, quality of institution, geographical and climate conditions).

The empirical exercise relates to the estimation of the long-run convergence process of per capita income in Europe (1977-2002) based on a spatial level (the NUTS2 EU regions) fine enough to properly model spatial effects (in terms of regional spillovers).

We organize the paper under the following headings: Section 2 is devoted to a detailed description of the data and the investigation of their spatial properties. Section 3 introduces the spatial panel approach to regional convergence and discusses the outcomes of the empirical analysis. A comparison of the results obtained with the main evidence available in the literature is also provided. Section 4 contains some discussion on how to test for spatial dependence in space-time series data. Some final remarks and indication for further researches conclude the paper.

2. PRELIMINARY DATA ANALYSIS

Spatial data availability is one of the greatest problems in the European context, although many progresses have been made in recent time by EURO-STAT. Thus, data availability remains insufficient and in many cases it is very difficult to gather harmonized data sets for consistent regional comparisons.

Data on the per capita GDP (millions of euro 1995) in logarithms used in the present application have been extracted from the Cambridge Econometrics European Regional database that is an extensive processing of the EURO-STAT REGIO database. Our choice is motivated by the fact that data from the REGIO presents many problems for the users. Firstly, the quality of the data is always variable across countries and time. Moreover, it is likely to have missing time observations at the NUTS2 level of spatial aggregation, and quantities are only expressed in current prices. In the Cambridge Econometrics dataset some rules have been followed to fill existing gaps and to extend the series to more recent years using national data when available¹. The length of the time series dimension is very important in evaluating growth dynamics, since convergence is a long-run process, and the use of short series may produce biased results². We include 125 regions of 10 European Countries: Belgium, Denmark, France, Germany, Luxembourg, Italy, Netherlands, Portugal, Spain and United Kingdom. Our sample ranges over the period from 1977 to 2002³.

In order to preliminarily test for global spatial autocorrelation in per-capita GDP in logarithm, we have calculated the Moran's I index for each year and its significance level (the value of the standard normal distribution and the relative p -value). The Moran's I index in matrix notation assumes the following form:

$$I_t(k) = \left(\frac{n}{S_0} \right) \mathbf{z}'_t \mathbf{W} \mathbf{z}_t (\mathbf{z}'_t \mathbf{z}_t)^{-1} \quad (1)$$

where \mathbf{z}_t is the vector of the n spatial observations for year t in deviation from the mean, and \mathbf{W} is a spatial weight matrix (Cliff and Ord, 1981). $S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$, with $i \neq j$, is the total sum of the weights. In spatial analysis, the weights reflect the degree of connectedness among spatial units and are generally related to geographical proximity. The value of the Moran's I statistics helps in determining whether regions that are neighbor are more similar to one another than would be expected under spatial randomness. The sign of the statistics discriminates between positive and negative spatial autocorrelation. As it can be observed from

¹ See the European Regional Prospect developed by Cambridge Econometrics for greater details on data treatment.

² We are well aware that 25 years might still not be considered a sufficiently long time span to evaluate convergence dynamics in Europe. Thus, it is somehow impossible to obtain a longer time series dimension for NUTS-2 data. In any case, we believe that the reduced time-span does not weaken our conclusions as it will be made clearer in the comments to the empirical part of the paper.

³ The great part of the works in the literature use data drawn from the REGIO dataset in empirical studies: Quah, 1996; Baumont *et al.*, 2002; Arbia and Paelink, 2003; among others.

Equation (1), the index can also be interpreted as the value of the coefficient from an OLS regression of \mathbf{z}_t over $\mathbf{W}\mathbf{z}_t$.

For the calculation of the index, as in all the following elaborations performed in the present work, we make use of a spatial weight matrix based on the contiguity criterion (the element of the matrix is equal to one if the two regions share a common border, and zero otherwise). Table 1 reports the computation of the Moran's I statistics for each year in our sample.

TABLE 1
Moran's I computed for the period 1978-2002 on the annual growth rate of per-capita income

YEAR	Moran's I	Z-value	Prob
1978	0.565	9.064	0.000
1979	0.502	8.069	0.000
1980	0.281	4.567	0.000
1981	0.160	2.657	0.000
1982	0.317	5.146	0.000
1983	0.172	2.855	0.000
1984	0.155	2.579	0.000
1985	0.148	2.478	0.000
1986	0.023	0.502	0.000
1987	0.259	4.225	0.000
1988	0.194	3.202	0.000
1989	0.207	3.406	0.000
1990	0.393	6.342	0.000
1991	0.797	12.715	0.000
1992	0.195	3.212	0.000
1993	0.474	7.617	0.000
1994	0.395	6.379	0.000
1995	0.273	4.453	0.000
1996	0.316	5.120	0.000
1997	0.234	3.825	0.000
1998	0.073	1.283	0.000
1999	0.185	3.061	0.000
2000	0.203	3.334	0.000
2001	0.437	7.045	0.000
2002	0.337	5.461	0.000

The results show that the Moran's I index is fairly stable across time. In fact, it assumes positive values during the entire sample period (1977-2002) this indicating the presence of positive global spatial autocorrelation in the sample. Inference is based on a permutation approach⁴ (10000 permutations), in which a reference distribution is calculated for spatially random simulated datasets. The Z-values reported in column two of Table 1 suggest that the null hypothesis of no spatial autocorrelation should be rejected (with the exception of 1986 and 1998) and that the estimation procedures have to be corrected in order to take over the presence of spatial autocorrelation. Finally, the results are also robust to different choices of the spatial weight matrix. In fact, we calculated the Moran's I using different specification of the weights⁵ but obtaining very similar evidence.

⁴ See Cliff and Ord, 1981

⁵ In particular, we have considered two more spatial weight matrices: inverse square distance matrix, and a binary spatial weight matrix with a simple distance-based critical cut-off. Results are available from the authors upon request.

3. THE SPATIAL PANEL DATA APPROACH TO ECONOMIC CONVERGENCE

As pointed out in the introduction, the main evidence available in the spatial econometrics regression approach to regional economic convergence concerns cross-sectional estimations. Abreu *et al.* (2005) provide an excellent survey of the empirical literature on growth and convergence that has addressed the importance of spatial effects.

In this section we will consider the use of panel data in testing the hypothesis of economic convergence among European regions. In order to correct the bias generated by omitted variables and heterogeneity in the classical cross-sectional regression, Islam (2003) suggested the use of panel data models which allow for technological differences across regions, or at least the unobservable and unmeasurable part of these differences, by modelling the regional specific effect. The inclusion of these effects also constitutes a specific advantage of spatial panel data with respect to cross-sectional spatial models. In fact, as it is well known, there are no tests to make a specific distinction between spatial dependence and heterogeneity⁶. As an example, the existence of a cluster can be originated either by spillover effects or from similarities between regions in some variables affecting growth⁷. As pointed out in Abreu *et al.* (2005), it is easier to differentiate between the two effects in a panel data perspective simply because omitted variables potentially presenting a spatial dimension can be captured by region-specific fixed effects.

In the remaining of the section we will present results from the estimation of three different models. We will take on the empirical analysis by estimating first a panel fixed-effect regression. We will then move to the two spatial specifications, namely the “spatial error” and the “spatial lag” panel data models.

A panel data version of the growth equation not including explicitly any form of spatial effects can be expressed in the following way:

$$\ln\left(\frac{y_{t,i}}{y_{t-1,i}}\right) = a_i + b \ln(y_{t-1,i}) + e_{t,i} \quad (2)$$

with $i = 1, \dots, N$ denoting regions, and $t = 1, \dots, T$, denoting time periods. The dependent variable $\ln\left(\frac{y_{t,i}}{y_{t-1,i}}\right)$ is the annual growth rate of the per capita income, and $\ln(y_{t-1,i})$ is the value of the per capita income at time $t-1$; a_i 's and β are parameters to be estimated. Let us stress again that a_i are time invariant and take over the region specific effects not explicitly included in the regression equation, or in other words, they reflect all those omitted variables that influence the growth process.

⁶ Anselin (2001) remarks that spatial dependence and spatial heterogeneity can be observationally equivalent in single cross-section models.

⁷ Variables such as geographical or physical characteristics of the regions, or those capturing differences in the level of institutions or technology, are all potential candidates for generating a cluster.

Equation (2) represents a fixed-effect panel data model, in that the a_i 's are fixed parameters estimated throughout a Least Square Dummy Variables procedure (see Baltagi, 2001). Following Islam (1995), many contributions have been produced concerning the problem of estimating the speed of convergence among regions using panel data models, also including various variants of the basic fixed-effect model (e.g. Canova and Marcet, 1995; Durlauf and Quah, 1999).

It should probably be mentioned here that we are well aware of some potential problems that may arise either from the definition of the dependent variable or from the time span considered. There is indeed a vast debate in the literature on whether or not annual growth rate are appropriate to study convergence dynamics. In fact, it can be objected that such a short time period tends to capture random adjustment towards the trend rather than long-term convergence. Furthermore convergence being a long-term phenomenon, the wider the time span the higher the probability is to be really tracing the true convergence path of the regional economies considered. In other words there is a sort of trade-off between this two aspects because a definition of the growth rate over a longer time period would mean less time observations in the analysis. For this reason, we decided to define the dependent variable in terms of annual growth rate even considering the difficulties to obtain long time series for data expressed at the regional level. However let us remark that our foremost objective is to prove that previous studies at regional level for the EU based on the simple β -convergence model are biased because they neglect both the fixed and the spatial effects. On the other hand, studies on panel data are biased because no spatial autocorrelation effects are considered. In this context, although we recognize the fact that convergence is a process that is likely to occur in the very long run (particularly within the EU), we are considering a time span that makes our results comparable with most of the empirical evidence available in regional convergence literature at the EU level.

A further consideration concerns the interpretation of the β -coefficient. Roughly speaking, panel data estimates produce evidences that are closer to the idea of conditional convergence or convergence given the differences in the region specific characteristic controlled for in the fixed effect term⁸.

Table 2 reports the results of the estimation of a fixed-effect panel data model (Equation (1)) based on the sample of the 125 EU NUTS2 regions described in the previous section. The dependent variable is the annual growth rate of the per capita GDP, and the only explanatory variable is the level of the income at the beginning of each period. In the most general specification, there are 125 different groups, each one corresponding to one of the European regions, and 26 observations for each group (1977-2002). Then, the total number of observations is 3250 for the entire sample. This number can be considered large enough to guarantee significant conclusions on our estimated model.

The value of the coefficient of the initial per capita income variable of Euro-

⁸ An analogous consideration is reported in Arbia and Paelinck (2003) for the case of an implicit specification of a spatial model. In that context the authors explicitly refers to spatial conditional convergence.

pean regions calculated over the entire time period is -0.019. The negative value assumed confirms then the hypothesis of convergence among the European regions considered in our sample. The value of the growth rate coefficient beta, that we have found using the fixed-effect estimator is smaller than those that can be found estimating a simple unconditional convergence model⁹. When the full dynamics of the phenomenon is accounted for, the speed of convergence for our specific sample of 125 regions is proved to be lower than that usually estimated using a cross section of observations. The approach based on Equation (1) is partial in that the presence of spatial dependence is not comprised in the previous specification.

TABLE 2
Convergence of per capita income in the 125 European regions (1977-2002) Estimation of the fixed-effect Model Equation (2) (numbers into brackets refer to the p-values)

Variable	Fixed-effect Model			
	Coefficient	Standard Error	T	P> t
Log of income	-0.019	0.002	-6.800	0.000
Constant	0.214	0.028	7.560	0.000
Sigma _z	0.018			
Sigma _ε	0.028			
Fraction of variance due to α _i	0.294			
F-test that all α _i =0	2.000 (0.000)			
	Goodness of fit			
R-square:	Within	0.014		
	Between	0.026		
	Overall	0.003		
Observation	3250			
Number of Groups	125			
Observation per Group	26			
Corr (α _i , X _b)	-0.903			

Many motivations can be brought to support the existence of spatial interrelations among regional economies in Europe and, hence, to theoretically justify the use of a spatial econometric approach. In fact, the neoclassical growth model that constitutes the basis of the growth regressions has been developed starting from the strong hypothesis of close economies. However, the unification process started some years ago in Europe (and that is still going on in these years) makes this last assumption very unrealistic. Indeed barriers to trade, to person and to factor flows have become considerably low. To better understand the implications of the openness hypothesis on the convergence process it is helpful considering the role of factor mobility, trade relations and technological diffusion (or knowledge spillovers). Factor mobility implies free movements of labour and capital in response to differentials in remuneration rates, which in turn depends on the relative factor abundance. Thus, capital will tend to flow from regions with a higher capital-labour ratio to those with a lower one, while labour movements will have opposite direction. As a consequence, regional economies with lower capital-labour ratios will show higher per capita growth rates.

⁹ The results from the cross-sectional model are not reported in the paper but are available upon request from the authors.

Credit market imperfections, finite lifetimes and adjustment costs for migration and investments will turn the speed of convergence to the steady-state to be higher than in the close economy case (Barro and Sala-i-Martin, 1995). The same outcome might be obtained by simply recovering the neoclassical growth model with the hypothesis of free trade relations rather than factor mobility producing the effect of speeding up the convergence rate.

A further pushing factor for the catching up process of poorer regions can be found in the practice of technology diffusion (or knowledge spillovers). In the presence of disparities in regional technological attainment, interregional trade can promote technological diffusion when progress is incorporated in traded goods (Grossman and Helpman, 1991; Barro and Sala-i-Martin, 1997). A broader interpretation of knowledge spillovers effects refers to positive knowledge external effects produced by firms at a particular location and affecting the production processes of firms located elsewhere. To be more precise, in this context it would be necessary to make a distinction between local and global geographic spillovers. When only local spillovers are present, the production process of firms located in one region only benefits from knowledge accumulation in that same region this most likely leading to regional divergence. On the contrary, a situation in which knowledge accumulation in one region improves the productivity of all firms wherever they are located is referred to as global spillovers effect. Hence global geographical spillovers are expected to favor regional convergence (Martin and Ottaviano, 1999; Kubo, 1995).

The best way to test the openness hypothesis discussed so far would consist of directly including interregional flows of labour, capital and technology in the growth regression model. It is quite clear, however, that such a direct approach is limited by data availability, especially with regards to capital and technology flows.

An alternative, but indirect, way to incorporate this typology of spatial effects is to start from the classical fixed-effect panel data model and account for spatial dependence by including a spatially lagged term of the dependent variable¹⁰. The previous model will then assume the following form:

$$\ln\left(\frac{y_{t,i}}{y_{t-1,i}}\right) = \alpha_i + \rho \sum_{j=1}^n w_{i,j} \ln\left(\frac{y_{t,i}}{y_{t-1,i}}\right) + \beta \ln(y_{t-1,i}) + \varepsilon_{i,t} \quad (3)$$

with $w_{i,j} \in \mathbf{W}$ a weight matrix as discussed in Section 2, ρ the spatial autoregressive coefficient, and $\varepsilon_{i,t}$ a zero mean error term assumed to be independently distributed under the hypothesis that all spatial dependence effects are captured by the spatially lagged variable term. This model takes the name of fixed-effect spatial

¹⁰ Elhorst (2003) mentions two problems that may arise when panel data models have a locational component. The first concerns spatial heterogeneity, which can be defined as parameters that may not be homogeneous throughout the data set, but varying with location. The second is represented by the spatial dependence that may exist between observations at each point in time. In the present work, we consider only the second aspect referring to a fixed-effect panel data model specification extended to spatial error correlation and leaving the treatment of spatial heterogeneity to a further development.

lag model and represents the extension of Model (2) to the case of panel data (Elhorst, 2001; 2003). Equation (3) is estimated via Maximum Likelihood as suggested by Elhorst (2003).

A second alternative to incorporate the spatial effects is to extend Equation (2) to the case in hand by leaving unchanged the systematic component and to model the error term by assuming, for instance:

$$\ln \left[\frac{y_{t,i}}{y_{t-1,i}} \right] = \alpha_i + \beta \ln(y_{t-1,i}) + \varepsilon_{t,i}$$

and (4)

$$\varepsilon_{t,i} = \delta \sum_{j=1}^x w_{i,j} \varepsilon_{t,j} + \eta_i$$

where $w_{i,j} \in \mathbf{W}$ and \mathbf{W} is again the spatial weight matrix, δ is the spatial autocorrelation coefficient of the error term, and the η_i are assumed to be normally distributed with zero mean, constant variance and a distribution independent from the explanatory variable. Such a model is called *fixed-effect spatial error model*. Again the parameters may be estimated by using maximum likelihood¹¹.

TABLE 3
*Convergence of per capita income in the 125 European regions (1977-2002) -
 Estimation of the fixed-effect Spatial Lag Model
 Equation (3) - (numbers into brackets refer to the p-values)*

Spatial Lag fixed-effect Model			
Variable	Coefficient	Asymptotic t-stat	P> t
Log of income	-0.010	-4.743	0.000
Spatially lagged growth rate	0.686	49.690	0.000
	Goodness of fit		
R-squared	0.527		
Sigma squared	0.000		
Log-likelihood	7597.569		
Observation	3125		
Number of variables	1		

TABLE 4
*Convergence of per capita income in the 125 European regions (1977-2002) -
 Estimation of the fixed-effect Spatial Error Model -
 Equation (4) - (numbers into brackets refer to the p-values)*

Spatial Error fixed-effect Model			
Variable	Coefficient	Asymptotic t-stat	P> t
Log of income	-0.033	-7.620	0.000
Spatially lagged growth rate	0.699	54.138	0.000
	Goodness of fit		
R-squared	0.535		
Sigma squared	0.000		
Log-likelihood	7615.201		
Observation	3125		
Number of variables	1		

¹¹ For a discussion of the asymptotic properties of the maximum likelihood estimation see Elhorst (2003).

In Tables 3 and 4 are reported the main results of the empirical analysis performed using Equations (3) and (4). As we said before, these specifications allow a more careful treatment of the unobserved factors that influence growth, and also a reduction of the bias that derives from not properly addressing the spatial dependence present in the data. Table 3 reports the results of the estimation of the fixed-effect spatial lag model in which a weighted average of the growth rate of the neighbour economies is added among the independent variables. The value of the estimated coefficient of the initial per-capita GDP level turns to be -0.010 for the sample period 1977-2202. The presence of the α_i 's parameters isolates the effect of the omitted variables, in terms of the different structural characteristics of the regional economies. The spatial autocorrelation coefficient, that is highly statistically significant, captures the effect of spatial autocorrelation in the form of any systematic pattern in the spatial distribution of the per-capita GDP growth rates.

The simultaneous presence of these two different factors produces, as it was expected, a value of the beta coefficient that is lower than in the fixed-effect model. From an economic point of view this result confirms the evidence obtained with the cross sectional estimates. The reduction of the coefficient of the model due to the inclusion of the spatial lag term (and, hence, the higher speed of convergence) confirms the positive effect of factor mobility, trade relationships, and the presence of spillovers on regional convergence.

A different consideration has to be made for the fixed-effect spatial error model (reported in Table 3). The value of the β coefficients is greater than that obtained in the classical fixed-effect model estimate even if it is still lower than those obtained with the unconditional β -convergence cross-sectional model and its spatially corrected versions. In this specification it is not possible to conclude that all the effect of omitted variable has been captured by the fixed-effect coefficients. Part of the explanatory power of the model can be still not explicitly considered, and, in particular, contained in the spatial autocorrelation coefficient appearing in the error term structure.

For these reasons, a spatial lag specification appears to be a more proper choice in studying convergence among EU regions because it clearly models the intuition of the spatial interaction effects among the different regions in the sample.

As a general conclusion we have to observe that our findings, based on a panel data specification of the convergence model, are in line with analogous studies found in the existing literature. To reinforce this conclusion in the reminder of this section we will review part of the extensive literature available on standard spatial econometric studies to regional convergence¹². However, it is important to

¹² Of course there exists many other interesting studies at the European regional level stemming from very different approaches (Pittau and Zelli, 2006; Pittau, 2005; Corrado *et al.*, 2006 among many others). However, we will limit our attention to all those works that exploit methodologies similar to the one discussed in the paper. This is because we believe it does not make that much sense to compare results from extremely different methodologies. Despite the differences in the methodologies, most of the evidence available in the literature confirms at various levels the hypothesis of convergence among the regional economies of Europe. However, the definition of convergence arising from those works might be slightly different from the one implied by the neo-classical formulation of the problem.

clarify at the beginning that the comparison of the different results is difficult and any substantive conclusion should be made with a great caution. In fact, as it was proved in Piras *et al.*, (2006), the results are extraordinarily sensitive to many different aspects such as the diverse methodology applied, the length of the time span considered or the sample size in terms of the number of regions that enters the analysis. Also the comparison between the coefficients from a spatial model with those from an a-spatial specification is not trivial (Abreu *et al.*, 2005).

As a general comment, it is remarkable that most of the evidence available in this stream of the literature predicts convergence according to the neo-classical definition.

Le Gallo *et al.*, (2003) analyze the consequences of spatial dependence on regional economic growth and convergence for a sample of 138 European Regions over the period 1980-1995. The specification adopted reveals a robust evidence of a substantial convergence process among the regions in their sample. They also find evidence that the two spatial specifications produce contrasting indications that are in line with our findings. In fact, also in their analysis, the spatial error model leads to a faster speed of convergence with respect to the a-spatial model, while the spatial lag specification supports a slower convergence process.

Badinger *et al.*, (2004) estimate the speed of income convergence for a sample of 196 regions over the period 1985-1999 by proposing a two-step procedure. After spatial filtering the data, they apply the standard GMM estimation for dynamic panel. Employing this procedure they obtain a speed of convergence about three times bigger than the one usually found in convergence studies. Baumont *et al.*, (2003) controls for both spatial dependence and heterogeneity by estimating a spatial regimes spatial error model over two distinct clubs identified as northern and southern Europe starting from a sample of 138 regions. The speed of convergence in the southern regions is well above the value for the northern subsample. Many other studies on convergence clubs (detected by mean of Explanatory Spatial Data Analysis) also certify the presence of a regional convergence process either between regions of the European Union or within regions of a single member state (Nieburh, 2001; Le Gallo and Dall'Erba, 2003; Dall'Erba *et al.* 2008; Ramajo *et al.* 2003; Carrington, 2003; Ramirez *et al.*, 2003).

4. THE CALCULATION OF THE MORAN'S I INDEX IN PANEL REGRESSIONS

Another interesting point we discuss in the present paper concerns the testing of the hypothesis of independence among residuals in a spatial panel data model. There are two obvious (although partial) approaches that can be followed. The first one concerns the test of spatial autocorrelation in the T different moments of time by using the classical Moran's I or LM tests (Anselin, 1988). The second approach refers to the test of temporal autocorrelation in the n locations considered and thus involves the computation of n distinct Durbin-Watson tests (Davidson and MacKinnon, 1993). A possible way of building a general procedure to test simultaneously the two features could be obtained in the following

way. Let us start from the familiar Moran's I expression that is more general and admits the Durbin-Watson procedure as a particular case (see e. g. Arbia, 2006). Let us rewrite the general expression in the following form:

$$I = b(\hat{\mathbf{e}}'\hat{\mathbf{e}})^{-1}(\hat{\mathbf{e}}'\mathbf{W}\hat{\mathbf{e}}) \quad (5)$$

where $\hat{\mathbf{e}}$ are the regression residuals, and b a normalizing factor such that

$$b = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \text{ with } w_{ij} \in \mathbf{W}.$$

In the case of a cross-section regression, \mathbf{W} is a $n \times n$ matrix, where n corresponds to the number of the spatial units considered. Conversely in the case of a panel regression the vector of residuals has a different dimension with respect to the spatial weight matrix. In this respect it is sufficient to build the weight matrix in a block diagonal form with the traditional spatial weight matrix repeated T times on the main diagonal. Formally the new space-time connectivity matrix $\mathbf{\Omega}$ can be expressed as

$$\mathbf{\Omega} = \begin{pmatrix} \mathbf{W} & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{W} & & & \dots \\ \dots & & \mathbf{W} & & \dots \\ \dots & & & \dots & \\ \mathbf{0} & & & & \mathbf{W} \end{pmatrix} \quad (6)$$

where \mathbf{W} are $n \times n$ connectivity matrices. The dimension of the $\mathbf{\Omega}$ matrix is now $nT \times nT$, as each block has dimension $n \times n$, and the number of blocks corresponds to the number of time periods. The computation of the Moran's I follows straightforwardly by replacing the \mathbf{W} matrix in Equation (5) with the $\mathbf{\Omega}$ matrix of Equation (6) and stacking the $n \times T$ matrix of space-time residuals in one single $nT \times 1$ column vector.

The asymptotic distribution for the Moran statistics, derived under the null hypothesis of no spatial dependence, is still normal as in the classical (purely spatial) formulation. However the expected value and the variance need to be derived explicitly in this situation. The previous expression accounts for spatial correlation in each time period. In those cases where the model considers both spatial and serial autocorrelation, the structure of the spatial weights matrix is different. In particular, the blocks above and below the main diagonal are also non-zero and the number of diagonals that are different from zero depends on the time periods considered in the serial autocorrelation term. For instance by limiting ourselves to temporal dependence at the first lag, we have:

$$\Omega = \begin{pmatrix} \mathbf{W} & \mathbf{W} & \dots & \dots & \mathbf{0} \\ \mathbf{W} & \mathbf{W} & \mathbf{W} & \mathbf{0} & \dots \\ \dots & \mathbf{W} & \mathbf{W} & \mathbf{W} & \dots \\ \dots & & & \dots & \mathbf{0} \\ \mathbf{0} & & \mathbf{0} & \mathbf{W} & \mathbf{W} \end{pmatrix} \quad (7)$$

that allows for simultaneous spatial and temporal (although only back to the first lag) correlation amongst residuals to be detected.

The derivation of the finite-sample and asymptotic properties of the space-time Moran statistics represents a field for further researches and is not undertaken here since it goes beyond the scope of the present paper. Along similar lines different approaches have been proposed by Anselin *et al.* (2008) for the LM test in spatial lag and spatial error panel data models and by Pesaran (2004) for a diagnostic test for unspecified spatial dependence in panels.

5. CONCLUSIONS AND FUTURE RESEARCH GUIDELINES

In the present paper we have considered the problem of regional economic convergence among European regions. Many works in literature study convergence by specifying fixed-effect models or cross-country regressions. Our investigation starts from the observation that both techniques impose strong a-priori restrictions on the model parameters. From one side, cross-sectional methods do not consider heterogeneity (unless considering models of “club convergence”), on the other hand the fixed-effect panel data approach incorporates heterogeneity only in the different intercepts for each region: all the differences in growth rates depend only on the different initial conditions of the spatial unit considered. In addition, both approaches neglect aspects connected with spatial dependence among regions. The methodology used in the present paper allows us to extend the traditional models by considering a specific treatment of both unexplained heterogeneity and spatial dependence.

An important result that we have obtained consists in the fact that, taking into account for the spatial dependence among spatial units in the form of a spatially lagged dependent variable, considerably improves the estimated values of the speed of convergence among the European regions. This result shows that the value of the fixed-effect coefficient is affected by the presence in the model of the positive effect of spatial dependence. The present paper may be considered as a point of departure for some future researches in regional convergence. An interesting possible advance could be based on the framework of dynamic panel data models extended to spatial error autocorrelation or to a spatially lagged dependent variable (Elhorst, 2001). Moreover, the use of semi-parametric techniques to allow the coefficients to vary among regions could be considered. The advantage of considering possible non-linearities within a spatial panel data

framework consists on the identification of different slopes together with systematic time-invariant regional effects. Thus, a larger flexibility would be guaranteed by this specification because regions may differ both in initial conditions and in their own growth path.

REAL, UIUC E DASTA,
Università di Chieti

GIANFRANCO PIRAS

DASTA,
Università di Chieti

GIUSEPPE ARBIA

ACKNOWLEDGMENT

The final version of the present paper was written while Gianfranco Piras was visiting the Regional Economic Application Laboratory at the University of Illinois at Urbana Champaign. A preliminary version was presented at the REAL seminar series, at the 44th European Congress of the European Regional Science Association, University of Porto, Porto, Portugal and at the 44th Annual Meeting of the Western Regional Sciences Association, San Diego, California. Comments from all the participants helped improving the paper. A special mention is due to Geoffrey Hewings for providing very useful suggestions on earlier versions of the paper.

REFERENCES

- M. ABREU, H.L.F. DE GROOT, R.J.G.M. FLORAX, (2005) *Space and growth: a survey of empirical evidence and methods*, "Région et Développement", 2005-21, pp. 13-44.
- L. ANSELIN, (1988) *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- L. ANSELIN, (2001) *Spatial Econometrics*, in B.H. BALTAGI (eds.), "A Companion to Theoretical Econometrics", Blackwell Publisher, Oxford.
- L. ANSELIN, J. LE GALLO, H. JAYET, (2008) *Spatial Panel Econometrics*, in L. MATYAS and P. SEVESTRE (eds.), "The Econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice" (3rd Edition), Kluwer, Dordrecht, pp. 625-660.
- G. ARBIA, (2004) *Alternative approaches to regional convergence exploiting both spatial and temporal information*, "Revista Estudios de Economía Aplicada", 22, pp. 1-18.
- G. ARBIA, (2006) *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence*, Springer Verlag, Berlin.
- G. ARBIA, J.H.P. PAELINCK, (2003) *Spatial econometric modelling of Regional Convergence in Continuous Time*, "International Regional Science Review", 26(3), pp. 342-362.
- H. BADINGER, W. MÜLLER, G. TONDL, (2004) *Regional Convergence in the European Union (1985-1999): A Spatial Dynamic Panel Analysis*, "Regional Studies", 38, pp. 241-253.
- B.H. BALTAGI, (2001) *Econometric analysis of panel data (2nd Edition)*, Wiley, Chichester, UK.
- R.J. BARRO, X. SALA-I-MARTIN, (1992) *Convergence*, "Journal of Political Economy", 100, pp. 223-251.
- R.J. BARRO, X. SALA-I-MARTIN, (1995) *Economic Growth*, McGraw Hill, New York.
- R.J. BARRO, X. SALA-I-MARTIN, (1997) *Technological diffusion, convergence and growth*, "Journal of Economic Growth", 2, pp. 1-16.

- C. BAUMONT, C. ERTUR, J. LE GALLO, (2003) *Spatial Convergence Clubs and the European Regional Growth Process, 1980-1995*, in B. FINGLETON (eds.), "European Regional Growth", Springer (Advances in Spatial Sciences).
- F. CANOVA, A. MARCET, (1995) *The poor stay poor: non-convergence across countries and regions*, Discussion Paper No 1265, Center for Economic Policy Research (CEPR), London.
- A. CARRINGTON, (2003) *A Divided Europe? Regional Convergence and Neighbourhood Effects*, "Kyklos", 56, pp. 381-394.
- A.D. CLIFF, J.K. ORD, (1981) *Spatial Processes: Models and Applications*, Pion, London.
- L. CORRADO, R. MARTIN, M. WEEKS, (2005) *Identifying and Interpreting Regional Convergence Clusters across Europe*, "The Economic Journal", 115(502), pp. C133-C160.
- S. DALL'ERBA, M. PERCOCO, G. PIRAS, (2008) *The European regional growth process revisited: increasing returns and spatial dynamic setting*, "Spatial Economic Analysis" (forthcoming).
- R. DAVIDSON, J.M. MACKINNON, (1993) *Estimation and inference in econometrics*, Oxford University Press, Oxford.
- S.N. DURLAUF, D. QUAH, (1999) *The new Empirics of Economic Growth*, in J. TAYLOR and M. WOODFORD (eds.), "Handbooks of Macroeconomics", Vol. 1A, North-Holland, Amsterdam.
- J.P. ELHORST, (2001) *Dynamic Models in Space and Time*, "Geographical Analysis", 33(2), pp. 119-140.
- J.P. ELHORST, (2003) *Specification and Estimation of Spatial Panel Data Models*, "International Regional Sciences Review", 26(3) pp. 244-268.
- J.P. ELHORST, (2005) *Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels*, "Geographical Analysis", 37, pp. 85-106.
- J.P. ELHORST, G. PIRAS, G. ARBIA, (2006) *Growth and Convergence in a Multi-Regional Model with Space-Time Dynamics*, paper presented at the First Workshop on Spatial Econometrics and Spatial Statistics, Rome, Italy.
- C. ERTUR, W. KOCH, (2007) *Technological interdependence and spatial externalities: theory and evidence*, "Journal of Applied Econometrics", 22(6) pp. 1033-1062.
- European Regional Prospect*, Cambridge Econometric Press.
- B. FINGLETON, (2003) *European regional growth*, Springer-Verlag, Berlin.
- M. FUJITA, P. KRUGMAN, A.J. VENABLES, (1999) *The Spatial Economy*, MIT Press, Cambridge.
- G.M. GROSSMAN, E. HELPMAN, (1991) *Trade, knowledge spillovers and growth*, "European Economic Review", 35(3), pp. 517-526.
- N. ISLAM, (1995) *Growth empirics: A panel data approach*, "Quarterly Journal of Economics", 4, pp. 1127-1170.
- N. ISLAM, (2003) *What have we learnt from the convergence debate?*, "Journal of Economic Surveys", 17(3), pp. 309-362.
- Y. KUBO, (1995) *Scale economies, regional externalities and the possibility of uneven regional development*, "Journal of Regional Sciences", 35, pp. 318-328.
- J. LE GALLO, S. DALL'ERBA, (2003) *Spatial Econometric Analysis of the Evolution of the European Convergence Process, 1980-1999*, Washington University, Economics Working Paper Archive at WUSTL, no. 0311001, Washington DC.
- J. LE GALLO, C. ERTUR, C. BAUMONT, (2003) *A spatial Econometric Analysis of Convergence Across European Regions, 1980-1995*, in B. FINGLETON (eds.), "European Regional Growth", Springer (Advances in Spatial Sciences).
- E. LOPEZ-BAZO, E. VAYA, A.J. MORA, J. SURINACH, (1999) *Regional economic dynamics and convergence in the European Union*, "Annals of Regional Science", 33, pp. 343-370.
- E. LÓPEZ-BAZO, E. VAYA, M. ARTÍS, (2004) *Regional externalities and growth: evidence from European Regions*, "Journal of Regional Science", 44, pp. 43-73.
- N.G. MANKIW, D. ROMER, D.N. WEIL, (1992) *A contribution to the empirics of economic growth*, "Quarterly Journal of Economics", May, pp. 407-437.

- T. MARTIN, G. OTTAVIANO, (1999) *Growing locations: Industry location in a model of endogenous growth*, "European Economic Review", 43, pp. 281-302.
- A. NIEBUHR, (2001) *Convergence and the effects of spatial interaction*, "Jahrbuch für Regionalwissenschaft (Review of Regional Research)", 21, pp. 113-133.
- M.H. PESARAN, (2004) *General diagnostic tests for cross-section dependence in panels*, WP, University of Cambridge, Cambridge, UK.
- G. PIRAS, G. ARBIA, J. LE GALLO, (2006) *A meta-analysis of regional economic convergence of the NUTS-2 European regions, 1977-2002*, paper presented at the 45th European Congress of the European Regional Sciences Association (ERSA)
- M. PITTAU, (2004) *Fitting Regional Income Distributions in the European Union*, "Oxford Bulletin of Economics and Statistics", 66.
- M. PITTAU, R. ZELLI, (2006) *Income Dynamics across European Regions: Empirical Evidence from Kernel Estimation*, "Journal of Applied Econometrics", 21(5), pp. 605-628.
- D. QUAH, (1996) *Regional convergence clusters across Europe*, "European Economic Review", 40, pp. 951-958.
- J. RAMAJO, M. MÁRQUEZ, M.D.M. SALINAS, (2003) *Spatial Patterns in EU Regional Growth: New Evidence about the Role of Location on Convergence*, Department of Applied Economics, University of Extremadura, Mimeo, Badajoz.
- M.T. RAMÍREZ, A.M. LOBOGUERRERO, (2002) *Spatial Dependence and Economic Growth: Evidence from a Panel of Countries*, Banco de la República, Borradores de Economía, no. 126, Bogotá.
- S.J. REY, B.D. MONTOURI, (1999) *US regional income convergence: a spatial econometric perspective*, "Regional Studies", 33(2), pp. 143-156.
- S.J. REY, M.J. JANIKAS, (2005) *Regional convergence, inequality, and space*, "Journal of Economic Geography", 5, pp. 155-176.
- R.M. SOLOW, (1956) *A contribution to the theory of economic growth*, "Quarterly journal of economics", LXX, pp. 65-94.
- T.W. SWAN, (1956) *Economic growth and capital accumulation*, "Economic Record", 32, November, pp. 334-361.

SUMMARY

Convergence in per-capita GDP across EU-NUTS2 regions using panel data models extended to spatial autocorrelations effects

Most of the empirical works in regional convergence are based on either cross-sectional or "a-spatial" panel data models. In this paper, we propose the use of panel data econometrics models that incorporate an explicit consideration of spatial dependence effects (Anselin, 1988; Elhorst, 2001; 2003). This allows us to extend the traditional convergence models to include a rigorous treatment of regional spillovers and to obtain more reliable estimates of the parameters.

We consider two models respectively based on the introduction of a spatial lag among the explanatory variables (the "spatial lag model") and imposing a spatial autoregressive structure to the stochastic component (the "spatial error model"). We apply such a modelling framework to the long-run convergence of per-capita GDP of 125 EU-NUTS2 regions observed yearly in the period 1977-2002. A comparison of the results obtained using the two spatial panel data specifications with the main evidence available in the literature is also provided.