ePub^{WU} Institutional Repository

Manfred M. Fischer and Peter Stumpner

Income Distribution Dynamics and Cross-Region Convergence in Europe. Spatial filtering and novel stochastic kernel representations

Paper

Income Distribution Dynamics and Cross-Region Convergence in Europe

Spatial filtering and novel stochastic kernel representations

Manfred M. Fischer and Peter Stumpner

Institute for Economic Geography and GIScience
Vienna University of Economics and Business Administration
A-1090 Vienna, Austria
manfred.fischer@wu-wien.ac.at

April 17, 2007

Abstract. This paper suggests an empirical framework for analysing income distribution dynamics and cross-region convergence in the European Union of 27 member states, 1995-2003. The framework lies in the research tradition that allows the state income space to be continuous, puts emphasis on both shape and intra-distribution dynamics and uses stochastic kernels for studying transition dynamics and implied long-run behaviour. In this paper stochastic kernels are described by conditional density functions, estimated by a product kernel estimator of conditional density and represented by means of novel visualisation tools. The technique of spatial filtering is used to account for spatial effects, in order to avoid misguided inferences and interpretations caused by the presence of spatial autocorrelation in the income distributions. The results reveal a slow catching-up of the poorest regions and a process of polarisation, with a small group of very rich regions shifting away from the rest of the cross-section. This is well evidenced by both, the unfiltered and the filtered ergodic density view. Differences exist in detail, and these emphasise the importance to properly deal with the spatial autocorrelation problem.

JEL Classification: C14, D30, O18, O47, R11

Keywords: Regional income, distribution dynamics, stochastic kernel estimation, spatial filtering, EU-27

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 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[1] 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[2]. 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[3]. Most research has, however, concentrated on the

[1] NUTS-2 regions vary considerably in size, but are nevertheless considered to be the most appropriate spatial units for modelling and analysis (Fingleton 2001). In most cases, the NUTS-2 regions are sufficiently small to capture subnational variations. But we are aware that NUTS-2 regions are formal rather than functional regions, and their delineation does not represent the boundaries of growth and convergence processes very well (Fischer and Stirböck 2006). This may cause nuisance spatial dependence in the data.

[2] This short observation period was enforced on us by the lack of reliable data in Central and Eastern Europe.

[3] Recent surveys of the new growth literature in general and the convergence literature in particular can be

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 (1993a), Friedman (1994) 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.

In this study, we follow the distribution dynamics approach, carried forward almost single handedly by Quah (1993a, 1996a, b, 1997a, b, c). This approach 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. 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.

Most of the applications of this approach have worked in a discrete state space set up[4] (see Quah 1996a, b, Fingleton 1997, 1999, 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[5] (Bulli 2001).

This paper avoids arbitrary discretisation of the state income space and its possible effects on

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.

[4] There are some few exceptions, most notably Quah (1997a, b), Magrini (2004), and Pittau and Zelli (2006).

[5] Bulli (2001) shows how to obtain a discrete state space Markov chain from a continuous state space Markov process and finds that this method is an accurate approximation to the distribution computed using a continuous state space procedure as performed in this paper.

the results by using stochastic kernels, the continuous equivalent of the transition probability matrix, to estimate the underlying regional income distribution and to analyse its evolution over time. Assuming that the income distribution at any point in time can be described by a density function, the stochastic kernel can be described by a conditional density function that can be estimated by a product kernel estimator of conditional density. Under the assumption of stationarity of the underlying process, the long-run (ergodic) out-of-sample limit of the distribution of regional incomes is estimated. The results are presented in terms of three-dimensional stacked conditional density plots and boxplots based on highest density regions, novel visualisation tools[6] introduced by Hyndman (1996).

The remainder of the paper is divided into two parts. The first, Section 2, provides an empirical framework to the study of distribution dynamics that avoids not only arbitrary discretisation of the state income space but also accounts for spatial dependence (i.e., autocorrelation). Spatial dependence can invalidate the inferential basis of the models, since the assumption of observational independence no longer holds[7]. This is achieved by combining stochastic kernel estimation with Getis' (1990) spatial filtering data view. We will refer to this as a spatial filter view of the continuous state income space.

The second part of the paper, Section 3, applies this framework to analyse income distribution dynamics and cross-region convergence in Europe, looking at evolving distributions of gross regional product per capita across 257 NUTS-2 regions in 27 countries from 1995 to 2003. A number of technical devices such as kernel smoothed densities, Tukey boxplots, cross-profile plots, continuous stochastic kernels and ergodic distributions are utilised – with and without taking a spatial filtering perspective – to identify empirical regularities in the data. The results highlight the importance to properly account for spatial autocorrelation in the data.

[6] These graphical devices highlight the conditioning and are more informative than three-dimensional perspective plots and contour plots that are generally very difficult to interpret in this context because their relationship to the conditional densities is not clear (Hyndman et al. 1996).

[7] Note that Fingleton (1997, 1999) discusses some of the problems that spatial dependence may induce for inferences when working in a discrete state space set up.

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 Section 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 distribution of regional incomes at time t, then the simplest scheme for modelling the dynamics of $\{F_t \mid t \text{ integer}\}$ is a first-order dependence specification of the following type

$$F_{t+1} = M\ F_t \tag{1}$$

where M is a mathematical operator that transforms one distribution at time t into another at time $t+1$, and tracks where points in F_t end up in F_{t+1}. Hence M encodes information on changes in the external shape of the distribution and intra-distribution mobility.

Evidently, Equation (1) is like a first-order autoregression from standard time series analysis, except its values are distributions rather than scalars or finite-dimensioned vectors, and it contains no explicit disturbance. By way of analogy with autoregression, there is no reason why the law of motion of F_t needs to be first-order, or why the underlying transition mechanism needs to be time-invariant. Nevertheless, Equation (1) is generally viewed as a useful step for analysing dynamics in $\{F_t\}$. Iteration yields a predictor for future cross-region distributions

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

Taking this to the limit as $\tau \to \infty$, one can characterise the likely long-run distribution of regional income. 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 is approximated by partitioning the set of possible income values into a finite number of intervals[8]. These intervals then constitute the states of a 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. In a continuous 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 Equation (1) may be viewed as a stochastic kernel or transition function, and 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 $\{(Y_1, Z_1), ..., (Y_n, Z_n)\}$, and the observations by $\{(y_1, z_1), ..., (y_n, z_n)\}$ 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

[8] Note that the arbitrary discretising grid used to construct the Markov chain transition matrix may be seen as a crude non-parametric estimator.

$$f_{t+\tau}(z) = \int_0^\infty g_\tau(z \mid y)\, f_t(y)\,\mathrm{d}y \tag{3}$$

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

$$f_\infty(z) = \int_0^\infty g_\tau(z \mid y)\, f_\infty(y)\,\mathrm{d}y. \tag{4}$$

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

2.2 Stochastic kernel estimation

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 $Z \mid (Y = y)$ is given by

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

The natural estimator[9] of this conditional density function (see Hyndman et al. 1996) is

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

where

[9] Hyndman et al. (1996) derive properties of this density estimator such as the mean square error, bias and variance, and show that this estimator yields a conditional mean function which is equivalent to the Nadaraya-Watson kernel smoother (see Hall et al. 1999).

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

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

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

the kernel estimator of $f_t(y)$ (see Hyndman et al. 1996). This estimator, though obvious, has not yet found much attention for analysing income distribution dynamics[10]. 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. $\|\,.\,\|_y$ and $\|\,.\,\|_z$ are distance metrics on the spaces Y and $Z,$ respectively. In this paper we use the standard euclidean distances, $\|\,.\,\|_y = |\,.\,|_y$ and $\|\,.\,\|_z = |\,.\,|_z$. The kernel conditional density estimator $\hat{g}_\tau\,(z\,|\,y)$ has two desirable characteristics which match those of the density being estimated: first, it is always non-negative, and, second, integrals with respect to z equal one.

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)

$$\int K(x)\,\mathrm{d}x = 1,\;\; \int x\,K(x)\,\mathrm{d}x = 0 \;\;\text{and}\;\; \sigma_K^2 = \int x^2\,K(x)\,\mathrm{d}x < \infty. \tag{9}$$

Popular choices for $K(x)$ are defined in terms of univariate and unimodal probability density functions. In this paper we use the Gaussian kernel[11] given by

[10] Exceptions include Pittau and Zelli (2006), and Basile (2006).

[11] 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.

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

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 problem[12] 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[13], a bootstrap bandwidth selection approach following the approach of Hall et al. (1999) for estimating conditional distribution functions, and a regression-based bandwidth selector[14] (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

[12] 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.

[13] 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.

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

$X_i\ (i=1,\dots,n)$ are independent, we say that there is no spatial structure. Independence implies the absence of spatial autocorrelation[15]. A violation of the independence assumption[16] may result in misguided inferences and interpretations (Rey and Janikas 2005).

This problem has been largely neglected in distribution analysis so far. One way to dealing with this 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 is X 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

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

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

[16] There is a broad agreement in regional science that the process of income dynamics and convergence is inherently endowed with a spatial dimension, and interactions or externalities across regions are likely to be the major sources of the violation of the assumption (see Abreu et al. 2004 for a survey of the existing evidence). It is moreover worth noting that the choice of the NUTS-2 level might give rise to a form of the modifiable areal unit problem (MAUP), well known in geography (see, for example, Getis 2005), that may induce (nuisance) spatial dependence.

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

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

$w_{ij}(\delta)$ denotes the (i, j)th element of a row-standardised binary spatial weight matrix with $w_{ij}(\delta) = 1$ if the distance[17] from region i to region 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 statistic[18] of Getis and Ord (1992) defined as

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

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 $G_i(d)$ 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).

---

[17] In this study distances are measured in terms of geodesic distances between regional centres.

[18] 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.

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 $g_\tau(\tilde{z} \mid \tilde{y})$ function:

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

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 kernel[19] that maps the distribution Y to the spatially filtered distribution $\tilde{Y} \mid Y$ so that

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

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 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 Section 3.1 we describe the data and the observation units. Kernel smoothed densities and Tukey boxplots are used in Section 3.2 to study the shape dynamics of the distribution[20]. Cross-profile plots, continuous stochastic kernels and implied ergodic

[19] Combining stochastic kernel estimation with the conditioning scheme suggested by Quah (1996b, 1997 a) 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.

[20] The distributions are weighted by the relative number of people in each region. One convenient interpretation is that it shows the distributions of individual incomes across people in Europe, assuming that within each

distributions are taken in Section 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)[21] and extracted from the Eurostat Regio database. We use Eurostat's purchasing power standardised per capita GRP to control for national differences in price levels[22].

By Europe we mean the European Union of 27 member states. The data used in this study refer to the time period from 1995 to 2003, the latest year for which GRP figures are available. 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[23]. Although varying considerably in size, NUTS-2 regions are those regions that are adopted by the 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

---

region individual personal incomes are equally distributed, and thus equal to the level of (per capita) income.

[21] 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).

[22] Eurostat does not estimate comparable regional price levels which would enable us to take into account regional differences in price levels within the same country.

[23] Note that nuisance spatial dependence will arise because NUTS-2 regions are formal rather than functional regions. In the case of some city NUTS-2 regions such as Hamburg and Île-de-France regional income tends to be overestimated, while in their surrounding regions underestimated.

regions used by the statistical office of the European Community for the production of regional statistics. Our sample includes 257 NUTS-2 regions[24] covering the 27 member states of the EU:

- *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 twelve 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. 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.

[24] We exclude the Spanish North African territories of Ceuta y Melilla, the Portuguese non-continental territories Azores and Madeira, and the French Départments d'Outre-Mer Guadeloupe, Martinique, French Guayana and Réunion.

Figure 1 to be positioned about here

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 1995, and the dashed line that in 2003. To read this type of figure, note that 1.0 on the horizontal axis indicates the European average of regional income, 2.0 indicates 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 mode[25] is located at about 110 percent of the European average, and the second mode at about 38 percent. The estimated densities reveal several changes over the observation period. The kernel estimated median value decreases by two percent, while the level of dispersion exhibits a small reduction. The kernel estimated standard deviation decreases by 3.3 percent from 0.393 in 1995 to 0.380 in 2003[26].

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. This finding may suggest an improvement in economic conditions of the

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

[26] Conventionally, this would be interpreted as evidence for σ-convergence. But note that this interpretation rests on implicit assumptions about the underlying data generating process and these are not satisfied in the presence of spatial effects.

poorest – generally Central and Eastern European – regions and a slow, gradual process of catching-up.

Figure 2 to be positioned about here

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 percent 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[27] from the rest of the regions. The figure, moreover, makes clear that the interquartile range is decreasing by more than 15 percent, and this falling is due to a decrease of Q_3 rather than Q_1.

[27] These represent Inner London, Brussels, Luxembourg, Hamburg, Île-de-France and Vienna.

The matching counterparts in Figure 1 and Figure 2 use exactly the same data. But they emphasise different empirical regularities. The bimodal shape is striking in Figure 1, but is far from obvious in Figure 2. The spreading out of the upper tail of the distribution is apparent in Figure 2. It appears in form of two smaller bumps in Figure 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 Figure 3 showing cross-profile dynamics[28]. 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.

Figure 3 to be positioned about here

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 Figure 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, Moravskoslezko (57th poorest in 1995) in the Czech Republic, Lüneburg (129th poorest) and Berlin (the 41st richest region) experienced economically

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

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.

Figure 4 to be positioned about here

Figure 4 shows the conditional kernel density estimate $\hat{g}_{\tau}(z \mid y)$ with fixed bandwidths $(h_y = 0.036, h_z = 0.023)$ [29] that describes the stochastic kernel across the 257 regions, averaging over 1995 through 2003. The stochastic kernel has been estimated for a five-year transition period, setting $\tau = 5.$ The figure displays the estimate, using Hyndman's (1996) visualisation tools. Figure 4(i) 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 $t+5$ 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 Figure 4(i), a large portion of the probability mass remains clustered along the main diagonal over the five-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 region boxplot, given in Figure 4(ii), makes this clearer. A highest density region (HDR) is the smallest region of the sample space containing a given probability. Figure 4(ii) shows a plot of the 50 percent and 99 percent high density regions[30], computed from the density estimates shown in Figure 4(i). Each vertical strip represents the conditional

[29] The bandwidths for the estimator were chosen according to Bashtannyk and Hyndman's three-step-strategy. See Section 2.2 for more details.

[30] An HDR boxplot replaces the box bounded by the interquartile range with the 50 percent HDR, the region bounded by the upper and lower adjacent values is replaced by the 99 percent 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.

density for one y value. The darker shaded region in each strip is a 50 percent HDR, and the lighter shaded region is a 99 percent 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 five-year transition horizon.

To read this type of boxplot note that strong persistence is evidenced when the main diagonal crosses the 50 percent 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 percent HDRs. Strong (weak) global convergence towards equality would manifest in 50 percent (99 percent) HDRs crossed by the horizontal line at 1.0. 50 percent HDRs consisting of two disjoint intervals would indicate a two-peaks property of the distribution.

The plot reveals persistence, mobility and polarisation features. Regions with an income range of 0.8 to 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 five-year transition horizon, indicating a very slow process of catching-up. Portions in the income range above 1.2 to 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 percent HDRs at the upper extreme of the income range. We see that regions starting with an income of 2.0 to 2.3 times the European average at time t are unlikely to remain there. Most see their Europe 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 to 2.6 times the average remains either unchanged or shifting away.

Figure 5 to be positioned about here

The evidence of Figure 4 is corroborated by the ergodic density function that is obtained by solving Equation (4). Figure 5 plots the estimated long-run (ergodic) density[31], $\hat{f}_{\infty}(z)$, implied by the estimated $g_{\tau}(z \mid y)$ 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 percent of the European relative per capita income has shifted upward into the 60 to 100 percentage range and shows a tendency to disappear.

Figure 5 shows that the estimate of the long-run distribution has twin peaks although the rich peak[32] is much smaller than the other. This peak accounts for about 97 percent of the regions clustered around the European average income while the rich peak represents a small cluster of relatively rich regions located at about 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[33] which turn out to be – with very few exceptions – regions in Central and Eastern Europe, and simultaneously a tendency towards polarisation – a clustering of the 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[34], obtained for a set of 110 regions covering twelve EU member countries over the time period from 1977 to 1996.

---

[31] 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.

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

[33] This suggests that in the long-run there is no development trap into which the poorest Central and Eastern European regions will be permanently condemned.

[34] It is worth mentioning that in our case the mode of the very rich regions is much smaller and more distant from the other mode.

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 capita) income[35]. This motivates a spatial filtering pass through the data to avoid inferences and interpretations, misguided by the violation of the independence assumption in the previous analysis.

Figure 6 and Figure 7 to be positioned about here

Figure 6 presents the spatially filtered counterpart of Figure 1, and shows that the lower income peak in Figure 1 is well explained by spatial effects. The filtered distributions in this figure are tighter and more concentrated than those in Figure 1. The boxplots in Figure 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 percent. 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 percent. From this, it is clear that the decline in standard deviation observed in Section 3.1 is caused by spatial dependence embedded in

[35] 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.

the income data.

Figure 8 to be positioned about here

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

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, Figure 8(i) 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 8(ii) reinforces this interpretation. The dominant feature in this figure appears to be intra-distribution mobility rather than persistence. Regions with an income less than 1.7 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 percent HDRs crossing the horizontal line at 1.0 and by the 50 percent 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.

Figure 9 to be positioned about here

Figure 9 provides stochastic kernel representations of five-year transition dynamics in the spatially filtered income space, using again a stochastic kernel estimator with fixed bandwidths $(h_{\tilde{y}} = 0.061,\ h_{\tilde{z}} = 0.047).$ This figure is the counterpart to Figure 4 for spatially filtered relative (per capita) regional incomes. Comparing the unfiltered and filtered kernels, one sees that fine details differ, but the global dynamics of the distribution remain roughly unchanged. There are the same polarisation, persistence and mobility features in both, but much more pronounced in the spatially filtered case. This indicates the importance of accounting for spatial dependence effects properly.

Figure 10 to be positioned about here

Additional insights can be gained by the long-run (ergodic) density function, $f_{\infty}(\tilde{z}),$ implied by the estimated $g_{\tau}(\tilde{z} \mid \tilde{y})$ function that can be found as solution to Equation (14). Figure 10 plots the estimate along with the counterpart, $\hat{f}_{\infty}(z),$ for the original (unfiltered) state income space. The solid line shows the estimated $f_{\infty}(\tilde{z}),$ while the dashed line the estimated $f_{\infty}(z).$ The figure highlights peak dynamics already observed in Figure 5 for the original income data and does this now without spatial effects as well. Comparing the estimated density functions, $f_{\infty}(\tilde{z})$ and $f_{\infty}(z),$ one sees some differences in detail that are worth noting. In particular, the peaks have not only become more pronounced, but also closer to each other. Such dynamics suggest economic mechanisms for growth different from more standard ones. The pattern suggests that regions above a certain threshold cluster around a higher income growth path, those below around a lower income growth path.

4 Concluding remarks

In this study, we followed a way to convergence analysis 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 region to the entire distribution as object of interest. The lack of an appropriate inferential theory restricts the work to a descriptive stage. We used product kernel estimators of conditional density to estimate the stochastic kernels. The properties of these estimators are unknown in the presence of spatial autocorrelation of the income series. To avoid misguided inferences and interpretations the paper suggests Getis' spatial filtering approach that is based on the autocorrelation observed with the use of the G_i local statistic and removes the spatial dependence embedded in the income variables.

The paper reveals that spatial effects matter, and the results highlight the importance of these effects in understanding regional income distribution dynamics. A substantial part of the features of the shape and intra-distribution dynamics can actually be attributed to spatial effects embedded in the income variable. The picture reveals emerges seems to give little support to the convergence predictions of the neoclassical model of growth. The results obtained strongly reject the hypothesis of absolute convergence, and suggest instead polarisation and divergence across the entire section, the opposite of catching-up, even though this appears to happen in the lower end of the tail of the distribution. Overtaking and catching-up occur simultaneously with persistence and polarisation. Polarisation manifests itself in the emergence of a twin peak structure in the filtered and unfiltered long-run regional income distributions. Differences exist in detail, and these emphasise the importance to properly deal with the spatial dependence (autocorrelation) problem.

Acknowledgements. This research was supported by the Austrian Science Fund (FWF) grant no. P19025-G11. We thank Lukas Leuprecht and Sascha Sardadvar for early help with the data preparation, and Thomas Seyffertitz for secretarial assistance. 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.

References

Abreu M., Groot H.L.F and Florax R.J.G. (2004): Space and growth: A survey of empirical evidence and methods, Tinbergen Institute, Working Paper No TI04-129/3. Available at SSRN: http://ssrn.com/abstract=631007

Bashtannyk D.M. and Hyndman R.J. (2001): Bandwidth selection for kernel conditional density estimation, *Computational Statistics & Data Analysis* 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

Bianchi M. (1997): Testing for convergence: Evidence from non-parametric multimodality tests. *Journal of Applied Econometrics* 12(4): 393-409

Boldrin M. and Canova F. (2001): Inequality and convergence in Europe's regions: Reconsidering European regional policies. *Economic Policy* 16(32): 205-253

Bulli S. (2001): Distribution dynamics and cross-country convergence: A new approach. *Scottish Journal of Political Economy* 48(2): 226-243

Cliff A.D. and Ord J.K. (1973): *Spatial Autocorrelation*. Pion, London

Durlauf S.N. and Quah D.T. (1999): The new empirics of economic growth. In: J.B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics, Volume 1*. Elsevier, Amsterdam, pp. 235-308

Fan J., Yao Q. and Tong H. (1996): Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika* 83(1): 189-206

Fingleton B. (ed) (2003): *European Regional Growth*. Springer, Berlin, Heidelberg and New York

Fingleton B. (2001): Equilibrium and economic growth: Spatial econometric procedures and simulations. *Journal of Regional Science* 41(1): 117-147

Fingleton B. (1999): Estimates of time to economic convergence: An analysis of regions of the European Union. *International Regional Science Review* 22(1): 5-34

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

Fischer M.M. and Stirböck C. (2006): Pan-European regional income growth and club-convergence. *The Annals of Regional Science* 40(4): 693-721

Friedman M. (1992): Do old fallacies ever die? *Journal of Economic Literature* 30(4): 2129-2132

Getis A. (2005): Spatial pattern analysis. In: K. Kempf-Leonard (ed), *Encyclopedia of Social Measurement*, Vol. 3. Academic Press, Amsterdam, pp. 627-632

Getis A. (1995): Spatial filtering in a regression framework: Examples using data on urban crime, regional inequality, and government expenditures. In: L. Anselin and R. Florax (eds.), *New Directions in Spatial Econometrics*. Springer, Berlin, Heidelberg and New York, pp. 172-188

Getis A. (1990): Screening for spatial dependence in regression analysis. *Papers of the Regional Science Association* 69: 69-81

Getis A. and Griffith D.A. (2002): Comparative spatial filtering in regression analysis. *Geographical Analysis* 34(2), 130-140

Getis A. and Ord J.K. (1992): The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24(3):189-206

Hall R.E. and Jones C.I. (1997): Levels of economic activity across countries. *The American Economic Review* 87(2): 173-177

Hall P., Wolff R.C. and Yao Q. (1999): Methods for estimating a conditional distribution function. *Journal of the American Statistical Association* 94(445): 154-163

Hyndman R.J. (1996): Computing and graphing highest density regions, *The American Statistician* 50(2), 120-126

Hyndman R.J. and Yao Q. (2002): Nonparametric estimation and symmetry tests for conditional density functions, *Journal of Nonparametric Statistics* 14(3), 259-278

Hyndman R.J., Bashtannyk D.M. and Grunwald G.K. (1996): Estimating and visualizing conditional densities, *Journal of Computational and Graphical Statistics* 5(4), 315-336

Islam N. (2003): What have we learnt from the convergence debate? *Journal of Economic Surveys* 17(3): 309-362

Johnson P.A. (2004): A continuous state space approach to “Convergence by parts”. Vassar College Economics Working Paper No. 54, Department of Economics, Vassar College

Jones C.I. (1997): On the evolution of the world income distribution. *The Journal of Economic Perspectives* 11(3): 19-36

LeGallo J. (2004): Space-time analysis of GDP disparities among European regions: A Markov chains approach. *International Regional Science Review* 27(2): 138-163

López-Bazo E., Vaya E., Mora A. J. and Surinach J. (1999): Regional economic dynamics and convergence in the European Union. *Annals of Regional Science* 33(3): 343-370

Magrini S. (2004): Regional (di)convergence. In: J. V. Henderson and J.-F. Thisse (eds), *Handbook of Regional and Urban Economics*. Elsevier, Amsterdam, pp. 2741-2796

Magrini S. (1999): The evolution of income disparities among the regions of the European Union. *Regional Science and Urban Economics* 29(2): 257-281

Ord J.K. and Getis A. (1995): Local spatial autocorrelation statistics: Distributional issues and an application. *Geographical Analysis* 27(4): 286-306

Paap R. and van Dijk H. K. (1998): Distribution and mobility of wealth of nations. *European Economic Review* 42(7): 1269-1293

Quah D. T. (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

Quah D. T. (1997a): *Regional cohesion from local isolated actions: I Historical outcomes*. Centre for Economic Performance, Discussion Paper No. 378. London School of Economics

Quah D. T. (1997b): *Regional cohesion from local isolated actions: II Conditioning*. Centre for Economic Performance, Discussion Paper No. 379. London School of Economics

Quah D. T. (1997c): Empirics for growth and distribution: Stratification, polarization, and convergence clubs. *Journal of Economic Growth* 2(1): 27-59

Quah D. T. (1996a): Regional convergence clusters across Europe. *European Economic Review* 40(3-5): 951-958

Quah D. T. (1996b): Empirics for economic growth and convergence. *European Economic Review* 40(6): 1353-1375

Quah D. T. (1993a): Empirical cross-section dynamics in economic-growth. *European Economic Review* 37(2-3): 426-434

Quah D. (1993b): Galtons fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics* 95(4): 427-443

Reichlin L. (1999): Discussion of ‘Convergence as distribution dynamics’, by D. Quah. In: R. Baldwin, D. Cohen, A. Sapir and A. Venables (eds), *Market, Integration, Regionalism, and the Global Economy*. Cambridge University Press, Cambridge, pp. 328-335

Rey S.J. (2001): Spatial empirics for economic growth and convergence. *Geographical Analysis* 33(3): 195-214

Rey S.J. and Dev B. (2006): σ-convergence in the presence of spatial effects. *Papers in Regional Science* 85(2): 217-234

Rey S.J. and Janikas M.V. (2005): Regional convergence, inequality and space. *Journal of Economic Geography* 5(2): 155-176

Silverman B.W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London and New York

Silverman B.W. (1981): Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society. Series B (Methodological)* 43(1): 97-99

Temple J. (1999): The new growth evidence. *Journal of Economic Literature* 37: 112-156

Wand M.P. and Jones M.C. (1995): *Kernel Smoothing*. Chapman and Hall, London

Appendix: List of regions used in the study

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
Switzerland	CH01	Région lémanique
	CH02	Espace Mittelland
	CH03	Nordwestschweiz
	CH04	Zürich
	CH05	Ostschweiz
	CH06	Zentralschweiz
	CH07	Ticino
Cyprus	CY00	Kypros / Kibris
Czech Republic	CZ01	Praha
	CZ02	Stredni Cechy
	CZ03	Jihozápad
	CZ04	Severozápad
	CZ05	Severovýchod
	CZ06	Jihovýchod
	CZ07	Stredni Morava
	CZ08	Moravskoslezko
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	Rheinhessen-Pfalz
	DEC0	Saarland

ctd.

Country	ID Code	Region
	DED1	Chemnitz
	DED2	Dresden
	DED3	Leipzig
	DEE1	Dessau
	DEE2	Halle
	DEE3	Magdeburg
	DEF0	Schleswig-Holstein
	DEG0	Thüringen
Denmark	DK00	Danmark
Estonia	EE00	Eesti
Spain	ES11	Galicia
	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
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
Greece	GR11	Anatoliki Makedonia, Thraki
	GR12	Kentriki Makedonia
	GR13	Dytiki Makedonia
	GR14	Thessalia
	GR21	Ipeiros
	GR22	Ionia Nisia
	GR23	Dytiki Ellada
	GR24	Sterea Ellada
	GR25	Peloponnisos
	GR30	Attiki
	GR41	Voreio Aigaio
	GR42	Notio Aigaio
	GR43	Kriti
Hungary	HU10	Közép-Magyarország
	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

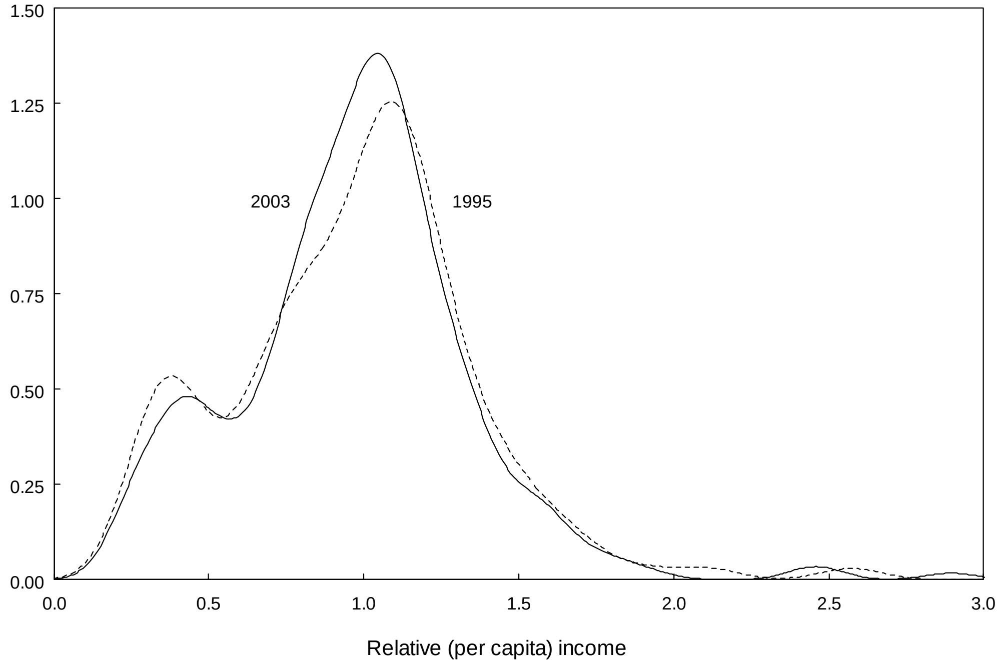
ctd.

Country	ID Code	Region
Iceland	IS00	Iceland
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
	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)
Norway	NO01	Oslo og Akershus
	NO02	Hedmark og Oppland
	NO03	Sør-Østlandet
	NO04	Agder og Rogaland
	NO05	Vestlandet
	NO06	Trøndelag
	NO07	Nord-Norge
Poland	PL11	Lódzkie
	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
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
Sweden	SE01	Stockholm
	SE02	Östra Mellansverige
	SE04	Sydsverige
	SE06	Norra Mellansverige
	SE07	Mellersta Norrland

ctd.

Country	ID Code	Region
	SE08	Övre Norrland
	SE09	Småland med öarna
	SE0A	Västsverige
Slovenia	SI00	Slovenija
Slovakia	SK01	Bratislavský kraj
	SK02	Západné Slovensko
	SK03	Stredné Slovensko
	SK04	Východné Slovensko
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

Figure 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.

Figure 2: Tukey boxplots of relative (per capita) regional income across 257 European regions

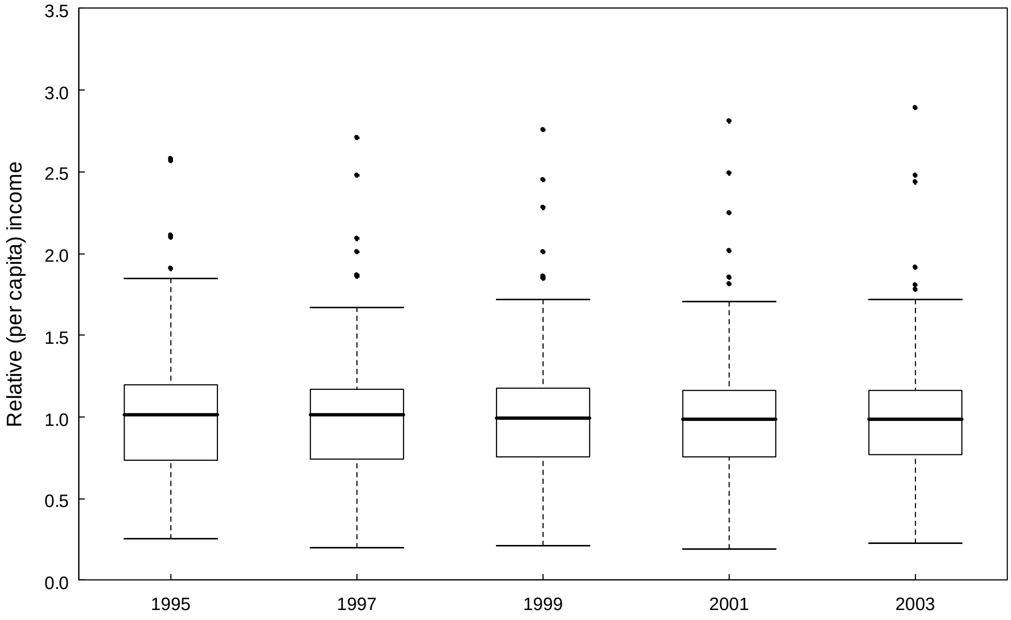


Figure 3: Cross-profile dynamics across 257 European regions, retaining the ranking fixed at the initial year, relative (per capita) income, advancing upwards: 1995, 1999 and 2003 (a guide to region codes can be found in the Appendix)

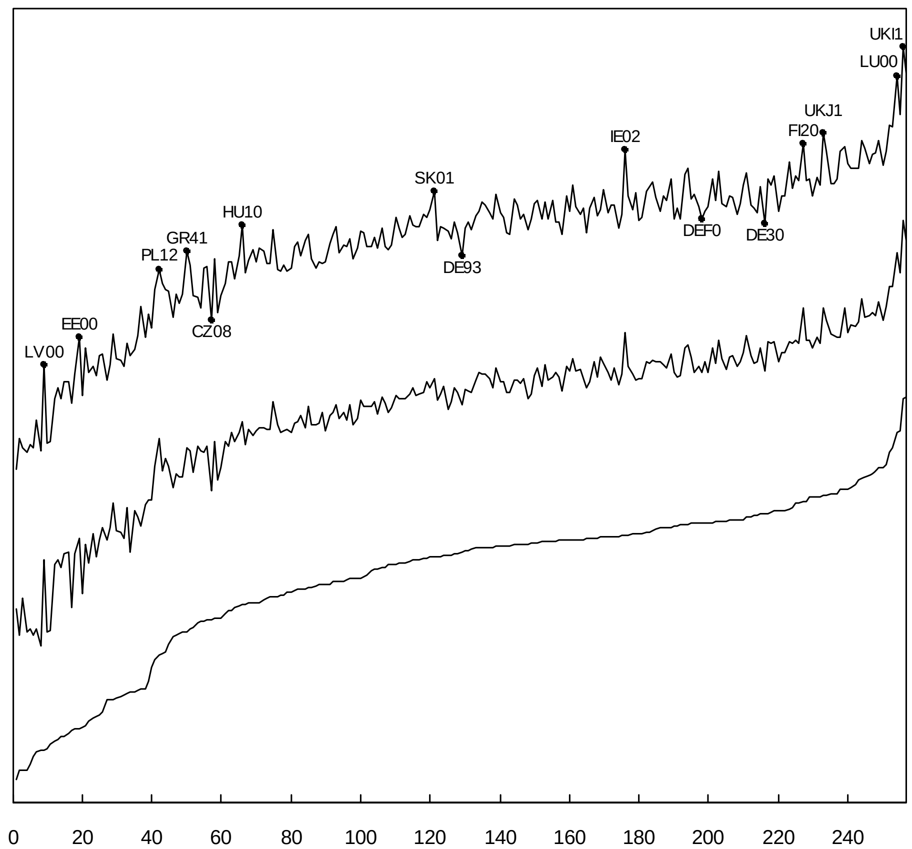
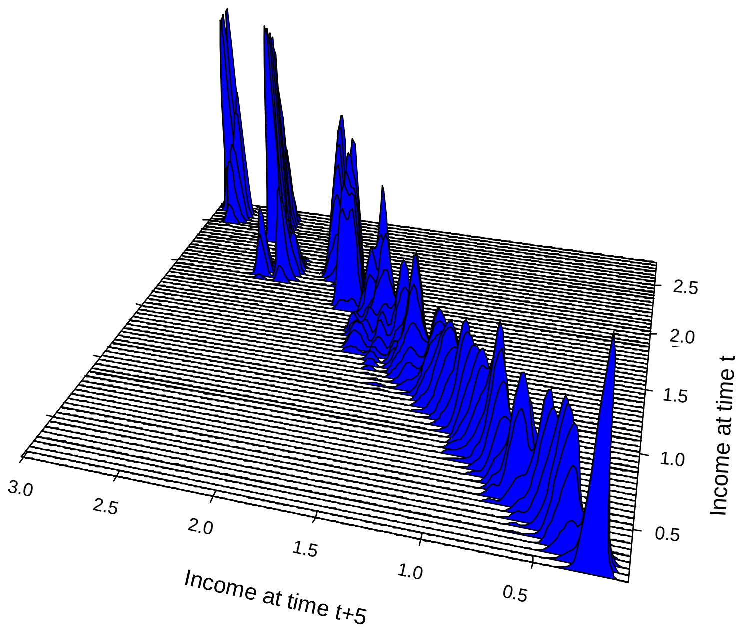
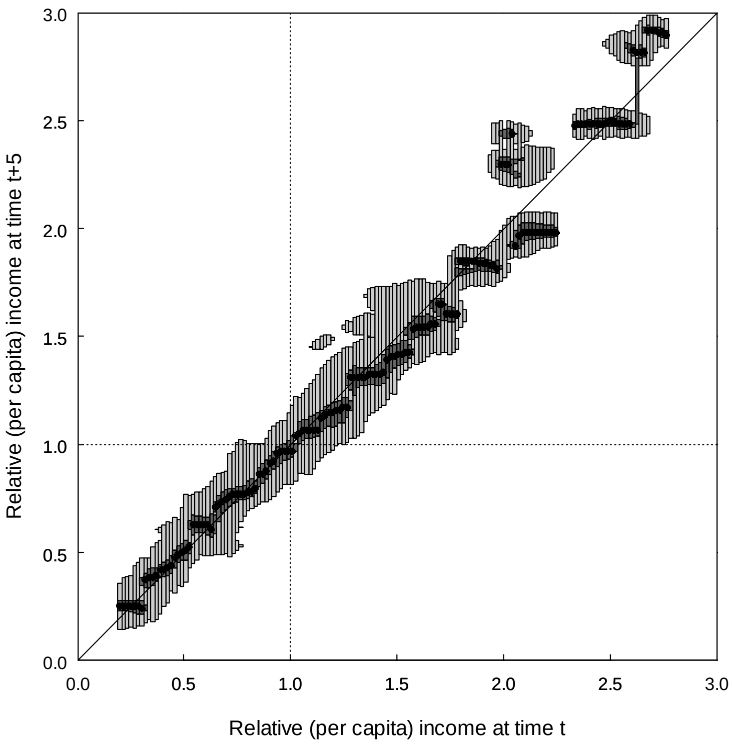


Figure 4: Relative income dynamics across 257 European regions, the estimated $g_5(z|y)$, see Equation (6): (i) the stacked density plot, and (ii) the highest density region boxplot



(i) Stacked density plot

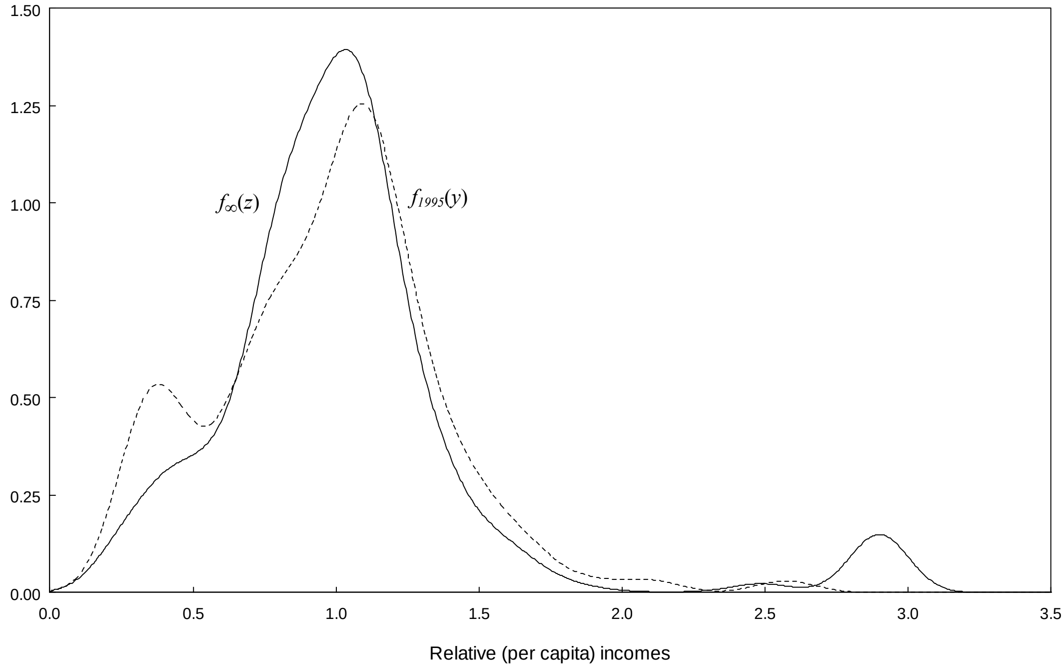


(ii) Highest density region boxplot

Notes: The lighter shaded region in each strip is a 99 percent HDR, and the darker shaded region a 50 percent HDR. The mode for each conditional density is shown as a bullet •.

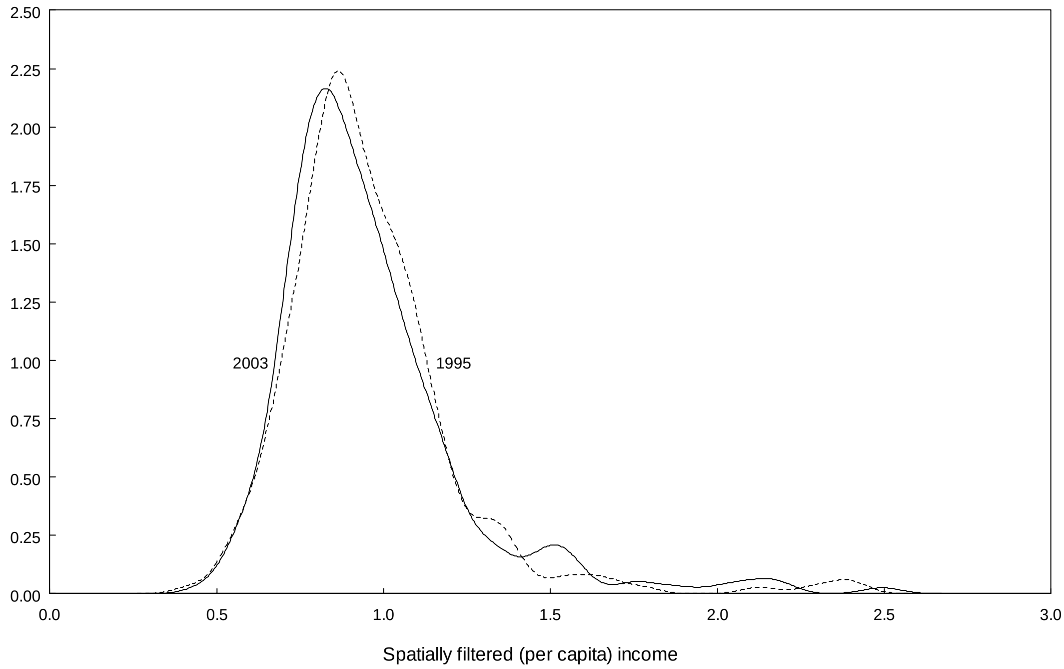
Technical notes: The conditional density $g_\tau\ (z|y)$ is estimated over a five-year transition horizon $\tau = 5$ between 1995-2003. Estimates are based on a Gaussian product kernel density estimator with bandwidth selection ($h_y = 0.036$, $h_z = 0.023$) based on the three-step-strategy suggested 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.

Figure 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 Equation (4).

Figure 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.

Figure 7: Tukey boxplots of relative (per capita) income, across 257 European regions: The spatial filtering view

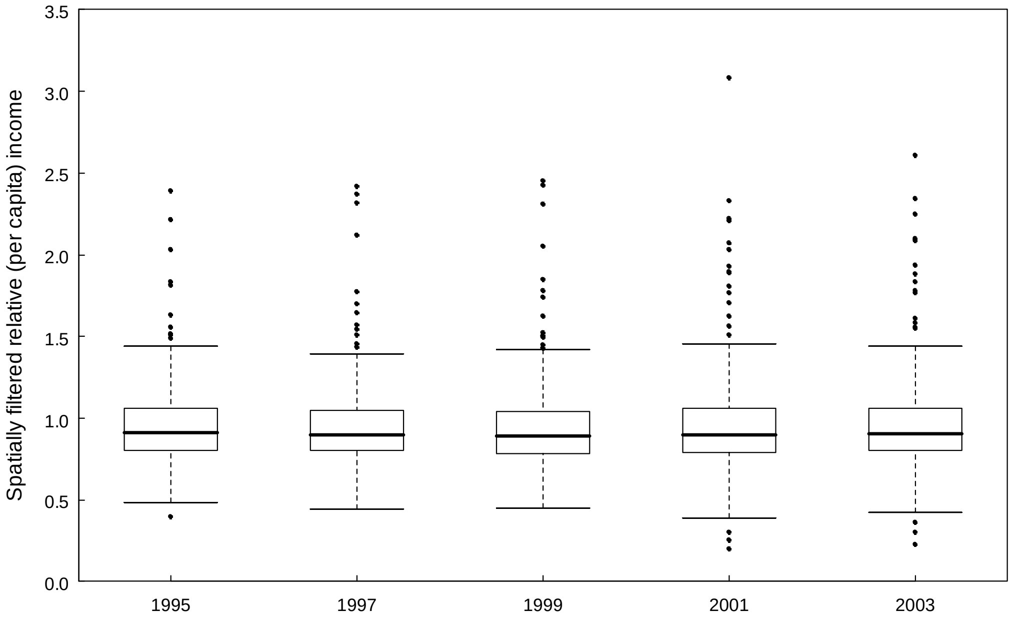
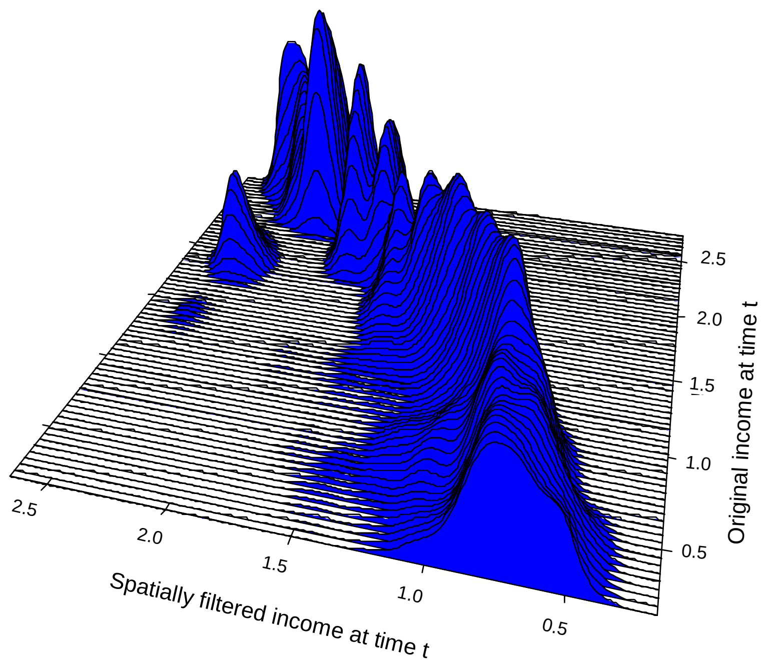
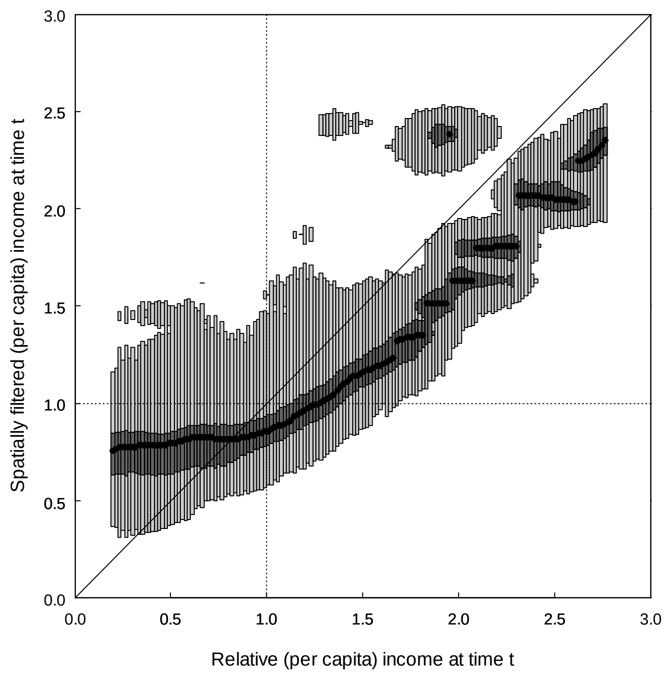


Figure 8: Stochastic kernel mapping from the original to the spatially filtered distribution, the estimated $g\ (\tilde{y}|y)$: (i) the stacked conditional density plot, and (ii) the highest density region plot



(i) Stacked conditional density plot

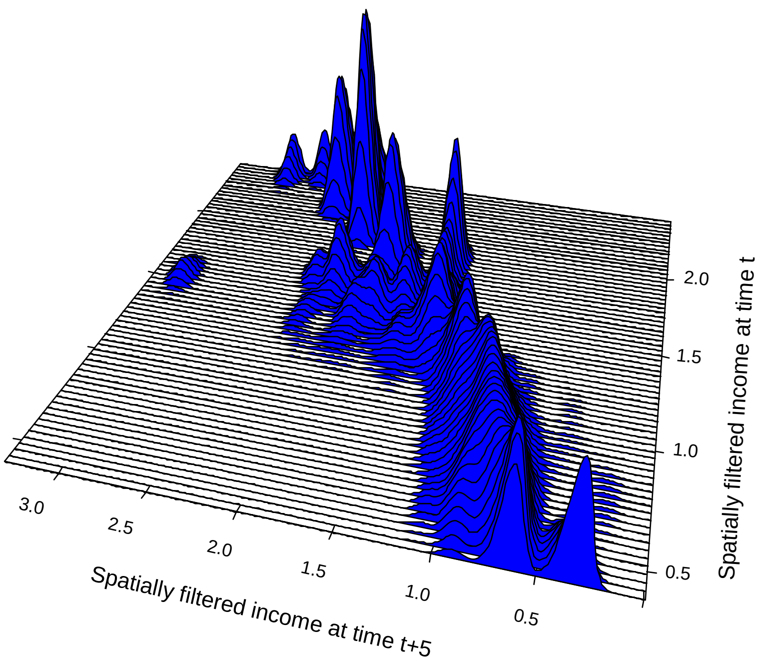


(ii) Highest density region boxplot

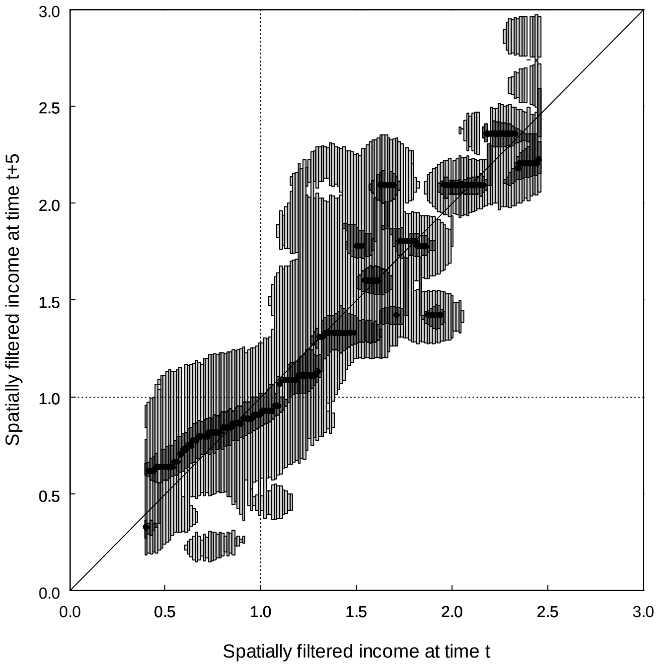
Notes: The lighter shaded region in each strip is a 99 percent HDR, and the darker shaded region a 50 percent HDR. The mode for each conditional density is shown as a bullet •.

Technical notes: The conditional density $g(\tilde{y}|y)$ is estimated over 1995-2003, 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.

Figure 9: The spatial filter view of relative income dynamics: The estimated $g_5(\tilde{z} \mid \tilde{y})$
(i) stacked density plot, and (ii) the highest density region boxplot



(i) Stacked conditional density plot

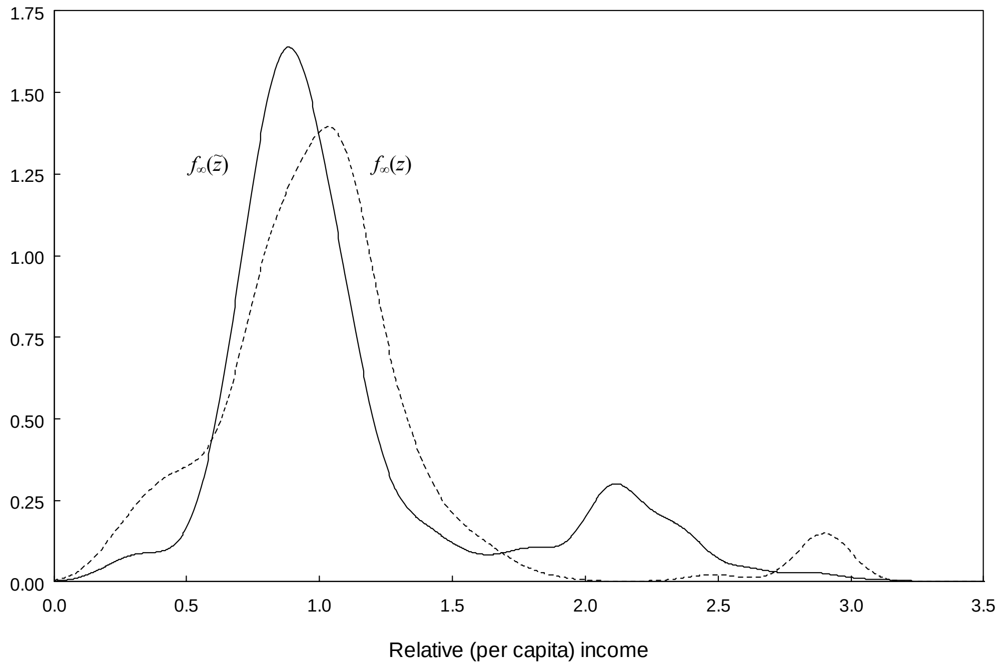


(ii) Highest density region boxplot

Notes: The lighter shaded region in each strip is a 99 percent HDR, and the darker shaded region a 50 percent HDR. The mode for each conditional density is shown as a bullet •.

Technical notes: The conditional density $g_\tau(\tilde{z}|\tilde{y})$ is estimated over a five-year transition horizon τ = 5 between 1995-2003, Estimates are based on a Gaussian product kernel density estimatior with bandwidth selection ($h_{\tilde{y}}$ = 0.034, $h_{\tilde{z}}$ = 0.021) 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 10: Ergodic income distributions, associated with a five-year transition horizon: The point estimates for the spatially filtered version $f_\infty(\tilde{z})$ and the unfiltered version $f_\infty(z)$



Notes: The solid line shows the point estimate for $f_\infty(\tilde{z})$ and the dashed line shows that for $f_\infty(z)$. The plots are densities calculated as solutions to Equation (4) and Equation (14), respectively.