Intra-distribution dynamics of regional per-capita income in Europe: evidence from alternative conditional density estimators

Roberto Basile

ISAE and University of Macerata

r.basile@isae.it,
ISAE (Institute for Studies and Economic Analyses)
Piazza dell'Indipendenza, 4, 00185 – Rome,
Tel. +39-06-44482874; Fax +39-06-44482249.

Abstract: This paper compares different conditional density estimators to analyze the cross-sectional distribution dynamics of regional per-capita incomes in Europe during the period 1980-2002. First, a kernel estimator with fixed bandwidth gives evidence of convergence. With a modified estimator with variable bandwidth and mean-bias correction, the dominant income dynamics is that of persistence and lack of cohesion: only a fraction of very poor regions improves its position over time converging towards a low relative income. An alternative graphical technique (more informative than the traditional contour plot) is also proposed to visualize conditional densities. Finally, a first-order spatial autoregressive model is applied to estimate the effect of spatial dependence on the evolution of income distribution.

1. Introduction

The interest in regional convergence has been growing intensively in the last decade. The most widely accepted method of testing the convergence hypothesis is the regression approach, known as the β-convergence approach. This method has been discussed from different points of view (see Durlauf *et al.*, 2005, for a review of the literature on economic convergence; and Magrini, 2004, for a survey focusing on regional convergence studies). One of the critical points is that this approach tends to concentrate on the behavior of the representative economy. In particular, it sheds light on the transition of this economy towards its own steady state, but provides no insight on the dynamics of the whole cross-sectional distribution of regional per-capita incomes. In fact, a negative association between the growth rates and the initial conditions can be associated with a rising, a declining and a stationary cross-section income dispersion. Clearly, a method that cannot differentiate between convergence, divergence and stationarity is of limited or no use. This failure is essentially a simple intuition of what is termed Galton's fallacy (Quah, 1993).

To overcome this problem, the combination of the β-convergence approach with the analysis of the evolution of the un-weighted cross-sectional standard deviation of the logarithm of per-capita income has been proposed. A reduction over time of this measure of dispersion is referred to as σ-convergence. However, concentrating on the concept of σ-convergence does not represent an effective solution: analyzing the change of cross-sectional dispersion in per-capita income levels does not provide any information on the intra-distribution dynamics. Moreover, a constant standard deviation is consistent with very

different dynamics ranging from criss-crossing and leap-fogging to persistent inequality. Distinguishing between these dynamics is, however, of essential importance.

More recently, moving from this picture, an alternative approach to the analysis of convergence has been suggested in order to solve such a problem. This method, known as the *intra-distribution dynamics* approach (Quah, 1996a, Quah 1996b, Quah 1996c, Quah 1997), examines directly how the whole income distribution changes over time and, thus, appears to be more informative than the convergence empirics developed within the regression paradigm.

The intra-distribution dynamics was generally analyzed through the application of Markov chain methodologies (Quah, 1996b; López-Bazo *et al*., 1999; Fingleton, 1997, 1999; Bulli, 2001) or, more recently, through the estimation of conditional densities using stochastic kernel estimators (Quah, 1997; Lamo, 2000; Pittau and Zelli, 2006; Magrini, 2004). All of the studies that make use of non-parametric stochastic kernel estimators provide contour plots of the conditional density to describe the law of motion of cross-sectional distributions. In this way, they treat the conditional density function as a bivariate density function, while it has been noticed that the conditional density function is a sequence of univariate functions. Furthermore, these studies scantly take account of the recent development in the statistical literature on conditional density estimation (Hyndman *et al*., 1996; Fan *et al.*, 1996; Hall *et al.*, 1999; Hyndman and Yao, 2002), which highlighted the strong bias problems associated with the widely used standard kernel estimator and has proposed new estimators with better statistical properties.

The aim of this paper is to explore alternative conditional density estimators and alternative graphical methods, both developed by Hyndman *et al*. (1996), to describe the

law of motion of cross-regional distributions of per-capita incomes in Europe. In particular, Hyndman *et al*. (1996) notice that the mean function of the kernel density estimator is equivalent to the Nadaraya-Watson kernel smoother. Because of the undesirable bias properties of this smoother, they propose a modified conditional density estimator with a mean equivalent to some other nonparametric regression smoothers that have better statistical properties in terms of mean-bias. This new estimator has smaller integrated mean square error than the standard kernel estimator.

The layout of the paper is the following. In Section 2, we review the most recent literature on the intra-distribution dynamics approach and on conditional density estimators. In Section 3, we report the estimation results obtained applying different estimators to data on per-capita GDP of European regions over the period 1980-2002. In Section 4 we examine the role of spatial dependence in affecting the cross-section distribution dynamics of per-capita GDP. Section 5 concludes and indicates some further possible developments.

2. Intra-distribution dynamics and density estimators

2.1 The transition dynamics approach

As pointed out in the introduction, many problems have been identified with respect to the regression approach to economic convergence and these drawbacks have pushed researchers to explore alternative methods. In particular, Quah (1993, 1996a, 1996b, 1996c, 1997) has suggested an interesting approach to the analysis of economic convergence based on the concept of transition dynamics. In a nutshell, this method consists of studying the dynamics of the entire distribution of the level of per-capita income of a set of economies. We will now review the basic ideas.

As a first step of the methodology, Quah (1993) suggests the development of a probability model describing how a given economy (a region or a country) observed in a given class of the income distribution at time t moves to another class of the income distribution in a subsequent moment of time $t+1$. Let assume the existence of h different income classes and T time periods and define F_t as the time invariant distribution of regional per-capita incomes at time t with ϕ_t the associated probability measure. The dynamics of ϕ_t can be modeled as a first-order auto-regressive process:

$$\phi_{t+1} = \mathbf{M}'\phi_t \qquad (1)$$

The matrix **M** is usually defined as the transition probability of a Markovian process. Each element of **M** describes the probability that an economy belonging to class i at time t will move to class j in the next period. Iterations of (1) yield a predictor for future cross-section distributions

$$\phi_{t+\tau} = \mathbf{M}'^{\tau}\phi_t \qquad (2)$$

since $\mathbf{M}'^{\tau}$ contains information about probability of moving between any two income classes in exactly τ periods of time.

Lòpez-Bazo *et al*. (1999) provide an example of application of the Markov chain approach to the case of European regions. However, even if intuitively appealing, this approach is not free of criticisms. In fact, it is worth noticing that the findings reached through the Markov chain methodology may be sensitive to the criterion used to define the transition probability matrix. Although some procedures have been suggested to determining the optimum number of states and boundaries between them (Magrini, 1999;

Bulli, 2001), usually the researchers decide arbitrarily. One way to solve this problem is to allow the number of cells of the Markov transition probability matrix to tend to infinity (Quah, 1997). If the process describing the evolution of the distribution is again assumed to be time-invariant and first-order Markov, than the relationship between the distribution at time $t+\tau$ and t can be written as

$$\phi_{t+\tau}(y)=\int_0^\infty f_\tau(y \mid x)\phi_t(x)dx \qquad (3)$$

where $f_\tau(y \mid x)$ is the probability density function of y (the per-capita income levels at time $t+\tau$) conditional upon x (the per-capita income levels at time t). In other words, the conditional density $f_\tau(y \mid x)$ describes the probability that a given region moves to a certain state of relative income (richer or poorer) given that it has a certain relative income level in the initial period. In this case convergence must be studied by visualizing and interpreting the shape of the income distribution at time $t+\tau$ over the range of incomes observed at time t.

The long run limit of the distribution of incomes across regions is the limit of (3) as τ tends to infinity. The resulting ergodic distribution is[1]:

$$\phi_\infty(y)=\int_0^\infty f_\tau(y \mid x)\phi_\infty(x)dx \qquad (4)$$

This function describes the long term behavior of the income distribution. Quah (2001) has, however, highlighted the imprecision in the estimates of the ergodic distribution and has

[1] See Johnson PA. 2004. *A continuous state space approach to "convergence by parts".* Department of Economics, Vassar College, Poughkeepsie, NY.

recommended that this distribution should not be read as forecast of what will happen in the future.

2.2 The kernel conditional density estimator

Operationally, the *transition dynamics approach* consists of estimating and visualizing the conditional density of Y given X, where Y is the regional per-capita income at time $t+\tau$ and X the regional per-capita income at time t. Denote the sample by $\{(X_1,Y_1),...,(X_n,Y_n)\}$ and the observations by $\{(x_1,y_1),...,(x_n,y_n)\}$; thus, the aim of the researcher is to estimate the density of Y conditional on $X=x$. Let $g_\tau(x,y)$ be the joint density of (X,Y), $h_\tau(x)$ the marginal density of X and $f_\tau(y\,|\,x)=g_\tau(x,y)/h_\tau(x)$ the conditional density of $Y|(X=x)$. The most obvious estimator of the conditional density is the kernel estimator, firstly proposed by Rosenblatt (1969). Recently, Hyndman *et al.* (1996) have further explored its properties. They define:

$$\hat{f}_\tau(y\,|\,x)=\frac{\hat{g}_\tau(x,y)}{\hat{h}_\tau(x)} \qquad (5)$$

where

$$\hat{g}_\tau(x,y)=\frac{1}{nab}\sum_{i=1}^{n}K\left(\frac{\|x-X_i\|_x}{a}\right)\left(\frac{\|y-Y_i\|_y}{b}\right)$$

is the estimated joint density of (X,Y) and

$$\hat{h}_\tau(x)=\frac{1}{na}\sum_{i=1}^{n}K\left(\frac{\|x-X_i\|_x}{a}\right)$$

is the estimated marginal density.[2]

Equation (4) can also be written as:

$$\hat{f}_{\tau}\left(y \mid x\right)=\frac{1}{b}\sum_{i=1}^{n} w_{i}\left(x\right) K\left(\frac{\left\|y-Y_{i}\right\|_{y}}{b}\right) \qquad (6)$$

where

$$w_{i}\left(x\right)=K\left(\frac{\left\|x-X_{i}\right\|_{x}}{a}\right)\Bigg/\sum_{j=1}^{n} K\left(\frac{\left\|x-X_{j}\right\|_{x}}{a}\right).$$

Equation (6) suggests that the conditional density estimate at $X=x$ can be obtained by summing the n kernel functions in the Y-space, weighted by $\left\{w_{i}\left(x\right)\right\}$ in the X-space. In other words, equation (6) can be interpreted as the Nadaraya-Watson kernel regression (or locally weighted averaging) of $K\left(\frac{\left\|y-Y_{i}\right\|_{y}}{b}\right)$ on X_i (see Hyndman and Yao, 2002). This estimator has two desirable properties: (*i*) it is always non-negative and (*ii*) integrals of the estimators with respect to y equal 1.

The two parameters a and b control the smoothness between conditional densities in the x direction (the smoothing parameter for the regression) and the smoothness of each conditional density in the y direction, respectively.[3] As usual, small bandwidths produce small bias and large variance whereas large bandwidths give large bias and small variance.

---

[2] $\left\|\cdot\right\|_{x}$ and $\left\|\cdot\right\|_{y}$ are Euclidean distance metrics on the spaces of X and Y respectively. $K(.)$ is a symmetric density function, known as the kernel function. Usually, the Epanechnikof kernel is used.

[3] It is worth noting that in the original Rosenblatt's estimator $a=b$.

The optimal bandwidths might be derived by differentiating the integrated mean square error function (IMSE) with respect to a and b and setting the derivatives to zero (Bashtannyk and Hyndman, 2001). However, this requires additional assumptions on the functional forms of both the marginal and the conditional densities. As a rule of thumb, it can be assumed that these densities are Gaussian or of some other parametric form.

The bandwidth a can either be fixed or it can vary as a function of the focal point x. When the data are not homogenously distributed over all the sample space (that is when there are regions of sparse data), a variable (or nearest-neighbor) bandwidth is recommended. In this case, we adjust *a(X)* so that a fixed number of observations m is included in the window. The fraction *m/n* is called the span of the kernel smoother.

2.3 A kernel conditional density estimator with mean-bias correction

Hyndman *et al*. (1996) have observed that the estimation of the conditional mean function obtained from the kernel density estimator (Equation 6) is equivalent to the Nadaraya-Watson kernel regression function:

$$\hat{m}(x) = \int y\hat{f}_\tau(y \mid x)dy = \sum_{i=1}^{n} w_i(x)Y_i \qquad (7).$$

As is well known, the Nadaraya-Watson smoother can present a large bias both on the boundary of the predictor space, due to the asymmetry of the kernel neighbourhood, and in its interior, if the true mean function has substantial curvature or if the design points are very irregularly spaced.

Given the undesirable bias properties of the kernel smoother, Hyndman *et al*. (1996) proposed an alternative conditional density estimator with a mean function equivalent to

that of other nonparametric regression smoothers having better properties than the Nadaraya-Watson approach.

The new class of conditional density estimators can be defined as

$$\hat{f}_{\tau}^{*}(y \mid x)=\frac{1}{b}\sum_{i=1}^{n} w_i(x) K\left(\frac{\left\|y-Y_i^{*}(x)\right\|_{y}}{b}\right) \tag{8}$$

where $Y_i^{*}(x)=e_i+\hat{r}(x)-\hat{l}(x)$, $\hat{r}(x)$ is an estimator of the conditional mean function $r(x)=E[Y \mid X=x]$, $e_i=Y_i-\hat{r}(x_i)$ and $\hat{l}(x)$ is the mean of the estimated conditional density of $e \mid (X=x)$.

Since the error term (e_i) has the same distribution of y_i except for a shift in the conditional mean, one may start by applying the standard kernel density estimator to the points $\{x_i, e_i\}$ and, then, adding the values of $\hat{r}(x)$ to the estimated conditional densities $\hat{f}_{\tau}^{*}(e \mid x)$ in order to obtain an estimate of the conditional density of $Y|(X=x)$. Since $\hat{l}(x)$ - the mean function of $\hat{f}_{\tau}^{*}(e \mid x)$ - is constant under certain conditions (homoskedastic and independent errors), the mean-bias of $\hat{f}_{\tau}^{*}(y \mid x)$ is simply the bias of $\hat{r}(x)$ and the integrated mean square error is reduced.

Obviously, setting $\hat{r}(x)=\hat{m}(x)=\sum_{i=1}^{n} w_i(x) Y_i$ (that is the Nadaraya-Watson smoother) implies that $\hat{f}_{*}(y \mid x)=\hat{f}(y \mid x)$. However, $r(x)$ can also be estimated by using many other

smoothers having better properties than the kernel regression estimator, $\hat{m}(x)$.[4] In other words, using the method developed by Hyndman *et al*. (1996), the mean function of $\hat{f}_{\tau}^{*}(y|x)$ is allowed to be equal to a smoother with better bias properties than the kernel regression. In this way, we obtain an estimate of the conditional density with a mean-bias lower than that of the kernel estimator.

2.4 Local linear conditional density estimators

Recently, alternative solutions to the excessive bias problem of the kernel density estimator have been suggested. For example, Fan *et al.* (1996) have proposed a local linear density estimator. Let

$$R(\beta_0,\beta_1;x,y)=\sum_{i=1}^{n}\left\{K\left(\frac{\|y-Y_i\|_y}{b}\right)-\beta_0-\beta_1(X_i-x)\right\}^2 K\left(\frac{\|x-X_i\|_x}{a}\right) \quad (9),$$

then $\hat{f}(y|x)=\beta_0$ is a local linear estimator, where $\hat{\beta}=(\hat{\beta}_0,\hat{\beta}_1)$ is that value of β which minimizes $R(\beta_0,\beta_1;x,y)$. When $\hat{\beta}_1$=0, this estimator is identical to (6). While this estimator has smaller bias than the Nadaraya-Watson estimator, it is not restricted to be non-negative. In order to solve this problem, Hyndman and Yao (2002) proposed an alternative estimation method, called the *local parametric estimator*, which is similar to the

[4] Using $\hat{r}(x)$ we often introduce an extra smoothing parameter, *c*. Notice that both *c* and *a* control smoothness in the *x* direction; *a* controls how quickly the conditional densities can change in shape and spread while *c* controls the smoothness of the mean of the conditional densities over *x*.

local logistic estimator proposed by Hall *et al*. (1999) and is a conditional version of the density estimator proposed by Loader (1996). It is simply defined as:

$$R_1\left(\beta_0,\beta_1;x,y\right)=\sum_{i=1}^{n}\left\{K\left(\frac{\left\|y-Y_i\right\|_y}{b}\right)-\exp\left(\beta_0-\beta_1\left(X_i-x\right)\right)\right\}^2 K\left(\frac{\left\|x-X_i\right\|_x}{a}\right)$$

(10).

This is equivalent to using local likelihood estimation for the regression of $K\left(\frac{\left\|y-Y_i\right\|_y}{b}\right)$ on X_i. This local linear density estimator can also be combined with the mean-bias-correction method described in section 2.3 in order to force the density function to have a mean equal to any pre-specified smoother. We will exploit this opportunity in the empirical analysis presented below.

3. Some evidence on regional convergence in Europe

3.1 Data, scatterplot smoothing and empirical strategy

We analyze the intra-distribution dynamics of regional per-capita incomes in Europe over the period 1980-2002. Per-capita income levels are normalized with respect to the EU average in order to remove co-movements due to the European wide business cycle and trends in the average values. The income variable is the total gross value-added (GVA) calculated according to the European System of integrated Accounts (ESA95). The total GVA figures are at constant prices 1995 and are converted to Purchasing Power Standards (PPS). However, only national PPS have been applied, since Eurostat does not possess comparable regional price levels that would enable us to take into account regional

differences in price levels. The number of NUTS2 regions included in the sample is 189 (see Appendix 1). Data are drawn from the Cambridge Econometrics Dataset.[5]

In order to estimate conditional density functions $f_{\tau}(y \mid x)$, evaluation at a large number of points is frequently required. For this reason, we fix $\tau = 15$ and exploit the panel structure of the dataset. Thus, Y and X are vectors of 1,512 observations (189 regions $\times$ 8 periods).

Figure 1 shows the scatterplot of relative per-capita income levels at time t and $t+15$. We can clearly observe three things: (1) data are distributed around the main diagonal, indicating a high degree of immobility; (2) at the extreme of the sample space data are sparser; (3) a few extreme observations appear on the right side of the scatterplot. These six points (included within a circle) refer to Groningen, a region often excluded from convergence analyses, since it always appears as an outlier.[6] However, in spirit of the distribution dynamics approach described by Quah (1997, p. 34), we did not exclude

[5] In alternative to the NUTS regions, some authors have used Functional Urban Regions (FURs) as units of analysis (Magrini, 1999) in order to take into account the spatial sphere of socio-economic influence of any basic unit. However, the main data sources (Eurostat and Cambridge Econometrics) only provide data at NUTS level.

[6] Groningen seems to have worsened its relative economic position in the second half of the eighties. However, the evolution of gas prices and changes in the way in which GDP in the energy sector was distributed between regions are well-kwon reasons for this feature. Thus, Groningen could not be considered as an economic outlier in strict sense and might be excluded from the analysis. However, in the present paper we decided to keep this region within the sample in order to show the potential effects of outliers on the estimate of conditional densities.

regions from the dataset just because they have "*performed extraordinary well or extraordinary poorly relative to the bulk of other macroeconomies*". They represent real people and real regions not just observations that might be useful to delete in statistical analysis. Rather, a researcher must endeavour to find estimation methods robust against outliers.

- Figure 1 about here -

In figure 1 we also superimpose the estimated fit of three different scatterplot smoothers: (a) the Nadaraya-Watson estimator ('dotted' curve) with a Gaussian kernel and a fixed bandwidth h=0.09; (b) the local linear regression smoother ('long-dashed curve') with a variable bandwidth (also known as the k-nearest-neighbor local linear smoother); and (c) the *lowess* ('solid' curve).[7] All bandwidth parameters have been selected by using the generalized cross validation method. In the cases (b) and (c) the span that defines the size of the neighborhood in terms of a proportion of the sample size is equal to 0.2 (the width of the smoothing windows always contain the 20% of the data).

As expected, the Nadaraya-Watson (or local averaging) smoother appears more sensitive than the other two smoothers to extreme observations (Groningen) and to the data sparseness at the boundary. Moreover, a difference between the local linear regression with

[7] The *lowess* can be interpreted as a tri-cube kernel scatterplot smoother, able to capture local fluctuations in the density function of the independent variable (Cleveland, 1979; Cleveland and Devlin, 1988). The combination of three features - nearest neighbours, smoothed weight function (the tricube kernel) and local expected value formed via locally weighted regressions - helps the *lowess* regression outperform many other scatterplot smoothers. In particular, a local linear smoother is, per se, not robust against outliers. Only, the *lowess* is very robust against 'far out' observations, since it down-weights large residuals.

variable bandwidth and the *lowess* emerges only at the extreme right side of the sample space, confirming that only the *lowess* is resistant against isolated points.

In the rest of this section we report the results of different conditional density estimators. First, we estimate $\hat{f}_{15}(y \mid x)$ using a kernel estimator with a constant bandwidth parameter a (equation 6). In this first step we compare two alternative graphical techniques for visualizing the conditional density estimators: the traditional perspective and contour plots, on the one side, and the new 'stacked' and '*HDR*' plots (described in section 3.2), on the other. Then, we estimate a conditional density using four alternative methods: (*i*) a kernel density estimator with variable bandwidth; (*ii*) a kernel density estimator with variable bandwidth and mean bias correction (equation 8); (*iii*) a local linear density estimator with variable bandwidth (equation 10); (*i*) a local linear density estimator with variable bandwidth and mean bias correction.[8]

3.2 New graphical methods for visualizing intra-distribution dynamics

All of the studies on intra-distribution dynamics which make use of nonparametric stochastic kernel density estimators provide three-dimensional perspective plots and/or the corresponding contour plots of the conditional density to describe the law of motion of cross-sectional distributions. In such a way, they treat the conditional density as a bivariate density function, while the latter must be interpreted as a sequence of univariate densities of relative per-capita income levels conditional on certain initial levels.

[8] All estimations were performed using the *R* software. In particular, we used the code *hdrcde* developed by Robert Hyndman and the code *locfit*.

Here we use new graphical methods for visualizing conditional density estimators developed by Hyndman *et al*. (1996) and Hyndman (1996). The first graphical technique, called the "*stacked conditional density plot*" (figures 3A), displays a number of conditional densities plotted side by side in a perspective plot.[9] It facilitates viewing the changes in the shape of the distributions of the variable observed at time $t+\tau$ over the range of the same variable observed at time t. In other words, like a row of a transition matrix, each univariate density plot describes transitions over 15 years from a given income value in period t. Hyndman *et al.* (1996) note that this plot is "*much more informative than the traditional displays of three dimensional functions since it highlights the conditioning*" (p.13).

The second type of plot proposed by Hyndman *et al*. (1996) is the "*highest conditional density region*" (*HDR*) plot (figures 3B-10). Each vertical band represents the projection on the xy plan of the conditional density of y on x. In each band the 25% (the darker-shaded region), 50%, 75% and 90% (the lighter-shaded region) *HDR*s are reported. A high density region is the smallest region of the sample space containing a given probability. These regions allow a visual summary of the characteristics of a probability distribution function. In the case of unimodal distributions, the *HDR*s are exactly the usual probabilities around the mean value; however, in the case of multimodal distributions, the *HDR* displays different disjointed subregions.

[9] Each univariate density plot is always non-negative and integrates to unity. Since the conditional density plot has been evaluated on an equispaced grid of 100 values over the range of x and y directions, figure 3A displays 100 stacked univariate densities.

The *HDR* plot is particularly important to analyze intra-distribution dynamics. If the 45-degree diagonal crosses the 25% or the 50% *HDR*s, it means that most of the elements in the distribution remain where they started (there is *'strong' persistence*); if it crosses only the 75% or the 90% *HDR*s, we can conclude in terms of *'weak' persistence*. If the horizontal line traced at the zero-value of the period t+15 axis crosses *all* the 25-50% (75-90%) *HDR*s, we can say that there is *'strong' ('weak') global convergence* towards equality. Finally, if *some* 25-50% (75-90%) *HDR*s are crossed by a horizontal line traced at any value of the *t+15* axis, we can say that there is *'strong' ('weak') local or 'club convergence'*.[10] Clearly, this method is particularly informative for the analysis of regional growth behavior, since it highlights the dynamics of the entire cross-section distribution. It remains important to analyze any other moment of the distribution (such as the mean and the variance) and any other central point. In particular, one may wish to analyze the modes, the values of y where the density function takes on its maximum values. In fact, especially when the distribution function is bimodal, the mean and the median are not very useful, since they will provide only a 'compromise' value between the two peaks. Thus, the modes may be considered as a form of robust nonparametric regression. In each figure, the highest modes for each conditional density estimate are superimposed on the *HDR* plots and shown as a bullet.

3.3 Empirical evidence

[10] The 'club convergence hypothesis' states that regions catch up with one another but only within particular subgroups.

Figure 2 shows traditional perspective and contour plots for the conditional kernel density estimate with fixed bandwidth, describing 15-year horizon evolutions of the distribution of per-capita income relative to the European average. As well-known, the selection of the bandwidth parameter is a crucial issue in the estimation of densities. Optimal bandwidths have been firstly selected using the method developed by Hyndman and Yao (2002) based on a combination of asymptotic properties of a local polynomial approximation of the conditional density and Silverman's normal reference rule. These optimal bandwidths, however, give evidence of under-smoothing, while multiplying the optimal bandwidths by 3 provides a better smoothing. Therefore, the final bandwidth a for the x direction is 0.149, while the final bandwidth b for the y direction is 0.091. This figure would suggest that over the period considered European regions have followed a convergence path.[11] In fact, using the standard terminology, we observe a clockwise shift in mass indicating some degree of intra-distribution mobility, which would imply that the richer regions became poorer and the poorer became richer. These findings appear consistent with those reported in previous work[12]. Moreover, as it is common in these kinds of analyses, a 'multiple-peaks' property manifests. In fact, we can observe some distinct local maxima (or 'basins of attraction'). Contour plot makes this clearer.

- Figure 2 about here -

[11] Using higher bandwidth than a = 0.149 and b = 0.091, the evidence of convergence is magnified. It is important to stress that the results of the intra-distribution dynamics approach based on the standard kernel density estimator strongly depend on the bandwidth parameters chosen.

[12] See, for example, Brasili C, Gutierrez L. 2004. Regional convergence across European Union. Development and Comp Systems 0402002, Economics Working Paper Archive EconWPA

The same estimation results discussed above are visualized in figure 3 using the alternative stacked density plot and the *HDR* plot. From this figure, we would learn that regions that at the beginning of the period had a per-capita income level lower (higher) than the EU average would be more likely to improve (worsen) their relative position over the next 15 years: the 25% *HDR*s associated with relative per-capita income levels at time *t* lower (higher) than 1.0 (that is the European average) are all above (below) the main diagonal. Again, this means that the poorer economies would be catching up with the richer ones. The *HDR* plot allows to identify (better than the standard contour plot) the presence of different 'convergence clubs'. The position of the highest modes and of the 25% *HDR*s would suggest local convergence at relative income levels of 0.7, 1.3, 1.8 and 2.2. Moreover, signs of bimodality would appear for very high levels of the distribution at time *t*: regions that at the beginning of the period had a very high income level would have experienced over time either a slowdown or a persistent behavior.

- Figure 3 about here -

The ergodic distribution of the standard kernel density is plotted in figure 4 along with the marginal density of relative per-capita incomes at time *t*. Both the initial and the stationary distributions display a picture in which one peak, just above the European average. The peak of the ergodic distribution, however, is higher than that of the initial density suggesting that some convergence is achieved in the long run.[13]

---

[13] The univariate density of relative per-capita income level at time *t* has been estimated using a Gaussian kernel density estimator with bandwidth parameter of 0.039 chosen according to Sheather and Jones (1991) procedure.

- Figure 4 about here -

However, looking more carefully at figure 3, we may observe that the plotted conditional density function does not fit the scattered points very well. In particular, we suspect that the sparseness of data at the boundaries and the presence of extreme points (Groningen) might have affected the entire estimated conditional density function, as well as they have affected the conditional mean function. Thus, alternative estimation methods are needed. First, we try with a kernel density estimator with a variable bandwidth to accommodate the problem of data sparseness (figure 5). The choice of a variable bandwidth substantially modifies the form of the conditional density function. In particular, the evidence of mobility (and of convergence) is now confined to the upper and lower tails of the distribution at time *t*, while regions that at the beginning of the period had a relative per-capita income between 0.7 and 1.5 did not change their relative position over time. The evidence of bimodality associated with very high initial income levels is still there. The S-shaped form of the modal regression function appears to fit the data better than in figure 3. However, we cannot ignore the role of the outlier in affecting the shape of the distribution, yet. An estimator robust against outliers is definitely needed.

- Figure 5 about here -

Thus, figure 6 reports the results based on the modified conditional kernel density estimator with mean function specified by a *lowess* smoother. As it can be observed, after a certain threshold (about 0.6 times the European average), the 45-degree diagonal crosses the 25% and 50% *HDR*s and the modal regression follows a straight line. This reveals a high degree of immobility or persistence: European regions tended to maintain their relative

positions over the study period. However, there is still some evidence of mobility at the left side of the sample space: the 25% *HDR*s and the relative modes lie above the main diagonal for values of regional income lower than the threshold. This means that very poor regions registered higher growth rates than the other regions between 1980 and 2002. Moreover, this group of regions seems to converge towards a common level of relative per-capita income of about 0.6 times the overall mean, in line with the club convergence hypothesis. The convergence within this poorer group is shown by the slope of the modal regression which is almost parallel to the horizontal axis. These results are more in line with those presented in Pittau and Zelli (2006) for the sample of NUTS2 regions belonging to the first twelve European Union countries.[14] The unimodal ergodic distribution obtained from the mean-bias corrected kernel density estimator is reported in figure 4. The peak of this distribution is slightly lower than that of the ergodic distribution obtained from the standard kernel density estimator, suggesting a higher degree of persistence or a lower degree of convergence in the long run.

- Figure 6 about here -

Finally, figures 7 and 8 provide the results of conditional local linear density estimators. However, these two figures do not add any information to the picture drawn above. Indeed, the results of the local linear density estimator with variable bandwidth (figure 7) are very

[14] It is worth noticing that, by using the mean-bias correction approach, we have also found a lower sensitivity of the estimates from the choice of the bandwidths. The optimal bandwidth parameters in this case have been multiplied by 2, but the results obtained using the original optimal bandwidths and by multiplying the optimal bandwidths by 5 are very similar.

similar to those of the kernel density estimator with variable bandwidth (figure 5), while the results of the local linear density estimator with variable bandwidth and mean bias correction (figure 8) are very similar to those of the kernel density estimator with variable bandwidth and mean bias correction (figure 6). Therefore, we consider figure 6 as the definitive picture drawn to represent the intra-distribution dynamics of regional per-capita incomes in Europe over the period 1980-2002.[15]

- Figures 7 and 8 about here -

4. Spatial conditioning

The picture of immobility drawn in Figure 6 is an instance of what Quah (1997, p.44) calls "*unconditional dynamics"*. This author also proposes a method to "*explain*" distribution dynamics, which is very different from "*discovering a particular coefficient to be significant in a regression of a dependent variable on some right-hand side variables*". This method called "*conditioning*" is based on "*an empirical computation that helps us understand the law of motion in an entire distribution*". The idea is to analyze income disparities after conditioning out the effect of some variables. In this last section of the paper, we explore the role of spatial dependence in explaining the evidence of persistence in the intra-distribution dynamics of regional per-capita income in Europe.

---

[15] In theory, the local linear density estimator should solve the mean-bias problem affecting the kernel density estimator. Thus, we asked Robert Hyndman whether these results are reasonable. He answered that the results reported in this paper are consistent with his experience, according to which the kernel density estimator with mean-bias correction gives more reliable findings than the local linear density.

The conditioning scheme adopted here is articulated in two steps, as in Quah (1997).[16] Firstly, a spatially filtered variable of regional per-capita incomes, $\tilde{y}$, is constructed by estimating a first-order spatial autoregressive model.[17] The filtered variable can be interpreted as that part of income of each region which is not explained by the spillover effects from the contiguous regions. Then, the conditional density function $f\left(\tilde{y}|x\right)$ is estimated. The idea is that if inter-regional spillovers play a key role in the regional growth process, the evidence of persistence disappears and some convergence emerges. Conversely, if the spatial contiguity is not influent, the conditional distribution of the transformed variable maintains its original characteristics.

Figure 9 shows the results obtained using a kernel density estimator with fixed bandwidth, while figure 10 reports the results obtained using a kernel density estimator with variable bandwidth and mean bias correction. Again, we can observe how bad the first

---

[16] It is far from the aim of this paper to analyse the effect of all potential factors conditioning the intra-distribution dynamics of regional per-capita incomes in Europe. Recently, some papers have proposed various techniques, based on a first-step growth regression equation, to analyse the influence of different variables jointly (see, for, example, Lamo, 2000).

[17] The filtered variable is the residual from the spatial autoregressive model $\ln y = \rho W \ln y + \varepsilon$, where $\ln y$ is the log of relative per-capita income, $W \ln y$ is its spatial lag with W being a 5-nearest neighbour weights matrix, and ρ the spatial autoregressive parameter. This model has been estimated for each year using the maximum likelihood procedure implemented in *matlab* by LeSage. The estimated $\hat{\rho}$ parameters range from 0.71 to 0.57. This method is different from that proposed by Quah (1997), which consists of calculating the ratio between the income level of the region and its spatial lag. This other method implies assuming a $\hat{\rho}$ parameter equal to one.

estimator fits the data. Comparing the more reliable figure 10 to figure 6, we can clearly identify some important changes in the conditional distribution of relative per-capita incomes. First, we observe that for initial values lower than the European average, the 75% *HDR*s are now closer to the horizontal line (the poor-regions' club is closer to the EU average). Thus, we can say that, without spillover effects, the probability for a poor region to migrate from a lower to a higher income class and to converge towards the average value would have increased: spatial dependence had a negative effect on regional convergence in Europe. Second, for initial values higher than the European average, the 25% *HDR*s are below the main diagonal, suggesting a lower persistence of these regions in the original income classes. The evidence of a higher degree of convergence in the case of spatially filtered data is also corroborated by the ergodic density function (figure 4).

- Figures 9 and 10 about here -

5. Conclusions

Different approaches have been used in the literature to analyze the process of regional income convergence. However, the intra-distribution dynamics approach, proposed by Quah (1997), is without any doubt one of the most reliable methods, since it examines directly how the whole income distribution changes over time. In particular, this methodology is much more informative than the regression approach that concentrates on the behavior of the representative economy (Magrini, 2004). All of the most recent studies on intra-distribution dynamics use the kernel density estimator to describe the law of motion of cross-sectional distributions of per-capita incomes. In particular, the empirical applications of the kernel stochastic approach to the case of European regions report

evidence of some degree of convergence: some mobility in the regional distribution of relative per-capita income occurs, in the sense that poor regions become richer and rich regions grow less rapidly. Other research has proposed the emergence of two distinct clubs of convergence: some rich regions are converging to a higher mean level of income, and some poor regions are also converging but to a lower level of income.

However, the kernel stochastic approach widely used in the literature to analyze the distribution of y (the per-capita income at time $t+\tau$) conditional on x (the per-capita income at time t) can be criticized from two different points of view. First, the kernel density estimator is usually implemented applying a *constant* bandwidth parameter in the x and y directions. These estimators have some undesirable bias properties that can affect the analysis of intra-distribution dynamics and, thus, may provide misleading evidence on the real convergence process. Secondly, the traditional method of visualizing the output of conditional density estimation is not adequate, since it actually displays the joint distribution.

In order to describe the law of motion of cross-sectional distributions of regional per-capita incomes in Europe during the period 1980-2002, in this paper we use an alternative kernel density estimator with two bandwidth parameters a and b (which control the smoothness between conditional densities in the x direction and the smoothness of each conditional density in the y direction, respectively) and an alternative graphical technique (the *Highest Density Regions* plot) for visualizing conditional density estimators. In particular, we use a kernel density estimator with variable bandwidth a and mean bias correction. This estimator, developed by Hyndman *et al.* (1996), has better properties than

the kernel density estimator with a constant bandwidth parameter generally used in the literature on intra-distribution dynamics.

Applying the new method to European data, we obtain interesting evidence that enriches the debate on the distribution dynamics. In particular, we obtain evidence of persistency: over the period 1980-2002 most of the regions appear to remain where they were at the beginning. Only a fraction of very poor regions improves its position over the time period converging towards a very low relative income level ('club convergence').

Finally, we have investigated the role of spatial dependence in affecting the observed pattern of regional growth, by combining the new methodology proposed here with standard spatial econometrics techniques. The results suggest that spatial dependence had a negative effect on regional convergence in Europe over the period 1980-2002: after conditioning out the effect of spatial dependence, there is still persistence, but the poor-regions' club is closer to the European average. In future work, we will take into account other determinants of growth following some recent contributions (Lamo, 2000). This analysis might be helpful in producing suggestions for a set of regional policies intended to reduce disparities.

Appendix 1: sample of NUTS2 regions

AT00	**Austria**	DEA2	Köln	FR51	Pays de la Loire	ITF6	Calabria	**UK00**	**United Kingdom**
AT11	Burgenland	DEA3	Münster	FR52	Bretagne	ITG1	Sicilia	UKC1	Tees Valley and Durham
AT12	Niederösterreich	DEA4	Detmold	FR53	Poitou-Charentes	ITG2	Sardegna	UKC2	Northumberland et al.
AT13	Wien	DEA5	Arnsberg	FR61	Aquitaine	**LU00**	**LUXEMBOURG**	UKD1	Cumbria
AT21	Kärnten	DEB1	Koblenz	FR62	Midi-Pyrénées	**NL00**	**Netherlands**	UKD2	Cheshire
AT22	Steiermark	DEB2	Trier	FR63	Limousin	NL11	Groningen	UKD3	Greater Manchester
AT31	Oberösterreich	DEB3	Rheinhessen-Pfalz	FR71	Rhône-Alpes	NL12	Friesland	UKD4	Lancashire
AT32	Salzburg	DEC0	Saarland	FR72	Auvergne	NL13	Drenthe	UKD5	Merseyside
AT33	Tirol	DEF0	Schleswig-Holstein	FR81	Languedoc-Roussillon	NL21	Overijssel	UKE1	East Riding et al.
AT34	Vorarlberg	**DK00**	**Denmanrk**	FR82	Prov.-Alpes-Côte d'Azur	NL22	Gelderland	UKE2	North Yorkshire
BE00	**Belgium**	**ES00**	**Spain**	FR83	Corse	NL31	Utrecht	UKE3	South Yorkshire
BE10	Bruxelles-Brussels	ES11	Galicia	**GR00**	**Greece**	NL32	Noord-Holland	UKE4	West Yorkshire
BE21	Antwerpen	ES12	Principado de Asturias	GR11	Anatoliki Makedonia	NL33	Zuid-Holland	UKF1	Derbyshire et al.
BE22	Limburg	ES13	Cantabria	GR12	Kentriki Makedonia	NL34	Zeeland	UKF2	Leicestershire et al.
BE23	Oost-Vlaanderen	ES21	Pais Vasco	GR13	Dytiki Makedonia	NL41	Noord-Brabant	UKF3	Lincolnshire
BE24	Vlaams Brabant	ES22	Navarra	GR14	Thessalia	NL42	Limburg	UKG1	Herefordshire et al.
BE25	West-Vlaanderen	ES23	La Rioja	GR21	Ipeiros	**PT00**	**Portugal**	UKG2	Shropshire et al.
BE31	Brabant Wallon	ES24	Aragón	GR22	Ionia Nisia	PT11	Norte	UKG3	West Midlands
BE32	Hainaut	ES30	Comunidad de Madrid	GR23	Dytiki Ellada	PT15	Algarve	UKH1	East Anglia
BE33	Liège	ES41	Castilla y León	GR24	Sterea Ellada	PT16	Centro	UKH2	Bedfordshire, Hertfordshire
BE34	Luxembourg	ES42	Castilla-la Mancha	GR25	Peloponnisos	PT17	Lisboa	UKH3	Essex
BE35	Namur	ES43	Extremadura	GR30	Attiki	PT18	Alentejo	UKI1	Inner London
DE00	**Germany**	ES51	Cataluña	GR41	Voreio Aigaio	**SE00**	**Sweden**	UKI2	Outer London
DE11	Stuttgart	ES52	Comunidad Valenciana	GR42	Notio Aigaio	SE01	Stockholm	UKJ1	Berkshire, Bucks and Oxon
DE12	Karlsruhe	ES61	Andalucia	GR43	Kriti	SE02	Östra Mellansverige	UKJ2	Surrey et al.
DE13	Freiburg	ES62	Región de Murcia	**IE01**	**IRELAND**	SE04	Sydsverige	UKJ3	Hampshire et al.
DE14	Tübingen	**FI00**	**Finland**	**IT00**	**Italy**	SE06	Norra Mellansverige	UKJ4	Kent
DE21	Oberbayern	FI13	Itä-Suomi	ITC1	Piemonte	SE07	Mellersta Norrland	UKK1	Gloucestershire et al.
DE22	Niederbayern	FI18	Etelä-Suomi	ITC2	Valle d'Aosta	SE08	Övre Norrland	UKK2	Dorset and Somerset
DE23	Oberpfalz	FI19	Länsi-Suomi	ITC3	Liguria	SE09	Småland med öarna	UKK3	Cornwall et al.

DE24	Oberfranken	FI1A	Pohjois-Suomi	ITC4	Lombardia	SE0A	Västsverige	UKK4	Devon
DE25	Mittelfranken	FI20	Åland	ITD1	Trentino Alto Adige			UKL1	West Wales et al.
DE26	Unterfranken	**FR00**	**France**	ITD3	Veneto			UKL2	East Wales
DE27	Schwaben	FR10	Île de France	ITD4	Friuli-Venezia Giulia			UKM1	North Eastern Scotland
DE50	Bremen	FR21	Champagne-Ardenne	ITD5	Emilia-Romagna			UKM2	Eastern Scotland
DE60	Hamburg	FR22	Picardie	ITE1	Toscana			UKM3	South Western Scotland
DE71	Darmstadt	FR23	Haute-Normandie	ITE2	Umbria			UKM4	Highlands and Islands
DE72	Gießen	FR24	Centre	ITE3	Marche			UKN0	Northern Ireland
DE73	Kassel	FR25	Basse-Normandie	ITE4	Lazio				
DE91	Braunschweig	FR26	Bourgogne	ITF1	Abruzzo				
DE92	Hannover	FR30	Nord - Pas-de-Calais	ITF2	Molise				
DE93	Lüneburg	FR41	Lorraine	ITF3	Campania				
DE94	Weser-Ems	FR42	Alsace	ITF4	Puglia				
DEA1	Düsseldorf	FR43	Franche-Comté	ITF5	Basilicata				

Bibliography

Bashtannyk D, Hyndman R. 2001. Bandwidth Selection for Kernel Conditional Density Estimation. *Computational Statistics and Data Analysis* **36**: 279-98.

Bulli S. 2001. Distribution Dynamics and Cross-Country Convergence: New Evidence. *Scottish Journal of Political Economy* **48**: 226-243.

Cleveland WS. 1979. Robust locally-weighted regression and scatterplot smoothing. *Journal of the American Statistical Association* **74**: 829-836

Cleveland WS, Devlin SJ. 1988. Locally-weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association* **83**: 596-610

Durlauf SN, Johnson PA, Temple JRW. 2005. Growth Econometrics. In *Handbook of Economic Growth*, Volume 1A, Aghion P, Durlauf SN (eds). North-Holland: Amsterdam.

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

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**: 385-403.

Fingleton B. 1999. Estimates of Time to Convergence: An Analysis of Regions of the European Union. *International Regional Science Review* **22**: 5-34.

Hall P, Wolff R, Yao Q. 1999. Methods for estimating a conditional distribution function. *Journal of American Statistical Association* **94**: 154–163.

Hyndman RJ. 1996. Computing and Graphing Highest Density Regions. *The American Statistician* **50**: 120-126.

Hyndman RJ, Bashtannyk DM, Grunwald GK. 1996. Estimating and visualizing conditional densities. *Journal of Computational and Graphical Statistics* **5:** 315-336.

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

Lamo A. 2000. On Convergence Empirics: Some Evidence for Spanish Regions. *Investigaciones Economicas* **24**: 681-707.

Loader CR. 1996. Local likelihood density estimation. *The Annals of Statistics* **24**: 1602–1618.

López-Bazo E, Vayá E, Mora AJ, Suriñach J. 1999. Regional economic dynamics and

convergence in the European Union. *The Annals of Regional Science* **33**: 343-370.

Magrini S. 1999. The Evolution of Income Disparities among the Regions of the European Union. *Regional Science and Urban Economics* **29**: 257-281.

Magrini S. 2004. Regional (Di)Convergence. In *Handbook of Regional and Urban Economics,* Henderson V, Thisse JF (eds.). North-Holland: Amsterdam.

Pittau MG, Zelli R. 2006. Income dynamics across EU regions: empirical evidence from kernel estimator. *Journal of Applied Econometrics* **21**: 605-628.

Quah D. 1993. Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics* **95**: 427–443.

Quah D. 1996a. Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *Economic Journal* **106**: 1045-55.

Quah D. 1996b. Empirics for economic growth and convergence. *European Economic Review* **40**: 1353–1375.

Quah D. 1996c. Convergence Emprics Across Economies with (Some) capital Mobility. *Journal of Economic Growth* **1**: 95-124

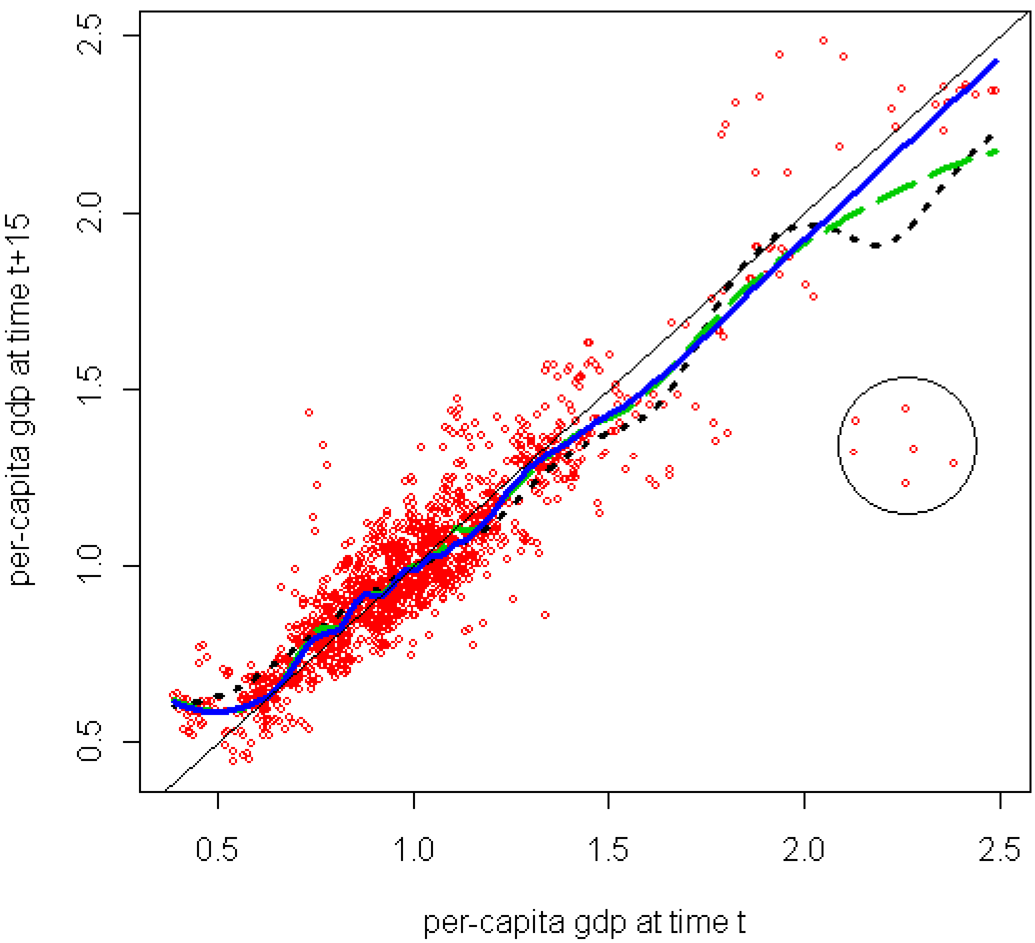
Quah D. 1997. Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth* **2:** 27-59.

Quah D. 2001. Searching for Prosperity: A Comment. *Carnegie-Rochester Conference Series on Public Policy* **55**: 305-19.

Rosenblatt M. 1969. Conditional Probability Density and Regression Estimators. In *Multivariate Analysis II,* Krishnaiah PR (ed.). Academic Press NY.

Sheather SJ, Jones MC. 1991. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of Royal Statistical Sociaty B* **53**:683-690.

Figure 1 – Regional per-capita income in Europe: comparing different scatterplot smoothers



Notes: the graph reports a scatterplot of relative per-capita income levels at time *t* and *t+15.* The estimated fits of three different scatterplot smoothers are superimposed: (a) the Nadaraya-Watson estimator ('dotted' curve); (b) the local linear regression smoother ('long-dashed curve') with variable bandwidth; and (c) the *lowess* ('solid' curve).

Figure 2 - Intra-Distribution Dynamics of regional per-capita income in Europe

Standard perspective plot (left hand side panel) and contour plot (right hand side panel) of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a kernel density estimator with fixed bandwidths (a = 0.149; b = 0.091)

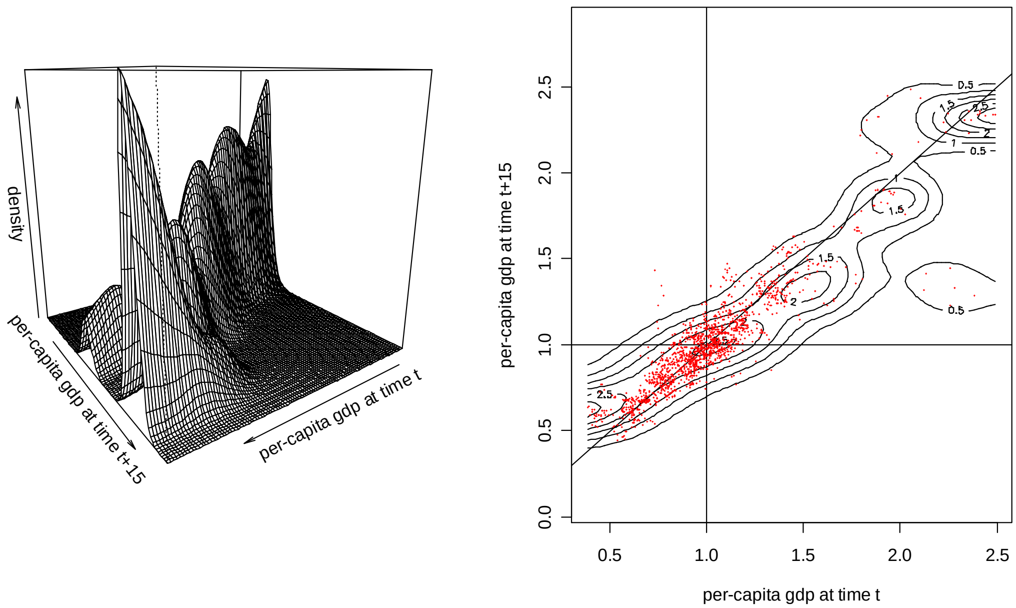
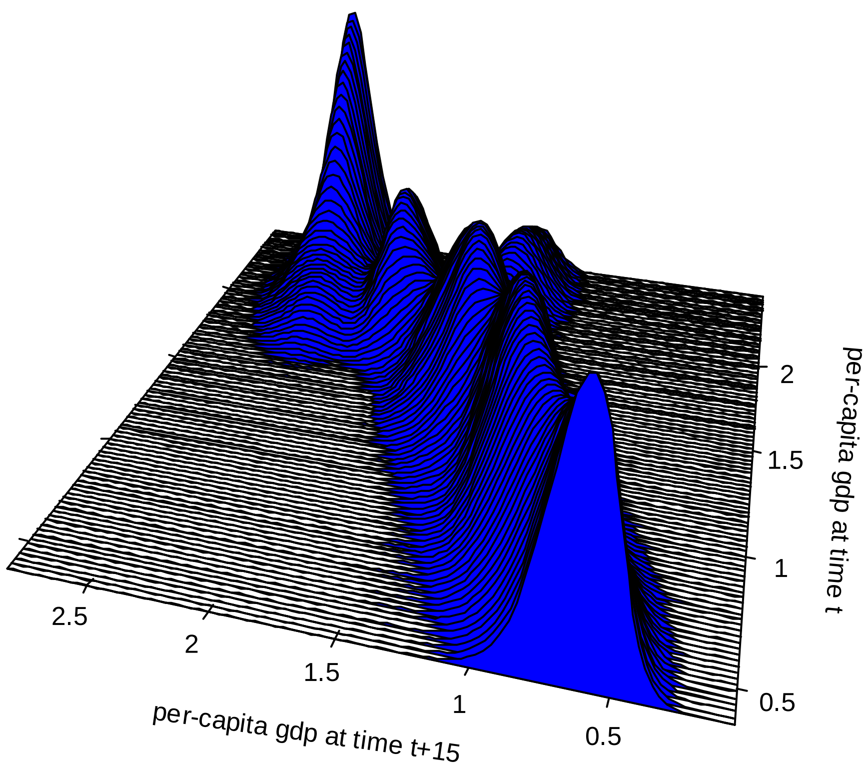


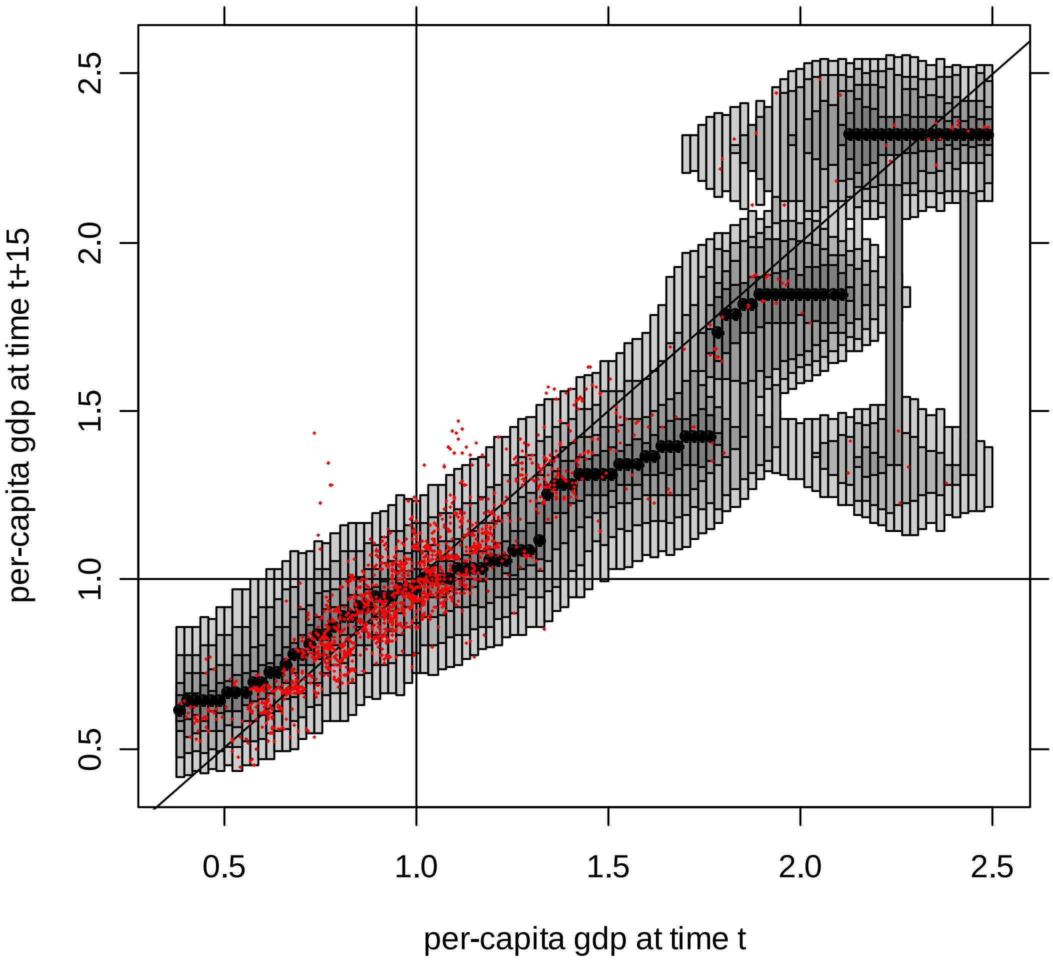
Figure 3 - Intra-Distribution Dynamics of regional per-capita income in Europe

Stacked density plot and HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a kernel density estimator with fixed bandwidths (a = 0.149; b = 0.091)

A - Stacked density plot



B - HDR plot



2.5
2.0
1.5
1.0
0.5
per-capita gdp at time t+15
0.5
1.0
1.5
2.0
2.5
per-capita gdp at time t

Figure 4 – Ergodic and initial distributions of regional per-capita income in Europe

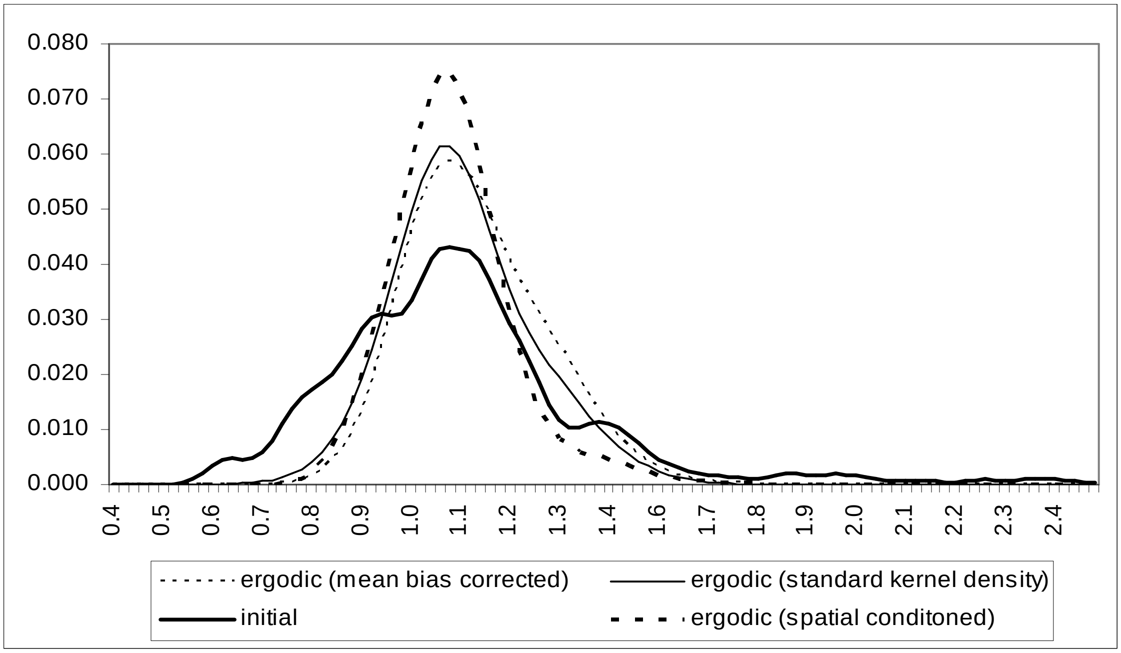


Figure 5 - Intra-Distribution Dynamics of regional per-capita income in Europe
HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a kernel density estimator with a variable bandwidth in the x direction (span = 0.3) and a fixed bandwidth in the y direction (b = 0.091)

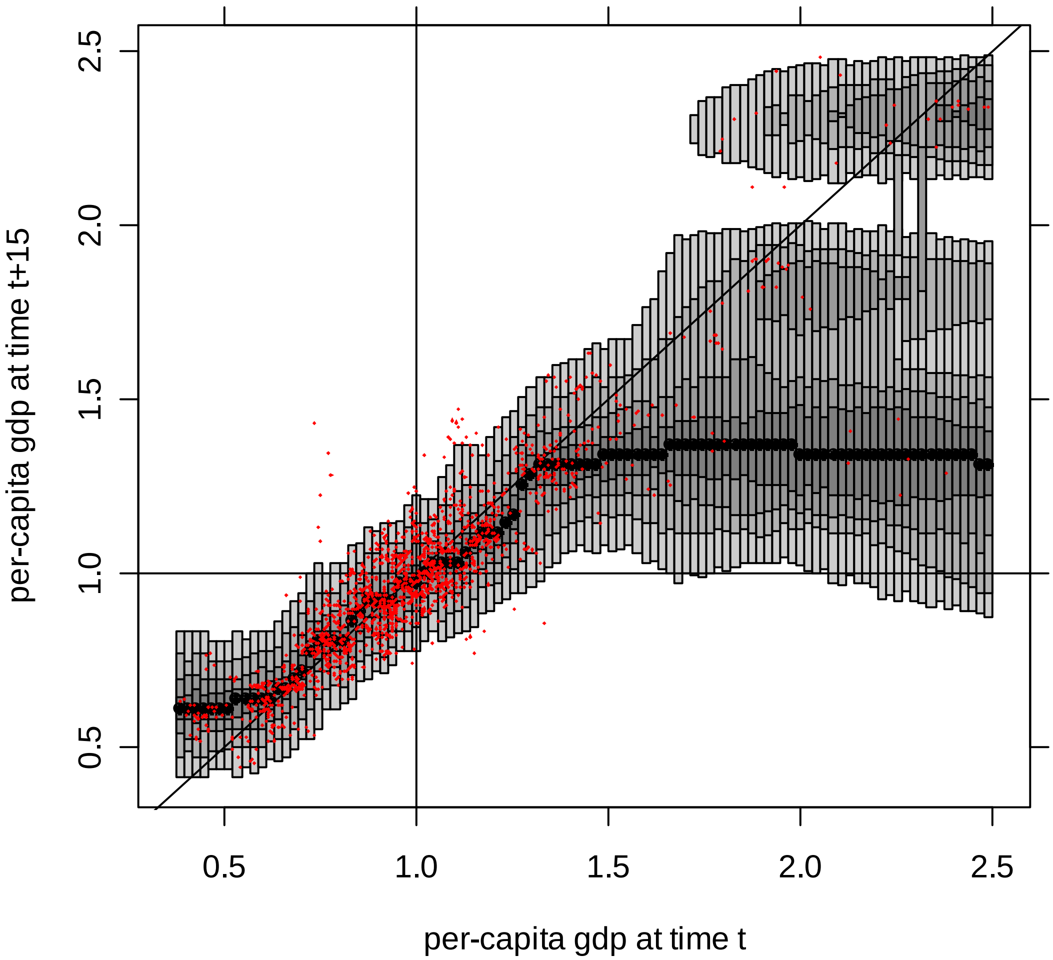


Figure 6 - Intra-Distribution Dynamics of regional per-capita income in Europe

HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a kernel density estimator with a variable bandwidth in the x direction (span = 0.3), a fixed bandwidth in the y direction (b = 0.091) and a mean function specified by a lowess smoother (span = 0.2)

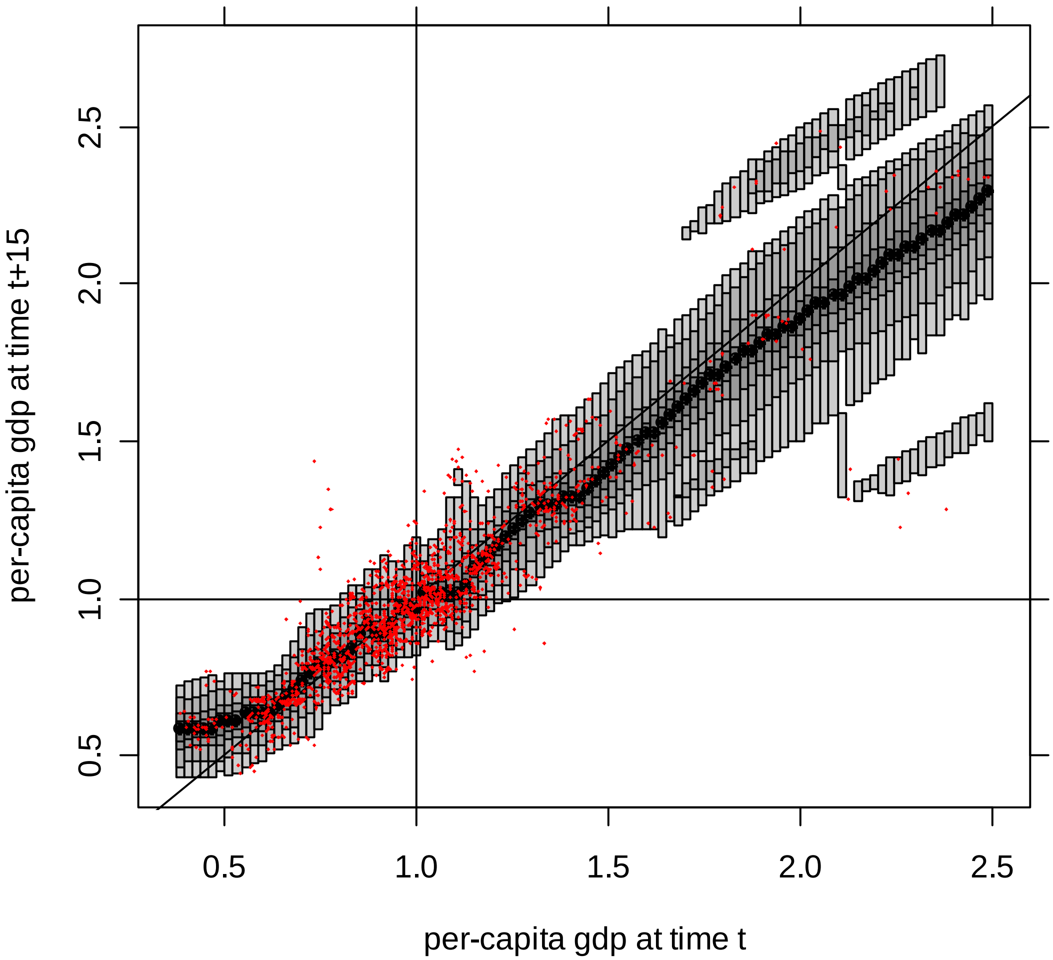


Figure 7 - Intra-Distribution Dynamics of regional per-capita income in Europe
HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a local linear density estimator with a variable bandwidth in the x direction (span = 0.3) and a fixed bandwidth in the y direction (b = 0.119) (no mean bias correction)

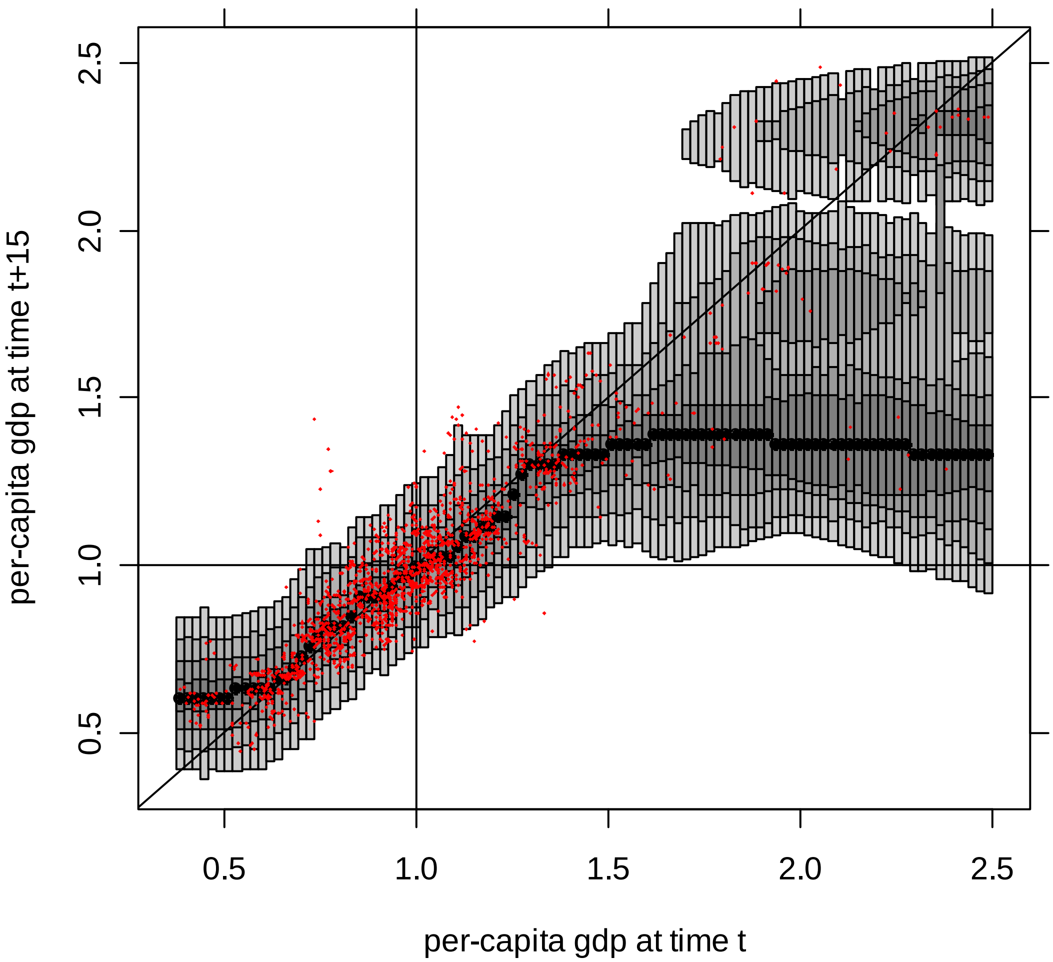


Figure 8 - Intra-Distribution Dynamics of regional per-capita income in Europe
HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a local linear density estimator with a variable bandwidth in the x direction (span = 0.3), a fixed bandwidth in the y direction (b = 0.119) and a mean function specified by a lowess smoother (span = 0.2)

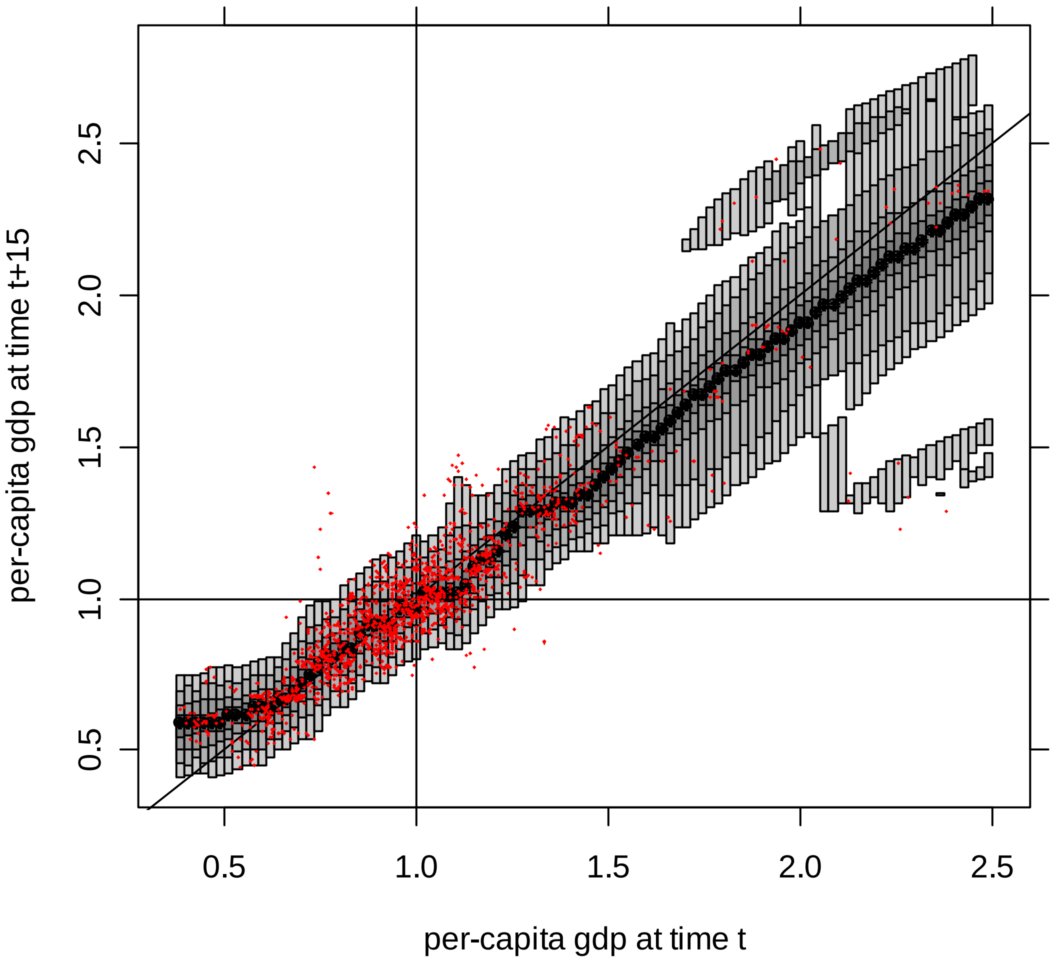


Figure 9 - Spatial conditioning

HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a kernel density estimator with different fixed bandwidths (a = 0.335; b = 0.168)

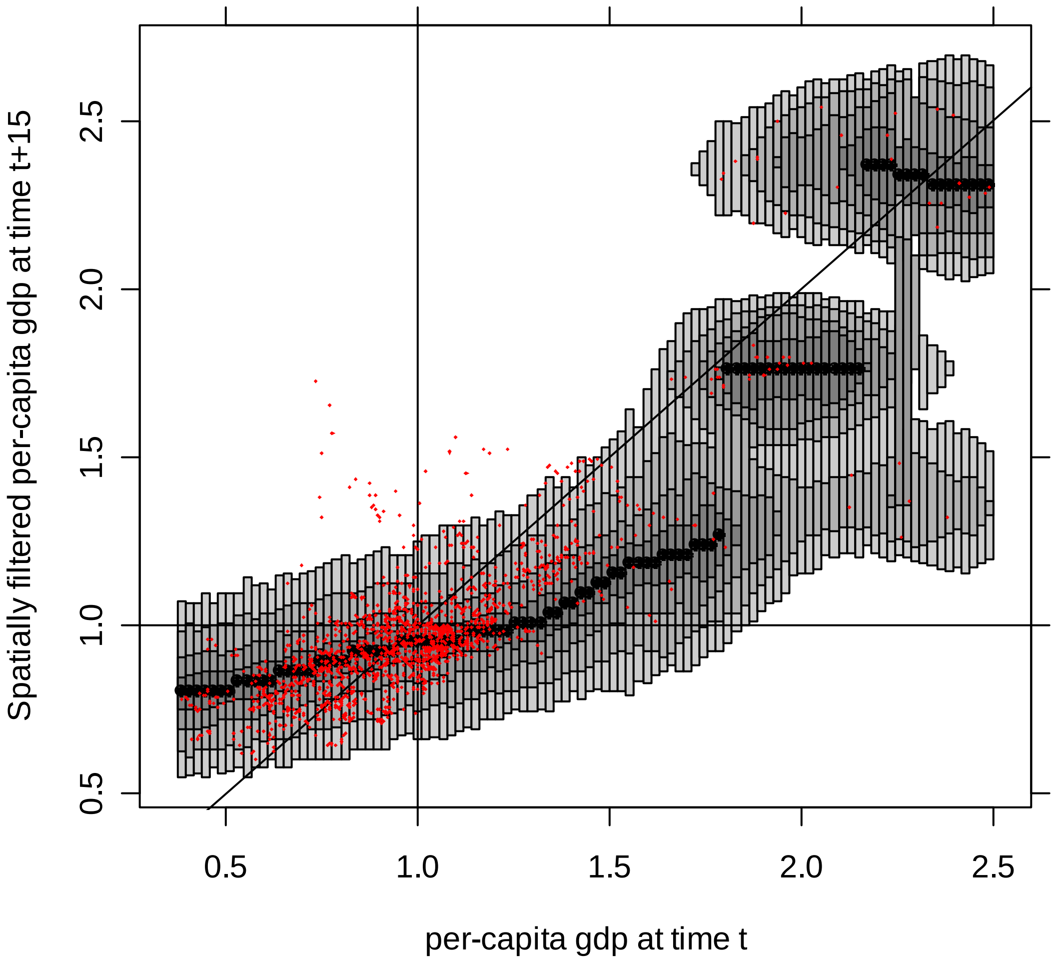


Figure 10 - Intra-Distribution Dynamics of regional per-capita income in Europe: spatial conditioning

HDR plot of conditional density for transitions of 15 years between 1980-2002. Estimates are based on a kernel density estimator with a variable bandwidth in the x direction (span = 0.3), a fixed bandwidth in the y direction (b = 0.168) and a mean function specified by a lowess smoother (span = 0.2)