Regional (Di)Convergence

Stefano Magrini*

University "Ca' Foscari" of Venice

1. Introduction

The question whether incomes are converging across regional economies has long attracted the attention of economists and decision makers. On the one hand, there is a widespread perception that persistent disparities in aggregate growth rates have led to sizable differences in welfare not only across countries but within them as well. On the other hand, the ample body of empirical research on the subject has not yet reached a common answer as to whether, and under which conditions, convergence actually takes place.[1]

The present paper aims at providing an overview of the key developments in the study of regional convergence, discussing the methodological issues that have arisen since the first attempts to analyse convergence and critically surveying the results that have been obtained for different regional systems.

---

* This essay is a draft of a chapter written for eventual publication in the *Handbook of Regional and Urban Economics*, Volume 4, edited by Vernon Henderson and Jacques-François Thisse. I am very grateful to Monica Billio, Donata Favaro, Dino Martellato, Raffaele Paci, Danny Quah and participants to the ESRC Workshop on "Cities and Geography" (Paris, December 2002) for helpful discussions. I would especially like to thank Paul Cheshire, Vernon Henderson and Jacques-François Thisse for support and constructive comments that played a major role in shaping and improving the essay.

[1] Interestingly, Williamson expressed similar concerns back in 1965 while introducing his empirical investigation into the relationship between regional inequalities and the process of national development.

In general, two broad threads of analysis can be identified. Within the first thread, the *regression approach*, a variety of methods has been developed to test the convergence predictions of the traditional neoclassical model of growth. Initially, following the seminal contribution by Baumol (1986) later refined by Barro (1991) and Barro and Sala-i-Martin (1991 and 1992), a large number of studies has made use of cross-sectional growth regressions to see whether regions are converging towards steady state paths and, if so, at what speed. Later, in order to control for unobserved heterogeneities that bias conventional cross-sectional convergence regressions and to deal with endogeneity concerns, panel data methods have been adopted. Other researchers have instead chosen to implement the regression approach by means of time series methods in which the definition of convergence relies on the notions of unit roots and cointegration. The first part of the chapter (Section 2) will therefore describe the main developments of this approach up to its most recent applications to regional datasets and discuss the many problems that still exist. A first underlying argument will be that the regression approach tends to concentrate on the behaviour of the representative economy. In other words, convergence analyses based on such an approach are, with few exceptions, uninformative: they can only shed light on the transition of this economy towards its own steady state whilst giving no information on the dynamics of the entire cross-sectional distribution of regional incomes. A second important point will be that most of the empirical work on regional convergence within the regression approach applies virtually the same empirical methods originally developed to analyse convergence across nations. However, regions and nations, being characterised by profoundly different degrees of openness are far from being interchangeable concepts. Thus, by totally overlooking this important difference, these empirical methods fail to properly account for spatial interaction effects.

The remainder of the chapter will therefore deal with these two issues. Sections 3 will start by considering the theoretical implications for convergence once openness is introduced into the neoclassical model of growth. In particular, it will be shown that the open-economy version of the neoclassical model predicts a faster speed of convergence than its closed-economy counterpart. Moreover, the existing evidence on the role that interregional flows brought in by openness may play in the explanation of regional convergence will be considered. The second part of the section will instead concentrate on the consequences of spatial interaction effects on convergence analyses from an econometric perspective and, after presenting the different sources of misspecification problems that have been identified in the spatial econometric literature, it will describe the ways in which these problems have (or have not) been addressed in regional convergence studies.

Section 4 will instead focus on an alternative approach to the analysis of convergence, the *distribution dynamics approach*, that examines directly how the cross-sectional distribution of per capita output changes over time, putting emphasis on both the change in its external shape and the intra-distribution dynamics. Examples are Markov chain methodologies or, more generally, approaches using stochastic kernels to describe the law of motion of cross-sectional distributions. A fundamental point will be that the distributional approach to convergence is not without problems of its own. However, despite these problems, the distributional approach to convergence – particularly when based on nonparametric stochastic kernel estimations – appears to be generally more informative than convergence empirics within the regression approach, and therefore represents a more promising way forward. Thus, an application of this methodology to data on per capita income for European regions over the period 1980-1995 is carried out at the end of the chapter. In particular, this analysis makes it possible not only to characterise regional convergence dynamics in Europe but also, using a spatial conditioning scheme, to evaluate the role of spatial factors in these dynamics. Finally, the adoption of a set of functionally defined regions highlights the risks from the use of datasets on administrative regions such as European NUTS. The boundaries of these regions are in fact the result of political and historical factors which are country-specific so that not only do they bear no relationship to the socio-economic factors that form the basis of a functional region, but they also vary from country to country. As a result, data for administratively defined regions are likely to be characterised by significant nuisance spatial dependence that, if not taken into adequate consideration, runs the risk of concealing important features of regional distribution dynamics.

2. The 'regression approach'

2.1 Theoretical foundations

The traditional neoclassical model of growth, originally set out by Solow (1956) and Swan (1956), and, following the work of Ramsey (1928), subsequently refined by Cass (1965) and Koopmans (1965), has provided the theoretical background for a vast body of empirical analyses on income convergence. The standard model and its main empirical implications for the convergence debate are well known so just a brief recap is offered in what follows.

Consider an economic system in which physical capital, K, and labour, L, are used in order to produce a homogeneous consumption good:[2]

$$Y = F\left(K, \widetilde{L}\right)$$

where $\widetilde{L} \equiv L \cdot A(t)$ is the effective amount of labour input and $A(t)$, the labour-augmenting technical change, grows exponentially at the exogenously given rate μ: $\mathrm{A}(t) = A(0)\,\mathrm{e}^{\mu t}$. Defining quantities per unit of effective labour as $\widetilde{y} \equiv Y/\widetilde{L}$ and $\widetilde{k} \equiv K/\widetilde{L}$, the (twice differentiable, homogeneous of degree 1, increasing, jointly concave in all its arguments and strictly concave in each) production function becomes

$$\widetilde{y} = f\left(\widetilde{k}\right). \tag{1}$$

Two accumulation frameworks are possible. In the Solow-Swan approach, an exogenously given fraction of output is saved and invested in new physical capital while the rest of output is consumed. Alternatively, in the Cass-Koopmans approach, rational households with perfect foresight choose the consumption path, and thus the saving path, by maximising intertemporal utility subject to a flow budget constraint:

$$\dot{\widetilde{k}} = f\left(\widetilde{k}\right) - \widetilde{c} - (\delta + n + \mu)\widetilde{k} \tag{2}$$

where δ is the rate of capital depreciation and n is the rate of population growth.

---

[2] Mankiw *et al*. (1992) add human capital to the basic Solow-Swan framework; since this feature does not affect our main points it is largely ignored in what follows.

The system exhibits saddle-path stability under either accumulation frameworks (Barro and Sala-i-Martin 1995, Durlauf and Quah 1999), so that the economy converges to a steady state equilibrium in which the level of income per capita, consumption per capita and the capital-labour ratio all grow at the exogenous rate of technological progress while variables per unit of effective labour are constant. If the share of capital in total income is a constant, as in the case of Cobb-Douglas technology, it is easy to show that the growth rate experienced by the economy is negatively related to the level of the capital-labour ratio: the lower the capital-labour ratio and, therefore, the lower per capita output, the further the economy is from its balanced growth path, and the higher its growth rate.

Finally, we can turn to the cross-sectional dynamics which can be derived from the empirical implications of the neoclassical model of growth around the steady state. Considering observed per capita income $y = \tilde{y}A$, a Taylor series approximation of the system's dynamics around the deterministic steady state yields:

$$\log y(t) = \left[\log y(0) - \left(\log \tilde{y}^* + \log A(0)\right)\right] e^{-\beta t} + \left[\log \tilde{y}^* + \log A(0) + \mu t\right] \quad (3)$$

where the coefficient β, describing the speed with which the economy converges towards the steady state, can be shown to be inversely related to capital's share in income so that as this coefficient approaches unity, the convergence rate tends to zero. According to equation (3), the log of per capita income can hence be viewed as having two components: a convergence component (the first term of the right-hand side, involving $e^{-\beta t}$) and a levels component (the rest of the right-hand side). Fig 1, in which different steady state paths corresponding to two possible values for the sum $\log \tilde{y}^* + \log A(0)$ have been exemplified, shows that as long as this sum remains unobserved or unrestricted, any pattern of cross-sectional growth and convergence is consistent with the model. While economies 2 and 3 diverge away from each other, the rich economy 1 stays rich and the poor economy 4 remains poor: if the number of economies exceeds that of the underlying time paths, then a clustering in the cross-sectional distribution (twin-peakedness or club convergence) could arise.

2.2 Empirical implementation: cross-sectional method

In a seminal study, Baumol (1986) implemented a method of testing the neoclassical prediction of convergence based on a simple cross-sectional regression:

$$\log[y(t)/y(0)] = a + b \log y(0)$$

where the left-hand-side of the equation represents the growth rate over the period (0,*t*). Obviously, the negative value for the coefficient *b* found by Baumol is interpreted as evidence of convergence, as this would mean that the economies with low initial levels of per capita GDP have experienced the fastest growth rates. Barro and Sala-i-Martin (1991 and 1992) expanded and refined this approach. Firstly, they pointed out that the traditional neoclassical model predicts that the growth rate of an economy is inversely related to the distance from its steady state. Therefore, poor economies grow faster than rich ones only if they all share the same steady state. By contrast, if differences in technological levels and attitudes toward saving exist among economies, then these economies are characterised by different steady states and the negative relationship between the growth rate of per capita GDP and its initial level does not hold in a cross-sectional sample. Convergence towards the same steady state is then labelled by these authors as *absolute or unconditional convergence*, while the second type of convergence is labelled as *conditional convergence*.

The fundamental element of the empirical analyses carried out by Barro and Sala-i-Martin is derived from the logarithmic linearisation of the transitional dynamics of the traditional neoclassical model around the steady state considered above. Equation (3) implies that, starting from the initial time 0, the average growth rate over an interval of $t \geq 0$ time periods is given by

$$(1/t)\log[y(t)/y(0)] = \mu + \left(1-e^{-\beta t}\right)/t \cdot [\log \tilde{y}^* + \log A(0)] - \left(1-e^{-\beta t}\right)/t \cdot \log y(0)$$

so that, other things being equal, the average growth rate of per capita income depends negatively on the initial level of per capita income, conditioned on the steady state value of per capita income per effective worker, on the exogenous growth rate of technology and on the initial level of technology. Since the exact value of these is unknown, they suggested that they should test the following convergence equation:

$$(1/t)\log[y(t)/y(0)] = c - \left(1-e^{-\beta t}\right)/t \cdot \log y(0) + u(t) \tag{4}$$

where $u(t)$ is a random disturbance while the constant summarises the unobserved parameters. The key parameter to be empirically estimated in this approach is the speed of adjustment to the steady state, β, i.e. the rate at which the economies approach their steady state growth paths. As already seen, within the theoretical framework adopted this parameter crucially depends on the capital-share coefficient; as this coefficient tends to one, so that if diminishing returns to capital no longer apply, the rate of convergence tends to zero.

To test the neoclassical prediction that the growth rate of an economy is inversely related to the distance from its steady state, or β-convergence as Barro and Sala-i-Martin label it, data sets have to be conditioned on the steady state. These authors suggest two possible ways of overcoming the problem. The first is to identify a group of homogenous economic systems characterised by similar technological levels and institutional environments, thus fulfilling the conditions assuring convergence towards the same level of steady state income. In this case, unconditional (or absolute) convergence is expected and equation (4) can be applied directly. In practice, however, the assumption of independence across economies for the error term implicit in equation (4) is far from being realistic as disturbances tend to affect different groups of regional economies in different ways. If this is the case, $\log y(0)$ and $u(t)$ are not uncorrelated, and the least-squares estimations of β are biased. This problem is overcome by decomposing the error term u_t into two separate components. The equation describing the behaviour of an economic system around its steady state thus becomes:

$$(1/t)\log[y(t)/y(0)] = c - \left(1 - e^{-\beta t}\right)/t \cdot \log y(0) + \varphi s(t) + v(t) \qquad (5)$$

where $v(t)$ is an independent disturbance, $s(t)$ is an aggregate disturbance and φ measures its effect on the growth rate of the economy. Assuming that, cross-sectionally, φ is distributed independently of $v(t)$ and that $\mathrm{cov}[\log y(0),\varphi]=0$, the composite error term is not correlated with $\log y(0)$ and the least-squares estimate of β is not biased.

In other cases, when the group of economies differ in their fundamentals, the group will show multiple steady states and the neoclassical model invokes the concept of conditional convergence. From an operational point of view, this requires the introduction of additional explanatory variables in the cross-sectional regression (5), which represent proxies for the different steady states.

Examples of analyses of this type are abundant within a regional context. Following Barro and Sala-i-Martin (1991, 1992a and 1995) and Sala-i-Martin (1996), who reported the existence of unconditional convergence across U.S. states, Japanese prefectures and several European countries (Germany, UK, France, Italy and Spain) and conditional convergence across a group of European regions, a vast number of studies have reported unconditional or conditional β-convergence across groups of regional economies worldwide (see Sala-i-Martin 1996, Durlauf and Quah 1999, de la Fuente 2000, for reviews). So, while Shioji (1996) confirms earlier results for Japan, Holz-Eakin (1993), Garofalo and Yamarik (2002) and Vohra (1996), although using a human capital augmented version of the neoclassical growth model (Mankiw *et al.*, 1992), report evidence of convergence within the U.S., and Cashin (1995) suggests that there exists β-convergence across the seven states of Australia. Similarly, several empirical studies, following comparable methodologies, confirm the original findings by Coulombe and Lee (1993) that unconditional convergence across Canadian provinces cannot be rejected (Coulombe and Lee 1995, Lee and Coulombe 1995, Coulombe and Day 1999, Coulombe and Tremblay 2001). It is also interesting to note that, together with the general support to β-convergence, another empirical regularity seems to emerge from this group of studies: the estimated value of β, the speed with which economies converge to their steady state, is rather small (around 2 per cent per year) and rather stable across different samples.

Moving now to European countries, studies of β-convergence have been carried out for regions in Austria (Hofer and Wörgötter, 1997), West Germany (Niebuhr 2001, Herz and Röger 1995, Funke and Strulik, 1999), Spain (de la Fuente and Vives 1995, de la Fuente 1996), Italy (Fabiani and Pellegrini 1997, Paci and Pigliaru 1995), UK (Chatterji and Dewhurst 1996), and Greece (Siriopoulos and Asteriou 1998), to cite just a few. It has to be noted, however, that while evidence of this type of convergence is reported in most cases, wide variations in the estimated values of the rate of convergence are found in different countries. When the attention is shifted to the whole of Europe, similar to

Barro and Barro and Sala-i-Martin's analyses, Member State dummy variables (as proxies for differences in countries' steady states) and other variables (to allow for industry structure differences between regions) are generally considered and conditional convergence across various groupings of European NUTS regions is again often found (Button and Pentecost, 1995 and 1999; Armstrong 1995 a, b and c; Neven and Gouyette, 1995; Martin, 2001; Cuadrado-Roura *et al*. 2000b; Fagerberg and Verspagen, 1996; Tondl, 1999 and 2001, Maurseth, 2001). However, it is also generally emphasised that there have been profound changes in the pattern of convergence over time: while conditional β-convergence was rather strong up to the end of the 1970s, it came to a halt during most of the 1980s and then re-emerged, although at quite a slow pace. Moreover, the results are not only sensitive on the choice of countries being considered and the level of NUTS regions employed, but β-convergence estimates are also somewhat sensitive to the choice of the additional explanatory variables. Overall, the general impression is that β-convergence is much weaker in Europe than in other areas, and is governed by a considerable country-specific component.

Researchers have identified a number of problems with cross-sectional regression analyses (see, e.g., Durlauf and Quah, 1999 and Temple, 1999 for surveys), the most important of which can be briefly examined. The first limitation of the cross-sectional regression approach is that, despite the fact that it is directly derived from the traditional neoclassical model, it does not test the validity of this model against alternative and conflicting ones. As clearly pointed out by several authors (Romer, 1993 and 1994; Fagerberg 1994, Paci and Pigliaru, 1997; Durlauf and Quah, 1999; amongst many others), dynamics such as those illustrated in Figure 1 are implicit in widely different theoretical interpretations of the growth process. Specifically, these interpretations range from the closed-economy, human capital-augmented version of Solow's traditional neoclassical model (Mankiw *et al*., 1992) to theories of technological diffusion, either within the neoclassical tradition – as the endogenous growth models (Aghion and Howitt, 1992 and 1998; Barro and Sala-i-Martin, 1995 and 1997; Grossman and Helpman, 1990 and 1991; Helpman, 1993; Lucas 1988; Rivera-Batiz and Romer, 1991; and Romer, 1987) – or within the evolutionary tradition – as the literature on the technological gap (Gerschenkron 1962, Abramovitz, 1979, 1986; Fagerberg, 1988;

Verspagen, 1991; and, for an adaptation which allows for spatial proximity and localised technological spillovers; Caniëls, 2000; Caniëls and Verspagen, 2001). Moreover, a set of theoretical models explicitly develops cross-sectional dynamics which conform to the behaviour depicted in Figure 1. In the first of such models (Quah, 1996d) – in which ideas are an important engine of growth and specialisation in production makes it possible to exploit economies of scale – economies endogenously select themselves into coalitions or convergence clubs depending on the initial distribution of characteristics across economies. In the second group (Azariadis and Drazen, 1990; Durlauf, 1996; Galor and Zeira, 1993; Murphy *et al.*, 1989; Quah, 1996a), nonconvexities in the aggregate production function associated with threshold effects lead to long-run dependence from initial conditions and polarization effects. A first natural conclusion therefore is that, if the aim of a researcher is to provide evidence to discriminate between different growth theories, cross-sectional regressions are of limited use. The regression techniques so far discussed at best produce results which are not inconsistent with neoclassical growth theories. But since they are also consistent with other explanations, they do not constitute a test of traditional neoclassical theory in any scientific sense. Moreover, under the neoclassical model, the conventional cross-sectional growth equation is (approximately) linear. In contrast, in many endogenous growth models it is highly nonlinear and, as shown in Bernard and Durlauf (1996), a linear specification is unable to discriminate between these models.

A second important line of criticism has concentrated on the informative content of cross-sectional regressions. First of all, several researchers (Friedman 1992; Quah 1993b; amongst others) emphasise the analogy between regressions of growth rates over initial levels and Galton's fallacy of regression towards the mean. In other words, they demonstrate that a negative relationship between growth rates and initial values does not indicate a reduction in the cross-sectional variance and, moreover, that it is also possible to observe a diverging cross-sectional distribution even when such a negative relationship holds.[3] In other words, standard convergence empirics are, at best,

[3] The fact that a positive β coefficient is a necessary but not a sufficient condition for a reduction in the cross-sectional dispersion is acknowledged by the proponents of the cross-sectional regression approach. A positive value for β is thus interpreted as indicating the existence of forces reducing the cross-sectional distribution while ongoing disturbances are seen as forces pushing in the opposite direction. The practical value of this interpretation is

uninformative as they concentrate on the behaviour of a representative economy. Even if the law of motion of an economy is actually independent of the behaviour of other economies, the best the traditional convergence approach can do is describe how this economy converges to its own steady state. However, this approach is completely silent on what happens to the entire cross-sectional distribution of economies. In contrast, in the presence of nonconvexities in the production function associated with threshold effects or interdependencies such as those described in coalition models, the traditional convergence approach is not only uninformative with regard to growth and convergence dynamics but can also be misleading. Within the standard neoclassical approach, dynamics such as those depicted in Figure 1 essentially depend on differences in one or more structural characteristics of each economy, regardless of the starting conditions. In contrast, within theoretical models with nonconvexities or models with club formation, these dynamics could be the result of differences in initial conditions across economies with similar structural characteristics. Thus, if a conditioning explanatory variable is not actually determining an economy's economic position as in the standard neoclassical approach but, rather, is evolving endogenously as a response to initial factors determining club membership, a traditional researcher would incorrectly attribute growth and convergence to the conditioning variable and never discover the true growth determinants.[4]

2.3 Empirical implementation: panel data methods

A second tactic to implement the regression approach is to resort to panel data methods, thus combining cross-sectional and dynamic information. Proponents of this approach argue that it has a clear advantage over cross-sectional regressions. As previously noted, conditional cross-sectional convergence analyses must allow for steady state income determinants in order to provide consistent estimates. Given that some of these determinants might be unknown or unmeasurable – and thus constitute nuisance

---

however somewhat dubious since even if information about these shocks was used in a cross-sectional regression, still a positive value for β would not imply that the variance of the cross-sectional distribution is decreasing.

[4] A similar concern is expressed by de la Fuente (2000) who notes that in practice the difference between conditional and unconditional convergence is not totally transparent. If we find that a number of explanatory variables enter significantly in equation (5) we would be tempted to conclude that convergence is only conditional since there are significant differences across economies in their underlying "fundamentals". However, if these variables change over time and tend to converge, it might well be that income is unconditionally converging in the long run.

parameters – it is argued that the only way to obtain consistent estimates is to use panel data methods.

The simplest fixed effects panel data model of the convergence process would then be:

$$\log\left[y(t)/y(t-1)\right] = c_0 + c_1(t) - b\log y(t-1) + u(t)$$

showing that the original constant c is now decomposed into an unobservable economy-specific effect (which is constant over time and determines the region's steady state) c_0, and a time-specific effect, c_1, affecting all economies. For the estimation, the least squares dummy variable estimator (Hsiao, 1986) was initially applied. However, since this estimator is consistent only for a large number of observations over time (Nickell, 1981), the most widely adopted alternative is represented by the 2-step GMM estimator suggested by Arellano (1988) and Arellano and Bond (1991) and introduced into the growth literature by Caselli, Esquivel and Lefort (1996). Starting from an autoregressive model with unobserved individual-specific effects, the approach requires taking the first-differences of the regression equation to remove unobserved time-invariant country-specific effects, and using levels of the series lagged two periods or more as instruments for the equation in first-differences, thus alleviating measurement error and endogeneity biases.

The results from convergence analyses adopting these panel data methods are generally at odds with those from cross-sectional regression studies. For example, in contrast with Barro and Sala-i-Martin's findings, Lall and Yilmaz (2001) find no evidence of absolute convergence among U.S. states. Moreover, the estimated rate of mean reversion appears to be considerably higher than in previous estimates. When European regions are employed, de la Fuente (2000) finds annual convergence rates between 26% and 39% within the five largest E.U. countries, depending on the estimation procedure adopted. Similarly, Tondl (1999 and 2001) reports a convergence rate of approximately 20%, and Cuadrado-Roura *et al.* (2000a) a rate of 17%. Canova and Marcet (1995), via a Bayesian-motivated parameterisation of the individual effects, find a convergence rate of about 23%, with each European region converging to its own steady state. Moreover, they find that individual effects do differ across economies implying that poorer regions stay poor. As for individual E.U. countries, Funke and Strulick (1999) report an average convergence rate of about 10% among German *Länder* using a Bayesian approach, while de la Fuente (1996) estimates a convergence rate of 12.7 % for Spain using a fixed effect model and subsequently confirms this estimate (de la Fuente, 2002) using a standard fixed effects model and a hybrid model with structural variables and fixed effects.

In general, therefore, estimates of the convergence rate via conventional panel data methods are substantially higher than cross-sectional estimates. However, it should be noted that Bond *et al*. (2001) have recently emphasised that the first-differenced GMM estimator may be subject to a large finite-sample bias when the time series are persistent – as is usually the case with output series – and short, so that lagged levels of the variables are weak instruments for subsequent first-differences. To overcome the problem they suggest using a system GMM estimator (Arellano and Bover, 1995; Blundell and Bond, 1998), i.e. a system combining the usual equations in first-differences with equations in levels in which the instruments are lagged first-differences. Applying this estimator to the same data set employed by Caselli *et al*. (1996), Bond *et al*. (2001) find a convergence rate of approximately 2% for both the basic Solow model and its human capital-augmented version. In other words, they re-establish the low convergence rate common to cross-sectional regression studies, and interpret the considerably higher estimates commonly found using first-differenced GMM estimators as arising from the substantial finite-sample bias of this estimator in the presence of weak instruments.

The value of panel data methods appears controversial. From an econometric point of view, the advantages over cross-sectional regressions are apparent: unobserved heterogeneities that bias conventional cross-sectional convergence regressions can be controlled for, and lags of the regressors can be used as instruments to deal with endogeneity concerns. However, if conditioning out for individual heterogeneities might represent an improvement from an econometric point of view, it appears a disadvantage from a conceptual one: conditioning out economy-specific heterogeneities means giving up any attempt to uncover what happens to the entire cross-sectional distribution as it is exactly these heterogeneities that explain who is rich and who is poor and how this pattern evolves over time. In other words, both the problem of open-ended alternatives and, more importantly, the failure to reveal any insight into how the entire cross-sectional distribution of economies evolves already noted during the discussion of cross-sectional regressions remain unsolved.[5]

2.4 Empirical implementation: time series methods

The last way to implement the regression approach is via time series methods in which the definition of convergence relies on the notions of unit roots and cointegration.

One such method has been developed by Evans and Karras (1996 a and b) who use a panel data approach in which economies 1, 2, …, N are said to converge if deviations of $y_{1,t+k}$, $y_{2,t+k}$, …, $y_{N,t+k}$ from their cross-economy average $\bar{y}_t$ are expected, conditional on current information, to approach a constant value as k tends to infinity:

$$\lim_{k\to\infty} E\left(y_{i,t+k} - \bar{y}_{t+k} \middle| I_t\right) = \mu_i \tag{7}$$

which holds if, and only if, every $y_{i,t}$ is non-stationary but every $y_{i,t} - \bar{y}_t$ is stationary. Moreover, convergence is absolute if $\mu_i = 0$ for all i or conditional if $\mu_i \neq 0$ for some i, while divergence is found if, and only if, $y_{i,t} - \bar{y}_t$ is non-stationary for all i.

In operational terms, moving from equation (5) and supposing that only cross-sectional data are available on the additional variables representing proxies for the different steady states, we can obtain:

$$\Delta y_{i,t} = d_i - \eta_i\left(y_{c,t-1} - \tau_{t-1}\right) + \xi_{i,t} \tag{8}$$

where d_i is a parameter that incorporates the proxies for the different steady states, τ_t is a common trend of steady state per capita income level, and $\xi_{i,t}$ is a stationary error term with zero mean and finite variance. Moreover, by averaging across economies and subtracting each member of the resulting equation from the corresponding member of equation (8) gives:

$$\Delta\left(y_{i,t} - \bar{y}_t\right) = \left(d_i - \bar{d}\right) - \eta_i\left(y_{c,t-1} - \bar{y}_{t-1}\right) + \left(\xi_{i,t} - \bar{\xi}_t\right)$$

---

5 An exception is the analysis by Funke and Strulik (1999) who, using a Bayesian panel data technique similar to Canova and Marcet (1995), find evidence of persistence of inequality among West

Finally, since the error term component $\xi_{i,t} - \overline{\xi}_t$ may be serially correlated, convergence is analysed running the augmented Dickey-Fuller (ADF) regression:

$$\Delta\left(y_{i,t} - \overline{y}_t\right) = \delta_i + \rho_i\left(y_{c,t-1} - \overline{y}_{t-1}\right) + \sum_{r=1}^{q} \phi_{ir} \Delta\left(y_{i,t-1} - \overline{y}_{t-1}\right) + \upsilon_{it} \qquad (9)$$

where $\phi_{i,1}$, $\phi_{i,2}$, …, $\phi_{i,p}$ are parameters arising from the serial correlation, υ_{it} is a serially uncorrelated error term with zero mean and finite variance, and ρ_i is negative if the economies converge and non-negative if they do not converge. In particular, Evans and Karras (1996a) carry out an overall test of convergence by combining the information in the individual ADF statistics, on the grounds that this method, treating the data as a panel, is expected to have greater power than performing a separate unit root test for each economy (Levin *et al.*, 2002), and find strong evidence in favour of rapid conditional convergence for the 48 contiguous U.S. states over the period 1929-1991. A similar procedure is also applied by Funke and Strulik (1999) who report evidence of conditional convergence among West German *Länder* between 1970 and 1994.

Using a similar framework, Carlino and Mills (1993, 1996 a and b) carry out individual ADF tests with a time trend as well as a constant to allow for time-invariant equilibrium differentials in relative per capita incomes (i.e. conditional convergence). They find no evidence of convergence in per capita income and per capita earnings among U.S. regions and U.S. states during the 1929-1990 period as they are not able to reject the null hypothesis of unit root for any of the regions and only for 18 states. However, after exogenously allowing for a break in the rate at which the regions were converging in 1946, they are able to reject the null of a unit root for 3 regions and 29 states when using per capita income, and for 1 region and 19 states when using per capita earnings. These results, together with evidence on the amount of persistence of shocks in the time series using parametric and nonparametric methods and on a notion of cross-sectional convergence, are then interpreted as evidence for conditional convergence in per capita income and, to a much lesser extent, in per capita earnings. Moreover, Loewy and Papell (1996), incorporating endogenously determined break points, are able to reject the unit root hypothesis in seven regions, thus supporting Carlino and Mills' evidence

German *Länder* for the period 1970-1994.

on conditional convergence. On a similar vein, more recent evidence on convergence among U.S. regions is also found by Tomljanovic and Vogelsang (2001).

A different method, based on a pure time series model, has instead been developed by Bernard and Durlauf (1995) who model an economy's output series as satisfying

$$a(L)y_{i,t} = \mu_i + \varepsilon_{i,t}$$

where $a(L)$ has one root on the unit circle and $\varepsilon_{i,t}$ is a mean zero stationary process, thus allowing for both linear deterministic and stochastic trends. Convergence in output is then defined as the equality across economies of long-term forecasts of per capita income taken at a given fixed date. In particular, given the information I_t at time t, two economies i and j are said to exhibit stochastic convergence if the long run forecasts of output are equal, that is:

$$\lim_{k \to \infty} E\left(y_{i,t+k} - y_{j,t+k} \middle| I_t\right) = 0 \tag{10}$$

Similarly, economies p = 1, …, N converge if the long run forecasts of output for all economies are equal:

$$\lim_{k \to \infty} E\left(y_{1,t+k} - y_{p,t+k} \middle| I_t\right) = 0 \qquad \forall\, p \neq 1 \tag{10'}$$

thus making it possible to distinguish between convergence between pairs of economies and convergence for all economies.

An important feature of this dynamic definition of convergence is that its existence also implies the definition of convergence as catching-up (i.e., β-convergence). Indeed, if convergence as catching up between t and t+T is defined as entailing a decrease in the expected deviation in output between economies:

$$\lim_{k \to \infty} E\left(y_{i,t+T} - y_{j,t+T} \middle| I_t\right) < y_{i,t} - y_{j,t} \qquad \text{if } y_{i,t} > y_{j,t}$$

it can be shown that time series forecast convergence implies β-convergence, when growth rates are measured between dates t and t+T for some fixed T. In other words, time series analyses based on equations (10) and (10') appear to resort to a stricter notion of convergence than cross-sectional analyses discussed above (Bernard and Durlauf, 1996).

Testing for convergence requires checking for the compatibility of $y_{i,t} - y_{j,t}$ with a time invariant Wold representation of the form:

$$y_{i,t} - y_{j,t} = \kappa_{i,j} + \sum_{r=0}^{\infty} \pi_{i,j,r} \varepsilon_{i,j,t-r}$$

such that $\kappa_{i,j} = 0$ and $\pi_{i,j,r}$ is square summable. According to Bernard and Durlauf, convergence, as defined by equations (10) and (10'), thus requires that $y_{i,t} - y_{j,t}$ is a mean zero stationary process, which again can be verified applying standard unit roots and cointegration procedures (Bernard and Durlauf, 1995 and 1996). Following this strategy, Tsionas (2001) finds strong evidence against the hypothesis that per capita income in U.S. regions has converged over the period 1929-1997.

Several authors have stressed the existing discrepancies in the results obtained from similar datasets using different approaches and methods. Bernard and Durlauf (1996) argue that the discrepancies between cross-sectional and time series analyses could be partly explained by the fact that time series tests are based on a stricter notion of convergence than cross-sectional tests. Moreover, they emphasise that the two approaches take a different view of the data: while time series methods assume regions are close to limiting distributions and convergence is interpreted as meaning that initial conditions have no effect on the expected value of output differences, cross-sectional methods assume regions are in transition towards a limiting distribution and convergence is interpreted to mean that initial output differences dissipate over a fixed time period. This interpretation is however challenged by Carvalho and Harvey (2002), who note not only that, while some unit root tests are sensitive to initial conditions, the ADF test is robust to initial values different from zero but also that, when the constant is dropped, its power increases the further the initial conditions are from equilibrium.

In order to explain the discrepancies deriving from the application of different time series methods, Nahar and Inder (2001) point to the inconsistencies in the links between the different definitions of convergence and the stationarity of output differences. In particular, as far as the method developed by Bernard and Durlauf is concerned, they note that certain non-stationary $y_{i,t} - y_{p,t}$ processes can meet their definition of convergence so that a test for stationarity of the process may fail to reject the null hypothesis of unit root and wrongly conclude that there is no convergence. A similar argument applies to the method developed by Evans and Karras, according to whom a necessary and sufficient condition for convergence, as defined by equation (7), is that every $y_{i,t}$ is non-stationary while every $y_{i,t} - \bar{y}_t$ is stationary. However, also in this case Nahar and Inder show that a non-stationary $y_{i,t} - \bar{y}_t$ process can meet the definition given in equation (7), thus implying that stationarity is not a necessary condition for convergence.

A different perspective is offered by Carvalho and Harvey (2002) who instead emphasise that while some unit root tests are sensitive to initial conditions, making them unsuitable when analysing whether regions are in the process of converging, the ADF *t*-test is the most robust and should be preferred. However, they note that even this test has virtually no power when a time trend is included, as in Carlino and Mills' studies. Moreover, they criticise the use of overall tests, as applied by Evans and Karras (1996 a and b) or Funke and Strulik (1999), as these tests do not take account of the cross-correlation between the series and remark that even individual tests, when based on equations such as (9) become uninformative when one region does not converge but all others do, so that $y_{i,t} - \bar{y}_t$ is non stationary for all economies. As a consequence, fitting a model similar to (9), they run pairwise ADF tests to the differences between all the annual series for the eight U.S. census regions between 1950 and 1999, finding that convergence is confined to the six poorest regions while the two richest ones, New England and Mid East, diverged.

Generally, Carvalho and Harvey maintain that running unit root tests to decide whether convergence is taking place is at best of limited value and that a better strategy could entail fitting multivariate unobserved components (structural) time series models so making it possible to gather information about both cycles and convergence (Carvalho and Harvey, 2002; Harvey and Carvalho, 2002). In particular, they develop a model combining unobserved components with a second-order error correction mechanism thus allowing a decomposition into trend, cycle and convergence components. Fitting this model to annual (1950-1999) and seasonally adjusted quarterly (1969:1-1999:4) data on per capita income of U.S. census regions, these authors show that absolute convergence is confined to the group of six poorest U.S. regions while the two richest ones have diverged, especially during the last two decades. Moreover, they show that the convergence processes characterising the group of poorest regions are neither monotonic across time nor homogenous across space, thus casting serious doubts against the results obtained via cross-sectional and panel data approaches to convergence.

3. Factor Mobility and Spatial Interaction

In the previous section, we concentrated our attention on the theoretical foundations of the regression approach to convergence and on the different ways in which this approach has been developed and implemented. Before moving to an alternative approach to convergence that answers some of the critical issues already raised, it is however necessary to focus on an aspect that has been so far largely neglected, i.e. the role of spatial interaction effects.

The traditional neoclassical model of growth has been developed starting from the assumption that the economies are fundamentally closed. This comes from the fact that the model was originally intended to explain the evolution of a single economy's growth rate over time. Only later has the model been employed for explaining differences of per capita income growth across different economies; but despite this change of perspective, the original assumption has been retained and transferred to empirical analyses on international convergence.

Moreover, as documented in the previous section, virtually the same empirical methods originally developed to analyse convergence across nations have been widely used to test for convergence processes at a sub-national level. Unfortunately, this attention to regional growth and convergence appears to be fundamentally motivated by the fact that regions offer data sets which are new, larger and more homogenous (Blanchard, 1991) without a proper recognition of the fact that regions and countries are far from being interchangeable concepts. So, while the assumption of a closed economy may be defensible for countries, it is clearly implausible for the regions within a country, where barriers to trade and to factor flows are considerably less. As a consequence, among the many issues of interest, at least two need to be emphasised:

i. what are the implications for convergence once openness is introduced into the theoretical framework?

ii. since interregional imply the existence of spatial dependence effects, how do these effects influence the empirical results?

3.1 Implications of interregional flows

The simplest way to start answering the first question is to consider the open-economy version of the traditional neoclassical model developed by Borts and Stein (1964) within the Solow-Swan framework. The basic hypotheses of this version are equivalent to those for a closed economic system. In addition, it is assumed that a factor's remuneration rate, determined by marginal productivities in perfectly competitive markets, can differ across regions, that factors can move freely in response to these differentials in rates of remuneration, and, finally, that agents of different regions have access to similar technologies and share roughly similar preferences.

In such a setting, equation (2), which represents the fundamental differential equation of the model, becomes:

$$\dot{\tilde{k}} = f\left(\tilde{k}\right) - \tilde{c} - \left(\delta + n + \mu\right)\tilde{k} + G_K\left(r - r^o\right) - G_L\left(w - w^o\right)$$

where r (w) and r^o (w^o) represent the rates of return to capital (labour) within and outside the region, while G describes the interregional flows of factors as a function of differential rates of return. In particular, G shows a positive, negative, or zero value according to whether there are positive, negative, or no interregional differentials in the rates of return.

Again, given the assumption of common technology and preferences, each regional economy converges to the common steady state equilibrium in which variables per unit of effective labour are constant. Similar to the closed-economy case, off steady state capital and labour rates of return may differ between regions only if the existing levels of capital-labour ratio differ. Moreover, the conventional neoclassical assumptions about the production function imply that the marginal product of capital is higher in regions with a lower capital-labour ratio, whereas the marginal product of labour is higher in regions with a higher capital-labour ratio. Since there are no constraints to interregional flows of capital and labour, capital will tend to flow from the regions with a higher capital-labour ratio to the regions with a lower level of capital per worker while labour will tend to flow in the opposite direction. As a consequence, regions with a higher capital-labour ratio will be characterised by a negative value for G_K and a positive value for G_L, thus reinforcing the convergence to the steady state predictions of the closed-economy model. Finally, as is clear from equation (6), while transitions to the steady state are taking place, the regions with lower capital-labour ratios will show higher income per (effective) worker rates of growth. In such a setting, if reaction functions $G(\cdot)$ are such that adjustment in either capital or labour markets is instantaneous, the speed of convergence would be infinite. To alleviate this paradoxical implication, the original setting then has to be modified introducing adjustment costs for investment and migration, so that the rate of convergence to the steady state is higher than in the closed economy case, but with a finite value. It should be noted, however, that this framework has a particularly unappealing feature in the present context: the capital market is not integrated in the sense that residents cannot borrow at a common, countrywide interest rate r. However, eliminating this feature leads to a further paradoxical result within the traditional neoclassical setting (Barro and Sala-i-Martin, 1995): consumption per unit of effective labour tends to zero and assets become negative for all regions except the most patient one, i.e. the one in which the preference parameters are such that per capita consumption grows at the slowest rate. This region will asymptotically own everything and consume all the overall output. As argued by Barro and Sala-i-Martin (1995), eliminating such implications in an open-economy setting might require a model combining credit market imperfections, finite lifetimes, and adjustment costs for investment. Even leaving aside issues concerning the appropriateness of such a modification within a regional context, the predicted rate of convergence, albeit finite, would nevertheless be higher than the rate predicted within a closed-economy setting.

Convergence in inter-regional per capita income within the traditional neoclassical setting can also be reinforced by trade relations rather than factor mobility. Even in the absence of factor mobility, progressive equalization in commodity price and specialization of regional productive structures according to relative factor abundance result from inter-regional trade, thus leading to factor price equalization (the traditional Samuelson factor price trade equalization theorem). Moreover, in the presence of disparities in regional technological attainment, inter-regional trade can promote technological diffusion when technological progress is incorporated in traded goods, thus providing yet another possibility for poorer economies to converge with richer ones (Nelson and Phelps 1966, Grossman and Helpman 1991, Segerstrom 1991, Barro and Sala-i-Martin 1995).

To sum up, the traditional neoclassical model describes an inherent tendency for the economic system to reach a situation of equilibrium not only for the regional markets but for the relationships between the region and the rest of the economic system as well. The regional economies that form the system described by the authors are populated by people sharing similar technological systems. The obvious implication is that these regional economies also share the same steady state. Within this context, therefore, any differences in regional economic growth are fundamentally the result of two combined sources: (i) the internally financed growth of the stock of capital per worker, and (ii) a progressive reduction of an initial interregional misallocation of resources, brought in by openness. As a combination of these two sources, speed of convergence to the steady state is faster than in the closed-economy case.

It is now tempting to see whether the open version of the neoclassical model can accommodate the outcomes of regression analyses. Obviously, this question might appear of limited relevance to those who share the belief that estimates from these analyses are severely biased either because of the reasons discussed in the previous section or, as will be explained below, because they are obtained starting from the implausible assumption that geographical units are fundamentally closed. However, even from such a radical perspective, this could still prove a valuable effort since it forces us to consider explicitly some mechanisms that, according to the neoclassical model, drive convergence across regions, other than the internally financed accumulation of capital.

As is widely recognised in the literature, convergence rate estimates from cross-sectional regressions are unenthusiastically low, even in international studies. As already mentioned, the coefficient describing the rate of convergence to the steady state is inversely related to capital's share in income. Indeed, the low speed of convergence usually found in cross-sectional regression studies, about 2 per cent, requires capital's share in income to be close to 0.7 or 0.8, a value much higher than the conventional value of 0.4 given in national income accounts. To account for these findings, Mankiw *et al*. (1992) and Barro and Sala-i-Martin (1992) suggest a modified version of the closed-economy traditional neoclassical model in which the usual production function is extended to allow for human capital. Indeed, by thinking of capital in a broad sense that includes human capital elements, the labour share is reduced to a value that thus is consistent with the cross-sectional evidence on the speed of convergence. However, as pointed out at the outset of this section, this way of reconciling theoretical predictions with cross-sectional empirical results appears inadequate in the present context.

At a first sight, more promising results are offered by estimates obtained from panel data approaches, which tend to be markedly higher than those found in cross-sectional analyses. In particular, these higher estimates have explicitly been interpreted by some authors (for example, Caselli *et al*., 1996) as supporting the validity of the open economy version of the neoclassical growth model versus its closed economy counterpart. However, there appear to be at least a couple of reasons for being sceptical about such an interpretation. As previously pointed out, Bond *et al*. (2001) argue that the high estimated rates of convergence commonly found using first-differenced GMM estimators arise from the substantial finite-sample bias of this estimator in the presence of weak instruments. Using a system GMM estimator, they re-establish the low convergence rate common to cross-sectional regression studies. On the other hand, convergence rate estimates via panel data methods are quite similar across widely different regional data sets, which are known to be characterised by very different degrees of openness, and are not materially higher than those found in international studies.

Probably, a more productive way of dealing with the issue is to look for more explicit evidence on the role played by the different interregional flows on convergence. Unfortunately, though, this type of study is not abundant. Focussing on the role of migration, Barro (1991) and Barro and Sala-i-Martin (1995) include the net migration rate as an explanatory variable in the growth regressions for U.S. states, Japanese prefectures and the regions of five European countries. The expectation was that, holding migration rates constant, the estimates of the rate of convergence should be reduced, while the size of the reduction would provide a direct measure of the actual role played by migration on convergence. Contrary to these expectations, however, they found that the estimates of the rate of convergence were not significantly affected by the introduction of migration rates, even when instruments were employed in order to allow for the likely endogeneity of the net migration rates. These results, together with the findings that the rate of net migration tends to respond positively to the initial level of per capita income, are then interpreted as suggesting that migration plays a small part in the explanation of convergence, while the bulk of the explanation is left to the internally financed adjustment of capital-labour ratios.

Quite a different implication emerges from the studies of Blanchard (1991) and Blanchard and Katz (1992). Using a simple model of a system of small open economies producing different bundles of goods and characterised by high factor mobility, Blanchard (1991) shows not only that U.S. states per capita incomes tend to converge towards a stable stochastic steady state distribution, but also that a crucial element of this convergence process is represented by labour mobility. Within a similar framework, the latter study tends to confirm the conclusion about the importance of migration over other forms of adjustment among U.S. states and further qualifies this conclusion suggesting that migration tends to be determined more by changes in unemployment than by changes in relative wages. Obviously, grounded as they are on the application of a VAR technique to very short samples, these strong conclusions are open to question on econometric grounds (Hall, 1991). Keeping this in mind, it is nonetheless somewhat surprising that what seems to be the main message of this work, i.e. that mobility of factors is crucial to the understanding the dynamics of regional growth, has gained so little attention in the literature on regional convergence. Moreover, this message appears to have rather more force for all those regional systems which are known to be characterised by a much lower degree of labour mobility than the U.S. and Europe is obviously the first case that comes to mind. There is ample evidence supporting such concerns. While relative unemployment and wages affect interregional migration in Great Britain (Pissarides and McMaster, 1990), and in West Germany (Decressin, 1994), the resulting regional adjustment to shocks is very low. Barro and Sala-i-Martin (1991 and 1995) find that the rate of net migration tends to respond positively to the initial level of per capita income in the case of U.S. states or Japanese prefectures but that this relation is much weaker for European regions. Moreover, Eichengreen (1993), comparing labour mobility in the U.S., Britain and Italy, finds that the elasticity of migration to unemployment differentials is twice as large in the U.S. than either European countries and even larger in the case of relative wages. A similar result is found by Bentivogli and Pagano (1999), who also note that, consequently, wage and unemployment differentials are generally greater and more persistent in the European Union than in the United States.

However, it should be noted that although labour migration could be rather imperfect as an adjustment mechanism within Europe, other mechanisms might substitute. For example, Eichengreen (1993) suggests these mechanisms could include interregional capital mobility and government policy, Decressin and Fatas (1995) point to changes in labour-force participation, while Cheshire and Magrini (2002) find evidence that, in those parts of the European Union where urbanisation is more dense, changes in commuting patterns can play a significant role as an alternative source of spatial adjustment.

3.2 Implications of Spatial Interaction Effects

Having suggested that interregional flows brought in by openness may play an important role in the explanation of regional convergence, we can now move to the second question outlined at the outset of this section and see, from an econometric perspective, the consequences of spatial interaction effects on convergence analyses.

In general, two broad sources of misspecification problems have been identified in the spatial econometric literature: spatial dependence and spatial heterogeneity (Anselin 1988). Spatial dependence (or spatial autocorrelation) arises from a lack of independence across spatially organised observational units (Cliff and Ord, 1973). In particular, Anselin and Rey (1991) distinguish between *substantive* and *nuisance* spatial dependence. Substantive spatial dependence reflects the existence of spatial interaction effects, such as technological spillovers or factor mobility, which are substantive components of the evolution of income disparities across regions. Nuisance spatial dependence, instead, may result from measurement problems such as a mismatch between the spatial pattern of the process under study and the boundaries of the observational units. The second source of misspecification problems, spatial heterogeneity, reflects a general instability of a behavioural relationship across observational units.

As emphasised by Rey and Montuori (1999), the literature on spatial econometrics offers a rich set of procedures for testing for the presence of spatial effects (Anselin, 1988; Anselin, 1995; Anselin and Bera, 1998; Anselin and Florax, 1995; Anselin and Rey, 1991; Anselin *et al.*, 1996; Getis and Ord, 1992). Moreover, within the cross-sectional regression approach, there exists a number of estimators for models that treat spatial effects explicitly.

A first form of substantive dependence can be incorporated into the traditional cross-sectional specification through a spatial lag of the dependent variable, i.e., the spatial lag model. If W is a row-standardised matrix of spatial weights describing the structure and intensity of spatial effects, based on equation (4), the spatial lag model would be

$$(1/t)\log\left[y(t)/y(0)\right] = c - \left(1 - e^{-\beta t}\right)/t \cdot \log y(0) + \lambda_1 W \log\left[y(t)/y(0)\right] + u(t) \quad (11)$$

where λ_1 is a spatial autoregressive parameter of the spatially lagged dependent variable. This specification can be interpreted as a way of controlling for spatial dependence in regional growth due to the convergence mechanism operating on spatially autocorrelated initial incomes (Anselin and Bera, 1998), or to spatial interaction in the data generating process arising when a region's growth rate is related not only to its own starting level of income but, indirectly through the effect on income growth, to those in other regions as well following a distance decay pattern (Anselin *et al*., 1998). Ordinary least squares to the spatial lag model are inconsistent and alternative estimators based on maximum likelihood and instrumental variables should be employed (Anselin, 1988). A second form of substantive dependence reflects spatial autocorrelation in the starting levels of income and can be dealt with a spatial cross-regressive model in which a spatial lag of initial per capita incomes is added to the original specification:

$$(1/t)\log[y(t)/y(0)] = c - \left(1 - e^{-\beta t}\right)/t \cdot \log y(0) + \lambda_2 W \log y(0) + u(t) \tag{12}$$

Since both the initial levels and the spatial lag of per capita income are exogenous, estimation of a spatial cross-regressive model can be based on OLS.

As for nuisance dependence, in the presence of this form of spatial interaction, the error term in the cross-sectional regression models becomes non-spherical:

$$\varepsilon(t) = (I - \lambda_3 W)^{-1} u(t) \qquad \text{with} \qquad \varepsilon(t) \sim N\left(0, \sigma^2 I\right)$$

where λ_3 is a scalar spatial error coefficient.[6] As a consequence, estimation via OLS will lead to unbiased estimates for the convergence parameter but biased estimates of its variance, thus generating potentially misleading inferences. In this case, inference should be based on the spatial error model

$$(1/t)\log[y(t)/y(0)] = c - \left(1 - e^{-\beta t}\right)/t \cdot \log y(0) + (I - \lambda_3 W)^{-1} \varepsilon(t) \tag{13}$$

estimated via maximum likelihood or general method of moments. From a spatial process perspective, another particularly interesting consequence of nuisance dependence is highlighted by Rey and Montuori (1999). In this instance, a random shock affecting a particular region affects the growth rates of all other regions through the spatial transformation $(I - \lambda_3 W)^{-1}$. Put it in a different way, movements away from a steady state growth path may not be a function of region-specific shocks alone, but of shock spillovers from other parts of the system as well.

---

6 Anselin (1982) shows that the matrix $(I - \lambda_3 W)$ is invertible when $-(1/\omega_{\max}) < \lambda < 1$, where $\omega_{\max}$ is the largest negative eigenvalue (in absolute value) of W.

As we already noted, conventional cross-sectional regression analyses that allow for the role of spatial effects are exceptions rather than the norm. Perhaps, the most comprehensive study is that of Rey and Montuori (1999). Focussing on the experience of 48 coterminous U.S. states between 1929 and 1994, they find strong evidence of positive spatial dependence in both levels and growth rates of per capita income, i.e., spatial clusters of states which are homogenous in terms of income levels and growth rates. Moreover, they find that the rich clusters tend to grow more slowly than poor clusters, a pattern that could be explained by the clustering of initial income levels together with a process of unconditional convergence. However, the estimation results for the different spatial dependence models in equations 11-13 make it possible to rule out such an explanation due to the presence of spatial error autocorrelation rather than the spatial lag. In addition, the analysis suggests that the traditional unconditional model suffers from misspecification due to omitted spatial dependence and that random shocks to individual states not only affect the state's dynamics toward the steady state but propagate throughout the system. Finally, they also find evidence that the indications of a structural change at the end of WWII in the rate of convergence of U.S. states (Carlino and Mills, 1996) tend to vanish when spatial dependence is taken into account.

In studying convergence among European NUTS regions, Armstrong (1995b), López-Bazo et al. (1999) and Rodríguez-Pose (1999) report the presence of significant spatial autocorrelation both for income levels and growth rates. These studies thus provide evidence for the European context also that traditional convergence analyses may suffer from a misspecification due to omitted spatial dependence. Following the standard convergence approach, Armstrong (1995b) adds national dummies as explanatory variables as in Barro and Sala-i-Martin (1991) but interprets them as a way to control for the influence of spatial factors. A similar route is followed by Rodríguez-Pose (1999), who, employing nationally weighted variables to eliminate the spatial autocorrelation of the error term, also reports a sharp reduction in the estimated rate of convergence. These specifications, however, despite being able to substantially reduce (or to eliminate) the presence of spatial autocorrelation in the error terms, appear debatable for two reasons: they are too restrictive, excluding spatial effects across borders, and they overlook the possibility of spatial structures within each member state.

A confirmation of the latter is indirectly provided by the study by López-Bazo *et al.* (1999), who, employing a more disaggregated regional data set, detect strong intra-national local spatial association in per capita income levels. Further evidence is provided by Niebuhr (2001), who, focussing on West German planning regions,[7] finds strong evidence of spatial dependence both in levels and growth rates of per capita Gross Value Added. This study, moreover, following an empirical strategy similar to Rey and Montuori (1999), confirms two of the findings of that study relating to U.S. states: (i) allowing for spatial effects results in a somewhat slower rate of convergence compared to that estimated following the traditional approach; (ii) spatial effects are not explained by a process of unconditional convergence coupled with the clustering of initial income levels. On the other hand, in contrast to the U.S. case where Rey and Montuori find evidence of nuisance spatial dependence, spatial dependence in West Germany appears to be of the substantive form. Niebuhr interprets this difference as a consequence of the different choice of observational units. As recalled above, nuisance spatial dependence may result from measurement problems such as a mismatch between the spatial pattern of the process under study and the boundaries of the observational units. Since U.S. states are large administrative areas while German planning regions are smaller functional regions which take commuting patterns into account, the author suggests the effects of an inadequate choice of the observational units might hide substantial dependence of income growth.

A similar call for greater attention to the issue of what spatial units are most appropriate for regional analysis has been recently made by other authors (Cheshire and Carbonaro, 1995; Cheshire and Hay, 1989; Cheshire and Magrini, 2000; Magrini, 1999). Due mainly to the availability of data, administratively defined regions are commonly used in empirical analyses. Within the European context, the typical example is represented by the Nomenclature of Territorial Units for Statistics (NUTS), a multi-level

---

[7] German planning regions (*Raumordnungsregionen*) are functionally defined and contain several German NUTS3 regions linked by intensive commuting.

classification characterised by a profound heterogeneity at every level, being the result of the unification of the regional systems already existing in E.U. Member countries. Suffice to say that NUTS-I level (the highest tier in the classification underneath the national level) comprises a heterogenous set of regions which include both large metropolitan areas alongside even larger regions containing several metropolitan areas and other regions containing just parts of one metropolitan region. However, two fundamental problems arise from the use of administratively defined regions in the present context. On the one hand, since output is measured at workplaces while population at residences, unless the definition of a region has been selected to abstract from commuting patterns, the measured levels of per capita income will be highly misleading. In addition, processes of decentralisation or recentralisation of residences relative to workplaces is likely to affect per capita income growth rates for administratively defined regions. The extent of these problems is exemplified in Table 1 that reports per capita GDP levels and growth rates for five NUTS-I metropolitan regions and for the corresponding Functional Urban Regions[8] (FURs).

Overall, once it is recognised that regions are naturally open to a range of economic flows and that, as a consequence, substantial interaction exists between them, the need of an explicit treatment of spatial interaction effects in regional convergence studies becomes apparent. The literature on spatial econometrics offers a number of estimators for models that treat spatial dependence explicitly but techniques for handling spatial dependence appear to be essentially confined to cross-sectional studies. Within the panel data approach, Badinger *et al*. (2002), in the absence of a direct estimator for dynamic panels with spatial dependence, propose a two-step procedure in which a system GMM for dynamic panels is used after a spatial filtering technique proposed by Getis and Griffith (2002) is employed in order to remove existing spatial correlation. Applying this procedure to a set of European NUTS-II regions over the period 1985-

---

[8] Functional Urban Regions have been derived by Hall and Hay (1980) and are broadly similar in concept to the (Standard) Metropolitan Statistical Areas used in the US. In particular, they are defined on the basis of core cities identified by concentrations of employment and hinterlands from which more commuters flow to the employment core than to any other subject to a minimum cut off. Cheshire and Hay (1989) provide a detailed description of their definition.

1999, they obtain a convergence rate estimate of about 6 per cent, hence substantially lower than estimates from previous panel data studies.

However, despite the obvious advantages of spatial econometric techniques in the present context, there remain reasons to be sceptical from a more conceptual standpoint. As we noted earlier, spatial dependence may arise from the existence of spatial interaction effects (substantive spatial dependence) or from measurement problems (nuisance spatial dependence). While filtering out the latter is clearly advisable, following a similar strategy for the former source of dependence appears somewhat more controversial. After all, substantive spatial dependence carries with it a lot of valuable information on the working of adjustment mechanisms within a system of open economies and filtering all this information out appears to be to abandon any attempt to explain the significant effect of the interaction across individual economies on convergence dynamics or throw light on spatial adjustment processes. In contrast, theoretical explanations of the working of a spatial economy are abundant and an alternative empirical strategy could be to look first at these theoretical explanations for guidance on how to define spatial variables capable of capturing adjustment mechanisms and, only at a later stage, turn to spatial filtering if tests point to the existence of further specification problems. Finally, it should be emphasised that the use of functionally defined regions could also prove useful as a strategy for minimising spatial nuisance dependence. This seems to be particularly important where the change in commuting patterns – rather than migration – represents an important source of spatial adjustment, as has been argued to be the case in densely urbanised areas of the European Union.

4. The Distributional Approach to Convergence

One of the fundamental messages conveyed in the second section of this chapter was that the regression approach, given its attention to the concept of β-convergence, tends to concentrate on the behaviour of the representative economy. In other words, with few exceptions, convergence analyses based on such an approach can only shed light on the transition of this economy towards its own steady state whilst giving no information on the dynamics of the entire cross-sectional distribution of income. On this basis, several authors have argued that the concept of β-convergence is irrelevant. To address these concerns, proponents of the regression approach suggest combining the analysis of β-convergence with an analysis of the evolution of the unweighted cross-sectional standard deviation of the logarithm of per capita income (Barro and Sala-i-Martin, 1991). A reduction over time of this measure of dispersion is then labelled σ-convergence. However, concentrating on the concept of σ-convergence does not appear to represent an effective solution: analysing the change of cross-sectional dispersion in per capita income levels gives no information on the intra-distribution dynamics. Moreover, as discussed above, a constant standard deviation is consistent with very different dynamics ranging from criss-crossing and leap-frogging to persistent inequality and poverty traps. Distinguishing between these dynamics is, however, of essential importance.

In what follows, we will therefore focus on an alternative approach for analysing income convergence, the distributional approach to convergence. The first part of the presentation will concentrate on its general features and the main methods proposed for its implementation. Later, given the discussion of the previous section, attention will be moved to the ways in which the role of space can be allowed for within this approach.

4.1 General Features of the Distributional Approach to Convergence

The distributional approach represents a radical departure from the regression approach: it examines directly how the cross-sectional distribution of per capita output changes over time, putting emphasis on both the change in its external shape and the intra-distribution dynamics. The approach, firstly suggested by Quah (1993a and b, 1994, 1996a and c, 1997) thus concentrates directly on cross-sectional distributions of per capita income, using stochastic kernels to describe their law of motion.

Let F_t denote the cross-sectional distribution at time t, and ϕ_t an associated probability measure. The simplest scheme for modelling the dynamics of $\{\phi_t : t \geq 0\}$ is a first order dependence specification:

$$\phi_t = T^*(\phi_{t-1}, u_t) = T^*_{u_t}(\phi_{t-1}) \tag{14}$$

where u_t is a sequence of disturbances, T^* an operator that maps the Cartesian product of probability measures at time t-1 and disturbances at time t, and $T^*_{u_t}$ absorbs the disturbance into the definition of the operator and encodes information of intra-distribution dynamics.

A first way to use equation (14) for the study of income convergence is to make the income space discrete, as a result of which the measures ϕ_t can be represented by probability vectors and $T^*_{u_t}$ simplifies into a transition probability matrix M_t whose rows and columns are indexed by the elements of the discretisation, and where each row reports the fraction of economies beginning from that row element and ending up in the different column elements.

Assuming that the underlying transition mechanism is time invariant, the model in equation (14) thus becomes a time-homogeneous (finite) Markov Chain. Then, iterations of (14) yield a predictor for future cross-sectional distributions

$$\phi_{t+s} = M'^s \phi_t \tag{15}$$

since the matrix M'^s contains information about probability of moving between any two income classes in exactly s periods of time. Moreover, taking (15) to the limit as $s \to \infty$, enables us to characterise the likely long-run or ergodic cross-sectional distribution of incomes via the ergodic row vector satisfying

$$\phi_\infty = M'\phi_\infty$$

Implications for the convergence debate are then drawn from the study of ϕ_{t+s} or of ϕ_∞: if they display a tendency towards a point mass, then we can conclude that there is convergence towards equality. If, on the other hand, ϕ_{t+s} and ϕ_∞ display a tendency towards a two-point or bimodal measure, one could interpret this as a manifestation of income polarization.

Different ways of partitioning the income space are obviously possible but very often subjectively chosen equi-sized cells or cells with variable upper endpoints (so as to get approximately the same number of occurrences in each class) are adopted. Applying this procedure to U.S. states, Quah (1996c) finds a high degree of mobility among classes and an ergodic distribution presenting no signs of bimodality. Different conclusions are reached for European NUTS regions by López-Bazo *et al*. (1999), who report evidence of a particularly high degree of persistence in lower income classes, indicating the existence of a poverty trap. Fingleton (1997, 1999) partitions the cross-sectional income space into four large classes and adopts various Markov chain log-linear models to investigate convergence among European NUTS-II. The results suggest that European regions are converging towards a limiting distribution characterised by sizeable differentials in per capita income levels and consistent with the existence of multiple steady states from which economies are continuously displaced by shocks. There is also some evidence suggesting that the limiting distribution of the Markov process had been attained in 1975.

One general problem with Markov chain methods is that they impose quite restrictive assumptions on the data generating process (Bickenback and Bode, 2001). In their attention to future and ergodic cross-sectional distributions predicted by means of the transition probability matrix M_t, these approaches assume that the data generating process is time invariant and satisfies the Markov property. Bickenback and Bode (2001) therefore propose chi-square tests of the Markov property and, using five income classes, suggest that the evolution of the income distribution across the 48 coterminous U.S. states between 1929 and 2000 has not followed a Markov process.

In addition, another significant difficulty comes from discretisation. Indeed, as commonly recognised in the literature, discretising a continuous first-order Markov process is likely to remove the Markov property. While Quah (1996c) suggests that the distortion arising from partitioning into five large cells is not likely to conceal the most important features of the process, Magrini (1999) adopts a procedure aimed at reducing the degree of arbitrariness in the discretisation by concentrating on histograms as approximations to continuous distributions and choosing the income grid optimally so as to minimise the (mean-squared or integrated absolute) error of approximation. By applying this procedure to a set of 122 European functionally defined regions, he reports a strong tendency towards polarisation in the cross-sectional distribution. Bulli (1999) however argues that discretisation of a continuous state-space Markov chain concentrating on the distribution of the process at some point in time is misleading, and recommends adopting a regenerative discretisation method originally employed in the Markov Chain Monte Carlo literature.

Given these critical remarks, a radical alternative is to get rid of discretisation altogether. In this case, the operator in equation (14) can be interpreted as a stochastic kernel (Quah, 1996a and 1997) and convergence can be studied analysing directly the shape of a three-dimensional plot of the stochastic kernel, thus also avoiding to impose restrictive assumptions on the data generating process. Figure 2 shows the nonparametric estimate of the three-dimensional stochastic kernel for the transition dynamics across 110 European NUTS regions and, in the lower part, the corresponding two-dimensional contour plot.[9] In particular, these plots describe how the cross-sectional distribution of per capita income relative to EU12 has evolved over the 1980-1995 period. The 45-degree diagonal in both graphs highlights persistence properties: if most of the graph were concentrated along this diagonal, then elements in the cross-sectional distribution remain where they started. In contrast, a 90-degree counter-clockwise rotation from that 45-degree diagonal indicates that substantial overtaking occurs, thereby suggesting that poor and rich economies periodically exchange their relative positions over the 15-years horizon under analysis. Finally, a tendency towards convergence to equality over this 15-years horizon in the cross-sectional distribution of per capita income would be signalled by a concentration of most of the graph around the 1-value of the 1995 axis and parallel to the 1980 axis.

As is evident from the figure, in the case of the European NUTS regions, despite a (very) slight counter-clockwise rotation for middle-low income regions suggesting that some degree of overtaking might be present between middle- and low-income regions,

---

[9] Following Paci (1997) and Paci and Pigliaru (1999), the set of regions includes different levels of NUTS regions on the grounds that the NUTS classification is not only quite heterogenous in socio-economic terms but also has, in some cases, no relationship with the administrative organisation of Member countries. In particular, the set adopted combines: NUTS-0 for Denmark, Luxembourg and Ireland; NUTS-1 for Belgium, Germany, Netherlands and UK; NUTS-2 for Italy, France, Spain, Portugal, and Greece (see Appendix A for the list of regions). GDP (adjusted for purchasing power parities and at 1990 prices) and population are based on Eurostat data and refined by CRENOS.

the fact that most of the graph is concentrated along the 45-degree diagonal indicates that persistence is the most evident feature across European regions over the 1980-1995 period. A different outcome is apparent if this same method is applied to data on U.S. state income levels. Johnson (2000) finds evidence of convergence in the cross-sectional income distribution, confirming results obtained by Quah (1996c) by means of the time-homogeneous (finite) Markov Chain methodology.

As emphasised at the outset, the distributional approach to convergence, studying both the shape and mobility dynamics of cross-sectional distributions of per capita income, appears to be generally more informative about the actual patterns of cross-sectional growth than convergence empirics within the regression approach. However, the work just described, while being able to formalise certain facts about the patterns of cross-sectional growth, does not provide an explanation for them. To address this issue, Quah (1996b, 1997a and b) proposes the application of a conditioning scheme. In technical terms, given a set of economies S, a conditioning scheme Ψ is defined (Quah, 1997a) as a collection of triples, one for each economy i in S at time t, where each triple is made of:

(i) an integer lag $\tau_i(t)$;

(ii) a subset $C_i(t)$ of S;

(iii) a set of probability weights $\omega_i(t)$ on S, never positive outside $C_i(t)$.

Within this scheme, the subset $C_i(t)$ identifies the collection of economies which are in some form of functional association, based on a theoretically motivated set of factors, with economy i and hence influence its evolution. Moreover, the set of probability weights $\omega_i(t)$ describe the relative strength of each member of the subset in affecting the evolution of i, while $\tau_i(t)$ represents the delay with which economy i is affected by the development of the economies in $C_i(t)$. Finally, if original observations on per capita incomes are represented by $Y = \{Y_i(t)\colon i \in S \text{ and } t \geq 0\}$, the conditional version $Y \mid \Psi = \widetilde{Y}$ is defined as follows:

$$\widetilde{Y}_i(t) \equiv Y_i(t) \big/ \hat{Y}_i(t)$$

where, for $j \in C_i(t)$,

$$\hat{Y}_i(t) \equiv \sum_j \omega_j(t) Y_i\left[t - \tau_i(t)\right].$$

In other words, observations in the conditional version $Y \mid \Psi = \widetilde{Y}$ are simply obtained normalising each region's observations by the weighted average of per capita income in functionally related regions.

Having defined the conditioning scheme, we can first of all see how a set of factors alters the cross-sectional distribution of income. For instance, suppose that inspection of kernel estimates of the cross-sectional distribution of per capita income at time t suggests the existence of bimodality, i.e. the presence of two convergence clubs. In this case, an interesting question would be whether this feature could be explained by a set of factors. In order to answer this question, the first step is to derive the conditioned version per capita income $Y \mid \Psi = \widetilde{Y}$, where conditioning is based on the chosen set of factors. At this point, to understand if this set of factors actually explains bimodality, all we need is an estimate of the stochastic kernel mapping the unconditional distribution to the conditioned one; then, if most of the graph is concentrated around the 1-value of the axis corresponding to conditioned data, and parallel to the unconditioned data axis, this indicates that the chosen set of factors are actually determining the observed bimodality.

In addition, conditioned income distributions can also give us information on dynamics. In this case, the effect of the set of factors on growth and convergence dynamics over a τ-year period starting at year t can be studied analysing directly the estimate of the stochastic kernel mapping the conditioned distribution at time t to the corresponding distribution at time $t+\tau$.

By means of this conditioning scheme, Quah (1997a) has emphasised the relevance of trade patterns and geographical spillovers for understanding cross-country patterns of economic growth and convergence. Moreover, in a different study (Quah, 1996b), he has also shown that while national macro factors and geographical spillovers must both be considered in order to explain observed distribution dynamics across European NUTS regions in the 1980s, the latter factor appears to play a particularly significant role. But before turning to spatial issues in more detail, it is important to conclude this general overview of the distribution analysis approach with a cautionary note on the use of kernel density estimates. If, as already mentioned, maintaining the income space continuous makes it possible to avoid the restrictive assumptions on the data generating process imposed by Markov chain methods, on the other hand, an important difficulty with the use of kernel density estimation is whether the observed features are actual features of the data as opposed to being artefacts of the natural sampling variability. While the main features of the data are unlikely to be affected by this problem, it must be said that a more rigorous solution has yet to be provided.

4.2 Spatial Interaction Issues Within the Distributional Approach

To avoid misguided inferences, the role of spatial effects has to be properly accounted for in this approach as with others. For example, Bickenback and Bode (2001) emphasise that, although the Markov chain approach requires spatial independence and spatial homogeneity, these assumptions are very rarely tested for.

Some evidence on the potential difficulties arising from the presence of spatial dependence is offered by Magrini (1999), who concentrates on the effects of nuisance spatial dependence, i.e. on spatial dependence arising from measurement problems such as a mismatch between the spatial pattern of the process under study and the boundaries of the observational units. In particular, modelling distribution dynamics as a time homogeneous (finite) Markov chain but choosing to discretise the income space optimally so as to minimise the integrated absolute error of approximation, strong evidence of per capita income absolute convergence among 169 NUTS-II regions over the 1980s is found. In contrast, when attention is shifted to 122 European FURs, i.e. on regions defined so as to minimise the extent of nuisance spatial dependence problems, a clear tendency towards divergence is reported, with six rich regions – Düsseldorf, Hamburg, Stuttgart, München, Frankfurt and Paris – growing away from all the others.

Remaining within Markov chain methods, a decisive step towards integrating local spatial statistics into these methods is taken by Rey (2001). Building on the conditioning scheme developed by Quah and presented above, Rey suggests a number of new spatially explicit measures that can be applied to the study of regional income convergence. Central to these new developments is the spatial Markov matrix, i.e. a modified traditional Markov matrix that conditions a region's transition probabilities on the income class of the region's neighbours. It thus summarises the space-time evolution of income distributions. Parallel to Quah's (1996b) results for European NUTS regions, application of the spatial Markov chain method to U.S. state income data shows that flows across geographically contiguous regions do matter for the evolution of regional income distributions as the upward and downward mobility rates are sensitive to the relative position of adjacent regions. In particular, Rey shows that the probability of a low income state moving upwards decreases as the income level of its neighbours also decreases; and mirroring this, the probability for a high income region moving downward increases as the income of adjacent regions gets lower.

However, despite the fact that a spatial transition matrix, taking substantial spatial dependence into explicit consideration, makes it possible to eliminate one potential source of misspecification within Markov chain methods, it is still true that these approaches impose quite restrictive assumptions on the data generating process and that a continuous first-order Markov process need no longer be even Markov when inappropriately discretised. Once more, a solution to this is represented by stochastic kernel estimation. Moreover, combining stochastic kernel estimation with the conditioning scheme suggested by Quah, it is also possible to evaluate the role played by space on growth and convergence dynamics across open economies. In order to address this issue within the European context, let S be a set of European regions, $y_i(t)$ denote per capita income in region i at time t and $\bar{y}_S(t)$ the corresponding European average value. Moreover, define $Y_i(t)$ as per capita income in region i and a time t relative to the European average. As a result, $Y = \{Y_i(t)$: i in S and $t \geq 0\}$ denotes the observations on regional per capita income relative to the European average. At this point, we can consider the following particular conditioning scheme Ψ. Set the time lag $\tau_i(t)$ equal to zero; moreover, let the subset $C_i(t) = C_i(0)$ identify the set of the five closest[10] regions surrounding i but excluding the region itself, and define $\omega_i(t) = \{1/5 \cdot \bar{y}_S(t)$ on $C_i(0)$ and 0 elsewhere$\}$. In other words, $\hat{Y}$ is the average per capita

[10] These are identified on the basis of great circle distances, using the main administrative city as the region's centre.

income in the five closest regions to i, and $\widetilde{Y} = Y \mid \Psi$ is per capita income in i relative to that in surrounding regions.

Having defined the conditioning scheme, it is now possible to assess the role played by spatial interaction among contiguous regions. Note in fact that a stochastic kernel mapping the unconditional distribution in Y to the conditional $Y \mid \Psi$ allows to confront the original distribution of regional (per capita) income relative to the European average to the spatially conditioned distribution, i.e. the distribution of regional (per capita) income relative to the average in each region's geographical neighbours. As a result, if local spatial factors account for a substantial part of the distribution of incomes across regions, then the stochastic kernel mapping Y to $Y \mid \Psi$ would depart from the identity map. Indeed, Figures 3 conveys precisely this message. In particular, these graphs show the stochastic kernel mapping the unconditional (original) distribution for European NUTS regions in 1980 to the spatially conditioned distribution in the same year. The evident counterclockwise shift in mass to parallel the *original* axis on value 1 of the *spatially conditioned* axis (compared to Figure 2) indicates that local spatial interaction flows do account for a large part of income inequality across European regions, thus confirming earlier results by Quah (1996b).[11] Moreover, in order to get information on the dynamics, Figure 4 provides stochastic kernel representations on the 1980-1995 transition in spatially conditioned incomes. As with unconditioned income data (Figure 2), persistence seems to dominate. Overall, then, the picture that emerges from the estimates presented here is that of a substantial degree of persistence in (relative) per capita income across European regions. Moreover, the use of spatially conditioned income data suggested that a substantial part of this finding can be attributed to spatial factors: once the effect of proximity is allowed for, convergence clearly manifests itself.

But, are these findings robust to the presence of nuisance spatial dependence? As discussed earlier, administratively defined regions are likely to misrepresent both the actual level and the growth rate of per capita income of the underlying economies and

[11] Note, however, that the conditioning scheme adopted here is slightly different from the scheme in Quah 1996b and 1997. In particular, the subset $C_i(t)$ here identifies the five closest regions to i rather than those physically contiguous.

muddle up truly spatial differences. In addition, as Table 1 bears witness, the incidence of nuisance spatial dependence appears to be particularly acute among European NUTS, mainly as a result of the profound degree of heterogeneity that characterises their definition. Further insights are provided in Appendices 3 and 4 which illustrate the growth dynamics of European NUTS and FURs over the period 1980-1995. In particular, Appendix C displays the growth rate of (annual average) per capita GDP for the 110 NUTS regions, grouping them into quintiles. Appendix D conveys the same sort of information for 122 European FURs.[12] The remarkably different dynamics that emerge thus suggest that, if we are to evaluate growth and convergence dynamics across regions correctly, the use of spatial units defined so as to abstract from commuting patterns is pretty much essential. Hence, Figure 5 provides stochastic kernel representations of transition dynamics across 122 Functional Urban Regions over the period 1980-1995. In general, the first thing to note is the previous findings of high persistence across European regions are broadly confirmed as most of the mass is concentrated along the 45-degree diagonal. However, in contrast to the case of the NUTS regions, a twin-peak property now manifests itself for FURs, with a group of richer regions growing away from the rest of the cross-sectional distribution. Hence, as noted elsewhere (Magrini, 1999), the use of data for administratively defined regions effectively runs the risk of concealing important features, as well as changes in those features, of the European regional distribution of income.

The next step is to analyse whether this twin-peak feature can be explained by spatial factors. As before, this can be done by means of the spatial conditioning scheme defined above. Figure 6 thus reports the stochastic kernel mapping the original distribution to the spatially conditioned distribution in 1980. While there still is a pronounced counterclockwise shift in mass to parallel the *original* axis, this shift appears somewhat less pronounced than that observed in Figure 3. Moreover, the twin-peak property still manifests itself. In other words, while geographic proximity of regions with a similar level of per capita income still accounts for a large part of the distribution of income across NUTS regions, this appears to be true to a lesser extent for FURs, i.e. when (at least part of) nuisance spatial dependence is removed via a functional definition of the

12 The full list of FURs in given in Appendix B.

regions. Finally, Figure 7 provides stochastic kernel representations on 15-year transitions in space-conditioned incomes for FURs. The message from unconditioned income (Figures 5) is somewhat amended but not overturned: clearly, high persistence manifests again, but the evidence of twin-peakedness becomes slightly weaker.

5. Conclusions

Do regions converge? At least on the face of it, the large body of empirical research on regional convergence overviewed in this chapter looks something of a disappointment when we try to formulate a decisive answer to this question. Indeed, profoundly different results are obtained from similar datasets using different approaches and methods and no obvious pattern seems to emerge even when attention is concentrated on a particular system of regions. However, not all approaches appear equally reliable and not all results equally convincing. Thus, while fully aware of the dangers from any generalisation, this last section will nonetheless make an effort to articulate a tentative answer by means of a personal interpretation of the main lessons that have so far emerged.

The first lesson is that typically the literature on regional convergence neglects the role of spatial interaction. The traditional neoclassical model of growth, that provides the theoretical framework for much of the empirical work on convergence, has been developed starting from the assumption that the economies are fundamentally closed. Moreover, virtually the same empirical methods originally developed to analyse convergence across nations, in which case the closed-economy assumption can questionably be retained, have been widely used to examine the existence of convergence processes at a sub-national level. However, regions and countries are far from being interchangeable concepts, and once this fact is recognised, two important consequences follow. From a theoretical point of view, convergence in an open-economy version of the neoclassical model of growth should be faster, and possibly more complete, than in the closed-economy case because the traditional source of convergence, the internally financed growth of the stock of capital per worker, is paralleled by interregional interaction that progressively reduces an initial misallocation of resources. Moreover, from an econometric point of view, the recognition that regions are naturally open to a range of economic flows and, consequently, that substantial interaction exists among them calls for an explicit treatment of spatial interaction effects in regional convergence studies. Regrettably, to date this call has gone largely unanswered.

The second lesson emerging from the examination of the different approaches developed for the analysis of income convergence is that empirical methodologies within what we have labelled the 'regression approach' suffer from substantial drawbacks, the most important of which relate to their informative content. Most applications of this approach in fact concentrate on the behaviour of the representative economy and are thus not only silent as to the cross-sectional distribution dynamics but can also be misleading as to the identification of the determinants of growth. There are nonetheless a few exceptions, particularly within time series methods. However, the lack of adequately extended series of data at the regional level hampers the general application of these methodologies. A viable alternative is represented by the 'distributional approach to convergence' that, using stochastic kernels to describe the law of motion of cross-sectional distributions of per capita income, puts emphasis on both shape and mobility dynamics and thus appears to be generally more informative on the actual patterns of cross-sectional growth than convergence empirics within the regression approach. In particular, two directions of empirical research on distribution dynamics strike us as promising. The first is represented by methodologies that allow the income state-space to be continuous and use nonparametric estimates of the stochastic kernel. These avoid some important drawbacks that characterise Markov chain methodologies. The second is the development of conditioning schemes for cross-sectional distributions that, used jointly with stochastic kernel estimates, provide an explanation for the patterns of cross-sectional growth.

We can now return to the question that motivates the chapter and look at the body of empirical research on regional convergence from the particular, and admittedly subjective, angle suggested from these broad lessons. The picture that emerges appears to lend little support to the convergence predictions of the traditional neoclassical model of growth, particularly when we focus on the U.S. case. Here, the traditional tenet is that the substantial lack of legal, cultural, linguistic and institutional barriers to factor movements should favour a process of rapid (and absolute) convergence across regions. Recent work based both on time series and distribution dynamics, however, strongly rejects the hypothesis of absolute convergence and suggests instead that the interregional distribution of per capita income is becoming polarised.

When we turn to the European case, a substantial lack of convergence emerges again but, compared to the U.S. case, this result is somewhat less controversial. Indeed, persistence in income disparities, rather than convergence, has been reported in many studies over a considerable period and the recognition of a European 'regional problem' has also meant that a substantial amount of resources have been spent in an attempt to mitigate its manifestations. Whether regional transfers taking place under structural and cohesion policies have proved to be an ineffective, misplaced or insufficient effort is obviously an important and intensely debated question, but a full account of the ongoing discussion on this issue would lead us way off the mark. Instead, returning to our original question, we can note that persistence is also confirmed by the inspection of the stochastic kernel estimates presented in Section 4, using data on two different sets of European regions.

However, the analysis presented in the latter section served two other purposes. First, it suggested that the use of administratively defined regions, such as the European NUTS, could lead to misleading inferences due to the presence of significant nuisance spatial dependence. In fact, the adoption of a set of functionally defined regions, i.e. of spatial units defined so as to reduce or eliminate nuisance spatial dependence, on the one hand confirms the high persistence across European regions but, on the other, suggests a process of polarisation, with a group of richer regions growing away from the rest of the cross-sectional distribution. Second, it revealed that a substantial part of the features of the cross-sectional (per capita) income distribution can actually be attributed to spatial factors. In particular, the use of a spatially conditioned distribution of income suggested that Europe is characterised by geographic clusters of regions with similar levels of per capita income and that once the effect of geographic proximity is allowed for, convergence tends to manifests itself. While valid in general, however, this finding is again sensitive to the presence of nuisance spatial dependence.

References

[1] Abramovitz, M. (1979), "Rapid Growth Potential and Its Realisation: The Experience of Capitalist Economics in the Postwar Period". In Malinvaud, E. (Ed.), *Economic Growth and Resources: Vol. 1. The Major Issues*, London: Macmillan.

[2] Abramovitz, M. (1986), "Catching Up, Forging Ahead, and Falling Behind". *Journal of Economic History*, 46 (2): 385-406.

[3] Aghion, P. and Howitt, P. (1992), "A Model of Growth Through Creative Destruction". *Econometrica*, 60 (2), 323-351.

[4] Aghion, P. and Howitt, P. (1998), *Endogenous Growth Theory*, Cambridge MA and London: MIT Press.

[5] Anselin, L. (1982), "A Note on Small Sample Properties of Estimators in a First-Order Spatial Autoregressive Model". *Environment and Planning A*, 14, 1023-1030.

[6] Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, London: Kluwer.

[7] Anselin, L. (1995), "Local Indicators of Spatial Association – LISA". *Geographical Analysis*, 27 (2), 93-115.

[8] Anselin, L. and Bera, A. (1998), "Spatial Dependence in Linear Regression Models". In Ullah, A. and Giles, D. (Eds.), *Handbook of Applied Economic Statistics*, New York: Marcel Dekker.

[9] Anselin, L. and Florax, R. J. G. M. (Eds.) (1995), *New Directions in Spatial Econometrics*, Berlin: Springer.

[10] Anselin, L. and Rey, S. J. (1991), "Properties of Tests for Spatial Dependence in Linear Regression Models". *Geographical Analysis*, 23 (2), 112-131.

[11] Anselin, L., Bera, A., Florax, R. J. G. M. and Yoon, M. (1996), "Simple Diagnostic Tests for Spatial Dependence". *Regional Science and Urban Economics*, 26 (1), 77-104.

[12] Anselin, L., Varga, A. and Acs, Z. J. (1998), "Geographic and Sectoral Characteristics of Academic Knowledge Externalities". Working Paper, Bruton Center for Development Studies, University of Texas.

[13] Arellano, M. (1988), "An Alternative Transformation for Fixed Effects Models with Predetermined Variables". Applied Economics Discussion Paper No. 57, Institute of Economics and Statistics, University of Oxford.

[14] Arellano, M. and Bond, S. (1991), "Some Test Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations". *Review of Economic Studies*, 58 (2), 577-297.

[15] Arellano, M. and Bover, O. (1995), "Another Look at the Instrumental Variable Estimation of Error-components Models". *Journal of Econometrics*, 68 (1), 29-51.

[16] Armstrong, H. W. (1995), *Trends and Disparities in Regional GDP per Capita in the European Union, United States and Austria*, Brussels: European Commission report 94/00/74/017. (a)

[17] Armstrong, H. W. (1995), "An Appraisal of the Evidence from Cross-sectional Analysis of the Regional Growth Process within the European Union". In H.

Armstrong and R. Vickerman (Eds.) *Convergence and Divergence among European Regions*, London: Pion. (b)

[18] Armstrong, H. W. (1995), "Convergence among Regions of the European Union, 1950-1990". *Papers in Regional Science*, 74 (2), 143-152. (c)

[19] Azariadis, C. and Drazen, A. (1990), "Threshold Externalities in Economic Development". *Quarterly Journal of Economics*, 105 (2), 501-526.

[20] Badinger, H., Müller, W. and Tondl, G. (2002), "Regional Convergence in the European Union (1985-1999): A Spatial Dynamic Panel Analysis". IEF Working Paper No. 47, Vienna University of Economics.

[21] Barro, R. J. and Sala-i-Martin, X. (1991), "Convergence across States and Regions". *Brooking Papers on Economic Activity*, 1: 107-182.

[22] Barro, R. J. and Sala-i-Martin, X. (1992), "Convergence". *Journal of Political Economy* 100 (2), 223-251. (a)

[23] Barro, R. J. and Sala-i-Martin, X. (1992), "Regional Growth and Migration: a Japanese-US Comparison". *Journal of the Japanese and International Economy*, 6 (4), 312-346. (b)

[24] Barro, R. J. and Sala-i-Martin, X. (1995), *Economic Growth*, New York: McGraw-Hill.

[25] Barro, R. J. and Sala-i-Martin, X. (1997), "Technological Diffusion, Convergence, and Growth". *Journal of Economic Growth*, 2 (1), 1-26.

[26] Barro, R. J., Mankiw, G. N. and Sala-i-Martin, X. (1995), "Capital Mobility in Neoclassical Models of Growth". *American Economic Review*, 85 (1), 103-115.

[27] Baumol, W. J. (1986), “Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show”. *American Economic Review*, 76 (5), 1072-1085.

[28] Birckenbach, F. and Bode, E. (2001), “Markov or Not Markov, This Should Be a Question”. Kiel Institute of World Economics Working Paper No.1086, University of Kiel.

[29] Bentivogli, C. and Pagano, P. (1999), “Regional Disparities and Labour Mobility: The EURO-11 versus the USA”. *Labour*, 13 (3), 737-760.

[30] Bernard, A. B. and Durlauf, S. N. (1995), “Convergence in International Output”. *Journal of Applied Econometrics*, 10 (2), 97-108.

[31] Bernard, A. B. and Durlauf, S. N. (1996), “Interpreting Tests of the Convergence Hypothesis”. *Journal of Econometrics*, 71 (1-2), 161-173.

[32] Blanchard, O. J. (1991), “Comments on [Barro and Sala-i-Martin, ‘Convergence across States and Regions’]”. *Brooking Papers on Economic Activity*, 1, 159-174.

[33] Blanchard, O. J. and Katz, L. F. (1992), “Regional Evolutions”. *Brooking Papers on Economic Activity*, 1, 1-75.

[34] Blundell, R. and Bond, S. (1998), “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”. *Journal of Econometrics*, 87 (1), 115-143.

[35] Bond, S., Hoeffler, H. and Temple, J. (2001), “GMM Estimation of Empirical Growth Models”. CEPR Discussion Paper No. 3048, London: CEPR.

[36] Borts, G. H. and Stein, J. L. (1964), *Economic Growth in a Free Market*, New York: Columbia University Press.

[37] Bulli, S. (2001), "Distribution Dynamics and Cross-country Convergence: A New Approach". *Scottish Journal of Political Economy*, 48 (2), 226-243.

[38] Button, K. and Pentecost, E. J. (1995), "Testing for Convergence of the EU Regional Economies". *Economic Enquiry*, 33 (4), 664-671.

[39] Button, K. and Pentecost, E. J. (1999), *Regional Economic Performance within the European Union*, Cheltenham: Elgar.

[40] Caniëls, M. C. J. (2000), *Knowledge Spillovers and Economic Growth*, Cheltenham: Edward Elgar.

[41] Caniëls, M. C. J. and Verspagen, B. (2001), "Barriers to Knowledge Spillovers and Regional Convergence in an Evolutionary Model". *Journal of Evolutionary Economics*, 11 (3), 307-329.

[42] Canova, F. and Marcet, A. (1995), "The Poor Stay Poor: Non-convergence across Countries and Regions". CEPR Discussion Paper No. 1265, London: CEPR.

[43] Carlino, G. and Mills, L. (1993), "Are U.S. Regional Incomes Converging?". *Journal of Monetary Economics*, 32 (2), 335-346.

[44] Carlino, G. and Mills, L. (1996), "Testing Neoclassical Convergence in Regional Incomes and Earnings". *Regional Science and Urban Economics*, 26 (6), 565-590. (a)

[45] Carlino, G. and Mills, L. (1996), "Convergence and the U.S. States: a Time-Series Analysis". *Journal of Regional Science*, 36 (4), 587-616. (b)

[46] Carvalho, V. M. and Harvey, A. C. (2002), "Growth, Cycles and Convergence in U.S. Regional Time Series". Cambridge Working Papers in Economics No. 0221, Cambridge University.

[47] Caselli, F., Esquivel, G. and Lefort, F. (1996), "Reopening the Convergence Debate: A New Look at Cross-country Growth Empirics". *Journal of Economic Growth*, 1 (3), 363-389.

[48] Cashin, P. (1995), "Economic Growth and Convergence across the Seven Colonies of Australia: 1861-1991". *Economic Record*, 71 (213), 132-144.

[49] Cass, D. (1965), "Optimum Growth in an Aggregative Model of Capital Accumulation". *Review of Economic Studies*, 32, 233-240.

[50] Chatterji, M. (1992), "Convergence Clubs and Endogenous Growth". *Oxford Review of Economic Policy*, 8 (4), 57-69.

[51] Chatterji, M. and Dewhurst, J. H. L. (1996), "Convergence Clubs and Relative Economic Performance in Great Britain: 1977-1991" *Regional* Studies 30 (1), 31-40.

[52] Cheshire, P. C. and Carbonaro, G. (1995), "Convergence-Divergence in Regional Growth Rates: An Empty Black Box?". In Armstrong, H.W. and Vickerman, R.W. (Eds.), *Convergence and Divergence Among European Regions*, London: Pion.

[53] Cheshire, P. C. and Hay, D. G. (1989), *Urban Problems in Western Europe: an economic analysis*, London: Unwin Hyman.

[54] Cheshire, P. C. and Magrini, S. (2000), "Endogenous Processes in European Regional Growth: Convergence and Policy". *Growth and Change*, 31 (4), 455-479.

[55] Cheshire, P. C. and Magrini, S. (2002), “The Distinctive Determinants of European Urban Growth: Does One Size Fit All?”. Research Papers in Environmental and Spatial Analysis No. 73, London School of Economics.

[56] Cliff, A. and Ord, J. (1973), *Spatial Autocorrelation*, London: Pion.

[57] Coulombe, S. and Day, K. (1996), “β-convergence, σ-convergence, and the Stationary-state Level of Regional Disparities: the Case of Canada”. University of Ottawa Research Paper No. 9608E.

[58] Coulombe, S. and Lee, F. (1993), “Regional Economic Disparities in Canada”. University of Ottawa Research Paper No. 9317E.

[59] Coulombe, S. and Lee, F. (1995), “Convergence across Canadian Provinces, 1961 to 1991”. *Canadian Journal of Economics*, 28 (4a), 886-898.

[60] Coulombe, S. and Tremblay, J.-F. (2001), “Human Capital and Regional Convergence in Canada”. *Journal of Economic Studies*, 28 (2-3), 154-180.

[61] Cuadrado-Roura, J. R., Mancha-Navarro, T. and Garrido-Yserte R. (2000), “Regional Productivity Patterns in Europe: An Alternative Approach”. *Annals of Regional Science*, 34 (3), 365-384. (a)

[62] Cuadrado-Roura, J. R., Mancha-Navarro, T. and Garrido-Yserte R. (2000), “Convergence and Regional Mobility in the European Union” 40th Congress of the European Regional Science Association, Barcelona. (b)

[63] De la Fuente, A. (1996), “On the Sources of Growth and Convergence: A Closer Look at the Spanish Regions”. CEPR Discussion Paper No. 1543, London: CEPR.

[64] De la Fuente, A. (2000), “Convergence Across Countries and Regions : Theory and Empirics”. CEPR Discussion Paper No. 2465, London: CEPR.

[65] De la Fuente, A. (2002), “On the Sources of Convergence: A Close Look at the Spanish Regions”. *European Economic Review*, 46 (3), 569-599.

[66] De la Fuente, A. and Vives, X. (1995), “Infrastructure and Education as Instruments of Regional Policy: Evidence from Spain”. *Economic Policy: A European Forum*, 0 (20), 11-51.

[67] Decressin, J. W. (1994), “Internal Migration in West Germany and Implications for East-West Salary Convergence”. *Weltwirschaftliches Archiv*, 130 (2), 231-257.

[68] Decressin, J. W. and Fatas, A. (1995), “Regional Labour Market Dynamics in Europe”. *European Economic Review*, 39 (9), 1627-1655.

[69] DeLong, J. B. (1988), “Productivity Growth, Convergence and Welfare: Comment”. *American Economic Review*, 78 (5), 1138-1154.

[70] Durlauf, S. N. (1996), “A Theory of Persistent Income Inequality”. *Journal of Economic Growth*, 1(1), 75-93.

[71] Durlauf, S. N. and Quah, D. T. (1999), “The New Empirics of Economic Growth”. In Taylor, J. and Woodford, M. (Eds.), *Handbook of Macroeconomics*, Amsterdam: North-Holland.

[72] Eichengreen, B. (1993), “Labour Markets and European Monetary Unification”. In Masson, P. R. and Taylor, M. P. (Eds.), *Policy Issues in the Operation of Currency Unions*, Cambridge, UK: Cambridge University Press.

[73] Evans, P. and Karras, G. (1996), "Convergence Revisited". *Journal of Monetary Economics*, 37 (2), 249-265. (a)

[74] Evans, P. and Karras, G. (1996), "Do Economies Converge? Evidence from a Panel of US States". *Review of Economics and Statistics*, 78 (3), 384-388. (b)

[75] Fabiani, S. and Pellegrini, G. (1997), "Education, Infrastructure, Geography and Growth: An Empirical Analysis of the Development of Italian Provinces". Temi di discussione No. 323, Roma: Banca d'Italia.

[76] Fagerberg, J. (1988), "International Competitiveness". *Economic Journal*, 98 (391), 355-374.

[77] Fagerberg, J. (1994), "Technology and International Differences in Growth Rates". *Journal of Economic Literature*, 32 (3), 1147-1175.

[78] Fagerberg J. and Verspagen, B. (1996), "Heading for Divergence? Regional Growth in Europe Reconsidered". *Journal of Common Market Studies*, 34 (3), 431-448.

[79] Fingleton, B. (1997), "Specification and Testing of Markov Chain Models: An Application to Convergence in the European Union". *Oxford Bulletin of Economics and Statistics*, 59 (3), 385-403.

[80] Fingleton, B. (1999), "Estimates of Time to Convergence: An Analysis of Regions of the European Union". *International Regional Science Review*, 22 (1), 5-34.

[81] Friedman, M. (1992), "Do Old Fallacies Ever Die?". *Journal of Economic Literature*, 30 (4), 2129-2132.

[82] Funke, M. and Strulik, H. (1999), “Regional Growth in West Germany: Convergence or Divergence?”. *Economic Modelling*, 16 (4), 489-502.

[83] Galor, O. and Zeira, J. (1993), “Income Distribution and Macroeconomics”. *Review of Economic Studies*, 60 (1), 35-52.

[84] Garofalo, G. A. and Yamarik, S. (2002), “Regional Convergence: Evidence from a New State-by-State Capital Stock Series”. *Review of Economics and Statistics*, 84 (2), 316-323.

[85] Gerschenkron, A. (1962), *Economic Backwardness in Historical Perspective*, Cambridge MA: Belknap.

[86] Getis, A. and Griffith, D. A. (2002), “Comparative Spatial Filtering in Regression Analysis”. *Geographical Analysis*, 34 (2), 130-140.

[87] Getis, A. and Ord, J. K. (1992), “The Analysis of Spatial Association by the Use of Distance Statistics”. *Geographical Analysis*, 24 (3), 189-206.

[88] Grossman, G. M. and Helpman, E. (1990), “Comparative Advantage and Long-Run Growth”. *American Economic Review*, 80 (4), 796-815.

[89] Grossman, G. M. and Helpman, E. (1991), *Innovation and Growth in the Global Economy*, Cambridge MA: MIT Press.

[90] Hall, P. and Hay, D. G. (1980), *Growth Centres in the European Urban System*, London: Heinemann.

[91] Hall, R. E. (1992), “Comments on [Blanchard and Katz, ‘Regional Evolutions’]”. *Brooking Papers on Economic Activity*, 1, 62-65.

[92] Harvey, A. and Carvalho, V. M. (2002), “Models for Converging Economies”. Cambridge Working Papers in Economics No. 0216, Cambridge University.

[93] Helpman, E. (1993), “Innovation, Imitation, and Intellectual Property Rights”. *Econometrica*, 61 (6), 1247-1280.

[94] Herz, B. and Röger, W. (1995), “Economic Growth and Convergence in Germany”. *Weltwirtschaftliches Archiv*, 131 (1), 133-143.

[95] Hofer, H. and Wörgötter, A. (1997), “Regional Per Capita Income Convergence in Austria”. *Regional Studies*, 31 (1), 1-12.

[96] Holz-Eakin, D. (1993), “Solow and the States: Capital Accumulation, Productivity, and Economic Growth”. *National Tax Journal*, 46 (4), 425-439.

[97] Hsiao, C. (1986), *Analysis of Panel Data*, Econometric Society Monographs, Cambridge: Cambridge University Press.

[98] Islam, N. (1995), “Growth Empirics: A Panel Data Approach”. *Quarterly Journal of Economics*, 110 (4), 1127-1170.

[99] Johnson, P. (2000), “A Nonparametric Analysis of Income Convergence across the US States”. *Economic Letters*, 69 (2), 219-223.

[100] Koopmans, T. (1965), “On the Concept of Optimal Economic Growth”. In *The Econometric Approach to Development Planning*, Amsterdam: North-Holland.

[101] Lall, S. and Yilmaz, S. (2001), “Regional Economic Convergence: Do Policy Instruments Make a Difference?”. *Annals of Regional Science*, 35 (1), 151-166.

[102] Lee, F. and Coulombe, S. (1995), “Regional Productivity Convergence in Canada”. *Canadian Journal of Regional Science*, 18 (1), 39-56.

[103] Levin, A., Lin, C.-F. and Chu, C.-S. J. (2002), "Unit Root Tests in Panel Data: Asymptotic and Final-Sample Properties". *Journal of Econometrics*, 108 (1), 1-24.

[104] Loewy, M. B. and Papell, D. H. (1996), "Are U.S. Regional Incomes Converging? Some Further Evidence". *Journal of Monetary Economics*, 38 (3), 587-598.

[105] López-Bazo, E., Vayá E., Mora, A. J. and Suriñach, J. (1999), "Regional Economic Dynamics and Convergence in the European Union". *Annals of Regional Science*, 33 (3), 343-370.

[106] Lucas, R. E. (1988), "On the Mechanics of Economic Development". *Journal of Monetary Economics*, 22 (1), 3-42.

[107] Magrini, S. (1999), "The Evolution of Income Disparities among the Regions of the European Union". *Regional Science and Urban Economics*, 29 (2), 257-281.

[108] Mankiw, G. N.; Romer, D.; and Weil, D.N. (1992), "A Contribution to the Empirics of Economic Growth". *Quarterly Journal of Economics*, 107 (2), 407-437.

[109] Martin, R. (2001), "EMU versus the Regions? Regional Convergence and Divergence in Euroland". *Journal of Economic Geography*, 1 (1), 51-80.

[110] Maurseth, P. B. (2001), "Convergence, Technology and Geography". *Structural Change and Economic Dynamics*, 12 (3), 247-276.

[111] Murphy, K. M., Shleifer, A. and Vishny, R. W. (1989), "Industrialization and the Big Push". *Journal of Political Economy*, 97 (4), 1003-1026.

[112] Nahar, S. and Inder, B. (2002), "Testing Convergence in Economic Growth for OECD Countries". *Applied Economics*, 34 (16), 2011-2022.

[113] Nelson, R. R. and Phelps, E. S. (1966), "Investment in Humans, Technological Diffusion, and Economic Growth". *American Economic Review*, 56 (2), 69-75.

[114] Neven, D. and Gouyette, C. (1995), "Regional Convergence in the European Community". *Journal of Common Market Studies*, 33 (1), 47-65.

[115] Niebuhr, A. (2001), "Convergence and the Effects of Spatial Interaction". *Jahrbuch für Regionalwissenschaft*, 21 (2), 113-133.

[116] Nickell, S. (1981), "Biases in Dynamic Models with Fixed Effects". *Econometrica*, 49 (6), 1417-1426.

[117] Paci, R. (1997), "More Similar and Less Equal: Economic Growth in the European Regions". *Weltwirtschaftliches Archiv*, 133 (4), 609-634.

[118] Paci, R. and Pigliaru, F. (1995), "Differenziali di Crescita tra le Regioni Italiane: un'Analisi Cross-section". *Rivista di Politica Economica*, 85 (10), 3-34.

[119] Paci, R. and Pigliaru, F. (1997), "Structural Change and Convergence: an Italian Regional Perspective". *Structural Change and Economic Dynamics*, 8 (3), 297-318.

[120] Paci R. and Pigliaru, F. (1999), "European regional growth: Do sectors matter?". In Adams, J. and Pigliaru, F. (Eds.), *Economic growth and change. National and regional patterns of convergence and divergence*, Cheltenham: Edward Elgar.

[121] Paci, R. and Pigliaru, F. (2001), "Technological Diffusion, Spatial Spillovers and Regional Convergence in Europe". CRENOS Working Paper No. 1/01.

[122] Pesaran, H. and Smith, R. (1995), "Estimating Long Run Relationship from Dynamic Heterogeneous Panels". *Journal of Econometrics*, 68 (1), 79-113.

[123] Pissarides, C. A. and McMaster, I. (1990), "Regional Migration, Wages and Unemployment: Empirical Evidence and Implications for Policy". *Oxford Economic Papers*, 42 (4), 812-831.

[124] Quah, D. T. (1993), "Empirical Cross-section Dynamics in Economic Growth". *European Economic Review*, 37 (2-3), 426-434. (a)

[125] Quah, D. T. (1993), "Galton's Fallacy and Tests of the Convergence Hypothesis". *Scandinavian Journal of Economics*, 95 (4), 427-443. (b)

[126] Quah, D. T. (1994), "One Business Cycle and One Trend from (Many,) Many Disaggregates". *European Economic Review*, 38 (3-4), 605-613.

[127] Quah, D. T. (1996), "Convergence Empirics Across Economies with (Some) Capital Mobility". *Journal of Economic Growth*, 1 (1), 95-124. (a)

[128] Quah, D. T. (1996), "Regional Convergence Clusters across Europe". *European Economic Review*, 40 (3-5), 951-958. (b)

[129] Quah, D. T. (1996), "Empirics for Economic Growth and Convergence". *European Economic Review*, 40 (6), 1353-1375. (c)

[130] Quah, D. T. (1996), "Ideas Determining Convergence Clubs". LSE Economics Department Working Paper, London School of Economics. (d)

[131] Quah, D. T. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs". *Journal of Economic Growth*, 2 (1), 27-59. (a)

[132] Quah, D. T. (1997), “Regional Cohesion from Local Isolated Actions: I. Historical Outcomes”. CEP Discussion Paper No. 378, London: CEP. (b)

[133] Ramsey, F. (1928), “A Mathematical Theory of Saving”. *Economic Journal*, 38 (152), 543-559.

[134] Rey, S. J. (2001), “Spatial Empirics for Economic Growth and Convergence”. *Geographical Analysis*, 33 (3), 194-214.

[135] Rey S. J. and Montuori, B. D. (1999), “US Regional Income Convergence: A Spatial Econometric Perspective”. *Regional Studies*, 33 (2), 143-156.

[136] Rivera-Batiz, L. A. and Romer, P. (1991), “International Trade with Endogenous Technological Change”. *European Economic Review*, 35 (4), 971-1004.

[137] Rodríguez-Pose, A. (1999), “Convergence or Divergence? Types of Regional Responses to Socioeconomic Change”. *Journal of Economic and Social Geography*, 90 (4), 363-378.

[138] Romer, P. (1986), “Increasing Returns and Long Run Growth”. *Journal of Political Economy*, 94 (5), 1002-1037.

[139] Romer, P. (1987), “Crazy Explanations for the Productivity Slowdown”. In Fisher, S. (Ed.), *NBER Macroeconomics Annual*, Cambridge MA: MIT Press.

[140] Romer, P. (1990), “Capital, Labor and Productivity”. *Brookings Papers on Economic Activity: Microeconomics*, 337-367. (a)

[141] Romer, P. (1990), “Endogenous Technological Change”. *Journal of Political Economy*, 98 (5), S71-S102. (b)

[142] Romer, P. (1990), "Are Nonconvexities Important for Understanding Growth?". NBER Working Paper No. 3271. (c)

[143] Romer, P. (1993), "Idea Gaps and Object Gaps in Economic Development". *Journal of Monetary Economics*, 32 (3), 543-573.

[144] Romer, P. (1994), "The Origins of Endogenous Growth". *Journal of Economic Perspectives*, 8 (1), 3-22.

[145] Sala-i-Martin, X. (1996), "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence". *European Economic Review*, 40 (6), 1325-1352.

[146] Segerstrom, P. S. (1991), "Innovation, Imitation, and Economic Growth". *Journal of Political Economy*, 99 (4), 807-827.

[147] Shioji, E. (1996), "Regional Growth in Japan". CEPR Discussion Paper No. 1425, London: CEPR.

[148] Siriopoulos, C. and Asteriou, D. (1998), "Testing for Convergence Across Greek Regions". *Regional Studies*, 32 (6), 537-546.

[149] Solow, R. M. (1956), "A Contribution to the Theory of Economic Growth". *Quarterly Journal of Economics*, 70 (5), 65-94.

[150] Summers, R. and Hestons, A. (1988), "A New Set of International Comparisons of Real Product and Price Levels: Estimates for 130 Countries, 1950-1985". *Review of Income and Wealth*, 34 (1), 1-25.

[151] Swan, T. W. (1956), "Economic Growth and Capital Accumulation". *Economic Record*, 32 (63), 334-361.

[152] Temple, J. (1999), “The New Growth Evidence”. *Journal of Economic Literature*, 37 (1), 112-156.

[153] Tomljanovich, M. and Vogelsang, T. J. (2001), “Are U.S. Regions Converging? Using New Econometric Methods to Examine Old Issues”. *Empirical Economics*, 27 (1), 49-62.

[154] Tondl, G. (1999), “The Changing Pattern of Regional Convergence in Europe”. *Jahrbuch fur Regionalwissenschaft*, 19 (1), 1-33.

[155] Tondl, G. (2001), *Convergence after Divergence? Regional Growth in Europe*, Wien: Springer.

[156] Tsionas, E. G. (2001), “Regional Convergence and Common, Stochastic Long-run Trends: A Re-examination of the US Regional Data”. *Regional Studies*, 35 (8), 689-696.

[157] Verspagen, B. (1991), “A New Theoretical Approach to Catching Up and Falling Behind”. *Structural Change and Economic Dynamics*, 2 (2), 359-380.

[158] Vohra, R. (1996), “How Fast Do We Grow?”. *Growth and Change*, 27 (1), 47-54.

[159] Williamson, J.G. (1965), “Regional Inequality and the Process of National Development: A Description of the Patterns”. *Economic Development and Cultural Change*, 13 (4), 3-45.

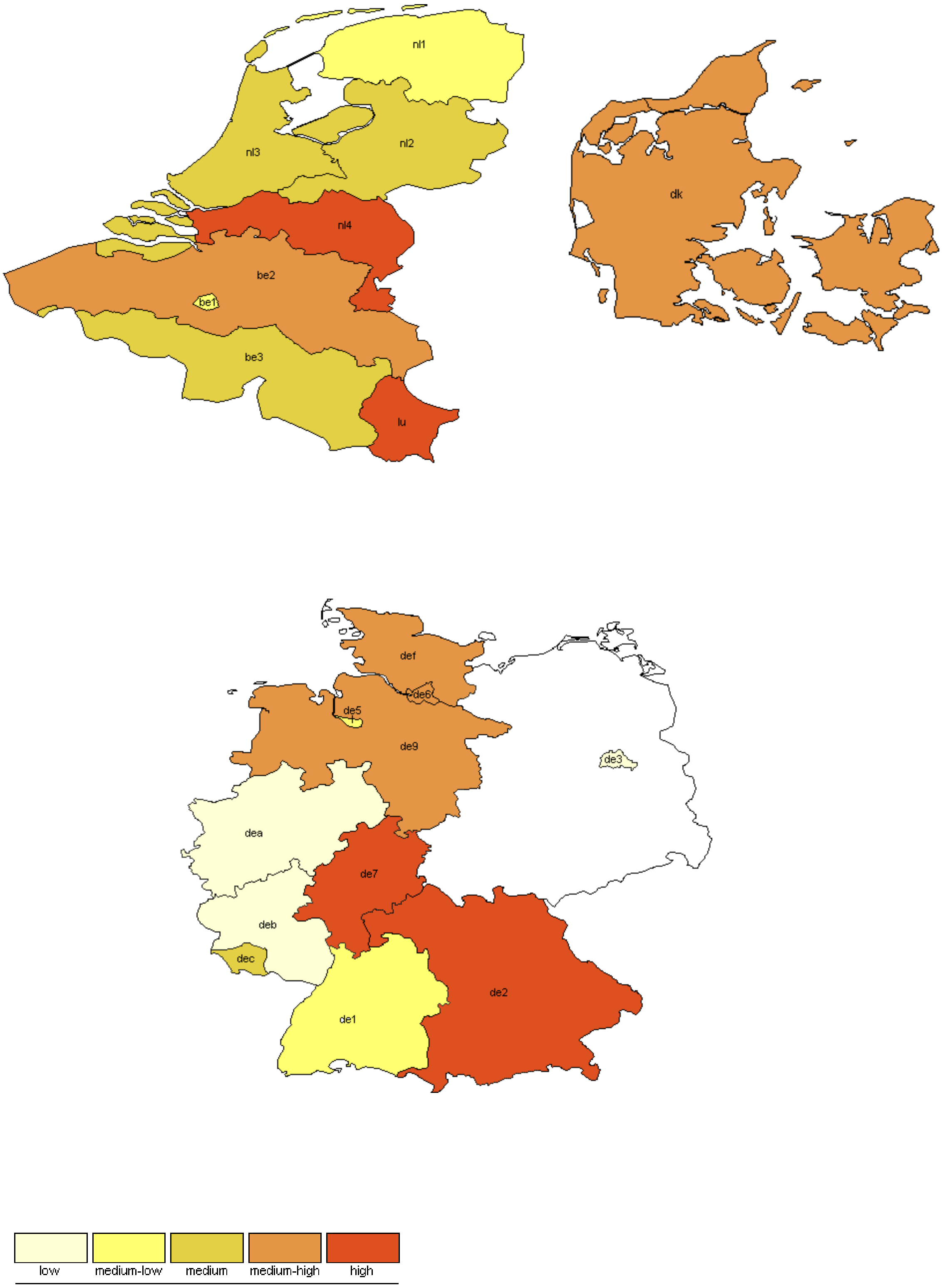
Appendix A: NUTS regions

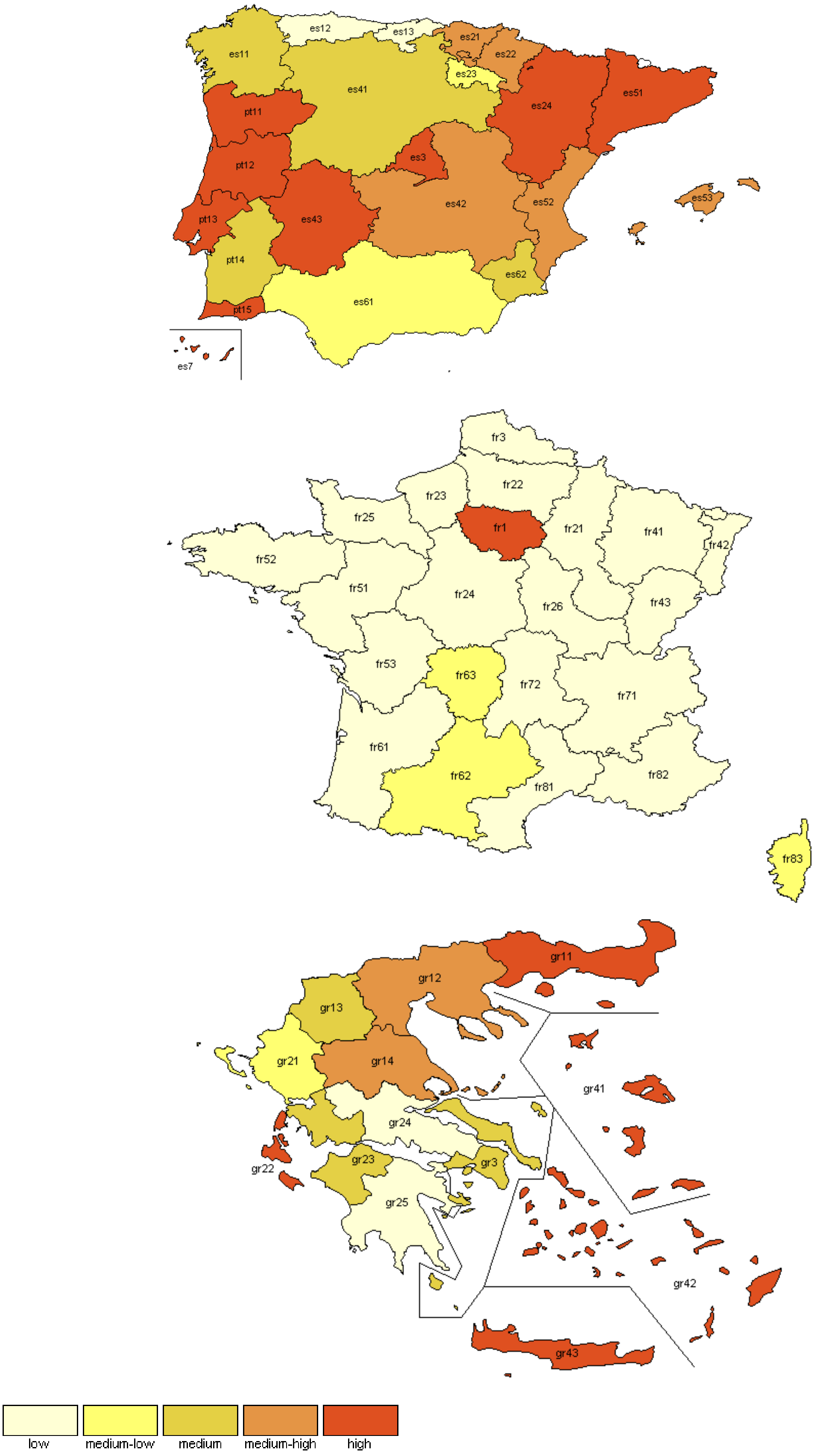
Code	Name	Code	Name	Code	Name
be1	Région Bruxelles-cap. Brussels hoofdstad gewest	es43	Extremadura	it33	Friuli-Venezia Giulia
be2	Vlaams Gewest	es51	Cataluña	it4	Emilia Romagna
be3	Région Wallonne	es52	Comunidad Valenciana	it51	Toscana
dk	Denmark	es53	Baleares	it52	Umbria
de1	Baden-Württemberg	es61	Andalucia	it53	Marche
de2	Bayern	es62	Murcia	it6	Lazio
de3	Berlin	es7	Canarias (ES)	it71	Abruzzo
de5	Bremen	fr1	Île de France	it72	Molise
de6	Hamburg	fr21	Champagne-Ardenne	it8	Campania
de7	Hessen	fr22	Picardie	it91	Puglia
de9	Niedersachsen	fr23	Haute-Normandie	it92	Basilicata
dea	Nordrhein-Westfalen	fr24	Centre	it93	Calabria
deb	Rheinland-Pfalz	fr25	Basse-Normandie	ita	Sicilia
dec	Saarland	fr26	Bourgogne	itb	Sardegna
def	Schleswig-Holstein	fr3	Nord - Pas-de-Calais	lu	Luxembourg
gr11	Anatoliki Makedonia, Thraki	fr41	Lorraine	nl1	Noord-Nederland
gr12	Kentriki Makedonia	fr42	Alsace	nl2	Oost-Nederland
gr13	Dytiki Makedonia	fr43	Franche-Comté	nl3	West-Nederland
gr14	Thessalia	fr51	Pays de la Loire	nl4	Zuid-Nederland
gr21	Ipeiros	fr52	Bretagne	pt11	Norte
gr22	Ionia Nisia	fr53	Poitou-Charentes	pt12	Centro (P)
gr23	Dytiki Ellada	fr61	Aquitaine	pt13	Lisboa e Vale do Tejo
gr24	Sterea Ellada	fr62	Midi-Pyrénées	pt14	Alentejo
gr25	Peloponnisos	fr63	Limousin	pt15	Algarve
gr3	Attiki	fr71	Rhône-Alpes	ukc	North East
gr41	Voreio Aigaio	fr72	Auvergne	ukd	North West
gr42	Notio Aigaio	fr81	Languedoc-Roussillon	uke	Yorkshire and The Humber
gr43	Kriti	fr82	Provence-Alpes-Côte d'Azur	ukf	East Midlands
es11	Galicia	fr83	Corse	ukg	West Midlands
es12	Principado de Asturias	ie	Ireland	ukh	Eastern
es13	Cantabria	it11	Piemonte	uki	London
es21	Pais Vasco	it12	Valle d'Aosta	ukj	South East
es22	Comun. Foral de Navarra	it13	Liguria	ukk	South West
es23	La Rioja	it2	Lombardia	ukl	Wales
es3	Comunidad de Madrid	it31	Trentino - Alto Adige	ukm	Scotland
es41	Castilla y León	it32	Veneto	ukn	Northern Ireland
es42	Castilla - La Mancha				

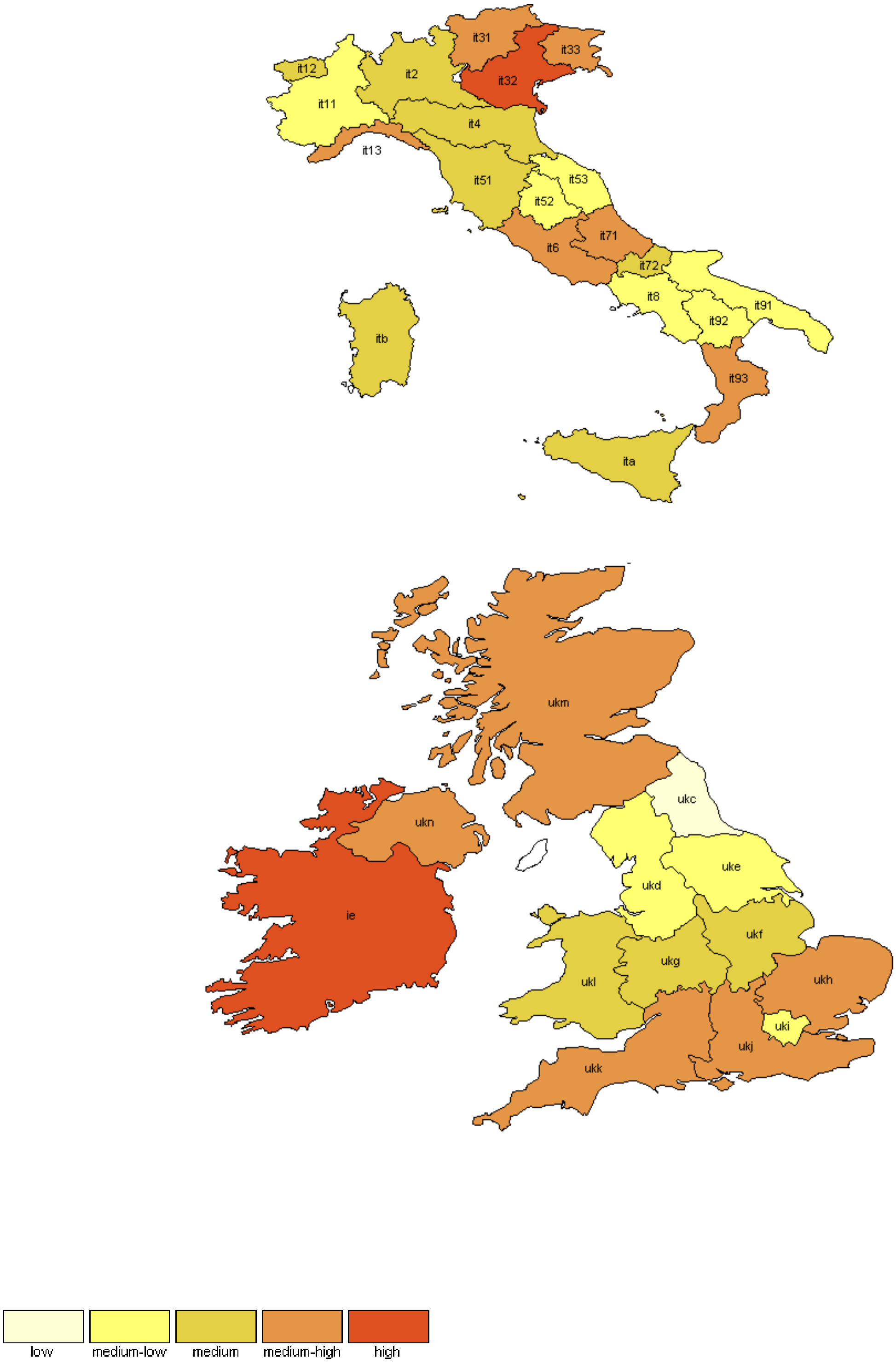
Appendix B: Functional Urban Regions

Code	Name
1	Antwerpen
2	Bruxelles-Brussel
3	Charleroi
4	Liège
5	Århus
6	Københavns
7	Aachen
8	Augsburg
9	Berlin
10	Bielefeld
11	Bochum
12	Bonn
13	Braunschweig
14	Bremen
15	Dortmund
16	Düsseldorf
17	Duisburg
18	Essen
19	Frankfurt
20	Hamburg
21	Hannover
22	Karlsruhe
23	Kassel
24	Köln
25	Krefeld
26	Mannheim
27	Mönchengladbach
28	München
29	Münster
30	Nürnberg
31	Saarbruecken
32	Stuttgart
33	Wiesbaden
34	Wuppertal
35	Athinai
36	Saloniki
37	Alicante
38	Barcelona
39	Bilbao
40	Cordoba
41	Gijon/Aviles
42	Granada
43	La Coruna
44	Madrid
45	Málaga
46	Murcia
47	Palma De Mallorca
48	Sevilla
49	Valencia
50	Valladolid
51	Vigo
52	Zaragoza
53	Bordeaux
54	Clermont-Ferrand
55	Dijon
56	Grenoble
57	Le Havre
58	Lille
59	Lyon
60	Marseille
61	Montpellier
62	Mulhouse
63	Nancy
64	Nantes
65	Nice
66	Orléans
67	Paris
68	Rennes
69	Rouen
70	St. Etienne
71	Strasbourg
72	Toulon
73	Toulouse
74	Valenciennes
75	Dublin
76	Bari
77	Bologna
78	Brescia
79	Cagliari
80	Catania
81	Firenze
82	Genova
83	Messina
84	Milano
85	Napoli
86	Padova
87	Palermo
88	Roma
89	Taranto
90	Torino
91	Venezia
92	Verona
93	Amsterdam
94	Rotterdam
95	S-Gravenhage
96	Utrecht
97	Lisboa
98	Porto
99	Belfast
100	Birmingham
101	Brighton
102	Bristol
103	Cardiff
104	Coventry
105	Derby
106	Edinburgh
107	Glasgow
108	Hull
109	Leeds
110	Leicester
111	Liverpool
112	London
113	Manchester
114	Newcastle
115	Nottingham
116	Plymouth
117	Portsmouth
118	Sheffield
119	Southampton
120	Stoke
121	Sunderland
122	Teesside

Appendix C: NUTS – per capita GDP (annual average) growth 1980-1995

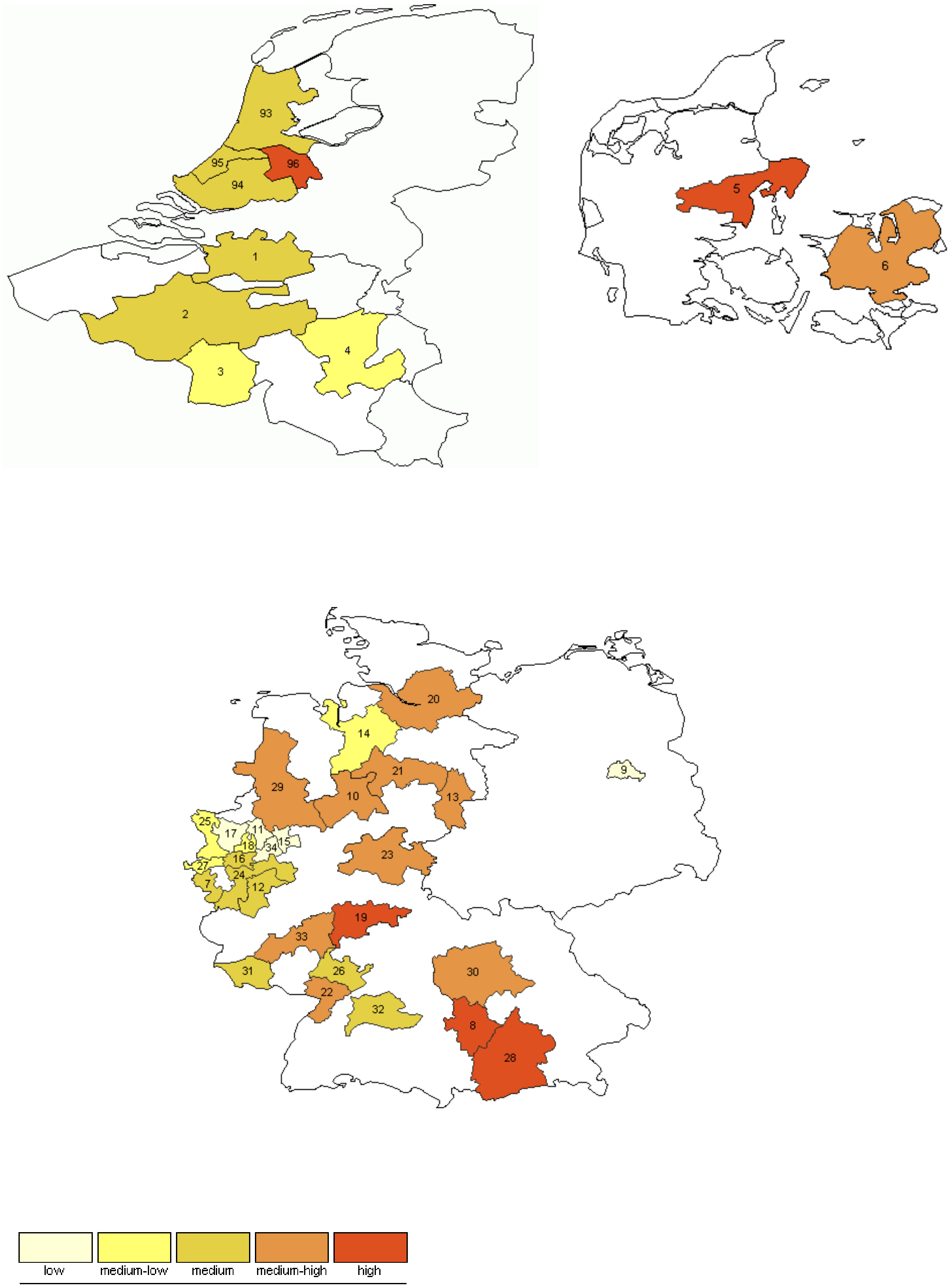


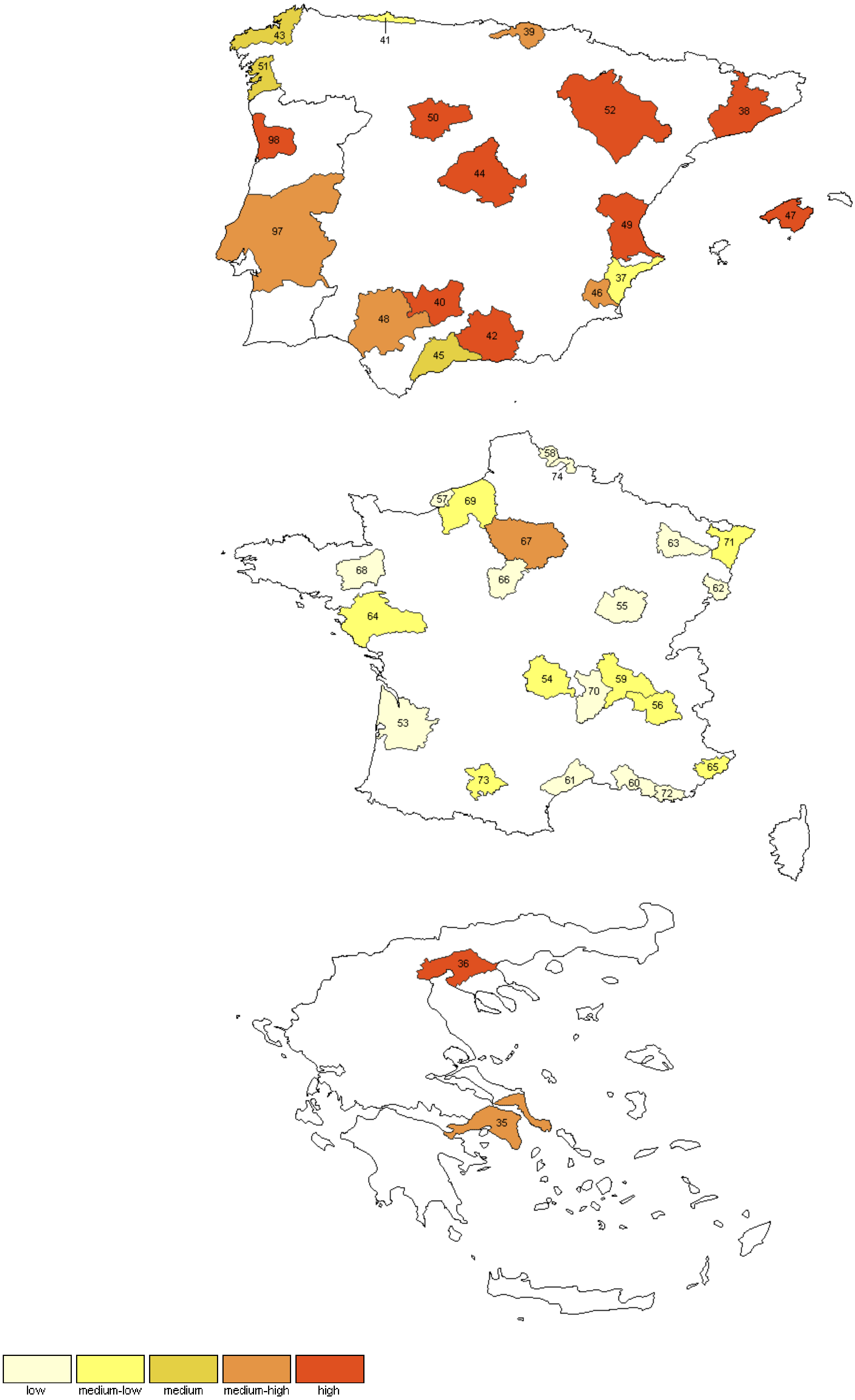

es11
es12
es13
es21
es22
es23
es24
es3
es41
es42
es43
es51
es52
es53
es61
es62
es7
pt11
pt12
pt13
pt14
pt15
fr1
fr21
fr22
fr23
fr24
fr25
fr26
fr3
fr41
fr42
fr43
fr51
fr52
fr53
fr61
fr62
fr63
fr71
fr72
fr81
fr82
fr83
gr11
gr12
gr13
gr14
gr21
gr22
gr23
gr24
gr25
gr3
gr41
gr42
gr43
low
medium-low
medium
medium-high
high



it12
it2
it31
it33
it32
it11
it4
it13
it51
it53
it52
it71
it6
it72
it8
it91
it92
itb
it93
ita
ukm
ukc
ukn
uke
ie
ukd
ukf
ukg
ukl
ukh
uki
ukj
ukk
low
medium-low
medium
medium-high
high

Appendix D: FURs – per capita GDP (annual average) growth 1980-1995




43
41
39
51
50
52
38
98
44
97
49
47
37
46
40
48
42
45
58
74
57
69
67
63
71
68
66
62
55
64
54
70
59
56
53
73
61
60
72
65
36
35
low
medium-low
medium
medium-high
high

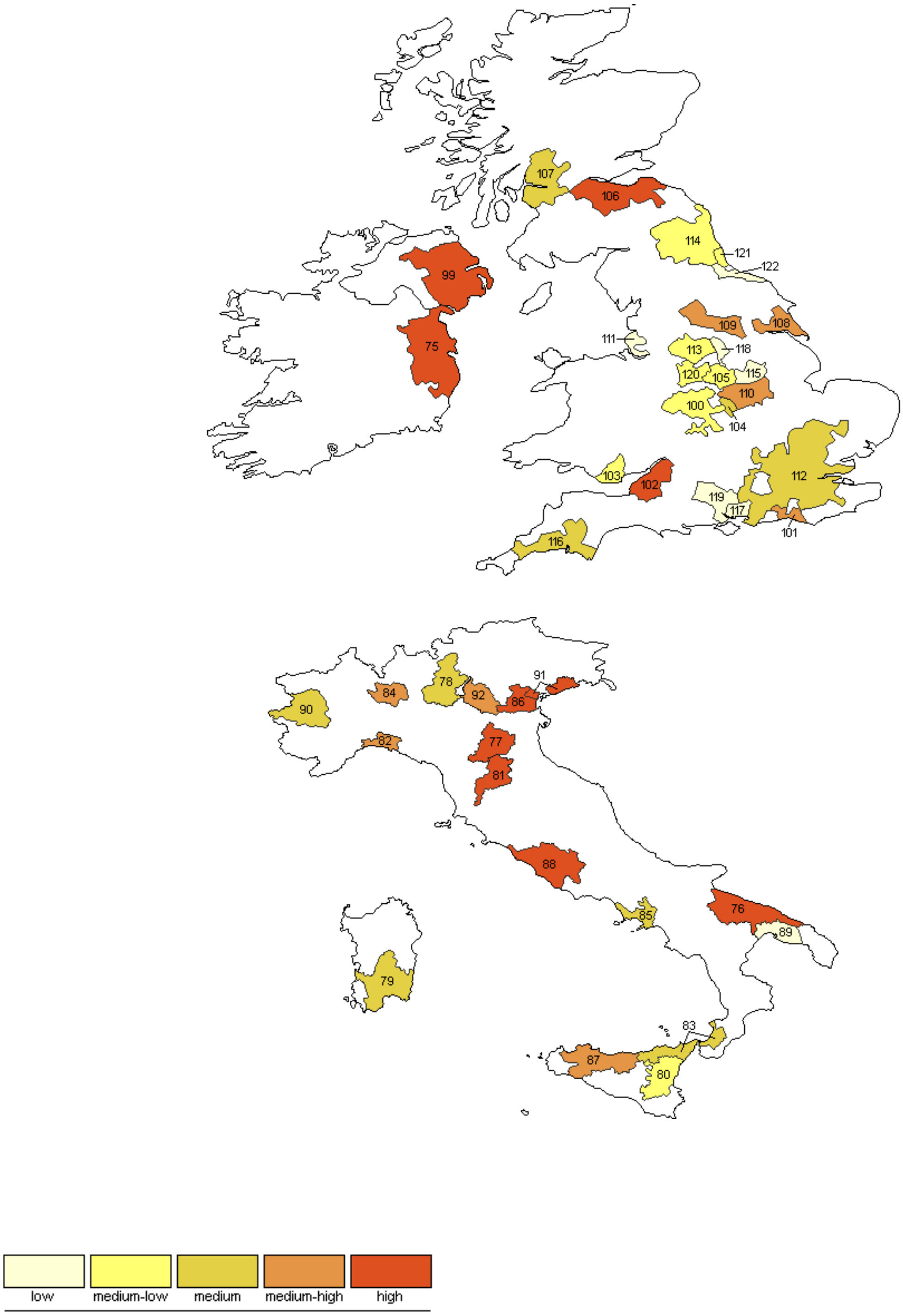

107
106
114
121
122
99
109
108
111
113
118
75
120
105
115
110
100
104
103
102
112
119
117
101
116
91
78
84
92
86
90
77
82
81
88
76
85
89
79
83
87
80
low
medium-low
medium
medium-high
high

Table 1: Per capita GDP in selected NUTS-I and FURs

	1995		1980		% change	
	FUR	NUTS	FUR	NUTS	FUR	NUTS
Bremen	16941	21990	16295	21155	4.0	3.9
Hamburg	21749	27946	19491	25053	11.6	11.5
Île de France/Paris	23675	25901	21701	21889	9.1	18.3
Brussels/Bruxelles	16002	24366	14742	23414	8.5	4.1
Greater London/London	17947	19394	17028	19420	5.4	-0.1
EU12	14603		13472		8.4	

Per capita GDP is measured in purchasing power parities at 1990 prices.

Sources: Eurostat and CRENOS for NUTS; estimates using Eurostat data and Cheshire and Hay (1989) definitions for FURs.

Fig. 1: Possible steady state paths in the neoclassical models

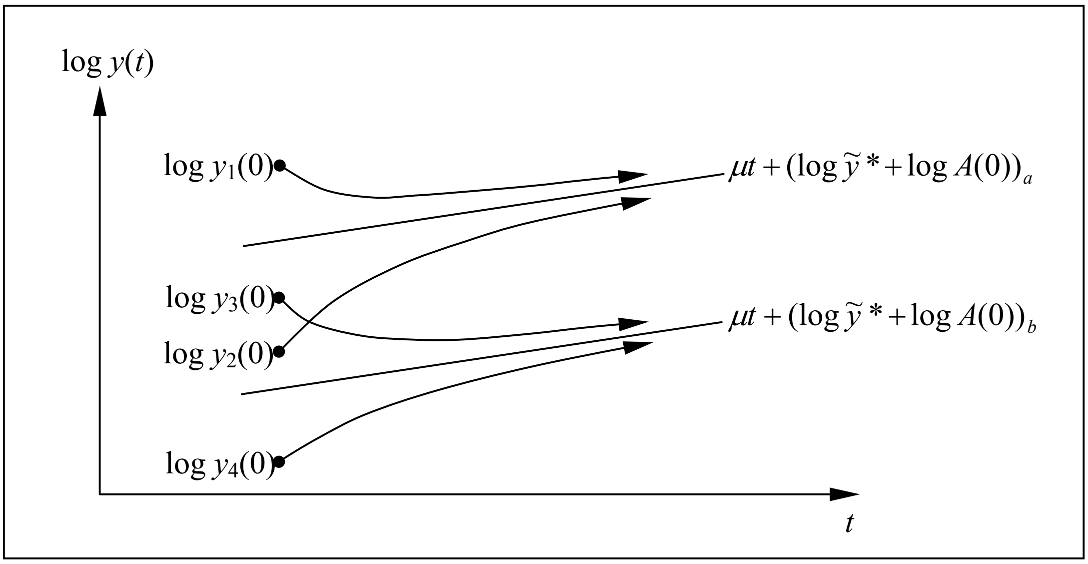
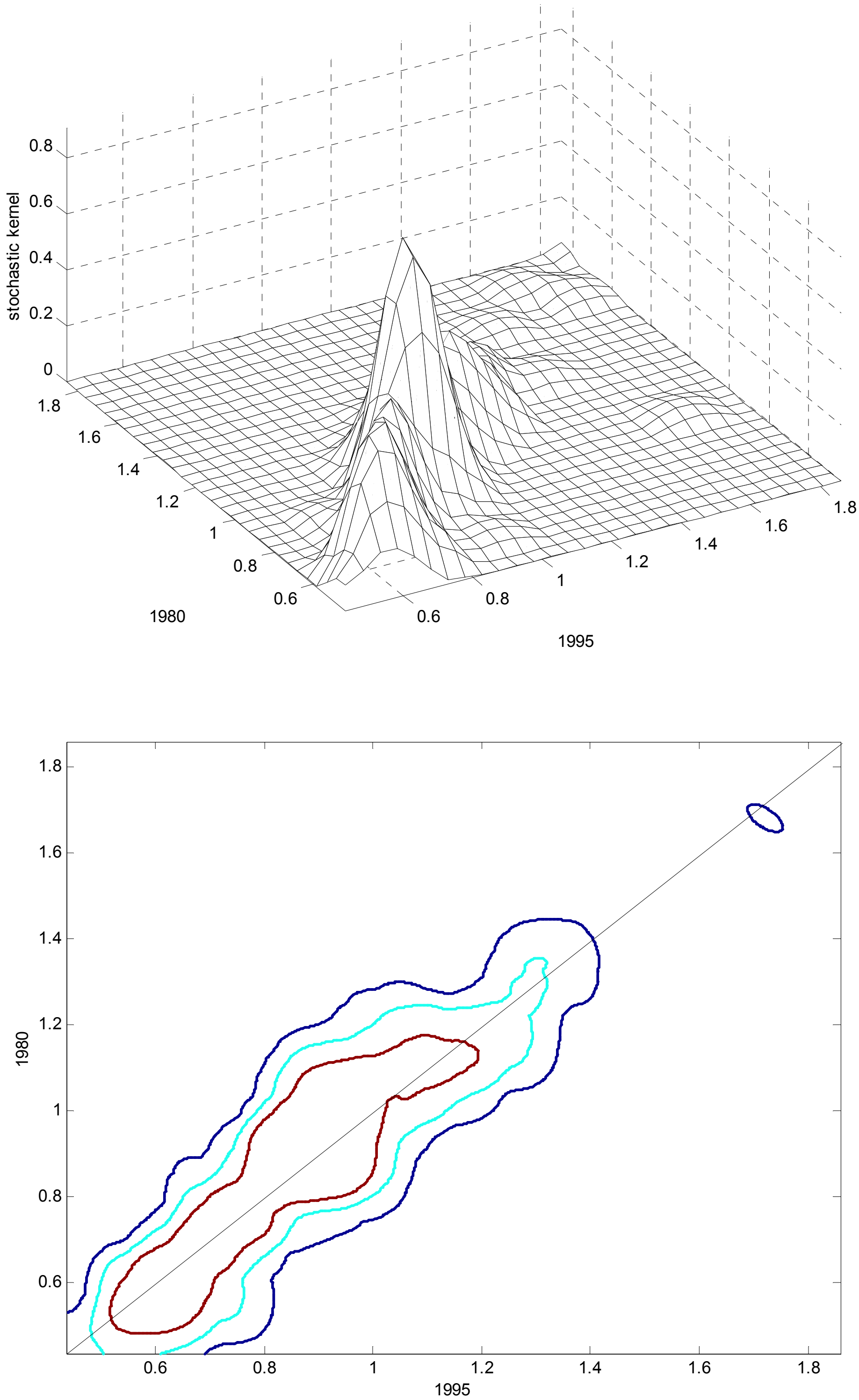
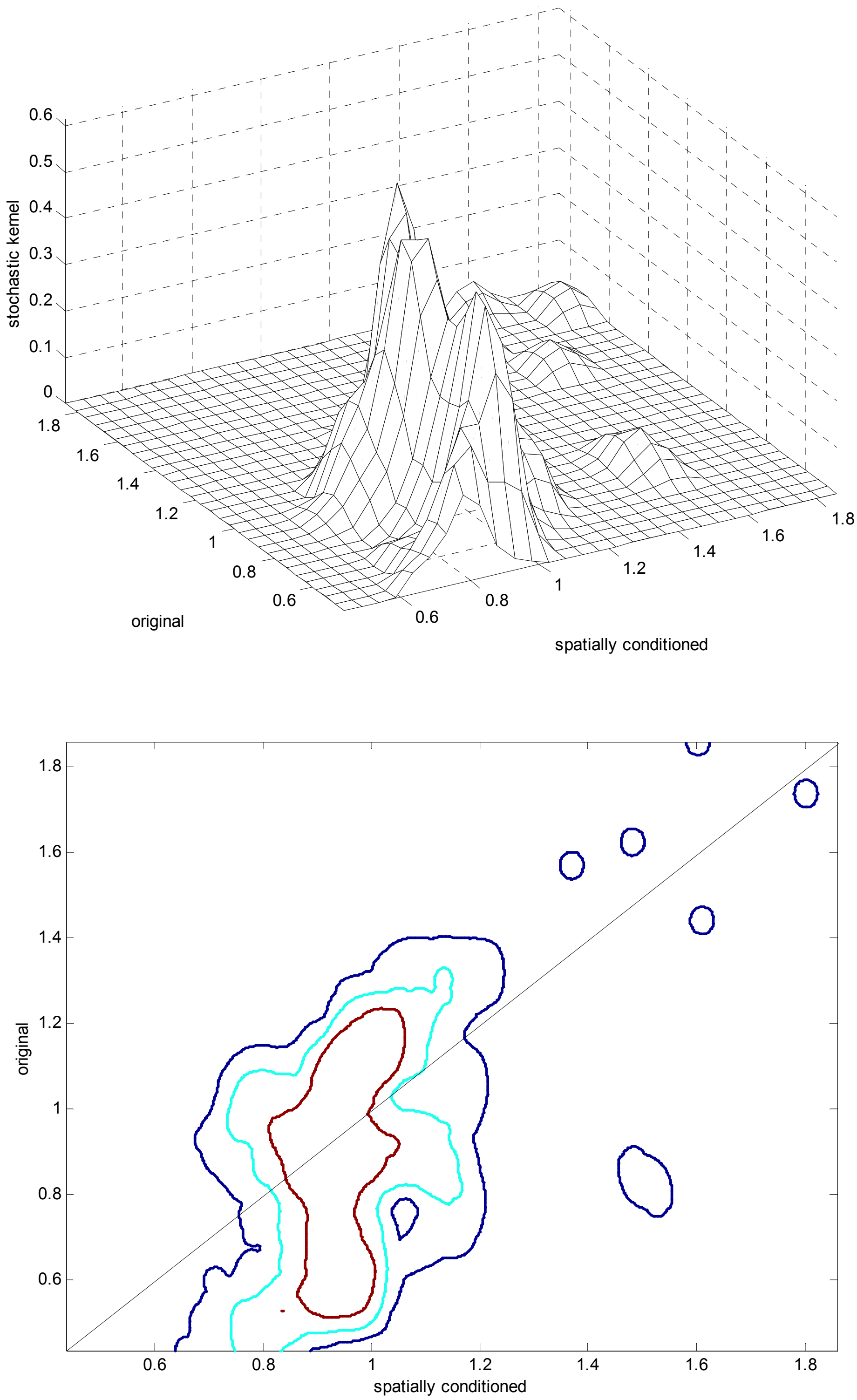


Fig. 2: Relative (per capita) income dynamics across selected NUTS Regions, 1980-1995



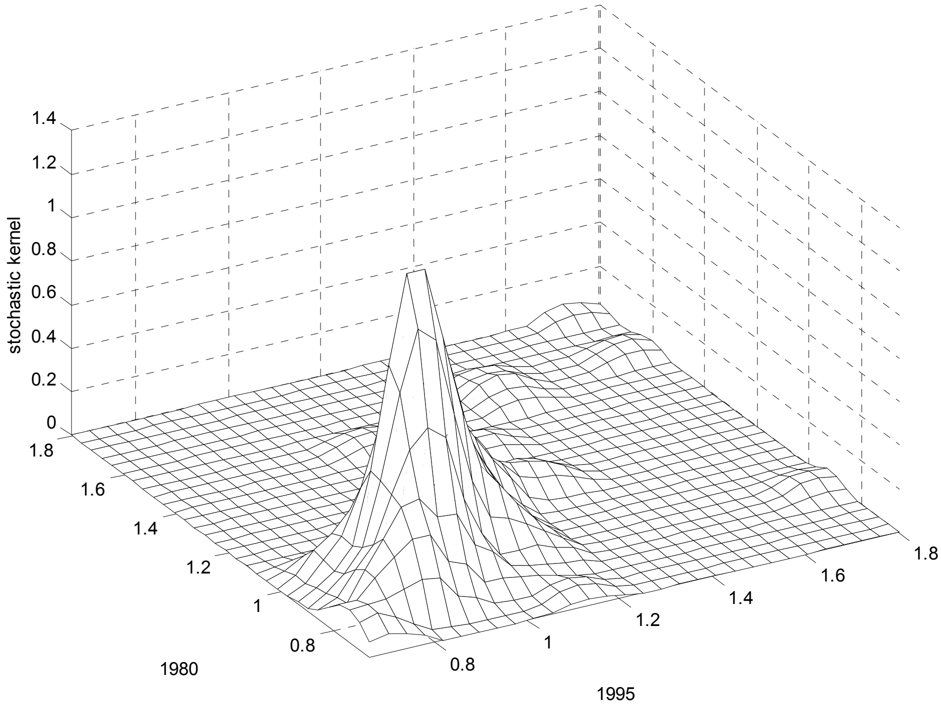
Epanechnikov kernel taken over a fifteen-year transition horizon. Contour plot at levels 0.05, 0.15, 0.3.

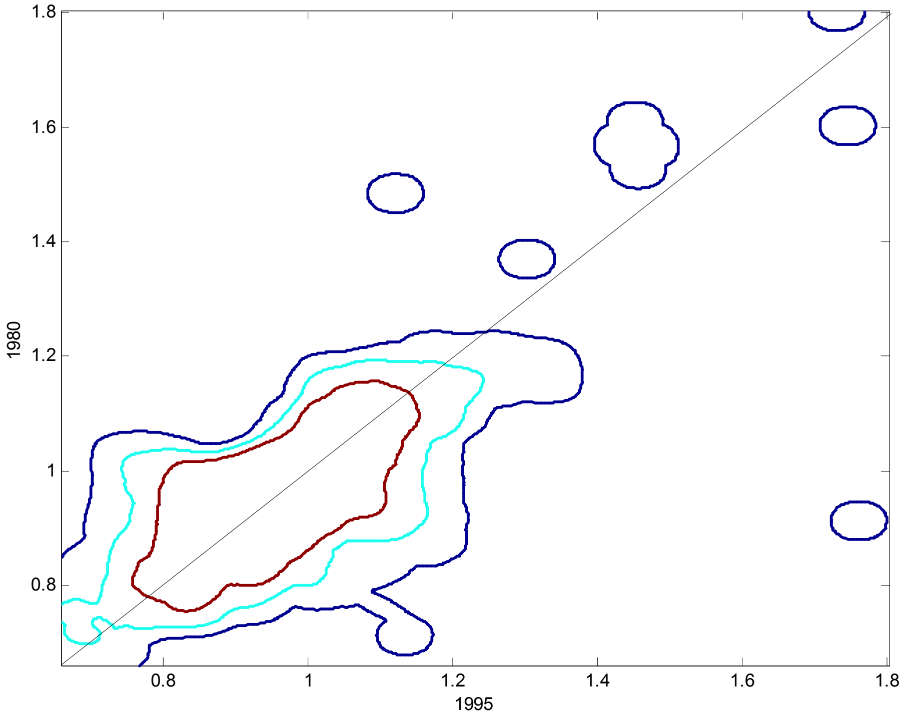
Fig. 3: Relative (per capita) income dynamics across selected NUTS Regions, 1980. Spatially Conditioned.



Epanechnikov kernel taken over original and spatially conditioned relative per capita income. Contour plot at levels 0.05, 0.15, 0.3.

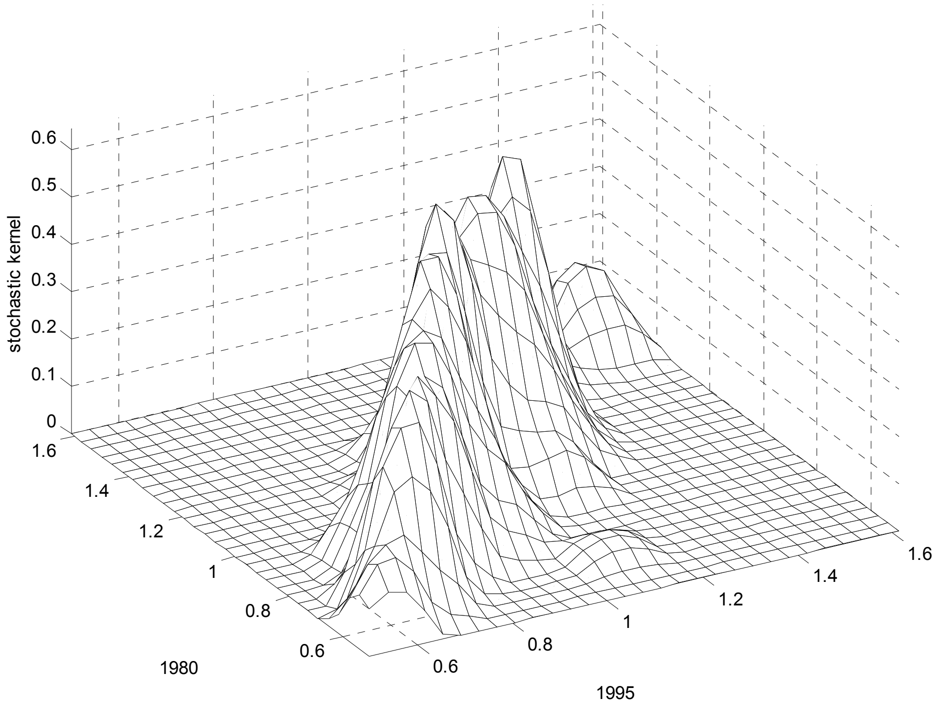
Fig. 4: Relative (per capita) income dynamics across selected NUTS Regions, 1980-1995. Spatially Conditioned.

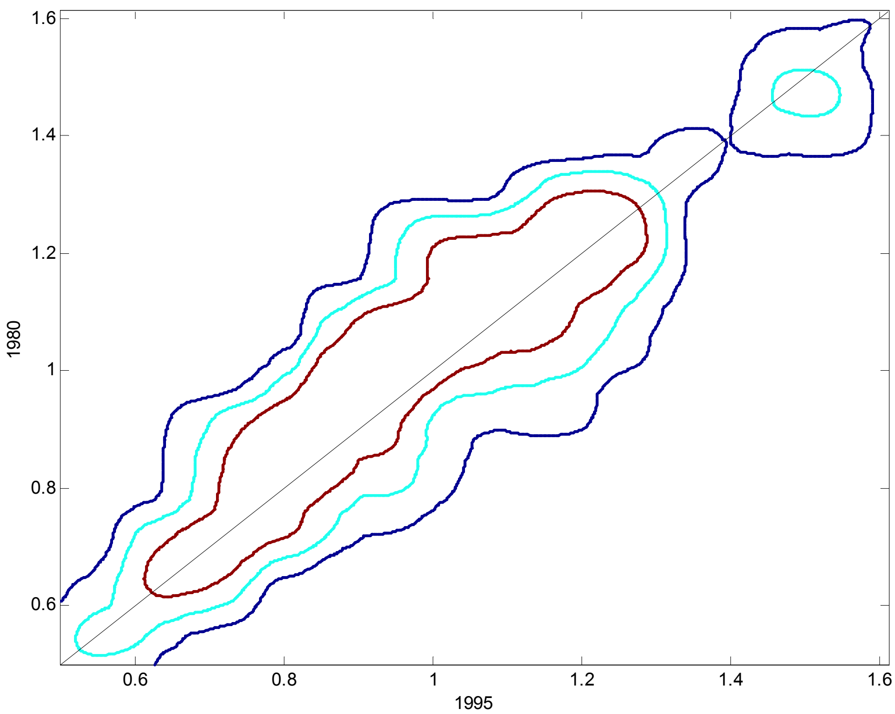




Epanechnikov kernel taken over a fifteen-year transition horizon. Contour plot at levels 0.05, 0.15, 0.3.

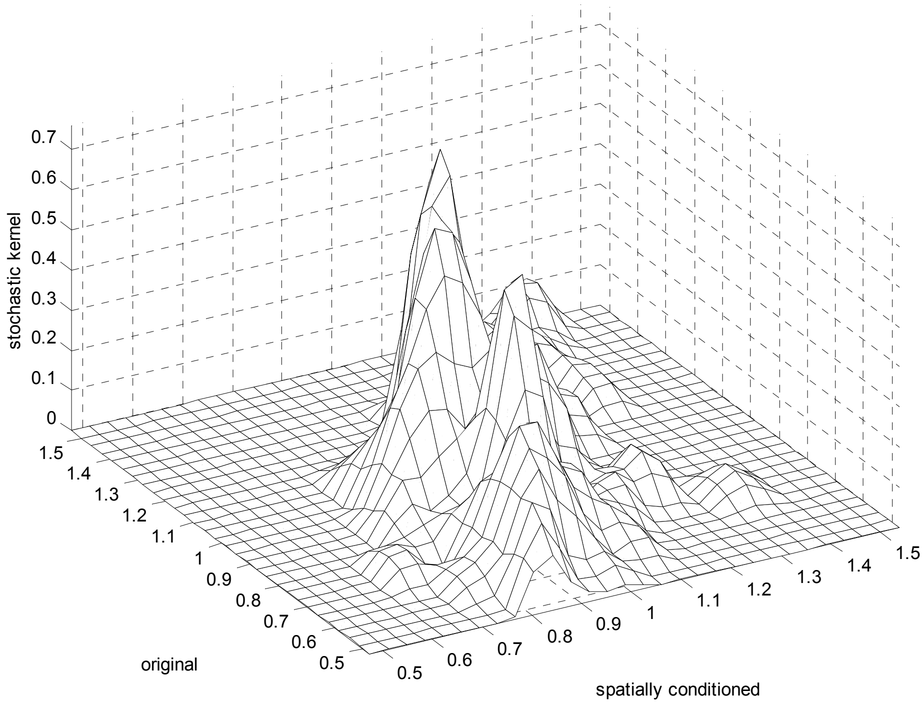
Fig. 5: Relative (per capita) income dynamics across FURs, 1980-1995

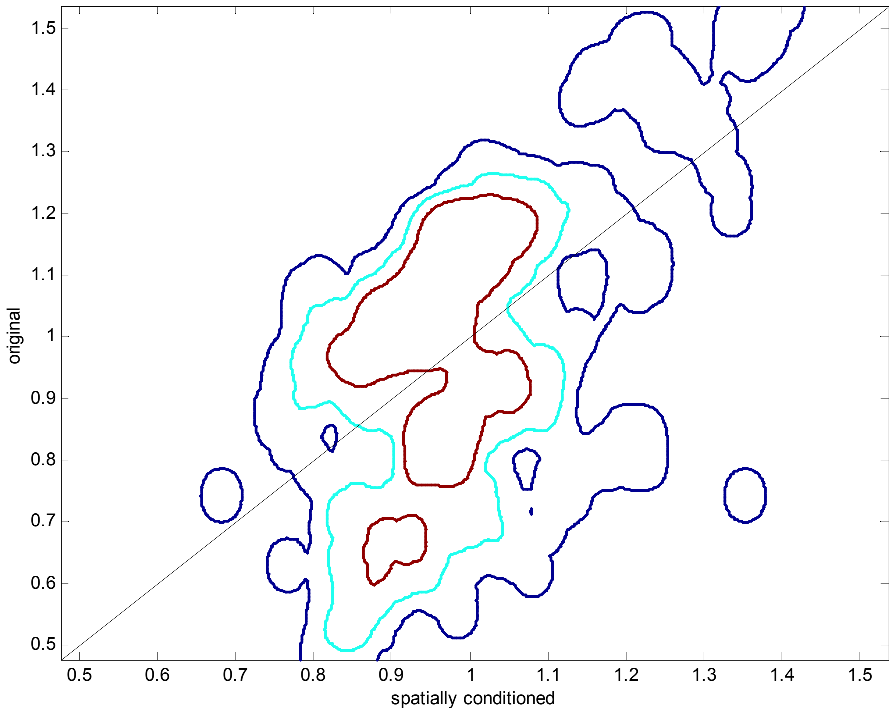




Epanechnikov kernel taken over a fifteen-year transition horizon. Contour plot at levels 0.05, 0.15, 0.3.

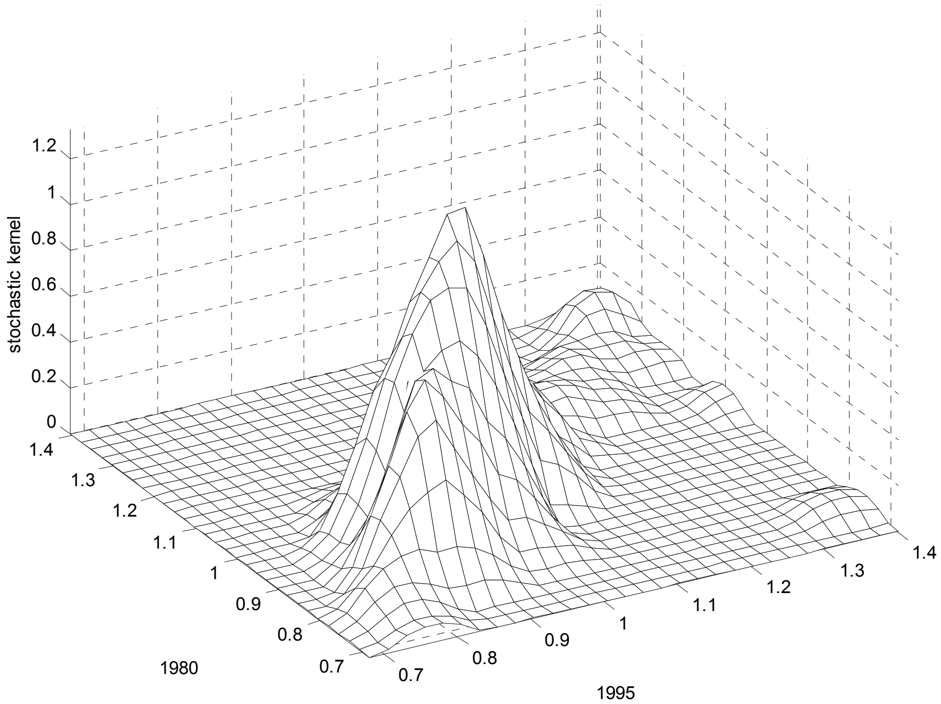
Fig. 6: Relative (per capita) income dynamics across FURs, 1980. Spatially Conditioned.

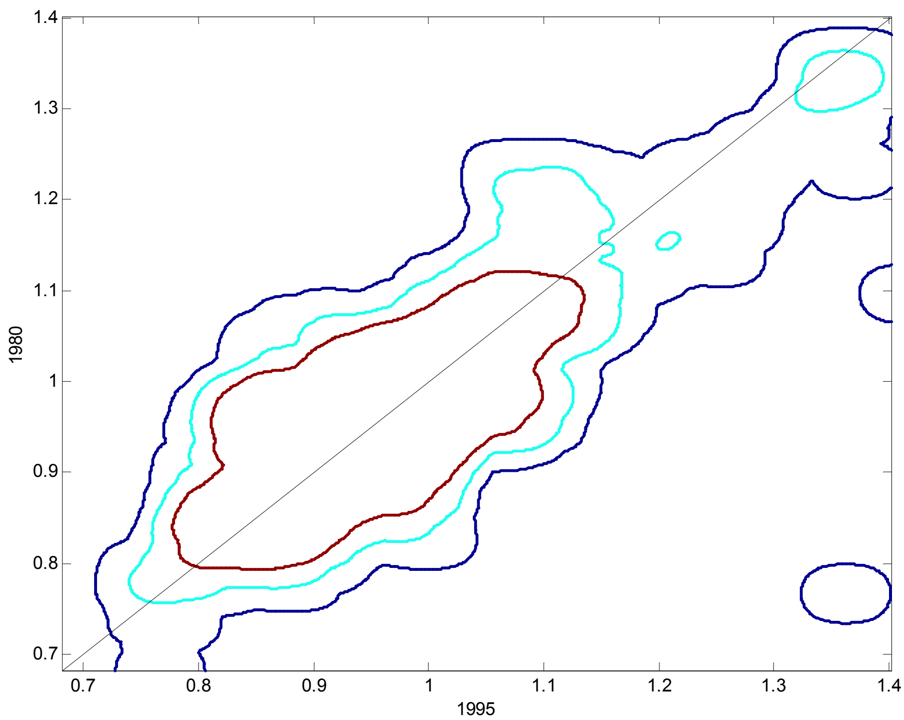




Epanechnikov kernel taken over original and spatially conditioned relative per capita income. Contour plot at levels 0.05, 0.15, 0.3.

Fig. 7: Relative (per capita) income dynamics across FURs, 1980-1995. Spatially Conditioned.





Epanechnikov kernel taken over a fifteen-year transition horizon. contour plot at levels 0.05, 0.15, 0.3.