

Income distribution dynamics and cross-region convergence in Europe

Spatial filtering and novel stochastic kernel representations

Manfred M. Fischer · Peter Stumpner

Received: 3 October 2007 / Accepted: 2 April 2008 / Published online: 6 June 2008
© Springer-Verlag 2008

Abstract This paper presents a continuous version of the model of distribution dynamics to analyse the transition dynamics and implied long-run behaviour of the EU-27 NUTS-2 regions over the period 1995–2003. It departs from previous research in two respects: *first*, by introducing kernel estimation and three-dimensional stacked conditional density plots as well as highest density regions plots for the visualisation of the transition function, based on Hyndman et al. (J Comput Graph Stat 5(4):315–336, 1996), and *second*, by combining Getis’ spatial filtering view with kernel estimation to explicitly account for the spatial dimension of the growth process. The results of the analysis indicate a very slow catching-up of the poorest regions with the richer ones, a process of shifting away of a small group of very rich regions, and highlight the importance of geography in understanding regional income distribution dynamics.

Keywords Regional income · Distribution dynamics · Stochastic kernel estimation · Spatial filtering · EU-27

JEL Classification C14 · D30 · O18 · O47 · R11

1 Introduction

Whether income levels of poorer regions are converging to those of richer is a question of paramount importance for human welfare (Islam 2003). In Europe interest in this question has been enhanced in recent years, with the entry of new countries to the European Union. This paper looks at evidence for regional income

M. M. Fischer (✉) · P. Stumpner
Institute for Economic Geography and GIScience, Vienna University of Economics and Business
Administration, Nordbergstr. 15/4/A, 1090 Vienna, Austria
e-mail: manfred.fischer@wu-wien.ac.at

convergence in Europe. By Europe we mean the European Union of 27 member states. The notion of convergence is a fuzzy term that can mean different things (see Quah 1999). In this paper we understand this notion in the sense of poorer regions catching-up with the richer. The observation units are NUTS-2 regions which the European Commission has chosen as targets for the convergence process and defined as the geographical level at which the persistence or disappearance of inequalities should be measured.

Measuring regional income and the extent to which convergence across regions—or what the European Commission calls regional cohesion—exists is a difficult issue. But per capita gross regional product (GRP) measured in purchasing power units seems like a natural definition if one is interested in an important determinant of average welfare. By focusing upon per capita GRP we are interested in the economic performance of regions and the claims that people living in those regions have over that wealth. Cohesion depends on the degree of equality in the distribution of per capita income and the extent to which there are processes of catch-up, in which less wealthy regions enjoy faster rates of income growth than more developed ones. The data were calculated on the basis of the 1995 European System of Accounts (ESA 95) and refer to the time period from 1995 to 2003, the latest year for which income data are available. This shorter time span makes apparent the need for a model, before we can speak of the underlying dynamic regularities in these data.

Empirical research on regional income convergence has proceeded in many directions, using different definitions and methodologies.¹ Most research has, however, concentrated on the cross-section regression approach to investigate β -convergence where β is the generic notion for the coefficient on the initial income variable in the growth-initial level regressions. A negative β is interpreted as evidence of convergence in terms of both income level and growth rate. But Quah (1993b), Friedman (1992) and others have emphasised that a negative β can just be an example of the more general phenomenon of reversion to the mean, and, by interpreting it as convergence, growth analysts falling into Galton's fallacy.

This study follows the tradition of the non-parametric approach that views the catching-up question as a question about the evolution of the cross-section distribution of income, and diverts attention from the individual or representative region to the entire distribution as object of interest (see, in particular, Quah 1993a, 1996a, b, 1997a, b, c). The distribution that is relevant here is the distribution of income *across* regions, not that within a given region. Purpose of the analysis is to find the law of motion that describes transition dynamics and implied long-run behaviour of regional income. In the spirit of Quah (1996a, b) we assume that each region's income follows a first-order Markov process with time-invariant transition probabilities. That is, a region's (uncertain) income tomorrow depends only on its income today.

¹ Recent surveys of the new growth literature in general and the convergence literature in particular can be found in Durlauf and Quah (1999), Temple (1999) and Islam (2003), while Fingleton (2003), Abreu et al. (2004), and Magrini (2004) survey the regional convergence literature, with region denoting a subnational unit.

Most of the applications of this approach have worked in a discrete state space set up (see Quah 1996a, b; Fingleton 1997, 1999; Paap and van Dijk 1998; López-Bazo et al. 1999; Magrini 1999; Rey 2001; LeGallo 2004 to mention some). This set up has several advantages, but the process of discretising the state space of a continuous variable is necessarily arbitrary. Experience from the study of income distributions shows that this arbitrariness can matter in the sense that statements on inferred dynamic behaviour of the distribution in question and the apparent long-run implications of that behaviour are sensitive to the choice of the discretisation (Jones 1997; Reichlin 1999). Indeed, it is well known that the Markov property itself can be distorted from inappropriate discretisation (Bulli 2001).

This paper avoids arbitrary discretisation of the income space and its possible effects on the results by using the stochastic kernel, the continuous equivalent of the transition probability matrix, as a suitable tool to overcome the problem. The remainder of the paper is divided into two parts. The first, Sect. 2, provides an empirical framework that extends current research by incorporating two novel techniques into the existing research: kernel estimation and graphical devices for the representation of the stochastic kernel (see Hyndman et al. 1996), and Getis' spatial filtering technique that enables to account for the effects of spatial autocorrelation. The second part of the paper, Sect. 3, applies this framework to analyse income distribution dynamics and cross-region convergence in Europe, looking at evolving distributions of purchasing power standardised per capita (relative) gross regional product across 257 NUTS-2 regions in 27 EU-countries from 1995 to 2003. Some concluding remarks are given in the final section.

2 The empirical framework

A distribution perspective to the study of income dynamics and cross-region convergence directs attention to the evolution of the entire cross-region income distribution, emphasising shape and intra-distribution dynamics, and long-run (ergodic) behaviour. Section 2.1 introduces a continuous version of the standard model of explicit distribution dynamics, pioneered by Quah (1993a), and argues that the stochastic kernel can be described as a conditional density function. In Sect. 2.2 we present a product kernel estimator for estimating this transition function, and briefly describe a three-step-strategy for solving the bandwidth selection problem, that appears to be crucial for estimation. Section 2.3 combines Getis' spatial filtering view with stochastic kernel estimation to account for the issue of spatial autocorrelation that may misguide inferences and interpretations if not properly handled.

2.1 A continuous version of the model of distribution dynamics

Let F_t denote the cross-section distribution of regional incomes at time t , then the simplest scheme for modelling the intra-distribution dynamics of $\{F_t | t \text{ integer}\}$ is a first-order Markov process with time-invariant transition probabilities. The distribution evolves according to

$$F_{t+1} = M F_t \quad (1)$$

where M maps the distribution from time t to time $t + 1$, and tracks where points in F_t end up in F_{t+1} . Iteration of Eq. (1) gives a prediction for future distributions of the ex-post probabilities

$$F_{t+\tau} = M^\tau F_t \quad \text{for } \tau > 0 \ (\tau = 1, 2, \dots). \quad (2)$$

In this framework, there are two goals, the estimation of M will give us information on persistence of regional income inequalities and the computation of the ergodic (steady-state) distribution. The latter provides information on the limiting behaviour of the regional income distribution. Convergence then might manifest in $\{F_{t+\tau}\}$ tending towards a point mass. A bimodal limit distribution can be interpreted as a tendency towards stratification into two different “convergence clubs”.

In the discrete version of the model, the operator M can be interpreted as the transition probability matrix of the Markov process. The operator is approximated by partitioning the set of possible income values into a finite number of intervals. These intervals then constitute the states of a (time-homogeneous) finite Markov process, and all the relevant properties of M are described by a Markov chain transition matrix whose (i, j) entry is the probability that a region in state i transits to state j in income space, in one time step. The inferred dynamic behaviour and the long-run implications of that behaviour are conditional on the discretisation chosen.

Regional income, however, is by nature a continuous variable, and hence discretisation may induce a possible bias. Instead of a state being a fixed interval we let the state be all possible interval, including the infinitesimal small ones. In this case one may think of the number of distinct cells to tend to infinity and then to continuum. The corresponding transition probability matrix then tends to a matrix with a continuum of rows and columns. In this case, the operator M in Eq. (1) may be viewed as a stochastic kernel or transition function that describes the (time-invariant) evolution of the cross-section distribution in time. Convergence can then be studied by visualising and interpreting the shape of the income distribution at time $t + \tau$ over the range of incomes observed at time t .

For notational convenience let Y and Z denote the variable (per capita) regional income at times t and $t + \tau$ ($\tau > 0$), respectively. The sample may be denoted then by $\{(Y_1, Z_1), \dots, (Y_n, Z_n)\}$, and the observations by $\{(y_1, z_1), \dots, (y_n, z_n)\}$ where n indicates the number of regions. We assume that the cross-region distribution of Y can be described by the density function $f_t(y)$. This distribution will evolve over time so that the density prevailing at $t + \tau$ is $f_{t+\tau}(z)$. If we continue to maintain the assumptions of time-invariance and first-order of the transition process, the relationship between the cross-region income distributions, at time t and τ -periods later, can be written as

$$f_{t+\tau}(z) = \int_0^{\infty} g_\tau(z|y) f_t(y) dy \quad (3)$$

where $g_\tau(z|y)$ is the conditional density function giving the τ -period ahead density of income z , conditional on income y at time t . Evidently, the (first-order) stochastic

kernel can be described by a conditional density function assuming that the marginal and conditional income distributions have density functions.

So long as $g_\tau(z|y)$ exists, the long-run (ergodic) density, $f_\infty(z)$, implied by the estimated $g_\tau(z|y)$ function can then be found as solution to

$$f_\infty(z) = \int_0^\infty g_\tau(z|y) f_\infty(y) dy. \quad (4)$$

In this paper we will use the solution procedure outlined in Johnson (2004) to estimate this long-run distribution of regional income per capita.

2.2 Kernel estimation of the conditional density function

If $f_{t,t+\tau}(y, z)$ denotes the joint density of (Y, Z) and $f_t(y)$ the marginal density of Y , then the conditional density of $Z|(Y = y)$ is given by

$$g_\tau(z|y) = \frac{f_{t,t+\tau}(y, z)}{f_t(y)}. \quad (5)$$

Probably, the most obvious estimator of this conditional density function² (see Hyndman et al. 1996) is

$$\hat{g}_\tau(z|y) = \frac{\hat{f}_{t,t+\tau}(y, z)}{\hat{f}_t(y)} \quad (6)$$

where

$$\hat{f}_{t,t+\tau}(y, z) = \frac{1}{n h_y h_z} \sum_{i=1}^n K\left(\frac{1}{h_y} \|y - Y_i\|_y\right) K\left(\frac{1}{h_z} \|z - Z_i\|_z\right) \quad (7)$$

is the kernel estimator of $f_{t,t+\tau}(y, z)$, and

$$\hat{f}_t(y) = \frac{1}{n h_y} \sum_{i=1}^n K\left(\frac{1}{h_y} \|y - Y_i\|_y\right) \quad (8)$$

the kernel estimator of $f_t(y)$ (see Hyndman et al. 1996). h_y and h_z are bandwidth parameters that control the degree of smoothing applied to the density estimate. h_y controls the smoothness between conditional densities in the y -direction, and h_z the smoothness of each conditional density in the z -direction. $\|\cdot\|_y$ and $\|\cdot\|_z$ are distance metrics on the spaces Y and Z , respectively. In this paper we use the standard euclidean distances, $\|\cdot\|_y = |\cdot|_y$ and $\|\cdot\|_z = |\cdot|_z$.

A multivariate kernel other than the product kernel might be used to define $\hat{g}_\tau(z|y)$. But the product kernel is simpler to work with, leads to conditional density estimators with several nice properties and is only slightly less efficient than other multivariate kernels (Wand and Jones 1995). The kernel $K(x)$, where x is variously y or z , is a real, integrable, non-negative, even function on \mathbb{R} concentrated at the origin so that (Silverman 1986)

² For alternative estimators see Hyndman and Yao (2002), and Basile (2006).

$$\int_{\mathbb{R}} K(x) dx = 1, \quad \int_{\mathbb{R}} x K(x) dx = 0 \quad \text{and} \quad \sigma_K^2 = \int_{\mathbb{R}} x^2 K(x) dx < \infty. \quad (9)$$

Popular choices for $K(x)$ are defined in terms of univariate and unimodal probability density functions. In this paper we use the Gaussian kernel³ given by

$$K(x) = \left(\sqrt{2\pi}\right)^{-1} \exp\left(-\frac{1}{2}x^2\right). \quad (10)$$

Whatever kernel is being used, bandwidth parameters chosen to minimise the asymptotic mean square error give a trade-off between bias and variance. Small bandwidths yield small bias but large variance, while large bandwidths lead to large bias and small variance. The problem of choosing, how much to smooth, is of crucial importance in conditional density estimation, and the results of the continuous state space approach to distribution dynamics strongly depend on the bandwidth parameters chosen.

In this study we follow Bashtannyk and Hyndman (2001) to solve this bandwidth selection problem⁴ by a three-step-strategy that combines three different procedures: a Silverman (1986) inspired normal reference rule that has proven useful in univariate kernel density estimation,⁵ a bootstrap bandwidth selection approach following the approach of Hall et al. (1999) for estimating conditional distribution functions, and a regression-based bandwidth selector⁶ (see Fan et al. 1996). *Step 1* involves finding an initial value for the smoothing parameter h_z using the rule with normal marginal density. Given this value of h_z , *Step 2* makes use of the regression-based bandwidth selector to find a value for h_y . In *Step 3* the bootstrap method is used to revise the estimate of h_z by minimising the bootstrap estimator of a weighted mean square error function. *Step 2* and *Step 3* may be repeated one or more times.

2.3 Spatial autocorrelation and stochastic kernel estimation

Stochastic kernel estimation rests on the implicit assumption that each region represents an independent observation providing unique information that can be used to estimate the transition dynamics of income. In essence, the cross-section observations at one point in time are viewed as a random sample from a univariate distribution, or in other words, X (where X stands variously for Y and Z) is assumed to be univariate and random. If the X_i ($i = 1, \dots, n$) are independent, we say that

³ On the basis of the mean integrated square error criterion, Silverman (1986) has shown that there is very little to choose between alternatives. In contrast, the choice of the bandwidths plays a crucial role.

⁴ It is well known that the selection of the bandwidth parameters rather than the choice between various kernels is of crucial importance in density estimation.

⁵ The rule is to assume that the underlying density is normal and to find the bandwidth which could minimise the integrated mean square error function.

⁶ For a given h_z and a given value z , finding $\hat{g}(z|y)$ is viewed here as a standard non-parametric problem of regressing $h_z^{-1} K(h_z^{-1}|z - Z_i|)$ on Y_i .

there is no spatial structure. Independence implies the absence of spatial autocorrelation.⁷ Spatial autocorrelation reflects a lack of independence between regions. This independence may arise from a variety of measurement problems, such as boundary mismatches between the NUTS-2 regions. But also interactions or externalities across regions such as, for example, knowledge spillovers, trade as well as commuting and migration flows are likely to be a major source of the violation of the assumption (see Abreu et al. 2004 for a survey of the existing evidence).

A violation of the independence assumption may result in misguided inferences and interpretations (Rey and Janikas 2005). This problem has been largely neglected in distribution analysis so far. One way⁸ to dealing with the problem involves the filtering of the variable X in order to separate spatial effects from the variable's total effects. While insuring spatial independence, this allows us to use the stochastic kernel to properly estimate the underlying regional income distribution and to analyse its evolution over time. The motivation for a spatial filter is simply that a spatially autocorrelated variable can be transformed into an independent variable by removing the spatial dependence embedded in it. The original variable, X , is hence partitioned into two parts, a filtered non-spatial variable, say \tilde{X} , and a residual spatial variable L_X . The transformation procedure depends on identifying an appropriate distance δ within which nearby regions are spatially dependent, and examining each individual observation for its contribution to the spatial dependence embedded in the original variable (Getis and Griffith 2002).

There have been several suggestions for identifying δ , but in this paper we adopt the Getis filtering approach (see Getis 1990, 1995) which is based on the local spatial autocorrelation statistic G_i (Getis and Ord 1992) to be evaluated at a series of increasing distances until no further spatial autocorrelation is evident. As distance increases from an observation (region i), the G_i -value also increases if spatial autocorrelation is present. Once the G_i -value begins to decrease, the limit on spatial autocorrelation is assumed to have been reached, and the associated critical δ identified. The filtered observation \tilde{x}_i is given as

$$\tilde{x}_i = \frac{x_i \left[\frac{1}{n-1} W_i \right]}{G_i(\delta)} \quad (11)$$

where x_i is the original income observation for region i , n is the number of observations and

$$W_i = \sum_{j=1}^n w_{ij}(\delta) \quad \text{for } j \neq i. \quad (12)$$

⁷ The controversy is not necessarily true (Ord and Getis 1995). Nevertheless, tests for spatial autocorrelation are typically viewed as appropriate assessments of spatial dependence. Moran's I and Geary's c statistics are typical testing tools.

⁸ Griffith's eigenfunction decomposition approach that uses an eigenfunction decomposition based on the geographic connectivity matrix used to compute a Moran's I statistic provides an alternative way (Griffith 2006).

with $w_{ij}(\delta) = 1$ if the distance⁹ from region i to region j ($i \neq j$), say d_{ij} , is smaller than the critical distance band δ , and $w_{ij}(\delta) = 0$ otherwise. $G_i(\delta)$ is the spatial autocorrelation statistic¹⁰ of Getis and Ord (1992) defined as

$$G_i(\delta) = \frac{\sum_{j=1}^n w_{ij}(\delta) x_j}{\sum_{j=1}^n x_j} \quad \text{for } i \neq j. \tag{13}$$

The numerator of (13) is the sum of all x_j within δ of i but not including x_i . The denominator is the sum of all x_j not including x_i .

Equation (11) compares the observed value of $G_i(\delta)$ with its expected value, $(n-1)^{-1}W_i$. $E[G_i(\delta)]$ represents the realisation, \tilde{X} , of the variable X at region i when no autocorrelation occurs. If there is no autocorrelation at i to distance δ , then the observed and expected values, x_i and \tilde{x}_i , will be the same. When $G_i(\delta)$ is high relative to its expectation, the difference $x_i - \tilde{x}_i$ will be positive, indicating spatial autocorrelation among high observations of X . When $G_i(\delta)$ is low relative to its expectation, the difference will be negative, indicating spatial autocorrelation among low observations of X . Thus, the difference between x_i and \tilde{x}_i represents the spatial component of the variable X at i . Taken together for all i , L_X represents a spatial variable associated, but not correlated, with the variable X . Thus, $L_X + \tilde{X} = X$ (Getis and Griffith 2002).

Combining this spatial filtering approach with stochastic kernel estimation as described in the previous section yields the long-run (ergodic) density, $f_\infty(\tilde{z})$, implied by the estimated $g_\tau(\tilde{z}|\tilde{y})$ function:

$$f_\infty(\tilde{z}) = \int_0^\infty g_\tau(\tilde{z}|\tilde{y}) f_\infty(\tilde{y}) \, d\tilde{y}, \tag{14}$$

where \tilde{y} and \tilde{z} denote the spatially filtered observations of Y and Z , respectively. To assess the role played by space on income growth and convergence dynamics across the regions, we consider a specific stochastic kernel¹¹ that maps the distribution Y to the spatially filtered distribution $\tilde{Y}|Y$ so that

$$g(\tilde{y}|y) = \frac{f(y, \tilde{y})}{f(y)} \tag{15}$$

where the stochastic kernel does not describe transitions over time, but transitions from unfiltered to spatially filtered regional income distributions, and, thus, quantifies the effects of spatial dependence. If spatial effects caused by spatial

⁹ In this study distances are measured in terms of geodesic distances between regional centres.

¹⁰ Getis and Ord (1992) and Ord and Getis (1995) show that the statistic $G_i(\delta)$ is asymptotically normally distributed as δ increases. When the underlying distribution of the variable in question is skewed, appropriate normality of the statistic can be guaranteed when the number of j neighbours is large.

¹¹ Combining stochastic kernel estimation with the conditioning scheme suggested by Quah (1996b, 1997a) is an alternative way to evaluate the role of spatial interactions among neighbouring regions. Conditioning means here normalising each region's observations by the (population weighted) average income of its neighbours. This approach removes substantive, but not nuisance spatial dependence effects.

interaction among regions and measurement problems would not matter, then the stochastic kernel would be the identity map.

3 Revealing empirics

This section applies the above framework to study regional income dynamics and convergence in Europe. In Sect. 3.1 we describe the data and the observation units. Kernel smoothed densities and Tukey boxplots are used in Sect. 3.2 to study the shape dynamics of the distribution. Cross-profile plots, continuous stochastic kernels and implied ergodic distributions are taken in Sect. 3.3 to investigate intra-distribution dynamics and long-run tendencies in the data. Section 3.4 proceeds to the spatial filtering view of the data to gain insights not affected by the spatial autocorrelation problem.

3.1 Data and observation units

We use per capita GRP over the period 1995–2003 expressed in ECUs, the former European currency unit, replaced by the Euro in 1999. The GRP figures were calculated on the basis of the 1995 European System of Integrated Economic Accounts (ESA 95)¹² and extracted from the Eurostat Regio database. We use GRP per capita in national PPS (purchasing power standards) as defined by Eurostat. These units are comparable to ECUs/Euros, with a slight correction.¹³

The time period is relatively short due to a lack of reliable figures for the regions in the new member states of the EU. This comes partly from the substantial change in measurement methods of national accounts in Central and East Europe (CEE) between 1991 and 1995. But more important, even if estimates of the change in the volume of output did exist, these would be impossible to interpret meaningfully because of the fundamental change of production from a centrally planned to a market system. As a consequence, figures for GRP are difficult to compare until the mid-1990s (Fischer and Stirböck 2006).

The observation units of the analysis are NUTS-2 regions.¹⁴ Although varying considerably in size, NUTS-2 regions are those regions that are adopted by the

¹² In order to deal with the widely known problem measuring Groningen's GRP figure we replaced its energy specific gross value added component by the average of the neighbouring regions (Drenthe and Friesland).

¹³ Figures given in PPPs are derived from figures expressed in national currency by using PPPs as conversion factors. These parities are obtained as a weighted average of relative price ratios in respect to a homogeneous basket of goods and services, both comparable and representative for each individual country. The use of national purchasing power parities is based on the assumption that there are no—or negligible—purchasing power disparities between the regions within individual countries. This assumption may not appear to be entirely realistic, but it is inevitable in view of the data available.

¹⁴ Note that the use of administratively defined regions, such as NUTS-2 regions, can lead to misleading inferences due to the presence of significant nuisance spatial dependence. In the case of Hamburg, for example, the NUTS-2 boundary is very narrowly drawn with respect to the corresponding functional region so that residential areas extend well beyond the boundary and substantial in-commuting takes place. This implies that per capita GRP is overestimated, while in the surrounding NUTS-2 regions underestimated.

European Commission for the evaluation of regional growth and convergence processes. NUTS is an acronym of the French for “the nomenclature of territorial units for statistic”, which is a hierarchical system of regions used by the statistical office of the European Community for the production of regional statistics. Our sample includes 257 NUTS-2 regions¹⁵ covering the 27 member states of the EU (see the Appendix for a description of the regions):

- *the EU-15 member states:* Austria (nine regions), Belgium (eleven regions), Denmark (one region), Finland (five regions), France (22 regions), Germany (40 regions), Greece (thirteen regions), Ireland (two regions), Italy (20 regions), Luxembourg (one region), Netherlands (twelve regions), Portugal (five regions), Spain (16 regions), Sweden (eight regions), UK (37 regions);
- *the 12 new member states:* Bulgaria (six regions), Cyprus (one region), Czech Republic (eight regions), Estonia (one region), Hungary (seven regions), Latvia (one region), Lithuania (one region), Malta (one region), Poland (16 regions), Romania (eight regions), Slovakia (four regions), Slovenia (one region).

3.2 Shape dynamics of the distribution

When studying income distribution dynamics across regions in Europe, one can consider incomes per region in absolute terms. Alternatively, one can study regional incomes normalised by the European average. Although there are merits to using the absolute income distribution, it is more natural to take relative incomes when considering changes in income distributions over time. Relative incomes allow us to abstract from overall changes in income levels.¹⁶ A natural approach to assess the shape dynamics of the distribution change over the observation period 1995–2003 is to estimate the cross-sectional distributions by using non-parametric kernel smoothing procedures, which avoid the strong restrictions imposed by parametric estimation. In this framework, if there is a bimodal density at a given point in time, indicating the presence of two groups in the population of regions, convergence implies a tendency of the distribution to move progressively towards unimodality.

Figure 1 plots the distribution of (per capita) GRP relative to the average of all 257 regions—what we call the Europe relative (per capita) income or simply the relative income. The plots are densities and can be interpreted as the continuous equivalent of a histogram, where the number of intervals has been let tend to infinity and then to the continuum. All densities were calculated non-parametrically using a Gaussian kernel with bandwidths chosen as suggested in Silverman (1986), restricting the range to the positive interval. The solid line shows the distribution in 2003, and the dashed line that in 1995. To read this type of figure, note that 1.0 on the horizontal axis indicates the European average of regional income, 2.0 indicates

¹⁵ We exclude the Spanish North African territories of Ceuta y Melilla, the Portuguese non-continental territories Azores and Madeira, and the French Départements d’Outre-Mer Guadeloupe, Martinique, French Guayana and Réunion.

¹⁶ This normalisation makes it possible to separate the global (European) effects on the cross-section distribution of European forces from the effects from regional-specific effects.

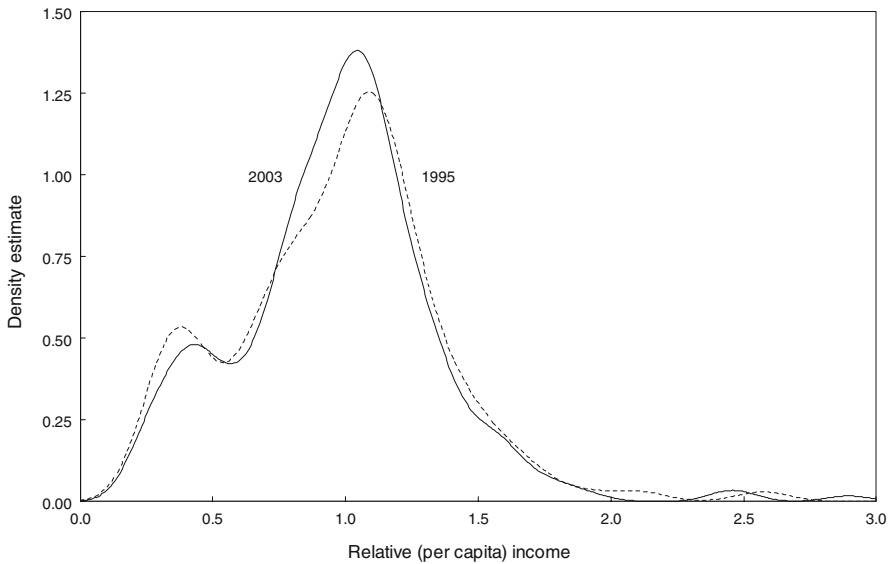


Fig. 1 Distributions of relative (per capita) regional income, 1995 versus 2003. *Notes:* The plots are densities calculated non-parametrically using a Gaussian kernel with bandwidth chosen as suggested in Silverman (1986), restricting the domain to be non-negative. The *solid line* shows the density for 2003 and the *dashed line* that for 1995

twice the average, and so on. The height of the curve over any point gives the probability that any particular region will have that relative income. Since the height of the curve at any particular point gives the probability, the area under the curve between, say 0.0 and 1.0, gives the total likelihood that a region will have a relative income that is between 0.0 and 1.0.

The figure shows a distribution with twin-peaks—to use the appellation coined by Quah (1993a)—in 1995, one corresponding to low income regions and the other to middle-income ranges, and a long tail with two smaller bumps at the upper end of the distribution. Technically, the income distribution is said to show a bimodal shape. The main mode¹⁷ is located at about 110% of the European average, and the second mode at about 38%. The estimated densities reveal several changes over the observation period. The kernel estimated median value decreases by 2%, while the level of dispersion exhibits a small reduction. The kernel estimated standard deviation decreases by 3.3% from 0.393 in 1995 to 0.380 in 2003.

Perhaps most remarkable is the change in the shape of the distributions. By 2003, the peaks have become closer together, and the richer peak has risen moderately at the expense of the poorer. We see this by noting that the area under the 2003 curve, that is between 0.5 and 1.1, is greater than the corresponding area under the 1995 curve, while the area that is to the left of 0.5 is smaller. The smaller peak seems to progressively collapse over time. This finding may suggest an improvement in

¹⁷ A mode is defined as a point at which the gradient changes from positive to negative.

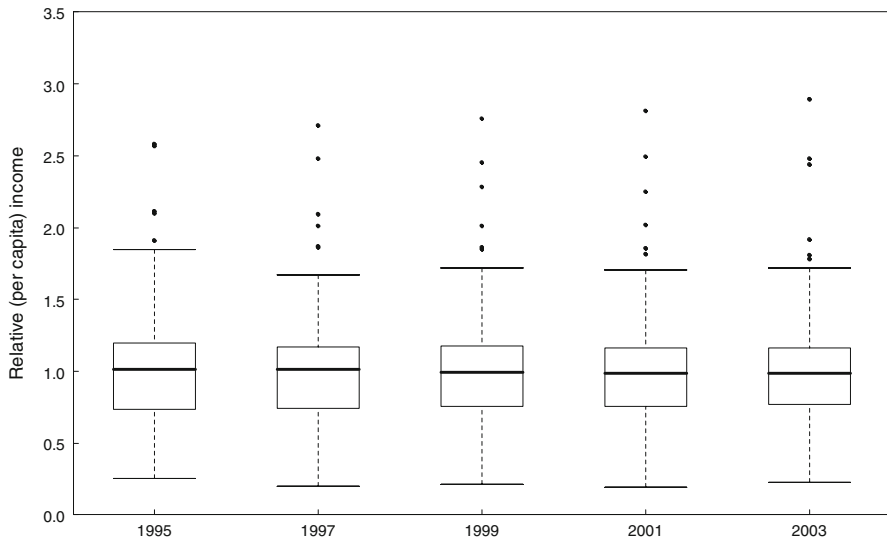


Fig. 2 Tukey *boxplots* of relative (per capita) regional income across 257 European regions

economic conditions of the poorest regions and reflect a trend, in some sense, of catching-up.

Figure 2 gives a sequence of Tukey boxplots for the 257 NUTS-2 regions. Recall that the units of income are PPS units scaled to the EU-27 average. Time appears on the horizontal axis, while the vertical axis maps relative per capita income values. To understand these pictures, recall the construction of a Tukey boxplot. Each boxplot includes a box bounded by Q_1 and Q_3 denoting sample quartiles. Thus, the box contains the middle 50% of the distribution. The thick line in the box locates the median. The upwards and downwards distances from the median to the top and bottom of the box provide information on the shape of the distribution. If these distances differ, then the distribution is asymmetric. Thin dashed vertical lines emanating from the box both upwards and downwards, reach upper and lower adjacent values, respectively. The upper adjacent value is the largest value observed that is not greater than the top quartile plus 1.5 times $(Q_3 - Q_1)$. The lower quartile is similarly defined, extending downwards from the 25th percentile. Dots indicate upper and lower outside values, that is, observations that lie outside the upper and lower adjacent values, respectively. These denote regions which have performed extraordinarily well or extraordinarily poorly relative to the set of other regions. Of course, upper and lower outside values might not exist. The adjacent values might already be the extreme points in a specific realisation.

There are no extraordinarily poorly performing regions, more accurately when regions performed especially badly, they were not alone. On the upside, by contrast, the figure shows several outstanding performers. At the beginning of the sample, five regions showed upper outside values, and by the end of the sample six outside values. The spreading apart in the regional income distribution has one distinct source, the pulling away of the upper outside values—representing Inner London,

Brussels, Luxembourg, Hamburg, Île-de-France and Vienna—from the rest of the regions. The figure, moreover, makes clear that the interquartile range is decreasing by more than 15%, and this falling is due to a decrease of Q_3 rather than Q_1 .

The matching counterparts in Figs. 1 and 2 use exactly the same data. But they emphasise different empirical regularities. The bimodal shape is striking in Fig. 1, but is far from obvious in Fig. 2. The spreading out of the upper tail of the distribution is apparent in Fig. 2. It appears in form of two smaller bumps in Fig. 1.

3.3 Intra-distribution dynamics and long-run tendencies

Thus far, we have considered only point-in-time snapshots of the income distribution across the regions. This section takes the next step in the analysis, and looks at the intra-distribution dynamics and then at the long-run (ergodic) tendencies. We start with Fig. 3 showing cross-profile dynamics.¹⁸ The vertical axis is the log of relative (per capita) incomes. Each curve in the figure refers to the situation at a given point in time. The lowest curve gives the cross-section of regions at time 1995 in increasing order. This ordering is then maintained throughout the time periods considered. Proceeding upwards, we see curves for 1999 and 2003. The character of the upper plots, thus, depends on 1995 when the ordering is taken.

In the plots, increasing jaggedness indicates intra-distribution mobility. In contrast, if each cross-profile would always monotonically increase over time, then income rankings were invariant. The most striking feature of Fig. 3 is not this comparative stability through time. It is the change in choppiness through time in the cross-profile plots indicated by local peaks. By 2003, we observe local peaks, for example, at the lower end of the distribution around regions ranked 9th, 19th, 42nd and 66th poorest in 1995, and at the upper end around regions ranked second and fourth richest. These turn out to be Latvia, Estonia, Mazowieckie (Warszawa) and Közép-Magyarország (Budapest), and Inner London and Luxembourg, respectively. By contrast, Moravskoslezsko (57th poorest in 1995) in the Czech Republic, Lüneburg (129th poorest) and Berlin (the 41st richest region) experienced economically significant relative declines by 2003. The cross-profile dynamics are informative. They illustrate when regions overtake one another, fall behind, or pull ahead. But they do not identify underlying dynamic regularities in the data. We thus turn to the stochastic kernel representation of intra-distribution dynamics next.

Figure 4 shows the conditional kernel density estimate $\hat{g}_\tau(z|y)$ with fixed bandwidths ($h_y = 0.036$, $h_z = 0.023$)¹⁹ that describes the stochastic kernel across the 257 regions, averaging over 1995 through 2003. The stochastic kernel has been estimated for a 5-year transition period, setting $\tau = 5$. The figure displays the estimate, using Hyndman's (1996) visualisation tools. Figure 4a presents the stochastic kernel in terms of a three-dimensional stacked conditional density plot in which a number of conditional densities are plotted side by side in a perspective plot. For any point y on the period t axis, looking in the direction parallel to the

¹⁸ The idea for this picture comes from Quah (1997a), and López-Bazo et al. (1999).

¹⁹ The bandwidths for the estimator were chosen according to Bashtannyk and Hyndman's three-step-strategy. See Sect. 2.2 for more details.

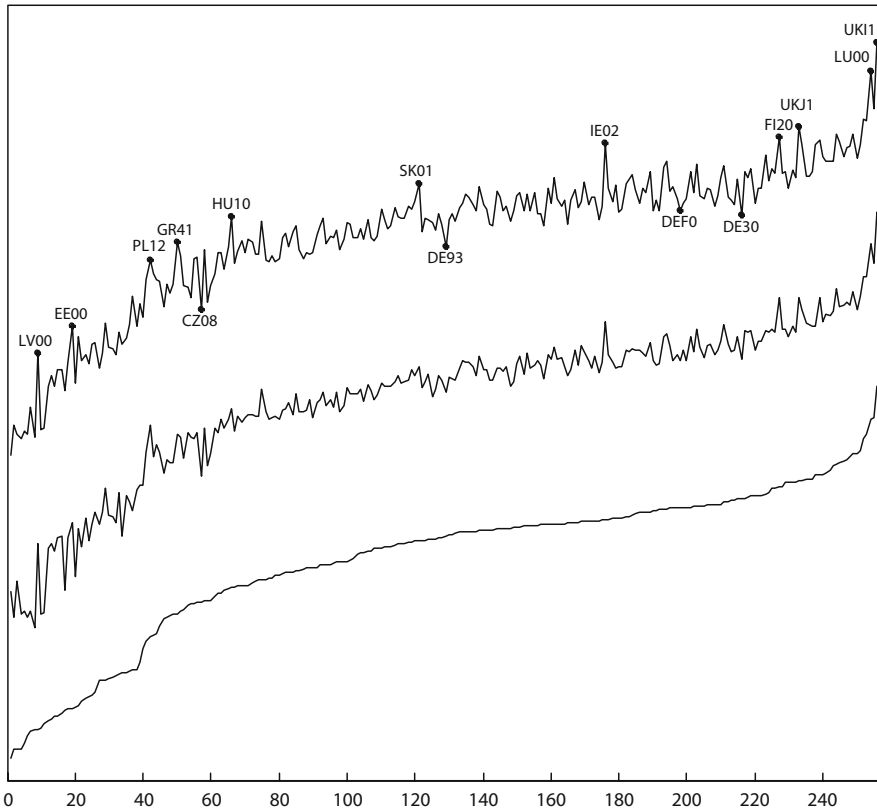


Fig. 3 Cross-profile dynamics across 257 European regions, retaining the ranking fixed at the initial year, relative (per capita) income, advancing upwards: 1995, 1999 and 2003 (a guide to region codes can be found in the Appendix)

$t + 5$ time axis traces out a conditional probability density. The graph shows how the cross-section income distribution at time t evolves into that at time $t + 5$. Just as with a transition probability matrix in a discrete set up, the 45-degree diagonal in the graph indicates persistence properties. When most of the graph is concentrated along this diagonal, then the elements in the cross-section distribution remain where they started. As evident from Fig. 4a, a large portion of the probability mass remains clustered along the main diagonal over the 5-year horizon, and most of the peaks lie along this line indicating a low degree mobility and modest change in the regional income distribution.

The highest density regions (HDRs) boxplot, given in Fig. 4b, makes this clearer. A HDR is the smallest region of the sample space containing a given probability. Figure 4b shows a plot of the 50 and 99% HDRs,²⁰ computed from the density

²⁰ An HDR boxplot replaces the box bounded by the interquartile range with the 50% HDR, the region bounded by the upper and lower adjacent values is replaced by the 99% HDR that roughly reflects the probability coverage of the adjacent values on a standard boxplot for a normal distribution. In keeping with the emphasis on highest density, the mode rather than the median is marked.

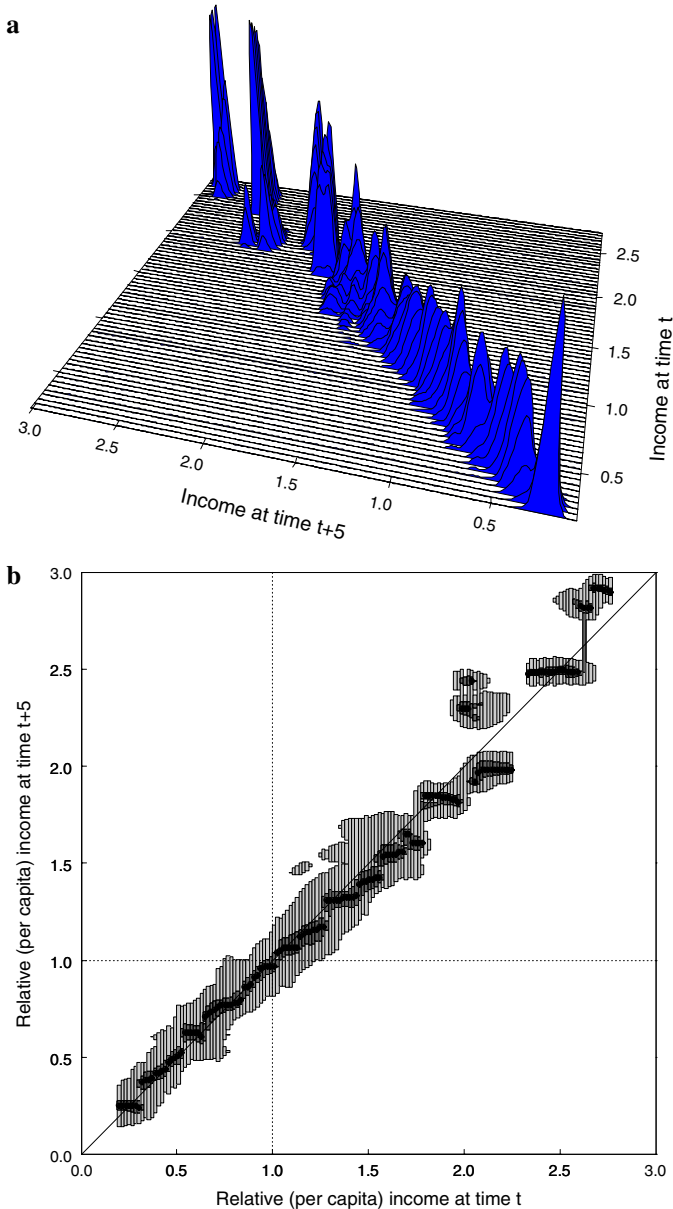


Fig. 4 Relative income dynamics across 257 European regions, the estimated $g_5(z|y)$, see Eq. (6): **a** stacked density plot, and **b** highest density regions boxplot. *Notes:* ad **b** The lighter shaded regions in each strip is a 99% HDR, and the darker shaded region a 50% HDR. The mode for each conditional density is shown as a bullet •. *Technical notes:* The conditional density $g_\tau(z|y)$ is estimated over a 5-year transition horizon $\tau = 5$ between 1995–2003. Estimates are based on a Gaussian product kernel density estimator with bandwidth selection ($h_x = 0.036$, $h_z = 0.023$) based on the three-step-strategy suggested by Bashtannyk and Hyndman (2001). The stacked conditional density plot and the high density region boxplot were estimated at 70 and 150 points, respectively. Calculations of the plots were performed using the R package HRDCDE, provided by Rob Hyndman

estimates shown in Fig. 4a. Each vertical strip represents the conditional density for one y value. The darker shaded region in each strip is a 50% HDR, and the lighter shaded region is a 99% HDR. The mode for each conditional density is shown as a bullet •. The vertical dashed line at 1.0 marks regions with income equal to the European average at time t , and the horizontal dashed line at 1.0 those with income equal to the average at $t + 5$. The 45-degree diagonal indicates intra-distribution persistence over the 5-year transition horizon.

To read this type of boxplot note that strong persistence is evidenced when the main diagonal crosses the 50% HDRs. It means that most of the elements in the distribution remain where they started. There is a low persistence and more intra-distribution mobility if that diagonal crosses only the 99% HDRs. Strong (weak) global convergence towards equality would manifest in 50% (99%) HDRs crossed by the horizontal line at 1.0. Fifty percent HDRs consisting of two disjoint intervals would indicate a two-peaks property of the distribution.

The plot not only reveals persistence, but also mobility and polarisation features. Regions with an income range of 0.8–1.2 times the European average show strong persistence. Some mobility occurs at the extremes of the distribution, more at the upper extreme than at the lower. Some portions of the cross-section in the income range below 0.8 times the average tend to slightly increase their relative position over the 5-year transition horizon, indicating a process of catching-up of the poorest regions with the richer ones. In contrast, portions in the income range above 1.2–1.8 times the average lose out their relative position, becoming relatively poorer. The boxplot also shows signs of polarisation, the opposite of catching-up. This is indicated by the disjoint intervals of the 50 and 99% HDRs at the upper extreme of the income range. We see that regions starting with an income of 2.0–2.3 times the European average at time t are unlikely to remain there. Most see their European relative income fall and others rise, with the result that this income class appears to vanish. The position of a small very rich group around 2.3–2.6 times the average remains either unchanged or shifting away.

The evidence of Fig. 4 is corroborated by the ergodic density function that is obtained by solving Eq. (4). Figure 5 plots the estimated long-run (ergodic) density,²¹ $\hat{f}_\infty(z)$, implied by the estimated $g_\tau(z|y)$ function for $\tau = 5$, along with the initial income distribution. The solid line shows the point estimate of the ergodic distribution and the dashed line the initial income distribution. Comparing these two distributions we see that the ergodic distribution is wider, both at the top and at the bottom. This reflects a shift in the mass of the distribution away from the lower end to the middle, and from the middle to the upper end. In particular, the peak in the initial distribution between 20 and 50% of the European relative per capita income has shifted upward into the 60–100% range and shows a tendency to disappear.

²¹ It is well known that the shape of the estimated ergodic density is sensitive to the bandwidths chosen in computing the underlying estimated joint density functions. Wider bandwidths tend to obscure detail in the shapes while narrower bandwidths tend to increase it but possibly spuriously so. It is important to note that smaller equiproportionate decreases and increases in bandwidths do not remove the tendency to bimodality in the ergodic density.

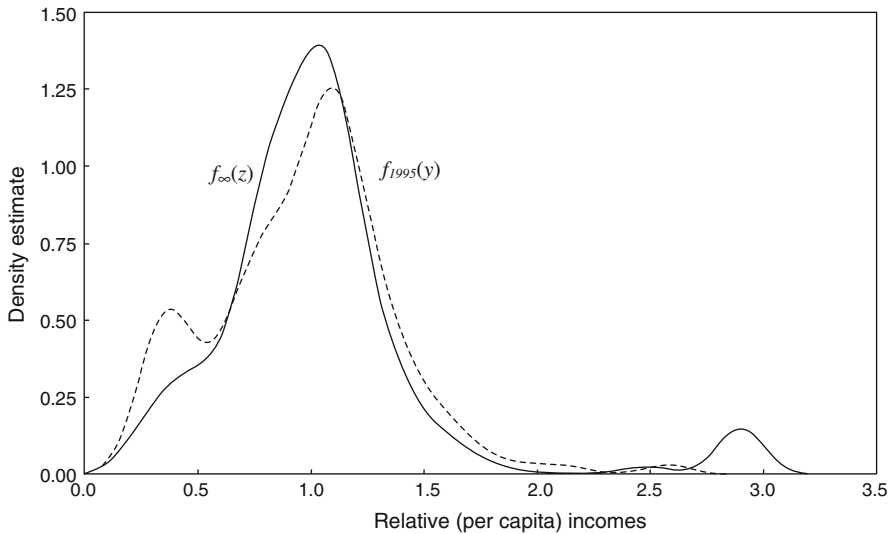


Fig. 5 The ergodic density $f_{\infty}(z)$ implied by the estimated $g_5(z|y)$ and the marginal density function $f_{1995}(y)$. *Notes:* The solid line shows the point estimate for $f_{\infty}(z)$ and the dashed line the estimate for the marginal density $f_{1995}(y)$. The ergodic function $f_{\infty}(z)$ has been found as solution to Eq. (4)

The stationary distribution across the 257 regions, plotted in Fig. 5, is distinctively bimodal. The dominant peak²² represents regions clustered just below the European average income, while a small group of relatively rich regions gathers around three times of the average European (per capita) income. The bimodal nature of the ergodic distribution in comparison with the initial income distribution provides indication for two types of processes at work over time: a gradual and slow catching-up of the poorest regions which turn out to be—with very few exceptions—regions in Central and Eastern Europe, and simultaneously a tendency towards polarisation—a small group of richer regions separating from the rest of the cross-section.

The bimodal shape of the ergodic distribution contradicts with Quah's (1996a) unimodal ergodic solution found in a discrete state space set up with a largely reduced set of 78 European regions over 1980–1989. The observation, however, is in line with Pittau and Zelli's (2006) findings, obtained for a set of 110 regions covering twelve EU member countries over the time period from 1977 to 1996.

To sum up this first pass through the data, we conclude that the data show a wide spectrum of intra-distribution dynamics. Overtaking and catching-up occur simultaneously with persistence and polarisation. Polarisation manifests itself in the emergence of a twin-peak structure in the long-run regional income distribution.

3.4 The spatial filtering perspective

Large significant and positive values of Moran's I reveal the presence of spatial association of similar values of neighbouring European regions in relative (per

²² The upper peak, however, is imprecisely estimated. Only few observations were actually made there, and the precision of the estimate is low.

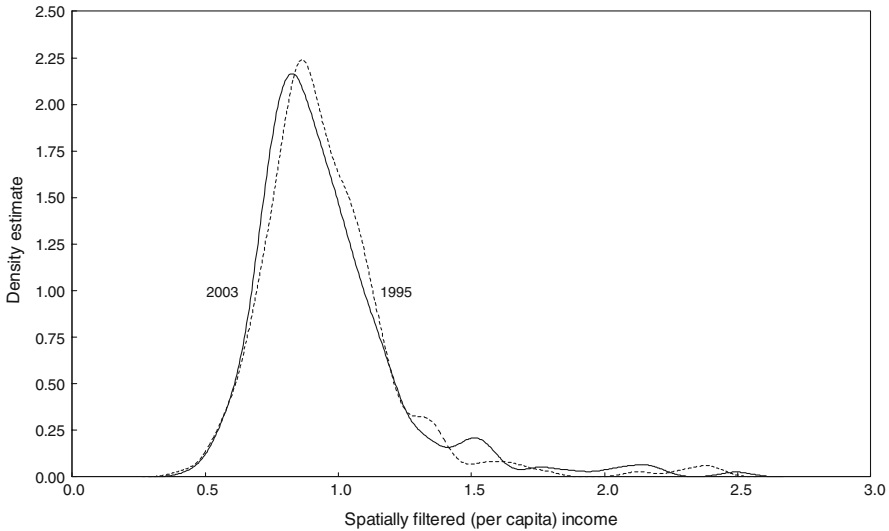


Fig. 6 Densities of relative (per capita) income, 1995 versus 2003: the spatial filtering view. *Notes:* The plots are densities calculated non-parametrically using a Gaussian kernel with bandwidth chosen as suggested in Silverman (1986), restricting the domain to be non-negative. The solid line shows the density for 2003 and the dashed line that for 1995

capita) income.²³ This motivates a spatial filtering pass²⁴ through the data to avoid inferences and interpretations, misguided by the violation of the independence assumption in the previous analysis.

Figure 6 presents the spatially filtered counterpart of Fig. 1. Comparing these densities with those in Fig. 1 indicates that the mode, which was situated at around 38% of the European average, has disappeared. Consequently, the economic performance of the regions is well explained by the neighbouring regions' performances, except may be for regions with very high relative (per capita) income.

The filtered distributions in this figure are tighter and more concentrated than those in Fig. 1. The boxplots in Fig. 7 make this particularly clear. Upper and lower outliers exist here, but the 25th and 75th percentiles are located close to the average income. Lower and upper adjacent values are compactly situated within about 0.5 and 1.5 times average income levels. The filtered distribution has a kernel estimated standard deviation of 0.262 in 1995, which increases to 0.283 in 1999, and then to 0.310 in 2003. The increase over the time 1995–2003 is 15%. The estimated standard deviations of the unfiltered data were found to be 0.393 in 1995 and 0.380 in 2003, indicating a slight decline by 3.3%. From this, it is clear that the evidence

²³ Using Moran's I , the spatial autocorrelation latent in each of the income variables ranges from $z(MI) = 8.86$ for the 1995 income variable to $z(MI) = 8.06$ for the 2003 income variable where $z(MI)$ denotes the z -score value of Moran's I . From this, it is clear that there is a strong spatial autocorrelation, and hence the assumption of spatial independence does not hold.

²⁴ Rather than use an individual δ for each observation, the modal value for δ was chosen for each income variable as recommended by Getis and Griffith (2002).

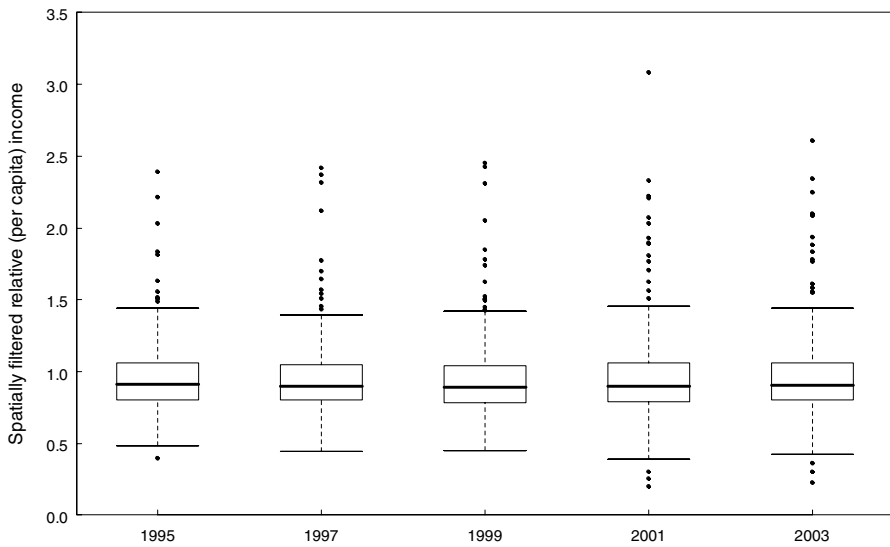


Fig. 7 Tukey boxplots of relative (per capita) income, across 257 European regions: the spatial filtering view

for σ -convergence found in Sect. 3.1 is caused by spatial dependence embedded in the income data.²⁵

More information on the role of spatial effects becomes evident when looking at the stochastic kernel in Fig. 8 that shows how the original (unfiltered) relative (per capita) income distribution is transformed into the spatially filtered one. Figure 8a displays the conditional kernel density estimate $\hat{g}(\tilde{y}|y)$ with fixed bandwidths ($h_y = 0.103$, $h_{\tilde{y}} = 0.052$) in terms of a three-dimensional stacked conditional plot as given in Fig. 8a, and an HDR boxplot in Fig. 8b.

If spatial effects account for a substantial part of the distribution, then the stochastic kernel mapping from the original (unfiltered) to the spatially filtered distribution would depart from the identity map. Indeed, Fig. 8a precisely conveys this message. The graph shows the kernel mapping the original to the filtered distribution in the same year. The evident clockwise reversal on the lower, but also on the higher part of the distribution indicates that spatial effects do account for a large part of income dynamics in Europe. Figure 8b reinforces this interpretation. The dominant feature in this figure appears to be intra-distribution mobility rather than persistence. Regions with an income less than 0.7 times the European average show a clear tendency towards cohesion. There are strong indications that the probability of the poorest regions to move up is negatively affected by the presence of spatial dependence effects. This is evidenced by the 99% HDRs crossing the horizontal line at 1.0 and by the 50% HDRs coming much closer to this line. However, while this is happening, the very highest parts of the income distribution show tendencies away from cohesion, and provide evidence for emerging twin peaks.

²⁵ See Rey and Dev (2006) for appropriate inference methods of σ -convergence in the presence of spatial effects.

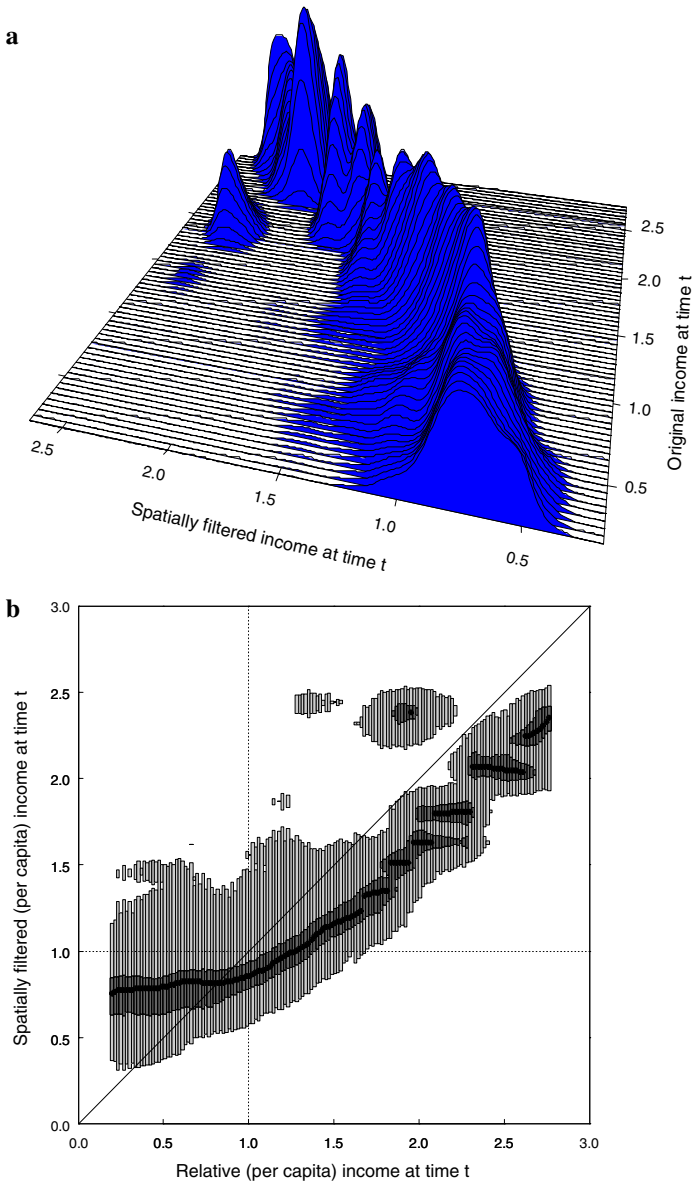


Fig. 8 Stochastic kernel mapping from the original to the spatially filtered distribution, the estimated $g(\tilde{y}|y)$: **a** stacked conditional density plot, and **b** highest density regions boxplot. *Notes:* ad **b** The lighter shaded region in each strip is a 99% HDR, and the darker shaded region a 50% HDR. The mode for each conditional density is shown as a bullet •. *Technical notes:* The conditional density $g(\tilde{y}|y)$ is estimated over a 5-year transition horizon $\tau = 5$ between 1995 and 2003. Estimates are based on a Gaussian product kernel density estimation with bandwidth selection ($h_y = 0.103, h_{\tilde{y}} = 0.052$) based on the three-step-strategy suggested by Bashtannyk and Hyndman (2001). The stacked conditional density plot and the high density region boxplot were estimated at 70 and 150 points, respectively. Calculations of the plots were performed using the R package HRDCDE, provided by Rob Hyndman, and spatial filtering, using the PPA package, provided by Arthur Getis

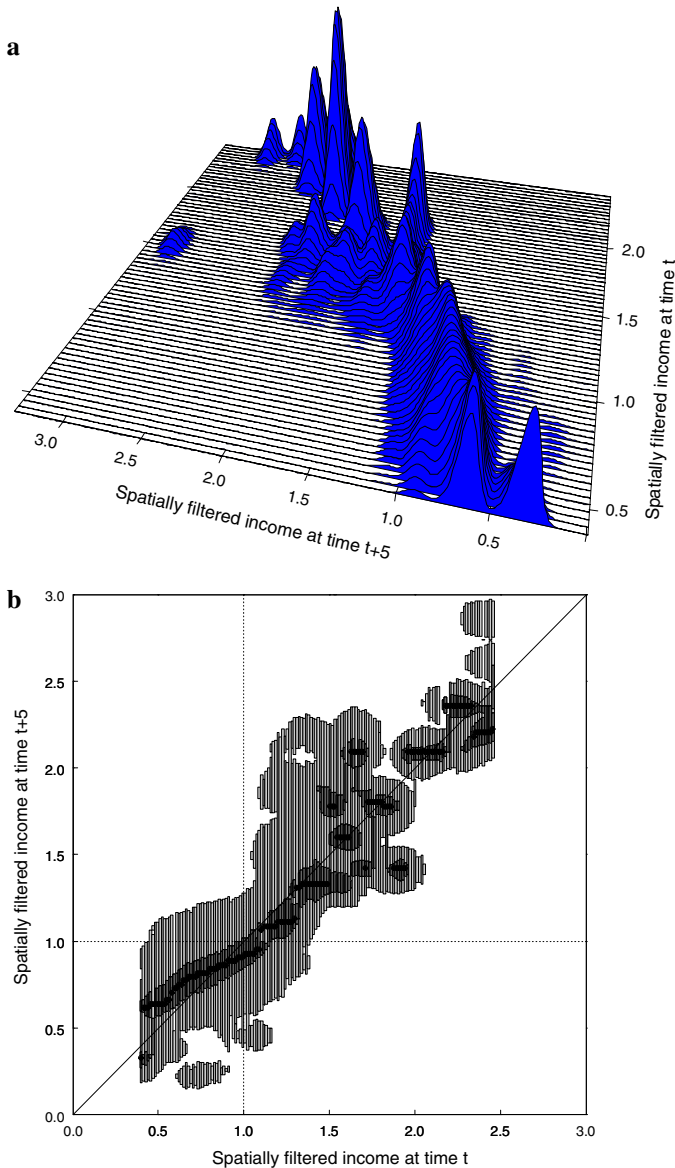


Fig. 9 The spatial filter view of relative income dynamics: The estimated $g_s(\bar{z}|\bar{y})$, **a** stacked density plot, and **b** highest density regions boxplot. *Notes:* **a** **b** The *lighter shaded region* in each strip is a 99% HDR, and the *darker shaded region* a 50% HDR. The mode for each conditional density is shown as a bullet \bullet . *Technical notes:* The conditional density $g_s(\bar{z}|\bar{y})$ is estimated over a 5-year transition horizon $\tau = 5$ between 1995–2003. Estimates are based on a Gaussian product kernel density estimator with bandwidth selection ($h_{\bar{y}} = 0.061, h_{\bar{z}} = 0.047$) based on the three-step-strategy suggested by Bashtannyk and Hyndman (2001). The stacked conditional density plot and the high density region boxplot were estimated at 70 and 150 points, respectively. Calculations of the plots were performed using the R package HRDCDE, provided by Rob Hyndman, and spatial filtering using the PPA package, provided by Arthur Getis

Figure 9 provides stochastic kernel representations of 5-year transition dynamics in the spatially filtered income space, using again a stochastic kernel estimator with fixed bandwidths ($h_{\bar{y}} = 0.061$, $h_{\bar{z}} = 0.047$). This figure is the counterpart to Fig. 4 for spatially filtered relative (per capita) regional incomes. Figure 9a presents the stochastic kernel in terms of a three-dimensional stacked conditional density plot, and Fig. 9b in terms of a HDRs boxplot. The picture that emerges from the estimates here is that of a substantial degree of intra-distribution mobility at the upper and lower tails of the income distribution. The remarkably different dynamics that emerge—in comparison to the unfiltered regional income case—suggest that—if we are to evaluate growth and convergence dynamics across regions correctly—the use of spatially filtered data is pretty much essential to avoid misleading interpretations.

4 Concluding remarks

The study follows the tradition of the non-parametric approach studying both the shape and mobility dynamics of cross-sectional distributions of relative (per capita) income that appears to be generally more informative about the actual patterns of cross-sectional growth than convergence empirics within the β -convergence regression approach. It differs from most of the previous work by going for a continuous kernel route which is more informative than research with discretely-defined income cells.

This paper incorporates two novel techniques into the continuous analysis: kernel estimation and more powerful graphical devices for the representation of the stochastic kernel, and Getis' spatial filtering technique to explicitly account for the spatial dimension of the growth process. The paper illustrates that the use of spatially filtered data is pretty much essential to evaluate growth and convergence dynamics across regions. The lack of an appropriate inferential theory, however, restricts the study to a descriptive stage.

The study has produced some interesting results. First, there is no development trap in the long-run into which the poorer Central and Eastern European regions will be permanently condemned. Second, the findings suggest a tendency of the cross-section distribution of regional per capita income to split up into two separate groups, where a small group of richer metropolitan regions is growing away from the rest of the European regions. This evidence is coherent to Pittau and Zelli's (2006) stationary distribution estimated on a sample of 110 EU-12 regions over the period 1977–1996. Third, spatial effects explain a substantial part of the income distribution, but not the emergence of the two-club regional world in the long-run. Growth theories now need to explain these facts. The distribution dynamics analysis carried out in this paper does not help further in this respect.

Acknowledgments The authors gratefully acknowledge the grant no. P19025-G11 provided by the Austrian Science Fund (FWF). They also thank two anonymous referees for their comments which improved the quality of the paper. The calculations were done using a combination of the R package HDRCDE, provided by Rob Hyndman, and the PPA package, provided by Arthur Getis. Special thanks to Roberto Basile for providing the original stimulus to carry out this study.

Appendix

NUTS is an acronym of the French for the “nomenclature of territorial units for statistics”, which is a hierarchical system of regions used by the statistical office of the European Community for the production of regional statistics. At the top of the hierarchy are NUTS-0 regions (countries) below which are NUTS-1 regions and then NUTS-2 regions. The sample is composed of 257 NUTS-2 regions located in 27 EU member states (NUTS revision 1999, except for Finland NUTS revision 2003). We exclude the Spanish North African territories of Ceuta and Melilla, and the French Départements d’Outre-Mer Guadeloupe, Martinique, French Guayana and Réunion, the Spanish North African territories of Ceuta y Mellila, and the Portuguese non-continental territories Azores and Madeira. Thus, we include the NUTS 2 regions listed in the table.

Country	ID code	Region
Austria	AT11	Burgenland
	AT12	Niederösterreich
	AT13	Wien
	AT21	Kärnten
	AT22	Steiermark
	AT31	Oberösterreich
	AT32	Salzburg
	AT33	Tirol
	AT34	Vorarlberg
Belgium	BE10	Région de Bruxelles-Capitale
	BE21	Prov. Antwerpen
	BE22	Prov. Limburg (B)
	BE23	Prov. Oost-Vlaanderen
	BE24	Prov. Vlaams Brabant
	BE25	Prov. West-Vlaanderen
	BE31	Prov. Brabant Wallon
	BE32	Prov. Hainaut
	BE33	Prov. Liège
	BE34	Prov. Luxembourg (B)
	BE35	Prov. Namur
Bulgaria	BG11	Severozapaden
	BG12	Severen tsentralen
	BG13	Severoiztochen
	BG21	Yugozapaden
	BG22	Yuzhen tsentralen
	BG23	Yugoiztochen
Cyprus	CY00	Kypros / Kibris

Appendix continued

Country	ID code	Region
Czech Republic	CZ01	Praha
	CZ02	Střední Čechy
	CZ03	Jihozápad
	CZ04	Severozápad
	CZ05	Severovýchod
	CZ06	Jihovýchod
	CZ07	Střední Morava
	CZ08	Moravskoslezsko
Denmark	DK00	Danmark
Estonia	EE00	Eesti
Finland	FI13	Itä-Suomi
	FI18	Etelä-Suomi
	FI19	Länsi-Suomi
	FI1A	Pohjois-Suomi
	FI20	Åland
France	FR10	Île de France
	FR21	Champagne-Ardenne
	FR22	Picardie
	FR23	Haute-Normandie
	FR24	Centre
	FR25	Basse-Normandie
	FR26	Bourgogne
	FR30	Nord-Pas-de-Calais
	FR41	Lorraine
	FR42	Alsace
	FR43	Franche-Comté
	FR51	Pays de la Loire
	FR52	Bretagne
	FR53	Poitou-Charentes
	FR61	Aquitaine
	FR62	Midi-Pyrénées
	FR63	Limousin
	FR71	Rhône-Alpes
	FR72	Auvergne
FR81	Languedoc-Roussillon	
FR82	Provence-Alpes-Côte d'Azur	
FR83	Corse	

Appendix continued

Country	ID code	Region
Germany	DE11	Stuttgart
	DE12	Karlsruhe
	DE13	Freiburg
	DE14	Tübingen
	DE21	Oberbayern
	DE22	Niederbayern
	DE23	Oberpfalz
	DE24	Oberfranken
	DE25	Mittelfranken
	DE26	Unterfranken
	DE27	Schwaben
	DE30	Berlin
	DE40	Brandenburg (Südwest and Nordost)
	DE50	Bremen
	DE60	Hamburg
	DE71	Darmstadt
	DE72	Gießen
	DE73	Kassel
	DE80	Mecklenburg-Vorpommern
	DE91	Braunschweig
	DE92	Hannover
	DE93	Lüneburg
	DE94	Weser-Ems
	DEA1	Düsseldorf
	DEA2	Köln
	DEA3	Münster
	DEA4	Detmold
	DEA5	Arnsberg
	DEB1	Koblenz
	DEB2	Trier
	DEB3	Rheinessen-Pfalz
	DEC0	Saarland
	DED1	Chemnitz
	DED2	Dresden
	DED3	Leipzig
	DEE1	Dessau
	DEE2	Halle
	DEE3	Magdeburg
	DEF0	Schleswig-Holstein
	DEG0	Thüringen

Appendix continued

Country	ID code	Region
Greece	GR11	Anatoliki Makedonia, Thraki
	GR12	Kentriki Makedonia
	GR13	Dytiki Makedonia
	GR14	Thessalia
	GR21	Ipeiros
	GR22	Ionia Nisia
	GR23	Dytiki Ellada
	GR24	Stereia Ellada
	GR25	Peloponnisos
	GR30	Attiki
	GR41	Voreio Aigaio
	GR42	Notio Aigaio
	GR43	Kriti
	Hungary	HU10
HU21		Közép-Dunántúl
HU22		Nyugat-Dunántúl
HU23		Dél-Dunántúl
HU31		Észak-Magyarország
HU32		Észak-Alföld
HU33		Dél-Alföld
Ireland	IE01	Border, Midlands and Western
	IE02	Southern and Eastern
Italy	IT31	Bolzano-Bozen e Trento
	ITC1	Piemonte
	ITC2	Valle d' Aosta/Vallée d' Aoste
	ITC3	Liguria
	ITC4	Lombardia
	ITD3	Veneto
	ITD4	Friuli-Venezia Giulia
	ITD5	Emilia-Romagna
	ITE1	Toscana
	ITE2	Umbria
	ITE3	Marche
	ITE4	Lazio
	ITF1	Abruzzo
	ITF2	Molise
	ITF3	Campania
	ITF4	Puglia
	ITF5	Basilicata

Appendix continued

Country	ID code	Region
Italy	ITF6	Calabria
	ITG1	Sicilia
	ITG2	Sardegna
Lithuania	LT00	Lietuva
Luxembourg	LU00	Luxembourg (Grand-Duché)
Latvia	LV00	Latvija
Malta	MT00	Malta
Netherlands	NL11	Groningen
	NL12	Friesland
	NL13	Drenthe
	NL21	Overijssel
	NL22	Gelderland
	NL23	Flevoland
	NL31	Utrecht
	NL32	Noord-Holland
	NL33	Zuid-Holland
	NL34	Zeeland
	NL41	Noord-Brabant
	NL42	Limburg (NL)
	Poland	PL11
PL12		Mazowieckie
PL21		Malopolskie
PL22		Slaskie
PL31		Lubelskie
PL32		Podkarpackie
PL33		Swietokrzyskie
PL34		Podlaskie
PL41		Wielkopolskie
PL42		Zachodniopomorskie
PL43		Lubuskie
PL51		Dolnoslaskie
PL52		Opolskie
PL61		Kujawsko-Pomorskie
PL62	Warminsko-Mazurskie	
PL63	Pomorskie	

Appendix continued

Country	ID code	Region
Portugal	PT11	Norte
	PT15	Algarve
	PT16	Centro (P)
	PT17	Lisboa
	PT18	Alentejo
Romania	RO01	Nord-Est
	RO02	Sud-Est
	RO03	Sud
	RO04	Sud-Vest
	RO05	Vest
	RO06	Nord-Vest
	RO07	Centru
	RO08	Bucuresti
Slovakia	SK01	Bratislavský kraj
	SK02	Západné Slovensko
	SK03	Stredné Slovensko
	SK04	Východné Slovensko
Slovenia	SI00	Slovenija
Spain	ES11	Galicía
	ES12	Principado de Asturias
	ES13	Cantabria
	ES21	País Vasco
	ES22	Comunidad Foral de Navarra
	ES23	La Rioja
	ES24	Aragón
	ES30	Comunidad de Madrid
	ES41	Castilla y León
	ES42	Castilla-La Mancha
	ES43	Extremadura
	ES51	Cataluña
	ES52	Comunidad Valenciana
	ES53	Illes Balears
	ES61	Andalucía
ES62	Región de Murcia	
Sweden	SE01	Stockholm
	SE02	Östra Mellansverige
	SE04	Sydsverige
	SE06	Norra Mellansverige

Appendix continued

Country	ID code	Region
Sweden	SE07	Mellersta Norrland
	SE08	Övre Norrland
	SE09	Småland med öarna
	SE0A	Västsverige
United Kingdom	UKC1	Tees Valley and Durham
	UKC2	Northumberland, Tyne and Wear
	UKD1	Cumbria
	UKD2	Cheshire
	UKD3	Greater Manchester
	UKD4	Lancashire
	UKD5	Merseyside
	UKE1	East Riding and North Lincolnshire
	UKE2	North Yorkshire
	UKE3	South Yorkshire
	UKE4	West Yorkshire
	UKF1	Derbyshire and Nottinghamshire
	UKF2	Leicestershire, Rutland and Northants
	UKF3	Lincolnshire
	UKG1	Herefordshire, Worcestershire and Warks
	UKG2	Shropshire and Staffordshire
	UKG3	West Midlands
	UKH1	East Anglia
	UKH2	Bedfordshire, Hertfordshire
	UKH3	Essex
	UKI1	Inner London
	UKI2	Outer London
	UKJ1	Berkshire, Bucks and Oxfordshire
	UKJ2	Surrey, East and West Sussex
	UKJ3	Hampshire and Isle of Wight
	UKJ4	Kent
	UKK1	Gloucestershire, Wiltshire and North Somerset
	UKK2	Dorset and Somerset
	UKK3	Cornwall and Isles of Scilly
	UKK4	Devon
	UKL1	West Wales and The Valleys
	UKL2	East Wales
	UKM1	North Eastern Scotland
UKM2	Eastern Scotland	
UKM3	South Western Scotland	
UKM4	Highlands and Islands	
UKN0	Northern Ireland	

References

- Abreu M, Groot HLF, Florax RJG (2004) Space and growth: a survey of empirical evidence and methods. Tinbergen Institute, Working Paper No TI04-129/3. Available at SSRN: <http://www.ssrn.com/abstract=631007>
- Bashtannyk DM, Hyndman RJ (2001) Bandwidth selection for kernel conditional density estimation. *Comput Stat Data Anal* 36:279–298
- Basile R (2006) Intra-distribution dynamics of regional per capita income in Europe: evidence from alternative conditional density estimators. Paper presented at the 53rd North-American meetings of the RSAI, November 2006, Toronto
- Bulli S (2001) Distribution dynamics and cross-country convergence: a new approach. *Scottish J Polit Econ* 48(2):226–243
- Durlauf SN, Quah DT (1999) The new empirics of economic growth. In: Taylor JB, Woodford M (eds) *Handbook of macroeconomics*, vol 1. Elsevier, Amsterdam, pp 235–308
- Fan J, Yao Q, Tong H. (1996) Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika* 83(1): 189–206
- Fingleton B (1997) Specification and testing of Markov chain models: an application to convergence in the European Union. *Oxf Bull Econ Stat* 59(3):385–403
- Fingleton B (1999) Estimates of time to economic convergence: an analysis of regions of the European Union. *Int Reg Sci Rev* 22(1):5–34
- Fingleton B (ed) (2003) *European regional growth*. Springer, Berlin
- Fischer MM, Stüböck C (2006) Pan-European regional income growth and club-convergence. *Ann Reg Sci* 40(4):693–721
- Friedman M (1992) Do old fallacies ever die? *J Econ Literat* 30(4):2129–2132
- Getis A (1990) Screening for spatial dependence in regression analysis. *Pap Reg Sci Assoc* 69:69–81
- Getis A (1995) Spatial filtering in a regression framework: examples using data on urban crime, regional inequality, and government expenditures. In: Anselin L, Florax R (eds) *New directions in spatial econometrics*. Springer, Berlin, pp 172–188
- Getis A, Griffith DA (2002) Comparative spatial filtering in regression analysis. *Geogr Anal* 34(2):130–140
- Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. *Geogr Anal* 24(3):189–206
- Griffith DA (2006) *Spatial autocorrelation and spatial filtering*. Springer, Berlin
- Hall P, Wolff RC, Yao Q (1999) Methods for estimating a conditional distribution function. *J Am Stat Assoc* 94(445):154–163
- Hyndman RJ (1996) Computing and graphing highest density regions. *Am Stat* 50(2):120–126
- Hyndman RJ, Yao Q (2002) Nonparametric estimation and symmetry tests for conditional density functions. *J Nonparam Stat* 14(3):259–278
- Hyndman RJ, Bashtannyk DM, Grunwald GK (1996) Estimating and visualizing conditional densities. *J Comput Graph Stat* 5(4):315–336
- Islam N (2003) What have we learnt from the convergence debate? *J Econ Surv* 17(3):309–362
- Johnson PA (2004) A continuous state space approach to “Convergence by parts”. Vassar College Economics Working Paper No. 54, Department of Economics, Vassar College
- Jones CI (1997) On the evolution of the world income distribution. *J Econ Perspect* 11(3):19–36
- LeGallo J (2004) Space-time analysis of GDP disparities among European regions: a Markov chains approach. *Int Reg Sci Rev* 27(2):138–163
- López-Bazo E, Vaya E, Mora AJ, Surinach J (1999) Regional economic dynamics and convergence in the European Union. *Ann Reg Sci* 33(3):343–370
- Magrini S (1999) The evolution of income disparities among the regions of the European Union. *Reg Sci Urban Econ* 29(2):257–281
- Magrini S (2004) Regional (di)convergence. In: Henderson JV, Thisse J-F (eds) *Handbook of regional and urban economics*. Elsevier, Amsterdam pp 2741–2796
- Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. *Geogr Anal* 27(4):286–306
- Paap R, van Dijk HK (1998) Distribution and mobility of wealth of nations. *Eur Econ Rev* 42(7):1269–1293
- Pittau MG, Zelli R (2006) Empirical evidence of income dynamics across EU regions. *J Appl Econ* 21:605–628

- Quah DT (1993a) Empirical cross-section dynamics in economic-growth. *Eur Econ Rev* 37(2–3):426–434
- Quah DT (1993b) Galtons fallacy and tests of the convergence hypothesis. *Scand J Econ* 95(4):427–443
- Quah DT (1996a) Regional convergence clusters across Europe. *Eur Econ Rev* 40(3–5):951–958
- Quah DT (1996b) Empirics for economic growth and convergence. *Eur Econ Rev* 40(6):1353–1375
- Quah DT (1997a) Regional cohesion from local isolated actions: I Historical outcomes. Centre for Economic Performance, Discussion Paper No. 378, London School of Economics
- Quah DT (1997b) Regional cohesion from local isolated actions: II Conditioning. Centre for Economic Performance, Discussion Paper No. 379, London School of Economics
- Quah DT (1997c) Empirics for growth and distribution: stratification, polarization, and convergence clubs. *J Econ Growth* 2(1):27–59
- Quah DT (1999) Regional cohesion from local isolated actions— historical outcomes, in study of the socio-economic impact of the projects financed by the cohesion fund—a modelling approach, vol 2. Office for Official Publications of the European Communities, Luxembourg
- Reichlin L (1999) Discussion of ‘Convergence as distribution dynamics’, by D Quah. In: Baldwin R, Cohen D, Sapir A, Venables A (eds) *Market, integration, regionalism, and the global economy*. Cambridge University Press, Cambridge, pp 328–335
- Rey SJ (2001) Spatial empirics for economic growth and convergence. *Geogr Anal* 33(3):195–214
- Rey SJ, Dev B (2006) σ -convergence in the presence of spatial effects. *Pap Reg Sci* 85(2):217–234
- Rey SJ, Janikas MV (2005) Regional convergence, inequality and space. *J Econ Geogr* 5(2):155–176
- Silverman BW (1986) *Density estimation for statistics and data analysis*. Chapman & Hall, London and New York
- Temple J (1999) The new growth evidence. *J Econ Literat* 37:112–156
- Wand MP, Jones MC (1995) *Kernel smoothing*. Chapman and Hall, London