ISTITUTO DI STUDI E ANALISI ECONOMICA

Using Spatial Panel Data in Modelling Regional Growth and Convergence

by

Giuseppe Arbia

University G.D'Annunzio, viale Pindaro 42, I-65127 Pescara, Italy
Tel: +39 085 4537 533 e-mail: arbia@unich.it

Roberto Basile

ISAE, Piazza Indipendenza 4, 00185 Rome, Italy
Tel: +39 06 4448 2874 e-mail: r.basile@isae.it
and
University of Macerata, Faculty of Economics

Gianfranco Piras

Tor Vergata University, via Columbia 2, I-00133 Rome, Italy
Tel: +39 06 7259 5625 e-mail: gianfrancopiras@gmail.com
and
Regional Economics Application Laboratory, University of Illinois at Urbana-Champaign, USA

Working paper n. 55
September 2005

Stampato presso la sede dell’Istituto

ISAE - Piazza dell’Indipendenza, 4 – 00185 Roma.
Tel. +39-06444821; www.isae.it

ABSTRACT

In this paper we use spatial dependence panel data models to analyse regional growth behaviour in Italy. Controlling for fixed-effects allows us to disentangle the effect of spatial dependence (or spatial interaction) from that of spatial heterogeneity and of omitted variables and, thus, to properly investigate the regional convergence process within the country.

Keywords: Regional convergence; Regional spill-over; Spatial dependence modelling; Spatial panel data models.

JEL Classification: C21, C23, R11.

NON-TECHNICAL SUMMARY

The empirical evidence currently available on regional β-convergence is almost based on cross-sectional regressions or panel data fixed-effects estimates. As it is well known, regional data cannot be regarded as independently generated because of the presence of similarities among neighbouring regions. As a consequence, the standard estimation procedures employed in many empirical studies can be invalid and lead to serious bias and inefficiency in the estimates of the convergence rate. This important issues has been neglected in both cross-sectional and fixed-effect studies.

Recently, some empirical studies have used the spatial econometric framework for testing regional convergence in Italy, Europe and the USA. All of them do not properly specify a conditional growth model. Indeed, all these studies start from the specification of a "minimal cross-section growth regression", that includes only the initial level of per-capita income (the so-called, "absolute convergence" model) and then show that the unconditional convergence model is mis-specified due to spatially autocorrelated errors. However, the use of the minimal specification of the growth model might imply that at least part of the estimated spatial dependence can be the effect of omitted explanatory variables rather than a genuine effect of spatial interactions.

In this paper we consider spatial dependence within a panel data fixed-effect approach. Controlling for fixed-effects allows us to disentangle the effect of spatial dependence from that of spatial heterogeneity and of omitted variables. We present the results of an empirical study of the long-run convergence of per-capita income in Italy (1951-2000) based on data aggregated at the NUTS 3 EU regions corresponding to the 92 Italian provinces.

Following the work by Elhorst (2003), we have estimated two different models of spatial panel data namely (a) the spatial lag model, which incorporates spatial dependence in the form of a spatial lag variable, and (b) the spatial error model, in which spatial effects are incorporated in the distribution of the error term. Our results show that controlling both for spatial effects and for spatial autocorrelation allows us to be more confident that spatial autocorrelation represents a genuine regional interaction effects, rather than absorbing heterogeneity and the effect of omitted variables, as in the standard cross-sectional models.

Following a consolidated evidence, we have considered a structural break in the growth path of Italian provinces at the beginning of the seventies. All models over the entire period and over two different subperiods (1951-1970 e 1970-2000) have confirmed this fact. In fact, the growth rate is very high during the first years and drops dramatically after 1970.

The speed of convergence estimated by using the spatial lag model is much lower than that obtained with the classical fixed-effect specification. A decrease in the β parameter referred to the initial condition, can be traced back to the introduction of a spatial lag term in the model, and indirectly confirms the positive effect of factor mobility, trade relationships, and knowledge spill-over on regional convergence. The results obtained using the spatial error specification are more difficult to be interpreted.

L’USO DI MODELLI DI ECONOMETRIA SPAZIALE PER DATI *PANEL* PER L’ANALISI DELLA CRESCITA E DELLA CONVERGENZA REGIONALE

SINTESI

In questo lavoro si usano modelli econometrici di dipendenza spaziale per dati panel al fine di analizzare il comportamento di crescita regionale in Italia. Il controllo degli effetti fissi consente di distinguere l’effetto della dipendenza spaziale (o dell’interazione spaziale) da quello dell’eterogeneità spaziale e delle variabili omesse. Ciò permette di analizzare in maniera appropriata il processo di convergenza regionale interno al Paese.

Parole chiave: Convergenza regionale; *Spill-over* spaziali; Modelli di dipendenza spaziale per dati *panel*

Classificazione JEL: C21, C23, R11.

CONTENTS

1 INTRODUCTION[1]

The empirical evidence currently available on regional β-convergence is almost based on cross-sectional regressions or panel data fixed-effects estimates (Magrini, 2003). As is well known, one of the drawbacks of the cross-section approach is its inability to solve the problem of omitted variables, especially in cross-region studies where the conditional convergence analysis is limited by data availability for key variables (e.g. variables like secondary school enrolment ratio or the investment to GDP ratio). A further limitation in using cross-sectional techniques concerns the need to impose full regional homogeneity in the parameters of the random process that describes the evolution of per-capita income.

The use of panel data methods can help in solving these two problems (omitted variables and homogeneity) (Islam, 1995). In particular they allow for unobserved heterogeneity, but confine differences across regions to the intercept of the model, assuming that the economies are characterized by a common convergence coefficient. Studies based on the fixed-effect approach generally obtain much higher convergence rates than those founded in cross-section regressions (Islam, 1995). However, high convergence rates are difficult to reconcile with the neoclassical growth theory. The main criticism moved to the fixed-effect approach concerns the fact that the use of annual growth rates catches up movements toward a trend more than the long period path of the variable even if this problem could be reduced by using the regional deviation from the national mean.

As it is well known, regional data cannot be regarded as independently generated because of the presence of similarities among neighbouring regions (Anselin, 1988; Anselin and Bera, 1998). As a consequence, the standard estimation procedures employed in many empirical studies can be invalid and lead to serious bias and inefficiency in the estimates of the convergence rate. This important issues has been neglected in both cross-sectional and fixed-effect studies.

Recently, some empirical studies have used the spatial econometric framework for testing regional convergence in Italy, Europe and the USA (Rey and Montouri, 2000; Arbia and Basile, 2005; Le Gallo et al. 2003). All of them

[1] A previous version of this paper was presented at the first seminar of Spatial Econometrics Jean Paelinck, Faculty of Economical Sciences, University of Zaragoza, Zaragoza, October 22-23. 2004. We are grateful to J.P.Elhorst and J.P. LeSage, for providing Matlab Routines, and to all the participants at the conference for helpful discussion and suggestions over a previous version of the paper. Errors remain solely responsibility of the authors.

do not properly specify a conditional growth model. Indeed, all these studies start from the specification of a “minimal cross-section growth regression”, that includes only the initial level of per-capita income (the so-called, “absolute convergence” model) and then show that the unconditional convergence model is mis-specified due to spatially autocorrelated errors. However, the use of the minimal specification of the growth model might imply that at least part of the estimated spatial dependence can be the effect of omitted explanatory variables rather than a genuine effect of spatial interactions.

The main idea of the present paper concerns the importance of considering spatial dependence within a panel data fixed-effect approach. Controlling for fixed-effects allows us to disentangle the effect of spatial dependence from that of spatial heterogeneity and of omitted variables. To our knowledge there is no empirical work in the literature accounting for spatial dependence in the panel data context that allows for spatial autocorrelation by including in the model a spatial structure for either the dependent variable or the error term.

In this paper, we present the results of an empirical study of the long-run convergence of per-capita income in Italy (1951-2000) based on data aggregated at a level of spatial resolution (the NUTS 3 EU regions corresponding to the 92 Italian provinces) that is fine enough to allow for spatial effects (for example, regional spill-over) to be properly modelled.

The paper is organized as follows. In Section 2, we present a review of the spatial econometric techniques that incorporate spatial dependence within the contest of a β-convergence modelling, including spatial panel data models. A description of the empirical data is given in Section 3.1. In Section 3.2 and 3.3, we report the results of an empirical analysis based on the 92 Italian provinces (European NUTS-3 level) and the per-capita income recorded in the period ranging from 1951 to 2000 and we show the different estimates of the speed of convergence and of the parameter half-life that can be obtained by using different model specifications. Finally, we discuss the results obtained and we outline possible extensions of the present work.

2 SPATIAL DEPENDENCE IN GROWTH STUDIES

The most popular approach in the quantitative measurement of convergence is based on the concept of β-regression model (Durlauf and Quah, 1999 for a review). Even if more recently, some alternative methods have been developed using intra-distribution dynamics approach (Quah, 1997; Rey, 2000) and the Lotka-Volterra predator-prey specification (Arbia and Paelinck, 2003). In the context of the β-regressions some important innovations have been introduced in the last few years. In particular, some authors introduced a panel fixed-effect specification to control for the effects of omitted variables (and of heterogeneity), while others introduced the effect of spatial dependence[2].

In this section, we review the classical β-convergence approach and we propose a new specification of the empirical growth regression which simultaneously includes fixed-effects to control for the effect of omitted (time invariant) variables (or heterogeneity) and spatial dependence.

2.1 β-convergence

The β-convergence approach moves from the neoclassical Solow-Swan exogenous growth model (Solow, 1956; Swan, 1956), assuming a closed economic system, exogenous saving rates and a production function based on decreasing productivity of capital and constant returns to scale. On this basis authors like Mankiw *et al.* (1992) and Barro and Sala-i-Martin (1992) suggested the following statistical model:

$$\ln\left[\frac{y_{T,i}}{y_{0,i}}\right] = \mu_i + \varepsilon_i \qquad (1)$$

where $\ln(y_{T,i} / y_{0,i})$ is the growth rate of the per-capita income over the entire period, as y_T is the value of per-capita income in the last period of time

---

[2] Other examples of the recent advances in quantitative measures of growth are given in Meliciani e Peracchi (2001), which use median unbiased estimators of the rate of convergence to the steady-state growth path allowing for unrestricted patterns of heterogeneity and spatial correlation across regions, and in Vayá *et al*(2004), which discuss, in particular, non linearity in the parameters of the spatial model. Finally, Basile and Gress (2004) propose a semi-parametric spatial auto-covariance specification of the conditional convergence model which simultaneously takes account for the problems of non-linearities and spatial dependence and notice a trade-off between the identification of non-linearities and the estimation of the spatial parameters.

considered (in our case T = year 2000), and y_0 is the value in the first period (1951); ε_i the error term. The formal expression for the systematic component μ_i is as follows:

$$\mu_i = \alpha + (1 - e^{-\lambda k}) \ln y_{0,i} \qquad (2)$$

with the parameter λ assuming the meaning of the "speed of convergence", measuring how fast economies converge towards the steady state. The assumption on the probability model implicitly made in this context is that ε_i is normally distributed $(0, \sigma^2)$ independently of $\ln y_{0,i}$. Finally, concerning the sampling model, it is assumed that $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ are independent observations of the probability model. Equation 1 is usually directly estimated through non-linear least-squares (Barro and Sala-i-Martin, 1995) or by re-parameterizing the statistical model setting $\beta = (1 - e^{-\lambda k})$ and estimating β by ordinary least squares. Absolute convergence is said to be present if the estimate of β is negative and statistically significant. If the null hypothesis $(\beta = 0)$ is rejected, we would conclude that poor regions grow faster than rich ones, and that they all converge to the same level of per-capita income.

2.2 Spatial dependence in cross-sectional models

The neoclassical growth model discussed in the previous section has been developed starting from the hypothesis that the economies are fundamentally closed. However, this hypothesis is too strong for regions within a country, where barriers to trade and to factor flows are considerably low. In order to understand the implications for convergence of the introduction of the openness hypothesis into the theoretical framework, we must consider the role of factor mobility, trade relations and technological diffusion (or knowledge spill-over).

Factor mobility implies that labour and capital can move freely in response to differentials in remuneration rates, which in turn depends on the relative factor abundance. Thus, capital will tend to flow from the regions with a higher capital-labour ratio to the regions with a lower capital-labour ratio, while labour will tend to flow in the opposite direction. Moreover, the regions with lower capital-labour ratios will show higher per-capita growth rates. Actually, if the adjustment process in either capital or labour is instantaneous, the speed of convergence would be infinite. By introducing credit market imperfections, finite lifetimes and adjustment costs for migration and investments in the model, the speed of convergence to the steady-state remains higher than in the closed economy case, but with a finite value (Barro and Sala-i-Martin, 1995). The same result can be obtained by introducing into the neoclassical growth model the hypothesis of free trade relations rather than factor mobility: convergence in interregional per-capita income will be higher than in the closed-economy version.

Another possibility for poor economies to converge towards the richer ones is represented by technological diffusion and knowledge spill-over. In the presence of disparities in regional technological attainment, interregional trade can promote technological diffusion when technological progress is incorporated in traded goods (Barro and Sala-i-Martin, 1997). A broader interpretation of knowledge spill-over effects refers to positive knowledge external effects produced by firms at a particular location and affecting the production processes of firms located elsewhere. However, when investigating regional convergence problem and studing the effect of geographical spill-over on growth, we must also distinguish between local and global geographic spill-over. With local spill-over, production processes of firms located in one region only benefit from the knowledge accumulation in that region. In this case, regional divergence is likely to be the outcome. With global geographical spill-over, we mean that knowledge accumulation in one region improves productivity of all firms wherever they are located. Thus, a global geographical spill-over effect contributes to regional convergence (Martin and Ottaviano, 1999, 2001; Kubo, 1995).

The speed of convergence towards the steady-state predicted by the open-economy version of the neoclassical growth model (as well as by the technological diffusion models) is faster than in the closed-economy version of the neoclassical growth model.

A direct way to test empirically wether a higher speed of convergence is to be expected in an open economy consists of 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 referred to capital and technology flows. Some attempts have also been made to test the role of migration flows on convergence. However, the results of these studies suggest that migration plays only a negligible role in the explanation of convergence (Barro and Sala-i-Martin, 1995).

An indirect way to control for the effects of interregional flows (or spatial interaction effects) on growth and convergence is through spatial dependence models. A first way to take spatial dependence into account is the so-called spatial autoregressive model (or SAR, Anselin and Bera, 1998; Arbia, 2005), where a spatial lag of the dependent variable is included on the right hand side of the statistical model. If $W = \sum_{j=1}^{n} w_{i,j}$ is a row-standardized matrix of spatial weights describing the structure and intensity of spatial effects, Equation 1 is re-specified as

$$\ln\left[\frac{y_{T,i}}{y_{0,i}}\right] = \alpha + \beta \ln y_{0,i} + \rho \sum_{j=1}^{n} w_{i,j} \ln\left[\frac{y_{T,i}}{y_{0,i}}\right] + \varepsilon_i \qquad (3)$$

where ρ is the parameter of the spatially lagged dependent variable $\sum_{j=1}^{n} w_{i,j} \ln\left[\frac{y_{T,i}}{y_{0,i}}\right]$ that captures the spatial interaction effect indicating the degree to which the growth rate of per-capita GDP in one region is determined by the growth rates of its neighbouring regions, after conditioning on the effect of $\ln y_{0,i}$. The error term is again assumed normally distributed and independently of $\ln y_{0,i}$ and of $\sum_{j=1}^{n} w_{i,j} \ln\left(\frac{y_{T,i}}{y_{0,i}}\right)$, under the assumption that all spatial dependence effects are captured by the lagged term. An alternative way to incorporate the spatial effects is via the spatial error model or SEM (Anselin and Bera, 1998; Arbia, 2005). In this specification we leave unchanged the systematic component and we model the error term in Equation 1 as a Markovian random field, for instance assuming that:

$$\varepsilon_i = \delta \sum_{j=1}^{n} w_{i,j} \varepsilon_i + \eta_i \qquad (4)$$

The error term η_i is assumed to be normally and independently distributed, with mean zero and constant variance (σ_η^2), independently of $\ln y_{0,i}$. As already noted, some empirical studies have previously used the spatial econometric framework to test for regional convergence. The most comprehensive studies are those of Rey and Montouri (2000), and Le Gallo, Ertur and Baoumont (2003). Both studies do not specify a conditional growth model. Indeed, all these studies start from the minimal growth regression model specification, which includes only the initial level of per-capita income as a predictor (the so-called, absolute convergence model) and then show that the unconditional convergence model is mis-specified due to spatially auto-correlated errors. However, the use of the minimal specification of the growth model might imply that at least part of the estimated spatial dependence actually absorb the effect of the omitted explanatory variables (or the effect of regional heterogeneity) rather than the effect of true spatial interactions. In order to overcome such a drawback here we propose a new specification of the growth regression model based on panel data containing spatial dependence and fixed-effects.

2.3 Panel data models

One of the major advantages of the panel data approach to convergence is that it can be helpful for the correction of the bias generated by omitted variables and heterogeneity in the classical cross-sectional regression (Islam, 2003). Panel data, in fact, allow for technological differences across regions, (or at least the unobservable and unmeasurable part of these differences), by modelling the regional specific effect. More formally, the panel version of the growth equation can be expressed in the following way:

$$\ln\left[\frac{y_{t+k,i}}{y_{t,i}}\right] = \alpha_i + \beta \ln y_{t,i} + \varepsilon_{t,i} \qquad (5)$$

with i ($i = 1,...,N$) denoting regions, and t ($t = 1,...,T$), denoting time periods. The dependent variable $\ln\left[\frac{y_{t+k,i}}{y_{t,i}}\right]$ is the annual growth rate of the per-capita income of Italian provinces, $\ln y_{t,i}$ is the value of the per-capita income at the beginning of each period over which the growth rate is calculated. It should be noted that α_i are time invariant and account for any individual-specific effect not included in the regression equation. Two different interpretations may be given of the parameters α_i, and two different models may be distinguished according to this interpretations. If the α_i is assumed to be fixed parameters Equation 5 is a fixed-effect panel data model. Conversely, if the α_i are assumed to be random, Equation 5 expresses the random effect panel data model. Generally, fixed-effect model is particularly indicated when the regression analysis is limited to a precise set of individuals (firms or regions). On the contrast, random effect is a more appropriate specification if we are drawing a certain number of individuals randomly from a larger population[3]. For this reason, since our data set consists of the observations over the 92 Italian

[3] For more details on the use of these two models for panel data we suggest to see specialistic books on panel data (like Baltagi 2001).

provinces, we have decided to employ a fixed-effect panel data model to test regional convergence. Following Islam (1995), a number of studies have tried to estimate the speed of convergence among regions using panel data sets and variant of fixed-effect model. In the main literature, there is a consistent evidence that the estimate of the speed of convergence from panel data with fixed-effects tend to be larger than the 2 percent-per-year number estimated from cross sections (Barro and Sala-i-Martin, 1995).

2.4 Spatial dependence in panel data growth models

The aim of the present paper is to account for spatial dependence in a panel data context.

The starting point is the classical fixed-effect panel data model, in which spatial dependence is accounted for by including a spatially lagged term of the dependent variable. According to this specification the model assumes the following expression:

$$\ln\left[\frac{y_{t+k,i}}{y_{t,i}}\right]=\alpha_i+\rho\sum\nolimits_{j=1}^{n} w_{i,j}\ln\left[\frac{y_{t+k,i}}{y_{t,i}}\right]+\beta\ln y_{t,i}+\varepsilon_{t,i} \qquad (6)$$

with $\sum\nolimits_{j=1}^{n} w_{i,j}$ the classical weight matrix[4] ρ is the so-called spatial-autoregressive coefficient, and $\varepsilon_{t,i}$ is the classical zero mean error term assumed independent 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 lag model*. The standard estimation method for the fixed-effect model is to eliminate the intercept term from the regression equation by expressing all variables in terms of deviation from their time average, and then using standard OLS estimators. In presence of spatial autocorrelation it is common practice in spatial econometric literature (Elhorst, 2003) to use maximum likelihood procedure to estimate the demeaned equation. The only difference is that ML estimators do not make corrections for the degree of freedoms. If the estimated value of the δ parameter is significantly positive (negative), we are in presence of positive (negative) spatial autocorrelation. An

[4] In the present paper we consider the classical binary connectivity matrix which assume the values of 1 if the two regions present a common border and zero otherwise.

alternative way to incorporate the spatial effects is to leave unchanged the systematic component and to model the error term, for instance assuming that

$$\begin{aligned} \ln\left[\frac{y_{t+k,i}}{y_{t,i}}\right] &= \alpha_i + \beta \ln y_{t,i} + \varepsilon_{t,i} \\ \varepsilon_{t,i} &= \delta \sum_{j=1}^{n} w_{i,j} \varepsilon_{t,i} + \eta_i \end{aligned} \tag{7}$$

where $\sum_{j=1}^{n} w_{i,j}$ is again the spatial weight matrix, δ is the spatial autocorrelation coefficient, and the η_i are assumed to be normally distributed indipendently of the explanatory variable with zero mean and known variance. This model is called *fixed-effect spatial error model*. Again the parameters can be estimated by using maximum likelihood.

3 EMPIRICAL EVIDENCE FROM ITALIAN PROVINCES

3.1 Data

The empirical study focuses on the case of Italian provinces, corresponding to the European NUTS-3 level in the official UE classification. The analysis is based on a newly compiled database on per-capita GDP for the 92 provinces over the period 1951-2000.

The provincial data on value added are based on the estimates made by the Istituto Guglielmo Tagliacarne using direct and indirect provincial indicators to disaggregate the regional product into provinces. These estimates are expressed in constant prices by using sectoral/regional value added deflators. The source of population data is ISTAT (National Institute of Statistics).

Italy is currently divided into 103 provinces, grouped into 20 regions. Over the period considered (1951-1999), however, the boundaries of some administrative provinces changed. For this reason, only the provinces that already existed in 1951 (92 units) have been considered for the empirical analysis.

3.2 Cross section results

In this section, we report the results of the empirical analysis run to test convergence of Italian provinces' per-capita incomes in the period considered (1951-2000). We start our empirical investigation from the OLS cross-sectional estimates of the unconditional mode of β-convergence (Equation 1). Previous empirical evidences suggest the existence of a break in the growth path of the Italian provinces at the beginning of the seventies. Here, we want to test if this break point is confirmed in our model and thus we estimate it over the entire period and then separately in two different sub-periods: the first ranging from 1951 to 1970, and the second from 1970 to 2000.

Tab. 1 Convergence of per-capita income in the 92 Italian provinces (1951-2000) – Unconditional Model – OLS Estimates

(number into brackets refer to the p-values)

	1951-2000	1951-1970	1970-1999
Constant	0.018 (0.697)	0.065 (0.495)	0.013 (0.774)
Income level	-1.032 (0.000)	-2.029 (0.000)	-0.179 (0.000)
Goodness of fit			
Adjusted R^2	0.435	0.418	0.001
Log Likelihood	-47.780	-113.004	-52.375
Schwartz Criterion	104.604	235.051	113.794
AIC	99.560	230.008	108.750
Observations	92	92	92
Regression Diagnostic			
Jarque-Bera normality test	1.422 (0.490)	2.491 (0.287)	1.480 (0.476)
Breusch-Pagan heteroschedasticity test	0.916 (0.338)	0.583 (0.445)	0.105 (0.745)
White robust heteroschedastic test	3.173 (0.204)	1.910 (0.384)	1.874 (0.391)
Morans'I spatial autocorrelation test	9.556 (0.000)	7.657 (0.000)	3.532 (0.000)
LM test (error)	79.283 (0.000)	50.218 (0.000)	9.550 (0.001)
LM test (lag)	25.261 (0.000)	11.475 (0.000)	8.155 (0.004)

Table 1 displays the OLS estimates of the unconditional β-convergence model for the 92 Italian provinces. In this specification the dependent variable is represented by the growth rate of the per-capita income computed over the entire period[5]. All the variables are scaled to the national average. Our results are consistent with the previous empirical findings related to the growth of Italian provinces. The OLS coefficient of the initial per-capita value added for the entire period is highly significant and negative, confirming the presence of absolute convergence over the years 1951-2000. Its value (-1.032) implies an annual rate of convergence of 1.43%. The time necessary for the economies to fill half of the difference from their steady states is about 67 years (see Tab. 2)[6].

If we split the entire period into two different sub-periods we obtain very different and interesting results. In particular, the evidence of two different paths in the growth of per-capita income of Italian province is confirmed. The coefficient of initial per-capita income estimated from 1951-1970 is still strongly significant and higher than the one computed over the entire period (-2.029). In the second period the estimated value (-0.179) is not significant. These results are confirmed by the evidence on the "speed of convergence" (2.56 % for the first subperiod, and only 0.18 % for the second compared with the overall 1.43% for the entire period). Furthermore, the "half-life" is about 33 years when computed in the first sub-sample up to 1970, and increases to 386 years in the second part of the sample period (it was 67 years in the whole period).

Table 1 also reports some diagnostics to identify misspecifications in the OLS cross-sectional model. Firstly, the value of the Jarque-Bera test is always far from significant, revealing that OLS errors can be considered normally distributed. Consequently, we can safely interpret the results of the various misspecification tests (heteroskedasticity and spatial dependence tests) that, as

---

5 The entire period growth rate is calculated as the difference in logs between the value at the end of the sample period and the value of the per capita income in the first period of our sample. This difference is dived by the number of periods (49 in our case), and multiplied by 100. This explains the magnitude of the coefficients reported in Table beta.

6 The "speed of convergence", interpreted as the annual rate of convergence, is measured by the following expression:

$$s = -\ln(1 + T\beta)/T.$$

The "half-life", defined as the time necessary for the economies to fill half of the gap from their steady states, is:

$$\tau = -\ln(2)/\ln(1+\beta).$$

For further details on the "sped of convergence" and on the "half-life" see Arbia (2005).

it is known, are based on such an assumption[7]. The value of the Breusch-Pagan statistics indicates that there is no heteroscedasticity, as it is also confirmed by the robust White statistics. The value of the log likelihood and the values of the Schwartz and AIC criterion are also reported. They clearly show how the OLS model fits much better to the data of the first period than to those of the second one (e.g. AIC from 108.750 in the first period to 230.008.in the second).

Tab. 2 Comparison of the "speed of convergence" and of the "half life" parameters estimated in the various model specifications

	1951-2000	1951-1970	1970-2000
CROSS SECTIONAL MODELS			
β-convergence			
Speed of convergence	1.437	2.562	0.183
Half-life	67	34	387
Spatial lag model			
Speed of convergence	1.302	2.413	0.251
Half-life	72	35	284
Spatial error model			
Speed of convergence	3.570	5.255	1.675
Half-life	41	20	51
PANEL DATA MODELS			
Fixed-effect Model			
Speed of convergence	0.149	0.412	0.137
Half-life	481	174	513
Fixed-effect Spatial Lag Model			
Speed of convergence	0.149	0.421	0.133
Half-life	481	170	528
Fixed-effect Spatial Error Model			
Speed of convergence	0.133	0.371	0.130
Half-life	536	192	541

[7] Heteroskedasticity tests have been carried out for the case of random coefficient variation (the squares of the explanatory variables were used in the specification of the error variance to test for additive heteroskedasticity).

In order to test for the presence of spatial dependence in the error term, three different tests were considered: the Moran's I and two different versions of the Lagrange Multiplier tests. The first version is very powerful against spatial dependence both in the form of error autocorrelation and spatial lag, but it does not allow to discriminate between the two alternative forms of misspecifications (see Anselin and Rey, 1991). Both LM tests have high values and are strongly significant, indicating significant spatial dependence, with an edge towards the spatial error (particularly in the entire period and in the first subperiod).

In conclusion, the previous results suggest that the OLS estimates suffer from a misspecification due to omitted spatial dependence: each region is not independent of the others, as it often happens in many empirical studies at a regional level. This evidence motivated to propose alternative specifications in order to remove residual spatial dependence.

Tables 3 and 4 display the results of the maximum likelihood estimates obtained by using the two models discussed in Section 2.2: the spatial lag and spatial error model. The parameters associated with the spatial error and the spatial lag terms are always highly significant. The fit of the spatial error model (based on the values of Akaike and Schwartz criteria) is always higher than that obtained in both the OLS and the spatial lag models. For the entire sample and the first period, the coefficient of the initial level of per-capita income decreases in the spatial lag model, while it increases in the spatial error model. A decrease in the absolute value of the parameter is due to the inclusion of the spatial lag term in the model, and thus indirectly confirms the positive effect of factor mobility, trade relations and knowledge spill-over on regional convergence. This result is in line with the dominant opinion that the quick convergence occurred in Italy after the second world war period and until the early 1970's was partly due to a technological transfer process (a strong convergence in term of labour productivity between Southern and Northern regions can indeed be observed in this period) and to a massive labour migration process.

The increase in the β-coefficient related to the initial per-capita income observed in the spatial error model for the first period can be given a rather different interpretation. In this second instance, indeed, the correction for spatial dependence tends to capture the effect of omitted variables (different from factor migration, trade and spill-over), which have a negative effect on growth (such as e.g. the crime rate). It is perhaps fair to say that the results obtained with the spatial error model in some way "obscure" the interpretation of the spatial dependence correction as a way to capture the effect of openness on regional convergence. This point suggests the need to use econometric panel data tools, (such as a fixed-effects model), which allow to properly capture the effects of omitted variables and thus to isolate the effect of spatial dependence.

The results of panel data fixed-effects spatial autocovariance models will be discussed in the next section.

Finally, Tables 3 and 4 show that (if compared with those obtained with OLS estimates of the unconditional model reported in Table 1) the coefficient of the initial per-capita income increases both in the spatial lag and in the spatial error specification. This suggests that over the second period regional spill-over and labour migration did not give any contribution to the regional convergence process, and thus spatial dependence parameters tend only to capture the effect of other omitted variables (which have conversely a negative impact on

Tab. 3 Convergence of per-capita income in the 92 Italian provinces (1951-2000) – Spatial Lag Model – Maximum Likelihood Estimates (numbers into brackets refer to the p-values)

	1951-2000	1951-1970	1970-1999
Constant	-0.047 (0.262)	-0.033 (0.713)	-0.007 (0.875)
Income level	-0.963 (0.000)	-1.936 (0.000)	-0.243 (0.140)
W-growth	0.384 (0.000)	0.295 (0.000)	0.268 (0.031)
Goodness of fit			
Adjusted R^2	0.544	0.481	0.086
Log Likelihood	-40.491	-109.357	-50.161
Schwartz Criterion	94.548	232.278	113.889
AIC	86.938	224.713	106.323
Observations	92	92	92
Regression Diagnostic			
Spatial Breusch-Pagan heteroschedasticity test	1.583 (0.208)	0.444 (0.504)	0.628 (0.427)
Likelihood ratio test spatial autocorrelation	14.577 (0.000)	7.294 (0.006)	4.427 (0.035)
LM test (error)	44.787 (0.000)	43.486 (0.000)	0.094 (0.758)

Tab. 4 Convergence of per-capita income in the 92 Italian provinces (1951-2000) – Spatial Error Model – Maximum likelihood Estimates

(numbers into brackets refer to the p-values)

	1951-2000	1951-1970	1970-1999
Constant	-0.110 (0.379)	-0.060 (0.793)	-0.121 (0.292)
Income level	-1.686 (0.000)	-3.324 (0.000)	-1.358 (0.000)
Lambda	0.820 (0.000)	0.755 (0.000)	0.660 (0.000)
Goodness of fit			
Adjusted R^2	0.442	0.424	0.012
Log Likelihood	-1.628	-83.149	-46.063
Schwartz Criterion	12.301	175.343	101.171
AIC	7.257	170.300	96.127
Observations	92	92	92
Regression Diagnostic			
Spatial Breusch-Pagan heteroschedasticity test	0.135 (0.713)	5.364 (0.020)	0.542 (0.461)
Likelihood ratio test spatial autocorrelation	92.303 (0.000)	59.707 (0.000)	12.622 (0.000)
Wald test	6.490 (0.010)	9.792 (0.001)	19.355 (0.000)
LM test (error)	6.169 (0.012)	6.474 (0.010)	32.897 (0.000)

growth, such as the crime rate). This interpretation is in line with the common opinion according to which in the period 1970-2000 the lack of regional convergence (in particular the lack of convergence between Northern and Southern regions) has been - at least in part - due to the reduction of the opportunities for technological catching-up and to the reduction of the economic stimulus to labour migration.

3.3 Panel data results

In this section we report the results obtained by adopting a panel data specification in the estimation of the process of convergence of per-capita income in the Italian provinces over the period 1951-2000. Our empirical analysis moves from the consideration that panel data models accounting for spatial autocorrelation lead to more reliable estimates of the parameters of interest in the evaluation of convergence by controlling for omitted variables and heterogeneity. Moreover, the joint presence of fixed and spatial effects allows us to assume that spatial dependence may capture only regional interaction effect rather than absorbing heterogeneity and the effect of omitted variables. Baltagi (2001) lists some of the benefits and of the limitations related to the use of classical panel data (Hsiao, 1985, 1986; Klevmarken, 1989; Solon, 1989). Firstly, they allow to control for individual heterogeneity. Furthermore, they are more informative with respect to pure time series or pure cross-sectional data, they present more variability and less collinearity, they provide more degrees of freedom and more efficient estimates. Finally, panel data offer clear the advantage that units are observed through times and this allows a semplification of many economic problem that would be more difficult or even impossible to study by using pure cross sections. Regional convergence is a very good example of such problems. Spatial panel data models has exactly the same advantages that classical panel data, and, in addition, they allow to control for spatial autocorrelation.

The general objective that lays behind the present study is that previous studies at provincial level for Italy that were carried out by using cross-sectional OLS estimates are biased because neither fixed-effect nor spatial effects were considered. On the other hand those studies based on panel data are also biased because the neglect spatial autocorrelation effects.

Table 5 reports the result of the estimation of a fixed-effect panel data regional convergence model of the kind discussed in Section 2.3. The dependent variable is the annual growth rate of the per-capita GDP, and the only predictor considered is the level of per-capita income at the beginning of the period considered. In the most general specification, there are 92 different groups, each one corresponding to one of the provinces, and 50 observation for each group (1951-2000). The number of time observations considered are thus 20 in the first part of the sample, and 30 in the second. Then, the total number of observations is 4600 for the entire sample, 1840 for the first period, and 2760 for the second. In each model specification the number of observations is far greater than the number of parameters to be estimated ensuring a large number of degrees of freedom.

Tab. 5 Convergence of per-capita income in the 92 Italian provinces (1951-2000) – Fixed-effect Model

(numbers into brackets refer to the p-values)

	1951-2000	1951-1970	1970-1999
Constant	-0.009 (0.000)	-0.042 (0.000)	-0.008 (0.000)
Income level	-0.144 (0.000)	-0.397 (0.000)	-0.135 (0.000)
$Sigma_{\alpha}$	0.370	0.112	0.035
$Sigma_{\varepsilon}$	0.046	0.062	0.027
ρ	0.393	0.763	0.626
F-test that all α_i=0	4.68 (0.000)	4.22 (0.000)	2.85 (0.000)
R^2 within	0.090	0.190	0.072
R^2 between	0.053	0.204	0.000
R^2 overall	0.017	0.032	0.003
Correlation (α_i,xβ)	-0.923	-0.965	-0.971
Observations	4600	1840	2760
Number of groups	92	92	92
Observations per group	50	20	30

The value of the coefficient β of the initial per-capita income level of Italian provinces computed over the entire time period is -0.144. It is negative and significant, showing the presence of convergence. It rises up to -0.397 if referred to the first period (-0.135 for the second) confirming the results obtained in the cross-sectional specification of the growth equation. The time necessary for the economies to fill half of the gap which separates them from their own steady state is, according this formulation, about 480 years over the entire sample. In the second subperiod it is sensibly higher, while it is of 170 years in the first period. Looking at the same problem from a different viewpoint, the speed of convergence in the first part of the sample is about 1/3 of the one calculated over the second period. With respect to the cross-sectional models, in this new specification the value of the β coefficients falls down. This happens because the value of the β coefficient is not influenced any more by the presence of omitted variables that are now captured by regional specific effect.

The last two empirical analyses that we report in this paper concern the estimation of a spatial version of the panel data model considered. As we have already remarked, an explicit treatment of the spatial dependence in a panel data model not only allows us to solve the problems connected with unobserved factors that influence growth, but also removes the bias introduced by spatial dependence in the error terms. Table 6 reports the results of the estimation of the fixed-effect spatial lag model (see Equation 6). The value of the estimated β coefficient is now about -0.129 for the entire sample period. The spatial autocorrelation coefficient (strongly significant) captures the effect of spatial residual correlation. The presence of the α parameters, conversely, isolates the effect of omitted variables, representing different structural characteristics of the regional economies. The simultaneous presence of these two different factors (spatial autocorrelation and fixed-effects) cause the estimates of the β coefficient to decrease with respect to the fixed-effect model, since it is now cleaned from the influence of both omitted variables and spatial autocorrelation. From an economic point of view this result confirms the evidence already obtained by using the cross sectional approach. The reduction of the absolute value of the β coefficient of the model (due to the inclusion of the spatial lag term) confirms the positive effect of factor mobility, trade relationships, and the presence of spill-overs on regional convergence.

Tab. 6 Convergence of per-capita income in the 92 Italian provinces (1951-2000) – Fixed-effect Spatial Lag Model

(numbers into brackets refer to the p-values)

	1951-2000	1951-1970	1970-2000
Spatially lagged income	0.347 (0.000)	0.344 (0.000)	0.274 (0.000)
Income level	-0.129 (0.000)	-0.359 (0.000)	-0.128 (0.000)
Adjusted R^2	0.207	0.307	0.157
Log-likelihood	7728.300	2497.948	6139.997
Observations	4600	1840	2760
Number of groups	92	92	92
Observations per group	50	20	30

Tab. 7 Convergence of per-capita income in the 92 Italian provinces (1951-2000) – Fixed-effect Spatial Error Model

(numbers into brackets refer to the p-values)

	1951-2000	1951-1970	1970-2000
Spatial autocorrelation coefficient	0.363 (0.000)	0.390 (0.000)	0.277 (0.000)
Income level	-0.144 (0.000)	-0.405 (0.000)	-0.131 (0.000)
Adjusted R^2	0.202	0.306	0.139
Log-likelihood	7741.612	2514.092	6139.676
Observations	4600	1840	2760
Number of groups	92	92	92
Observations per group	50	20	30

Different considerations have to be made in the case of the fixed-effect spatial error model. In this second instance the values of the coefficients do not particularly differ from those obtained with the classical fixed-effect model estimate. This implies that there is no advantage in modelling the error term as an autoregressive process with respect to the simple fixed-effect term. As a conseguence, our empirical investigation shows that the spatial lag specification has to be preferred to the spatial error modelling framework in studying the convergence of Italian provinces.

4 CONCLUDING REMARKS

Evidences currently available in the main literature on convergence are based mainly on cross-sectional and panel data models. Both of these frameworks present some problems. In particular, the presence of omitted variables and heterogeneity suggests the need to reformulate the cross-sectional growth regressions. Panel data models constitute an alternative way to solve these problems. In the present paper we have focused our attention to the biases deriving from the presence of spatial autocorrelation effects not directly considered in the models. In particular, while some extensions to the

SAR and SEM models are present in the literature on cross-sectional regional convergence (see e.g. LeGallo et al. 2003)), to our knowledge, this is the first case in which spatial autocorrelation effects are included in an econometric panel framework. Following the work by Elhorst (2003), we have estimated two different models of spatial panel data namely (a) the spatial lag model, which incorporates spatial dependence in the form of a spatial lag variable, and (b) the spatial error model, in which spatial effects are incorporated in the distribution of the error term. Our results show that controlling both for spatial effects and for spatial autocorrelation allows us to be more confident that spatial autocorrelation represents a genuine regional interaction effects, rather than absorbing heterogeneity and the effect of omitted variables, as in the standard cross-sectional models.

Our empirical analysis focused on the convergence of per-capita income in the Italian provinces over 40 years. Following a consolidated evidence, we have considered a structural break in the growth path of Italian provinces at the beginning of the seventies. All models over the entire period and over two different subperiods (1951-1970 e 1970-2000) have confirmed this fact. In fact, the growth rate is very high during the first years and drops dramatically after 1970.

The speed of convergence estimated by using the spatial lag model is much lower than that obtained with the classical fixed-effect specification. A decrease in the β parameter referred to the initial condition, can be traced back to the introduction of a spatial lag term in the model, and indirectly confirms the positive effect of factor mobility, trade relationships, and knowledge spill-over on regional convergence. The results obtained using the spatial error specification are more difficult to be interpreted.

The present paper may be considered as the point of departure for further researches in this field. Future extension of the present framework can be directed towards the use of dynamic spatial panel data models accounting for both temporal and spatial lagged effects. A further extension would be the use of non-parametric methods to allow non-linearity in the parameters of the growth process.

REFERENCES

Anselin L., 1988, Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht.

Anselin L., and Bera A.K., 1998, Spatial dependence in linear regression models with an introduction to spatial econometrics, in Hullah A. and D.E.A. Gelis (eds.) Handbook of Applied Economic Statistics, Marcel Deker, New York, 237-290.

Anselin L. and Rey S.J., 1991, Properties of test for spatial dependence, Geographical Analysis, 23, 112-131.

Arbia G., 2005, Introductory Spatial Economtrics with application to regional convergence, Springer-Verlag (Advances in Spatial Sciences), Berlin (Forthcoming)

Arbia G. and Basile R., 2005, Spatial Dependence and Non-linearities in Regional Growth Behaviour in Italy, Statistica, forthcoming.

Arbia G., and Paelinck J.H.P., 2003, Economic convergence or divergence? Modelling the interregional dynamics of EU regions 1985 – 1999, Geographical Systems, 5, 1-24.

Baltagi B.H., 2001, Econometric analysis of panel data, 2nd edition Chichester, UK: Wiley.

Barro R.J. and Sala-i-Martin X., 1992, Convergence, Journal of Political Economy, 100, 223-251.

Barro R.J. and Sala-i-Martin X., 1995, Economic Growth, McGraw Hill, New York.

Barro R.J. and Sala-i-Martin X., 1997, Technological diffusion, convergence and growth, Journal of Economic Growth, 2, 1-16.

Basile, R. and Gress, B., 2004, Semi-parametric Spatial Auto-covariance Models of Regional Growth Behaviour in Europe, paper accepted at the Third Spatial Econometric Workshop, Strasbourg, June 9, 2004.

Durlauf S.N., and Quah D., 1999, The New Empirics of Economic Growth, in J.B. Taylor e M. Woodford (eds.), Handbook of Macroeconomics, vol. IA, Cap. 4, North-Holland, Amsterdam.

Elhorst J.P., 2003, Specification and Estimation of Spatial Panel Data Models, International Regional Sciences Review, 26, 244-268.

Hsiao C., 1986, Analysis of panel data, Cambridge, UK: Cambridge University Press.

Islam N., 1995, Growth empirics: A panel data approach, Quarterly Journal of Economics, 4, 1127-1170.

Islam N., 2003, What have we learnt from the convergence debate?, Journal of Economic Surveys, 17, 309-362.

Klevmarken N.A., 1989, Panel Studies: what can we learn from them?, European Economic Review 33, 523-529.

Kubo Y., 1995, Scale economies, regional externalities and the possibility of uneven regional development, Journal of Regional Sciences, 35, 318-328.

LeGallo J., Ertur C., and Baoumont C., 2003, A spatial Econometric Analysis of Convergence Across European Regions, 1980-1995, in Fingleton, B (ed.), European Regional Growth, Springer-Verlag (Advances in Spatial Sciences), Berlin.

Magrini S., 2003, Regional (Di)Convergence, in V. Henderson and J.F., Thisse (eds.), Handbook of Regional and Urban Economics, Volume 4.

Mankiw N.G., Romer D. and Weil D.N., 1992, A contribution to the empirics of economic growth", Quarterly Journal of economic, 107, 407-437.

Martin T., and Ottaviano G., 1999, Growing locations: Industry location in a model of endogenous growth, European Economic Review, 43, 281-302.

Meliciani V., and Peracchi F., 2001, Convergence in per-capita GDP across European regions a reappraisal, Working Paper.

Quah D., 1997, Empirics for growth and distribution: stratification, polarization, and convergence clubs, Journal of Economic Growth, 2, 27-59.

Rey S.J., 2000, Spatial empirics for economic growth and convergence, Working Papers of San Diego State University.

Rey S.J. and Montouri B.D., 2000, US regional income convergence: a spatial econometric perspective, Regional Studies, 33, 143-156.

Solon G.S., 1989, The value of panel data in economic research, in D. Kesprzyk, G.J. Duncan, G. Kalton and M.P. Singh, eds., Panel Surveys, John Wiley, New York, 486-496.

Solow R.M., 1956, A contribution to the theory of economic growth, Quarterly journal of economics, 70, 65-94.

Swan T.W., 1956, Economic growth and capital accumulation, Economic Record, 32, 334-361.

Working Papers available:		
n. 30/03	F. KOSTORIS PADOA SCHIOPPA	Il mercato e le politiche economiche in Italia
n. 31/03	S. DE NARDIS C. VICARELLI	The Impact of Euro on Trade: the (Early) Effect Is not So Large
n. 32/03	S. LEPROUX	L'inchiesta ISAE-UE presso le imprese del commercio al minuto tradizionale e della grande distribuzione: la revisione dell'impianto metodologico
n. 33/03	G. BRUNO C. LUPI	Forecasting Euro-area Industrial Production Using (Mostly)\ Business Surveys Data
n. 34/03	C. DE LUCIA	Wage Setters, Central Bank Conservatism and Economic Performance
n. 35/03	E. D'ELIA B. M. MARTELLI	Estimation of Households Income from Bracketed Income Survey Data
n. 36/03	G. PRINCIPE	Soglie dimensionali e regolazione del rapporto di lavoro in Italia
n. 37/03	M. BOVI	A Nonparametric Analysis of the International Business Cycles
n. 38/03	S. DE NARDIS M. MANCINI C. PAPPALARDO	Regolazione del mercato del lavoro e crescita dimensionale delle imprese: una verifica sull'effetto soglia dei 15 dipendenti
n. 39/03	C. MILANA ALESSANDRO ZELI	Productivity Slowdown and the Role of the Ict in Italy: a Firm-level Analysis
n. 40/04	R. BASILE S. DE NARDIS	Non linearità e dinamica della dimensione d'impresa in Italia
n. 41/04	G. BRUNO E. OTRANTO	Dating the Italian Business Cycle: a Comparison of Procedures
n. 42/04	C. PAPPALARDO G. PIRAS	Vector-auto-regression Approach to Forecast Italian Imports
n. 43/04	R. DE SANTIS	Has Trade Structure Any Importance in the Transmission of Currency Shocks? An Empirical Application for Central and Eastern European Acceding Countries to EU
n. 44/04	L. DE BENEDICTIS C. VICARELLI	Trade Potentials in Gravity Panel Data Models

Working Papers available:		
n. 45/04	S. DE NARDIS C. PENSA	How Intense Is Competition in International Markets of Traditional Goods? The Case of Italian Exporters
n. 46/04	M. BOVI	The Dark, and Independent, Side of Italy
n. 47/05	M. MALGARINI P. MARGANI B.M. MARTELLI	Re-engineering the ISAE manufacturing survey
n. 48/05	R. BASILE A. GIUNTA	Things change. Foreign market penetration and firms' behaviour in industrial districts: an empirical analysis
n. 49/05	C. CICCONI	Building smooth indicators nearly free of end-of-sample revisions
n. 50/05	T. CESARONI M. MALGARINI G. ROCCHETTI	L'inchiesta ISAE sugli investimenti delle imprese manifatturiere ed estrattive: aspetti metodologici e risultati
n. 51/05	G. ARBIA G. PIRAS	Convergence in per-capita GDP across European regions using panel data models extended to spatial autocorrelation effects
n. 52/05	L. DE BENEDICTIS R. DE SANTIS C. VICARELLI	Hub-and-Spoke or else? Free trade agreements in the "enlarged" European Union
n. 53/05	R. BASILE M. COSTANTINI S. DESTEFANIS	Unit root and cointegration tests for cross-sectionally correlated panels. Estimating regional production functions
n. 54/05	C. DE LUCIA M. MEACCI	Does job security matter for consumption? An analysis on Italian microdata