

The Role of Spatial Dependence in Regional Convergence of Per-Capita Income

*Il Ruolo della Dipendenza Spaziale
nella Convergenza Regionale dei Redditi Pro-Capite*

Paolo Postiglione, Mirella Salvatore, Giuseppe Arbia

Dipartimento di Scienze,
Università “G. d’Annunzio” di Chieti

Riassunto: Negli ultimi anni si è registrato un aumento di interesse (Barro, Sala-i- Martin (1992), Quah (1997)) nei riguardi del tema della convergenza regionale dei redditi pro-capite. Questi autori, però, non considerano il ruolo dell’informazione spaziale nella definizione dei loro modelli. In questo lavoro vengono presentati alcuni modelli di analisi territoriale (Anselin, Bera 1998), che vengono utilizzati per la misurazione della convergenza economica tra regioni e che possono essere considerati estensioni di quelli presentati da Barro, Sala-i-Martin (1992). Viene anche proposta una applicazione di tali modelli all’analisi della convergenza tra 119 regioni dell’U.E.

Keywords: Regional economic convergence, Spatial dependence

1. Introduction

One of the most remarkable evidence in empirical economics is the difference in economic growth rates experienced by different countries and by different regions within a country. Economic theory has long been aware of this problem and various explanations have been provided in the past (Solow, 1956). However the studies on regional convergence have experienced an acceleration of interest recently due to the issues raised in Europe by the unification process. In the literature, the empirical measurement of regional convergence is still based on the computation of basic statistical measures in which the geographical characteristics of data play no role (see Barro, Sala-i-Martin, 1992). This approach has been already criticised for instance in Arbia (2002). In the present paper, we show that neglecting the spatial features of data is doomed to produce serious biases when quantifying concepts like regional convergence. It is well known that regional data cannot be regarded as independently generated because of the presence of spatial similarities among neighbouring regions (Anselin, Bera, 1998). As a consequence the standard estimation procedures used in many empirical studies can lead to biases and inefficiencies in the estimates.

2. Neoclassical versus spatial approach to regional convergence

The most popular approaches in the quantitative measurement of convergence are those based on the concepts of *s*-convergence and *b*-convergence (Barro, Sala-i-Martin, 1992). An alternative approach is based on transition matrices proposed by Quah (1997) and used by

Rey (1999) to describe convergence among regions. The *s*-convergence approach is based on the study of the time trend of the variance of the logarithms of the regional per-capita income. If there is a decreasing long-term trend then regions appear to converge to a common growth rate. The *b*-convergence approach moves from the neoclassical Solow-Swan *exogenous growth theory* (Solow, 1956) assuming exogenous saving rates and a production function based on decreasing productivity of capital and constant returns. On this basis Barro, Sala-i-Martin (1992) suggest the following statistical model:

$$\log(y_{it}/y_{i0})/T = a - [(1 - e^{-bT})/T] \log(y_{i0}) + e_i \quad (2)$$

where y_{it} is per-capita income of *i*-th economy at time *t*, y_{i0} is the initial level of per-capita income, (y_{it}/y_{i0}) is the growth rate of *i*-th region, *a* is the intercept, *T* is the length of the period over which the growth rate is computed, e_i is the error term and *b* is the *speed of convergence*. Using a little algebra and reparametrizing the model, Equation (2) can be written as:

$$\log(y_{it}/y_{i0}) = c + b \log(y_{i0}) + w_i \quad (3)$$

where $c + b \log(y_{i0})$ is the systematic component, $c = Ta$, $b = -(1 - e^{-bT})$ and w_i is the new error term. We will say that *b*-convergence is satisfied in a cross-section of economies, if we find a negative relation between the growth rate of per-capita income and the initial level of income, i.e. if $b < 0$. The assumption on the probability model implicitly made in the phase of estimation is that w_i is normally distributed $(0, \sigma^2)$ independently of $\log y_{it}$. Finally concerning the sampling model it is assumed that $\{w_1, w_2, \dots, w_n\}$ are independent observations of the probability model. Model (2) is usually estimated directly through non-linear least-squares (Barro, Sala-i-Martin, 1992) or by reparametrizing the statistical model as in Equation (3) and estimating *b* via Ordinary Least Squares, with $\mathbf{b}$ linked to *b* by the expression $\mathbf{b} = \log(b + 1)/(-T)$. However the sampling model of independence seems inadequate to the case considered because regional observations are likely to display positive spatial dependence (Arbia, 1989). A more correct statistical model that takes spatial correlation into account is the so-called *spatial lag model* (Anselin, Bera, 1998) in which spatial dependency is accounted for by including an autoregressive term of the dependent variable, so that we can reformulate model (3) as :

$$\log(y_{it}/y_{i0}) = c + b \log(y_{i0}) + rL[\log(y_{it}/y_{i0})] + w_i \quad (4)$$

where $L[.]$ is spatial lag operator, *r* is the spatial autoregressive coefficient, the error term is once again assumed normally distributed independently of $\log y_{it}$ and of $L[.]$.

In such a model $\{w_1, w_2, \dots, w_n\}$ are again considered independent observations of the probability model in the hypothesis that all spatial dependency effects are captured by the lagged term. The parameters of model (4) can be estimated via maximum likelihood (ML), that provides consistent and unbiased estimates also in the case of spatially distributed data. An alternative model to incorporate the spatial effects is to leave unchanged the systematic

component of (3) and to model the error term w_i as an autoregressive random field, for instance by assuming the model:

$$w_i = \rho L(w_i) + u_i \quad (5)$$

and reformulate a probability model for the u 's by assuming them normally distributed $(0, \sigma_u^2)$ independently of $\log y_{it}$ and randomly drawn. We call this second model *spatial error model* (Anselin, Bera, 1998). The parameters can be estimated by using ML or the Generalized Method of Moments (GMM) procedures.

3. Empirical analysis of regional convergence

In the empirical analysis we consider the per-capita income of 119 European regions for the period 1980/94. The data are collected by Eurostat in the REGIO database and refer to all units at level NUTS 2 of seven European countries (France, Italy, Spain, Belgium, Denmark, Netherlands and West Germany). We estimate the speed of convergence with three different models: the neoclassical approach (3) using OLS method; the *spatial lag model* (4) via ML and the *spatial error model* (5) using both ML and GMM methods. The data used satisfy the assumption of normality, hence a likelihood function can be derived and maximized. The main results are summarized in Table 1.

Table 1: Estimation of the speed of convergence between 119 European regions in 1980-94 using three different models.

MODEL	b	Speed of convergence b	AIC (LIK)	Spatial correlation parameters (p-value)	Spatial independence (p-value)
Neoclassical approach OLS	-0.0792 (0.0034)	0.59%	-222.131 (113.065)	----	$I = 8.4756$ (0.0000)
Spatial Lag model ML	-0.0471 (0.0285)	0.34%	-266.687 (136.343)	$\rho = 0.5915$ (0.0000)	$LR = 46.5561$ (0.0000)
Spatial error model ML	-0.1027 (0.0013)	0.77%	-273.429 (138.715)	$\delta = 0.6294$ (0.0000)	$LR = 51.2983$ (0.0000)
Spatial error model GMM	-0.0997 (0.0015)	0.75%	----	$\delta = 0.5855$ (0.0000)	----

The results highlight convergence: the estimates of the speed (b) are always significant, but they range from 0.34% to 0.77% by employing different models. In all the estimated models homoskedasticity is not rejected, but spatial independence cannot be accepted at the 95% significance levels. For this reason we move a step ahead to improve modelling. The graphical analysis of the dataset suggests the existence of two different spatial regimes for poor and rich European regions. For this reason we attempt the use of two different dummy variables. In the first case (dummy 1) we have 1 if the starting level income is below the European average and 0 otherwise. In the second case (dummy 2) we label poor (code 1) those regions having average income (from 1980 to 1994) below the European mean in the time period considered and rich (code 0) otherwise.

Table 2: Estimation of the speed of convergence between 119 European regions in 1980-94 with structural changes (b_i and b_1 are the parameters when the dummy is equal to $i=0,1$).

MODEL	b_0	Speed of convergence b_0	b_1	Speed of convergence b_1	AIC (LIK)	Chow-Wald
Spatial Error model dummy 1 (ML)	-0.0357 (0.4709)	0.26%	-0.2278 (0.0009)	1.85%	-273.917 (140.959)	5.1675 (0.0755)
Spatial Error model dummy 1 (GMM)	-0.0351 (0.4903)	0.26%	-0.2072 (0.0025)	1.66%	-----	4.1241 (0.1272)
Spatial Error model dummy 2 (ML)	-0.1065 (0.0189)	0.80%	-0.3069 (0.0001)	2.62%	-278.208 (147.604)	19.2098 (0.0007)
Spatial Error model dummy 2 (GMM)	-0.1361 (0.0110)	1.04%	-0.2522 (0.0000)	2.08%	-----	17.4899 (0.0002)
Spatial Lag model dummy 1 (ML)	-0.0014 (0.9804)	0.01%	-0.0399 (0.2848)	0.29%	-263.664 (136.832)	0.9700 (0.6156)
Spatial Lag model dummy 2 (ML)	-0.1249 (0.0097)	0.95%	-0.1162 (0.0008)	0.88%	-274.504 (142.252)	13.2662 (0.0013)

Based on the value of AIC, the *spatial error model* still achieves the best fit. Table 2 shows that the case of dummy 2 allows to identify an highly significant model with spatial structural change, based on the Chow-Wald test. In agreement with the theoretical convergence hypothesis, the structural change model identifies a higher speed (2.62%) for “poor” regions and a lower speed (0.80%) for the “rich” ones. In conclusion, the empirical analysis reveals that the neoclassical approach, suggested by Barro and Sala-i-Martin, suffers from a misspecification due to the presence of strong spatial dependence. The best model among the alternatives considered is the *spatial error model* with or without spatial structural change, but it is not enough to eliminate completely the spatial dependence. Further possible improvements includes, for example, spatial random effects model.

References

- Anselin L., Bera, A. K. (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics, in *Handbook of applied economic statistics*, Ullah, A. and Giles, D. E. A. (eds.), Marcel Dekker, New York, 237-290.
- Arbia, G. (1989) *Spatial data configuration in statistical analysis of regional economic and related problems*, Kluwer Academic Publisher, Dordrecht.
- Arbia G. (2002) The role of spatial effects in the empirical analysis of regional concentration, to appear on *International Journal of Geographical System*.
- Barro, R. J., Sala-i-Martin, X. (1992) Convergence, *Journal of Political Economy*, Vol. 100, n. 2, 223-251.
- Quah, D. (1997) Empirics for growth and distribution: stratification, polarization, and convergence clubs, *Journal of Economic Growth*, 2, 27-59.
- Rey, S.J. (1999) Spatial empirics for economic growth and convergence, *Working Papers of San Diego State University*.
- Solow, R. M. (1956) A contribution to the theory of economic growth, *Quarterly Journal of Economics*, LXX, 65-94.