Empirical Economics (2008) 34:105–121
DOI 10.1007/s00181-007-0168-8

ORIGINAL PAPER

A spatially filtered mixture of β-convergence regressions for EU regions, 1980–2002

Michele Battisti · Gianfranco Di Vaio

Received: 15 July 2006 / Accepted: 28 June 2007 / Published online: 26 October 2007

Abstract Assessing regional growth and convergence across Europe is a matter of primary relevance. Empirical models that do not account for structural heterogeneities and spatial effects may face serious misspecification problems. In this work, a mixture regression approach is applied to the β-convergence model, in order to produce an endogenous selection of regional growth patterns. A priori choices, such as North–South or centre-periphery divisions, are avoided. In addition to this, we deal with the spatial dependence existing in the data, applying a local filter to the data. The results indicate that spatial effects matter, and either *absolute*, *conditional*, or *club* convergence, if extended to the whole sample, might be restrictive assumptions. Excluding a small number of regions that behave as outliers, only a few regions show an appreciable rate of convergence. The majority of data show slow convergence, or no convergence at all. Furthermore, a dualistic phenomenon seems to be present inside some States, reinforcing the "diverging-convergence" paradox.

Keywords Regional growth · Convergence patterns · Mixture regression · Spatial effects

JEL Classification C21 · O40 · R11

A previous version of the paper was presented at the International Workshop on Spatial Econometrics and Statistics, May 25–27, 2006, Rome, Italy. We wish to thank all the participants for their useful comments. Our acknowledgements to Rolf Turner and two anonymous referees for technical advice.

M. Battisti · G. Di Vaio (✉)
Dipartimento di Scienze Economiche e Aziendali,
LUISS Guido Carli, Viale Pola 12,
00198 Rome, Italy
e-mail: gdivaio@luiss.it

1 Introduction

Assessing regional convergence across Europe, in terms of per capita income or product, is a relevant matter, not only to verify what the growth theories predict, but also to evaluate the effectiveness of the Cohesion Policies. The expectations of New Entrants, indeed, require feasible answers from the policy makers. After the recent enlargement, economic disparities increased dramatically. The Union's ten richest regions each have a GDP equal to 189% of the EU-25 average, while for the ten poorest regions this indicator is equal only to 36%. In the New Member States, 90% of the total population live in regions with a per capita level of income below 75% of the Community average, which is the admissibility threshold to receive the Objective 1 Structural Funds (Commission of the European Communities 2005).

In the last few years, there has been much debate on European integration, regional growth and convergence, and on Cohesion Policies (for a review see Funck and Pizzati 2003). On the one hand, the EC stresses the gains of integration and the positive role of regional policies, which sustain the economic growth of the regions lagging behind (European Commission 2001, 2004). On the other hand, some scepticism has been voiced (Boldrin and Canova 2001) and, consequently, some questions have been raised. Will EU citizens see their welfare equalised to Community averages? Or will their standard of living fall, subjected to growing inequalities? Will cohesion policies have a positive impact on growth and convergence? Or will these policies be ineffective, serving mainly as instruments of redistribution?

In this work we examine the regional convergence process across Europe, focusing on some relevant issues linked to the empirical analysis. In particular, we show that misspecification sources need to be carefully taken into account. The remainder of the paper is structured as follows: Sect. 2 reviews the literature on the European regional convergence and poses the problem of heterogeneity and spatial effects. Section 3 highlights the existence of spatial dependence among EU regions and, at the same time, presents the data. Section 4 introduces the proposed methodology. Section 5 shows the results and, finally, Sect. 6 gives the conclusions.

2 The empirics of regional convergence

Convergence hypotheses across countries or regions have been subjected to several theoretical interpretations. Following the taxonomy originally proposed by Galor (1996), three different definitions can be identified: (a) *unconditional* or *absolute* convergence, meaning that per capita incomes converge to a common level in the long-run, if structural homogeneities exist across the economies and their initial conditions do not matter;[1] (b) *conditional* convergence, meaning that per capita incomes converge to different levels in the long-run, if structural heterogeneities exist across the economies and their initial conditions do not matter; (c) *club* convergence, meaning that per capita

[1] This hypothesis seems to be the one that the EC is interested in, as Quah (1996b, p. 1048, note 4) already pointed out.

incomes converge to different levels in the long-run, if structural heterogeneities exist across the economies and their initial conditions do matter.

Much of the empirical analysis aimed at testing the validity of these hypotheses has been based on the measure of β-convergence, derived from the neoclassical growth model of Solow–Cass–Koopmans (see Barro and Sala-i-Martin 2004). As is well known, the measure refers to the tendency for poor economies to grow faster than rich ones, i.e. to "catch up", and is related to ranking dynamics in the sample distribution (Sala-i-Martin 1996a). Other measures are often used, which look at the reduction of distributional dispersion over time, such as sigma-convergence[2] (Barro and Sala-i–Martin 1992), or the evolution of the entire distribution, such as the transition matrix approach[3] (Quah 1993a, 1996a). These concepts, however, are less suited to the questions assessed in this work and to the methodology adopted, hence we prefer to focus on β-convergence.

Empirical β-convergence models usually take the form of a cross-country/region growth regression

$$g_i = a - bx_i + u_i, \quad i = 1, \ldots, n, \tag{1}$$

where $g_i = [\ln(y_{i,T}) - \ln(y_{i,0})]/T$ is the average growth rate of the economy i's per capita income between time 0 and T, a is a constant, $b = (1 - e^{-\beta T})/T$ is a convergence coefficient, $x_i = \ln(y_{i,0})$ is the log of the economy i's initial level of per capita income, and $u_i \sim N(0, \sigma^2)$ is an error term with the usual properties (see Durlauf et al. 2005). A positive value of the parameter β is supportive of convergence, and provides the rate at which the economy approaches the steady-state.[4]

Early empirical studies (Barro and Sala-i-Martin 1991; Sala-i-Martin 1996b) estimated Eq. (1) without control variables, based on two strong homogeneity assumptions. First, the constant term was considered inclusive of technological progress (γ_i), the steady-state value of effective per capita output ($\tilde{y}_i^*$), and initial efficiency ($A_{i,0}$), namely

(i) $a = a_i = \gamma_i + (1 - e^{-\beta_i T})/T \cdot \ln(A_{i,0} \cdot \tilde{y}_i^*), \quad \forall i.$

Second, the convergence coefficient was considered constant across the economies, that is

(ii) $b = b_i = (1 - e^{-\beta_i T})/T, \quad \forall i.$

Equation (1) estimated with assumptions (i) and (ii) can be seen as a test for type (a) convergence, a positive $\hat{\beta}$ implying that poor regions *unconditionally* grow faster

[2] As random shocks may produce criss-crossing or overshooting effects, β-convergence is a necessary but not sufficient condition for sigma-convergence (see Barro and Sala-i-Martin 2004).

[3] Some researchers prefer this approach, since it provides a more complete set of information. β-convergence, in fact, suffers from the so-called "Galton's fallacy", hence it may be consistent with a stationary distribution over time (Quah 1993b; Hart 1995).

[4] Estimates are usually obtained calculating $\hat{b}$ by ordinary least squares (OLS), and re-parameterizing $\hat{\beta} = -\ln(1 - \hat{b}T)/T$. Estimation may also be done with non linear least squares (NLS). However, the use of one technique rather than another does not lead to appreciable statistical discrepancies (on this point, see Abreu et al. 2005a).

than rich ones, at the same rate, towards a unique steady-state independently from the initial conditions. Barro and Sala-i-Martin (1991, p. 154), analysing a sample of 73 Western European regions over the period 1950–1985, found an "empirical regularity that the rate of β convergence is roughly 2% a year in a variety of circumstances [...] the half-life of this convergence process is 35 years".[5] In the last edition of their book, they concluded that "absolute β convergence is the norm for these regional economies" (Barro and Sala-i-Martin 2004, p. 496).

Tests of *absolute* β-convergence are plausible when the object of study is *within*-country convergence. In that case, regional economies share common steady-states, being affected by similar saving rates, preferences, governmental policies, property rights, infrastructures, and so on. In the case of *between*-country convergence, however, type (a) convergence appears quite unrealistic, since regions belonging to different countries may not show a common steady-state. As Solow (1999, p. 640) argued, "there is nothing in growth theory to require that the steady-state configuration be given once and for all [...] the steady-state will shift from time to time [and, we say, from space to space] whenever there are major technological revolutions, demographic changes, or variations in the willingness to save and invest".

If the determinants of steady-state are not constant across regions, it follows that $a_i \neq a, \forall i$, leading assumption i) to fail.[6] Mankiw et al. (1992) solved the problem by relaxing assumption (i) in two ways. First, in the presence of heterogeneity, $\tilde{y}_i^*$ can be included among a set of control variables added to Eq. (1). Second, if $A_{i,0}$ reflects not only the initial technology, but also resource endowment, climate, institution, and other region-specific factors affecting growth, it may be constituted by a common and a random component, $\ln(A_{i,0}) = \ln(A_0) + e_i$, where A_0 is the common factor and e_i is the specific effect. The error term is now equal to $u_i = (1 - e^{-\beta T})/T \cdot e_i + \varepsilon_i$, $\tilde{y}_i^*$ being independent of the error term. Assumption (i) is hence replaced by

$$\text{(iii)} \quad a = a_i = \gamma_i + (1 - e^{-\beta_i T})/T \cdot \ln(A_0), \quad \forall i.$$

Equation (1), with the assumptions (ii) and (iii), and the inclusion of control variables, implies homogeneity in the convergence parameter, the initial efficiency, and the technological progress. Steady-states determinants, on the contrary, are allowed to be heterogeneous. Estimation can be considered a test for type (b) convergence, with a positive $\hat{\beta}$ meaning that poor regions grow *conditionally* faster than rich ones, at the same rate, towards different steady-states.

Armstrong (1995), testing for *absolute* and *conditional* convergence, either *within* or *between*-country, on a sample of 85 EU regions over the period 1950–1990, found significant discrepancies between the two hypotheses. The rate of *between*-country *absolute* convergence was about 1% per annum, much slower than the 2% found by previous studies. In fact, a rate of 2% was only found during the post-War period, for

[5] The so-called half-life condition is given by $e^{-\beta T} = 1/2 \Rightarrow T = \ln(2)/\beta$. If the speed of convergence is equal to 2% per year, it follows that $T \cong 0.69/0.02 \cong 35$, hence the economy fills half the gap in about 35 years.

[6] Steady-state variables might be comprised in the error term, $u_i = (1 - e^{-\beta_i T})/T \cdot \ln(\tilde{y}_i^*) + \varepsilon_i$, where ε_i is a random component. However, if those variables were related to initial income levels, and they had an impact on growth, $\tilde{y}_i^*$ would be an omitted variable and the coefficient $\hat{b}$ would be biased (Bernard and Durlauf 1996).

within-country *conditional* convergence.[7] On the other hand, the years following the oil crisis saw a decrease of annual convergence rates ranging from 0.8 to 1%.

Type (a) and (b) convergence tests, however, have been criticized under many respects. From a general cross-country perspective, parameters such as heterogeneity, outliers, and measurement errors have been highlighted (Temple 1998, 2000). Looking at the European regional experience, some researchers (Martin 1998; Petrakos et al. 2005) sustained that the convergence process does not follow a homogeneous pattern of growth.[8] In this case, testing for types (a) or (b) convergence would be misleading, if the "true" convergence process saw the regions converging at different rates towards different income levels.

The existence of structural heterogeneities may be compatible, for instance, with the presence of multiple regimes in cross-country growth behaviour identified by Durlauf and Johnson (1995), or with the convergence clubs—the "twin-peaks"—in world income distribution detected by Quah (1997). On the one hand, country-specific constraints on the adoption of technologies may affect the efficiency of regional economies and produce structural heterogeneities, as verified world-wide by Durlauf et al. (2001). In this case, assumption (iii) does not hold, because

(iv) $a \neq a_i = \gamma_i + (1 - e^{-\beta_i T})/T \cdot \ln(A_0), \quad \forall i.$

On the other hand, growth models similar to the one developed by Azariadis and Drazen (1990) assume that spillovers due to physical or human capital accumulation cause threshold effects, which produce shifts in the aggregate production function, leading to multiple, locally stable, steady-state equilibria—i.e. to different convergence "clubs" (see Durlauf and Quah 1999). A threshold value in the income level, $\bar{y}$, implies $\beta_i = \beta_1$, if $y_{i,0} \prec \bar{y}$, and $\beta_i = \beta_2$ otherwise. Hence, assumption ii) needs to be replaced by

(v) $b \neq b_i = (1 - e^{-\beta_i T})/T, \quad \forall i.$

If initially poor regions converge towards a lower income level,[9] then estimates of Eq. (1) with assumptions (iv) and (v) can be seen as tests of type (c) convergence. A positive $\hat{\beta}$ indicates that poor regions grow faster than rich ones, at different rates, towards different steady-states depending on their initial conditions.

The existence of *club* convergence across Europe has been recognized by several authors. Early studies imposed exogenous assumptions on the number of clubs, to emphasize geographical and distributional factors, such as North–South, centre–periphery, or rich–poor divisions. Neven and Gouyette (1995) split a sample of 142 EU regions, over the period 1980–1989, into a Northern and a Southern club. They found a very low rate of 0.53% *absolute* convergence for the whole sample, and no statistically significant convergence inside either of the two clubs. Only when country-specific effects are controlled for, the rate of convergence takes on significant values, comprised between 1.1 and 1.8%.

7 Country-specific dummies are used to control for heterogeneity in steady-states. The common practice of employing dummy variables is due to lack of data at a regional level.

8 Convergence *between* States—towards the outside—but not *within*—towards the inside—has been defined as the "diverging-convergence" phenomenon.

9 Falling into a "poverty trap".

One should, however, consider more fully endogenous criteria to detect the presence of clubs, rather than making exogenous choices which arbitrarily assign structural heterogeneities to different clubs. Canova (2004), for instance, adopted a predictive density approach to find convergence clubs in a sample of 144 EU regions, over the period 1980–1992. Avoiding a priori assumptions, he identified four homogeneous clubs, with different convergence rates and steady-states values highlighting North–South or poor–rich dimensions, the initial conditions influencing the probability of belonging to a club.

Furthermore, many authors have argued that, due to geographical spillovers, the distribution of regional per capita income across Europe tends to be influenced by physical location (Quah 1996c; López-Bazo et al. 1999; Le Gallo and Ertur 2003). Ertur et al. (2006) treated spatial problems in the context of club convergence. In a sample of 138 EU regions, over the period 1980–1995, they found that per capita income levels were highly spatially correlated. In particular, an exploratory spatial analysis (ESDA) revealed a division between Northern-rich regions and Southern-poor ones. Assuming the existence of heterogeneity across two different spatial regimes, and taking the spatial autocorrelation into account, they found no convergence in the Northern club, and an annual convergence rate of 2.9% in the Southern one.

The procedure followed by Ertur et al. (2006) is based on an exogenous assumption that structural parameters are heterogeneous across regions, due to their geographical locations. To our knowledge, a procedure that merges together an endogenous identification of convergence paths and spatial dynamics is not yet available either in the theoretical or the empirical literature. Working in this direction, we implement a strategy that leads to an endogenous selection of convergence regimes once spatial dependence effects have been taken into account. It can be considered a first step for future research.

3 Spatial dependence across European regions

Spatial dependence, if not properly modelled, leads to serious misspecification problems in linear regressions (Anselin 1988, 2001; Anselin and Bera 1998). In the cross-sectional growth framework, in which the observations are spatially organized, the existence of geographical spillovers may violate the assumption that the error terms from neighbouring regions are independent (Rey and Montouri 1999). The common practice is to explicitly incorporate in the regression a spatial component, in the form of a spatial error or a spatial lag (Arbia 2006). Another approach, as we will see later, is to filter out the spatial dependence.

A simple check of spatial dependence, in its weaker version of spatial autocorrelation, can be performed by means of Moran's I statistic. As is well known, the statistic can be expressed as

$$I = \frac{n}{q} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} x_i x_j}{\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j},$$

where w_{ij} is an element of a binary spatial weight matrix $\mathbf{W}$, x_i is a specific variable for observation i, n is the number of observations, q is a scaling factor equalling the sum of all the elements of the matrix. In this paper we use a row-standardized binary matrix, based on the k-nearest neighbouring regions, the elements of which are

$$\begin{cases} w_{ij}(k) = 0 & \text{if } i = j \\ w_{ij}(k) = 1 & \text{if } d_{ij} \leq d_i(k) \\ w_{ij}(k) = 0 & \text{if } d_{ij} > d_i(k) \end{cases}$$

where $d_i(k)$ is a critical cut-off distance, defined for each observation i, ensuring that every single region of the sample has the same number (k) of neighbours.

Abreu et al. (2005b) has shown that contiguity-based matrices are the most popular choice in the literature. In the case of the European regions, however, these kind of matrices leave the islands unconnected to the continent, hence distance-based specifications have been preferred in applied works (Le Gallo and Ertur 2003; Le Gallo and Dall'erba 2006). We chose a specification with $k = 20$, since it cancels out the spatial autocorrelation in the filtered series.[10] Such a specification is also able to link Cyprus with the Greek regions, which in turn are connected to Italy; Ireland with the UK, which is connected to continental Europe; Sicily and Sardinia with continental Italy; Corsica with the continental French regions.[11]

Data on per capita GDP, expressed in Euros at 1995 prices, are taken from Cambridge Econometrics, *European Regional Database*, 2004. We work with two samples. The larger sample includes 242 NUTS-2 regions from EU-25,[12] covering the period 1991–2002, while the smaller one comprises 190 NUTS-2 regions from EU-15, covering the period 1980–2002. Figure 1 shows standardized scatter-plots for the smaller sample based on Moran's I of (a) the log of per capita GDP, and (b) the average growth rate of per capita GDP.[13] A highly positive spatial correlation of per capita income levels among the European regions is clearly evident. The majority of the observations fall into the high–high (HH) or the low–low (LL) quadrant. In fact, rich (poor) regions are surrounded by rich (poor) regions. Spatial correlation of growth rates is also positive, albeit in a weaker form. Over the whole period, the spatial dynamic of the growth rates is unable to offset the spatial concentration of the economic activity. At the end of the period, the physical location of income levels is agglomerated as it was at the beginning. Moran' s I of per capita GDP in 2002 equals 0.6, only 1% point less than its value in 1980.

From the above statistics, we would expect that spatial dependence matters for the study of β-convergence in Europe. As a preliminary analysis we estimate Eq. (1) for both samples by standard OLS, and test the existence of spatial autocorrelation among the regression residuals. We split the EU-15 sample into two sub-periods of equal length, 1980–1991 and 1991–2002. The breakdown is useful for at least two

[10] However, other weight matrices, with $k = 10, 15$, produced very similar results in the mixture. See the next section for the filtering procedure.

[11] The weight matrix was obtained using the GeoDa software package (Anselin 2005).

[12] Inclusive of German Ex-Länder. The list of regions is available upon request.

[13] To save space, we do not show here the figures for the larger sample. The results, however, are very similar.

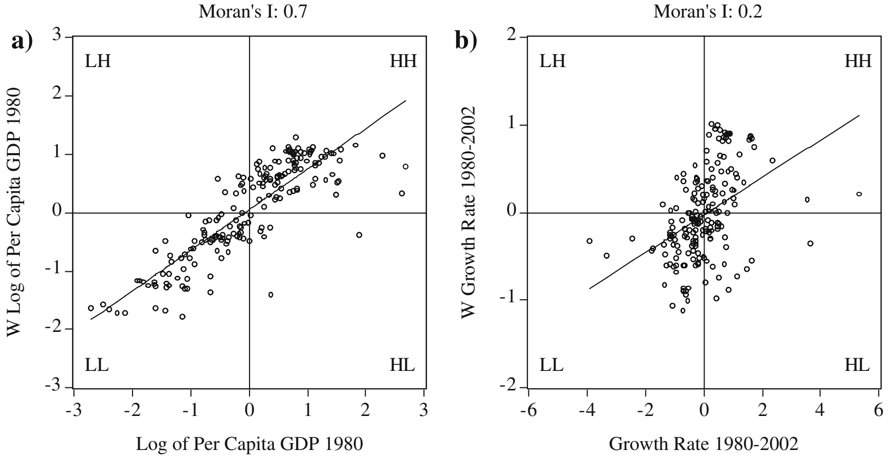

Fig. 1 Moran scatterplots (standardized) of **a** per capita GDP, 1980, **b** average growth rate of per capita GDP, 1980–2002. 190 NUTS-2 regions from EU-15. **Data Source** Cambridge Econometrics, *European Regional Database*, 2004

Table 1 *Absolute* β-convergence

	EU-15			EU-25
	1980–2002	1980–1991	1991–2002	1991–2002
$\hat{a}$	0.076 (0.011)	0.040 (0.016)	0.102 (0.015)	0.093 (0.012)
$\hat{b}$	0.006 (0.001)	0.002 (0.002)	0.009 (0.002)	0.008 (0.001)
Half-life	108 years	310 years	69 years	79 years
R^2	0.12	0.01	0.15	0.17
Log-likelihood	686	623	642	700
Obs	190	190	190	242
Diagnostics for spatial autocorrelation				
Moran's I err-u*	0.12 (0.00)	0.16 (0.00)	0.17 (3.00)	0.07 (0.00)
Moran's I err-f*	−0.03 (0.24)	−0.01 (0.87)	−0.02 (0.42)	−0.02 (0.44)

* P values within parentheses
Standard errors within parentheses. OLS estimates and diagnostics for spatial autocorrelation
Data source Cambridge Econometrics, *European Regional Database*, 2004

reasons. First, in the Nineties major institutional changes, such as the implementation of Cohesion policies and the establishment of the adhesion criteria to EMU, may have had an impact on the convergence process of the EU regions. Second, such a breakdown makes it possible to compare the smaller sample with the larger one, to see if the regions which first joined the Union experienced different convergence patterns. The results are shown in Table 1.

Over the whole period, the convergence coefficient for the EU-15 sample is highly significant. Its magnitude, however, is only about one fourth of the "empirical norm" of 2%. Actually, the convergence rate equals about 0.6% per annum, leading to a half-life of 108 years.[14] The poorest regions are thus supposed to fill half the gap with the richest ones in over a century. Looking at the two sub-periods, it can be clearly seen that the bulk of convergence is given by 1990s. During 1980s, the convergence coefficient is very slow and not statistically significant, pointing to a lack of convergence. Enlarging the sample to the EU-25 regions does not change the picture very much, since convergence decreases only slightly when compared to the EU-15 sample.

Finally, the bottom part of Table 1 shows the spatial autocorrelation diagnostics. The test is based on Moran's I statistic applied to regression residuals. Under some regularity conditions, the distribution of the test corresponds to the standard normal (Anselin, 1988). The results highlight a significant spatial autocorrelation among the residuals,[15] for both samples, either over the whole period or for the two sub-periods. Interestingly, during 1990s, the spatial dependence is higher for the EU-15 regions, meaning that per capita income is less concentrated in the enlarged Europe. This preliminary spatial analysis shows that spatial phenomena can be relevant for the study of EU regional convergence, and that they should be taken into account in the cross-sectional framework, in order to avoid misspecification problems. *Absolute* β-convergence, tested by standard OLS with no spatial specification, suffers from many shortcomings which invalidate its ability to explain the regional growth processes.

4 The spatially filtered mixture of regressions

Given the spatial influence highlighted in the previous section, our interest here lies in an endogenous determination of heterogeneity in regional convergence patterns, once the spatial dependence in the data has been properly treated. We avoid a priori restrictions, such as geographical (North–South or centre–periphery), or exploratory (based upon spatial association indices) divisions. To this end, we use a spatially filtered mixture regression approach (for mixture densities, see Titterington et al. 1985; Wedel and Kamakura 1998; MacLachlan and Peel 2000). Previous attempts to apply mixture densities, or mixture regressions, to convergence analysis are found in the works of Paap and Van Djik (1998), Tsionas (2000), and Bloom et al. (2003). Those studies, however, do not deal with spatial related questions, and differ from ours as regards several other aspects.

Let us begin by considering spatial dependence. As we have seen in the previous section, if spatial dependence is present across a sample, OLS estimates are biased or not efficient. In the case of spatial effects influencing the errors in Eq. (1), statistical inference based on OLS is not reliable, because assumption of errors independence from neighbouring regions may be violated. We treat potential sources of misspecification in the β-convergence framework by isolating the spatial correlation by means of

[14] The result may not be very interesting in terms of policy implications. On how to calculate $\hat{\beta}$ see note 4 in Sect. 2.

[15] The test was carried out on the residuals of the unfiltered series (err-u), as well as on the filtered series (err-f) obtained with the filtering procedure described in the next section. Once spatial correlation is filtered out from the series, the test become not significant.

a local filter.[16] Two filtering procedures have been shown to give the same empirical results in "cleaning up" the spatial effects from geographically organized variables (Getis and Griffith 2002).

We use the Getis local filter $Gi(d)$ (Getis 1995) constructed as, for each spatial unit i,

$$G_i(d) = \frac{\sum_{j=1}^{n} w_{ij}(d)x_j}{\sum_{j=1}^{n} x_j}, \quad i \neq j$$

where x is the original unfiltered variable and w_{ij} is the element of the spatial weight matrix related to the j neighbouring regions comprised within the distance threshold d. The filtered variable x^F (where F stands for filtered) can then be obtained as

$$x_i^F = x_i \frac{\left[\sum_{j=1}^{n} w_{ij}(d)\right]}{(n-1)} \bigg/ G_i(d),$$

while the residual spatial component can be defined as $x^S = x - x^F$ (S stands for the spatial component). Applying this filter to the series makes the OLS estimates consistent.

We test different specifications of $\mathbf{W}$ ($k = 10, 15, 20$) until Moran's I on regression residuals becomes not significant. The specification with $k = 20$ gives the desired outcome.[17]

A similar two-step procedure, with a spatial filtering in the first step and a panel regression in the second, is implemented, for example, by Badinger et al. (2004). Our approach uses the same first step, where the spatial dependence is "cleaned up", while the second step is based on the application of the mixture regression model to the filtered variables. At the end of the procedure we obtain a transformed version of Eq. (1)

$$g_i^F = a - bx_i^F + u_i \tag{2}$$

the least squares estimation of which is consistent.[18]

In the second step we employ the mixture regression model to detect the existence of regional convergence patterns. Assume that the "true" density function of a population is a mixture of several functions, one for each pattern with different parameters, weighted by the probability of belonging to a specific pattern. If the population is divided into k groups,[19] the number of groups being unknown and the sum

[16] In a previous version of this work we used a global filter obtained by means of a spatial parameter estimated in a spatial error model (see Anselin and Bera 1998). Results about convergence rates and patterns identification were quite similar, but that procedure is less consolidated in the literature and it causes inference problems in the second step deriving from estimating mixtures of equation with generated regressors (see Pagan 1984).

[17] See the row labelled Moran's I err-f in Table 1.

[18] We estimate Eq. (1) without control variables, because we admit heterogeneity in steady-states, allowing for different intercepts across clubs.

[19] We refer to patterns, groups, or regimes without distinction.

of the probabilities of belonging to one of these being equal to one for each observation, then, according to the total probability theorem, the conditional distribution function is

$$f(g_i^F|\theta) = \sum_{s=1}^{k} \psi_s f_s(g_i^F|\theta_s) \tag{3}$$

where g_i^F is the dependent variable, ψ_s is the probability of belonging (a priori) to a regime s, with $s = 1, \ldots, k$, and θ is a vector of parameters. Once ψ_s has been estimated, the subsequent probability that observation i comes from s has to be computed by Bayes theorem.

Consider the function of the filtered variable g_i^F as normal. The density function conditional to belonging to the regime k is

$$f(g_i^F|s = k, \theta) = (2\pi\sigma_k^2)^{1/2} e^{-\left[g_i^F-(a_k-b_k x_i^F)\right]^2/2\sigma_k^2}. \tag{4}$$

In this way, the component represented by $(a_k - b_k x_i^F)$ gives us a linear predictor that replaces the population mean of the group. From Bayes rule, it is straightforward to extract the unconditional probability of g_i^F for $s = k$ as a joint probability, that is the product of the conditional probability and the marginal probability of belonging to a club. This latter is equal to ψ_k, hence the joint probability is $\psi_k f(g_i^F|s = k, \theta)$. Summing all the values of s gives the unconditional density of g_i^F

$$f(g_i^F, \theta) = \sum_{s=1}^{k} \psi_s (2\pi\sigma_s^2)^{1/2} e^{-\left[g_i^F-(a_s-b_s x_i^F)\right]^2/2\sigma_s^2}. \tag{5}$$

The vector of parameters θ, which also contains the weights ψ_s, is unknown. A simple way to solve this type of missing data problem is through the Expectation–Maximization (EM) algorithm. The solution is to find an initial value of the parameters, then compute the density for these parameters, and re-compute the final θ, by maximization of the log-likelihood. The algorithm therefore has two alternated steps: in the first (expectation), it computes the density function for the chosen parameters, while in the second (maximization), it derives the estimation of the parameters a_s, b_s, and σ_s^2. In the case of a linear mixture regression, De Sarbo and Cron (1988) show how the second step is equivalent to performing k weighted least squares regressions, where the weights are the roots of the probabilities of belonging to a club.

We begin with random starting probabilities, then we update the probabilities step by step. This strategy could have two types of shortcoming. First, the results might depend on the initial probabilities. Second, the maximization of the log-likelihood could converge in a local optimum. To avoid these problems, and to be reasonably confident that our estimates do not correspond to a local maximum, we tried 500 different starting values. We chose the highest value of log-likelihood, which also helps to determine the number of components in the mixture, as we will see below.

Finally, each region is attributed to a regime, if the probability of belonging to that regime is higher than the probability to belong to the others. A last point refers to inference considerations. Since standard errors in the EM algorithm are not used to iterate (Wedel and Kamakura 1998), when the algorithm converges to the final value, an estimation of the covariance matrix is given by the Fisher information matrix. In the case of the EM algorithm, Louis (1982) finds this observed information matrix as the difference between two matrices—the total information matrix and the missing data information matrix. Turner (2000) shows computational details for mixture regression models.[20]

5 Results

This section shows the results obtained with the two-step methodology described above. The choice of the mixture's components number is based upon two complementary rules. On the one hand, we look at the improvements given by additional components in the log-likelihood, through a set of information criteria. On the other hand, we consider a meaningful interpretation of the data. The latter rule regards either the existence of appreciable differences in the significance of the parameters, or the certainty of the regions' attribution to specific regimes.[21] The logic of the former rule is to penalize the increase of the components in the mixture, in order to avoid an excessive—and useless—number of parameters. The AIC (Akaike Info Criterion) is the less restrictive criterion, so it generally selects less parsimonious specifications. On the contrary, the BIC (Bayesian Info Criterion) is the most restrictive one. The MAIC (Modified Akaike Info Criterion) falls in-between (for a detailed description, see Hawkins et al. 2001).

According to the criteria reported in Table 2, a specification with only one component is a choice not supported by the data, while a three-components specification is the better approximation for all the samples, except for the EU-15, 1991–2002, for which all the criteria indicate a choice of two regimes.[22] Table 3 reports the results obtained by the mixture regression, applied to the spatially filtered variables. Generally, the two-components specification selects a small group with very fast convergence rates, ranging from about 1.3–3.2%, becoming in the Nineties a divergence regime. Such a group, comprising from 12 to 26% of the data, seems to behave like an "outliers bin", since it collects regions with particular growth experiences. The majority of the regions fall into a regime characterized by a very slow convergence rate, equal to about

[20] The observed information matrix can be computed as the difference between the total information matrix and the missing data information matrix. This matrix is then inverted to extract the square roots of the elements from the main diagonal, in order to obtain the standard errors of the parameters. The matrix dimension is given by the number of parameters minus one, because one of the weights is a linear combination of the others.

[21] In some cases the attribution may be less precise (i.e. there are regions with a 100% probability of belonging to a group, and regions with only a 51% probability). Generally, we find about 90% of regions that are attributed with a high difference with respect to the alternative regime.

[22] To save space, we do not show the tables with four and five components. The criteria, however, do not record any improvement with respect to the three components specification.

Table 2 Decision criteria for the mixture's components number

	Log-likelihood	AIC	MAIC	BIC
EU-15: 1980–2002				
1 Component (OLS)	693	−1,380	−1,377	−1,370
2 Components	730	−1,446	−1,439	−1,423
3 Components	736	−1,450	−1,439	−1,415
EU-15: 1980–1991				
1 Component (OLS)	629	−1,252	−1,249	−1,243
2 Components	660	−1,306	−1,299	−1,283
3 Components	669	−1,316	−1,305	−1,281
EU-15: 1991–2002				
1 Component (OLS)	637	−1,268	−1,265	−1,258
2 Components	653	−1,293	−1,286	−1,270
3 Components	657	−1,292	−1,281	−1,256
EU-25: 1991–2002				
1 Component (OLS)	732	−1,458	−1,455	−1,447
2 Components	750	−1,485	−1,478	−1,461
3 Components	754	−1,487	−1,476	−1,448

Data source Cambridge Econometrics, *European Regional Database*, 2004

0.5% per year. This regime does not show substantial differences across samples and periods. The half-lives, in all cases, exceed a century, being this way not particularly interesting in terms of cohesion.

The three-components specification makes it clear that the slow convergence regime is constituted by a smaller group of fast convergence rates, ranging from about 1.2 to 4.7%, and a larger regime with slower or absent convergence.[23] This latter comprises one half or two thirds of the regions in the samples, depending on the periods considered. The convergence regime shrinks during the Nineties, a decade that saw the fostering of the EU integration process due to the Maastricht Treaty and adhesion to EMU. Overall results, interestingly, seem to be in line with recent empirical investigations on the subject (see Meliciani and Peracchi 2006).

As a final step, we proceed with a visual inspection of the regions, allocated by the mixture to the different regimes, net of spatial effects (Fig. 2). As an illustration, the EU-25 sample is shown. Regions in dark grey make up the faster convergence group, the light grey group is the slower convergence regime, while white indicates the regions which do not converge. The map depicts the existence of dualistic phenomena in many States. Rich and poor regions in Ireland, the UK, France, Germany, Spain, Italy, as well as in many of the New Member States, fall into opposite regimes that are not converging between themselves. Such phenomena reinforce the "diverging-converging" paradox (i.e. convergence between States, but not within).

[23] In the case of EU-15, 1980–1991, the convergence rate remains similar to the two-components specification, being equal to about 0.6% per year.

Table 3 Spatially filtered mixtures regressions

	2 Components		3 Components		
	Regime 1	Regime 2	Regime 1	Regime 2	Regime 3
EU-15: 1980–2002					
$\hat{a}$	0.324 (0.120)	0.062 (0.015)	0.340 (0.120)	0.183 (0.023)	0.037 (0.020)
$\hat{b}$	0.032 (0.013)	0.005 (0.002)	0.034 (0.013)	0.018 (0.002)	0.002 (0.002)
Half-life (years)	5	131	11	30	339
Weight	12%	88%	12%	22%	66%
EU-15: 1980–1991					
$\hat{a}$	0.267 (0.108)	0.067 (0.021)	0.454 (0.048)	−0.072 (0.127)	0.075 (0.023)
$\hat{b}$	0.027 (0.012)	0.005 (0.002)	0.047 (0.005)	−0.010 (0.014)	0.006 (0.003)
Half-life (years)	22	135	10	–	112
Weight	26%	74%	15%	19%	66%
EU-15: 1991–2002					
$\hat{a}$	−0.061 (0.145)	0.064 (0.025)	−0.161 (0.179)	0.255 (0.111)	0.036 (0.030)
$\hat{b}$	−0.008 (0.015)	0.005 (0.003)	−0.019 (0.02)	0.025 (0.012)	0.002 (0.003)
Half-life (years)	–	135	–	24	343
Weight	20%	80%	14%	24%	62%
EU-25: 1991–20e02					
$\hat{a}$	0.144 (0.063)	0.069 (0.016)	0.222 (0.085)	0.137 (0.032)	0.012 (0.024)
$\hat{b}$	0.013 (0.007)	0.006 (0.002)	0.021 (0.009)	0.012 (0.003)	−0.000 (0.002)
Half-life (years)	49	112	29	54	–
Weight	23%	77%	16%	38%	46%

Standard errors within parentheses
Data source Cambridge Econometrics, *European Regional Database*, 2004

6 Conclusions

In this work we analysed regional convergence patterns, trying to avoid potential sources of problems due to spatial effects, parameter heterogeneity and outliers. The methodology adopted here shows that neither *absolute*, *conditional*, nor *club* convergence are the best hypotheses for explaining regional growth in Europe, over the period 1980–2002. Summarizing, a common specification for the whole sample is an assumption too much restrictive, convergence rates are far removed from the "empirical norm" of 2% per year, and spatial effects matter.

Once the data have been spatially filtered, the mixture endogenously identifies multiple, a-spatial, growth regimes. In the case of a three-components specification, generally one regime behaves as an "outlier bin", the other shows a sustained convergence rate, and the third, comprising the majority of the sample, shows no convergence at all. Many regions, whether inside "poor" or "rich" States, fall into the non convergence regime, where agglomeration factors, and increasing returns, might play a role. Such a mechanism reinforces the paradox of the so-called "diverging-convergence",

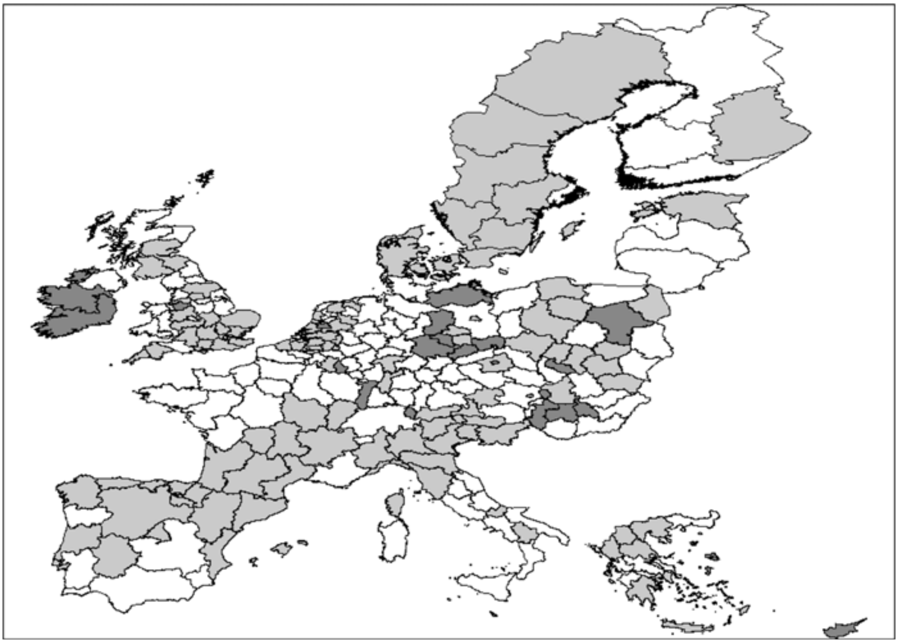

Fig. 2 Convergence patterns, net of spatial influence, EU-25: 1991–2002*. **Dark grey* fast convergence, *light grey* slow convergence, *white* no convergence

that is, the convergence between States but not within. Furthermore, a North–South division does not emerge, except in Italy's case, while a core-periphery dynamic seems a more plausible scenario.

Finally, since the main intent of this work was to take into account misspecification sources in the β-convergence framework, policy prescriptions cannot be easily drawn. However, some implications may be discussed. Since convergence rates do not vary very much between the two sub-periods, cohesion policies do not make a substantial difference. If anything, 1990s see an expansion of the non convergence area. Looking at the enlarged sample, regions belonging to the New Member States show different trends. In conclusion, over the period considered, regional growth dynamics does not seem to have followed a common pattern towards the convergence of per capita income across Europe.

References

Abreu M, De Groot HLF, Florax RJGM (2005a) A meta-analysis of β-convergence: the legendary 2%. J Econ Surv 19:389–420

Abreu M, De Groot HLF, Florax RJGM (2005b) Space and growth: a survey of empirical evidence and methods. Région et Développement 21:13–44

Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht

Anselin L (2001) Spatial econometrics. In: Baltagi BH (ed) A companion to Theoretical econometrics. Blackwell, Oxford, pp 310–330

Anselin L (2005) Exploring spatial data with GeoDa: a workbook. Center for spatial integrated social science, Santa Barbara

Anselin L, Bera AK (1998) Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A, Giles DEA (eds) Handbook of applied economic statistics. Marcel Dekker, New York, pp 237–289

Arbia G (2006) Spatial econometrics: statistical foundations and applications to regional convergence. Springer, Berlin

Armstrong HW (1995) Convergence among regions of the European Union, 1950–1990. Pap Reg Sci 74:143–152

Azariadis C, Drazen A (1990) Threshold externalities in economic development. Q J Econ 105:501–526

Badinger H, Müller WG, Tondl G (2004) Regional convergence in the European Union, 1985–1999: a spatial dynamic panel analysis. Reg Stud 38:241–253

Barro RJ, Sala-i-Martin X (1991) Convergence across states and regions. Brookings Pap Econ Act 1:107–182

Barro RJ, Sala-i-Martin X (1992) Convergence. J Polit Econ 100:223–251

Barro RJ, Sala-i-Martin X (2004) Economic growth, 2nd edn. MIT, Cambridge

Bernard AB, Durlauf SN (1996) Interpreting tests of the convergence hypothesis. J Econom 71:161–173

Bloom DE, Canning D, Sevilla J (2003) Geography and poverty traps. J Econ Growth 8:355–378

Boldrin M, Canova F (2001) Inequality and convergence in Europe's regions: reconsidering European regional policies. Econ policy 16:207–253

Canova F (2004) Testing for convergence clubs in income per capita: a predictive density approach. Int Econ Rev 45:49–77

Commission of the European Communities (2005) Communication from the Commission. Third progress report on cohesion: towards a new partnership for growth, jobs and cohesion. 17.5.2005 COM(2005) 192 final. Brussels

De Sarbo WS, Cron WL (1988) A maximum likelihood methodology for clusterwise linear regression. J Classif 5:249–282

Durlauf SN, Johnson PA (1995) Multiple regimes and cross-country growth behaviour. J Appl Econom 10:365–384

Durlauf SN, Quah DT (1999) The new empirics of economic growth. In: Taylor JB, Woodford M (eds) Handbook of macroeconomics, vol 1. North Holland, Amsterdam, pp 235–308

Durlauf SN, Kourtellos A, Minkin A (2001) The local Solow growth model. Eur Econ Rev 45:928–940

Durlauf SN, Johnson PA, Temple JRW (2005) Growth econometrics. In: Durlauf SN, Aghion P (eds) Handbook of economic growth. North Holland, Amsterdam, pp 555–677

Ertur C, Le Gallo J, Baumont C (2006) The European regional convergence process, 1980–1995: do spatial regimes and spatial dependence matter? Int Reg Sci Rev 29:3–24

European Commission (2001) Unity, solidarity, diversity for Europe, its people and its territory. Second report on economic and social cohesion. Office for Official Publications of the European Communities, Luxembourg

European Commission (2004) A new partnership for cohesion. Convergence competitiveness cooperation. Third report on economic and social cohesion. Office for Official Publications of the European Communities, Luxembourg

Funck B, Pizzati L (eds) (2003) European integration, regional policy, and growth. World Bank, Washington D.C.,

Galor O (1996) Convergence? Inferences from theoretical models. Econ J 106:1056–1069

Getis A (1995) Spatial filtering in a regression framework: examples using data on urban crime, regional inequality, and government expenditures. In: Anselin L, Florax R (eds) New directions in spatial econometrics. Springer, Berlin, pp 172–88

Getis A, Griffith DA (2002) Comparative spatial filtering in regression analysis. Geogr Anal 34:130–140

Hart PE (1995) Galtonian regression across countries and the convergence of productivity. Oxf Bull Econ Stat 57:287–293

Hawkins DS, Allen DM, Stromberg AJ (2001) Determining the number of components in mixture of linear models. Comput Stat Data Anal 38:15–48

Le Gallo J, Ertur C (2003) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. Pap Reg Sci 82:175–201

Le Gallo J, Dall'erba S (2006) Evaluating the temporal and spatial heterogeneity of the European convergence process, 1980–1999. J Reg Sci 46:269–288

López-Bazo E, Vayá E, Mora AJ, Suriñach J (1999) Regional economic dynamics and convergence in the European union. Ann Reg Sci 33:343–370

Louis TA (1982) Finding the observed information matrix when using the EM algorithm. J R Stat Soc B 44:226–233
MacLachlan G, Peel D (2000) Finite mixture models. Wiley, New York
Mankiw NG, Romer D, Weil DN (1992) A contribution to the empirics of economic growth. Q J Econ 107:407–437
Martin P (1998) Can regional policies affect growth and geography in Europe? World Econ 21:757–774
Meliciani V, Peracchi F (2006) Convergence in per-capita GDP across European regions: a reappraisal. Empir Econ 31:549–568
Neven D, Gouyette C (1995) Regional convergence in the European community. J Common Market Stud 33:47–65
Paap R, van Dijk HK (1998) Distribution and mobility of wealth of nations. Eur Econ Rev 42:1269–1293
Pagan A (1984) Econometric issues in the analysis of regressions with generated regressors. Int Econ Rev 25:221–247
Petrakos G, Rodríguez-Pose A, Rovolis A (2005) Growth, integration, and regional disparities in the European union. Environ Plann A 37:1837–1855
Quah DT (1993a) Empirical cross-section dynamics in economic growth. Eur Econ Rev 37:426–434
Quah DT (1993b) Galton's fallacy and tests of the convergence hypothesis. Scand J Econ 95:427–443
Quah DT (1996a) Empirics for economic growth and convergence. Eur Econ Rev 40:1353–1375
Quah DT (1996b) Twin peaks: growth and convergence in models of distribution dynamics. Econ J 106:1045–1055
Quah DT (1996c) Regional convergence clusters across Europe. Eur Econ Rev 40:951–958
Quah DT (1997) Empirics for growth and distribution: stratification, polarization, and convergence clubs. J Econ Growth 2:27–59
Rey SJ, Montouri BD (1999) US regional income convergence: a spatial econometric perspective. Reg Stud 33:143–156
Sala-i-Martin X (1996a) The classical approach to convergence analysis. Econ J 106:1019–1036
Sala-i-Martin X (1996b) Regional cohesion: evidence and theories of regional growth and convergence. Eur Econ Rev 40:1325–1352
Solow RM (1999) Neoclassical growth theory. In: Taylor JB, Woodford M (eds) Handbook of macroeconomics, vol 1. North Holland, Amsterdam, pp 637–667
Temple JRW (1998) Robustness tests of the augmented Solow model. J Appl Econom 13:361–375
Temple JRW (2000) Growth regressions and what the textbooks don't tell you. Bull Econ Res 52:181–205
Titterington DM, Smith AFM, Makov UE (1985) Statistical analysis of finite mixture distributions. Wiley, Chichester
Tsionas EG (2000) Regional growth and convergence: evidence from the United States. Reg Stud 34:231–238
Turner RT (2000) Estimating the propagation rate of a viral infection of potato plants via mixtures of regressions. Appl Stat 49:371–384
Wedel M, Kamakura WA (1998) Market segmentation: conceptual and methodological foundations. Kluwer, Dordrecht