The Regional Economics Applications Laboratory (REAL) is a unit of the University of Illinois focusing on the development and use of analytical models for urban and regional economic development. The purpose of the **Discussion Papers** is to circulate intermediate and final results of this research among readers within and outside REAL. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of the University of Illinois. All requests and comments should be directed to Geoffrey J. D. Hewings, Director, Regional Economics Applications Laboratory, 607 South Matthews, Urbana, IL, 61801-3671, phone (217) 333-4740, FAX (217) 244-9339.
Web page: www.uiuc.edu/unit/real

ANALYZING INTRA-DISTRIBUTION DYNAMICS: A REAPPRAISAL

by

Giuseppe Arbia, Roberto Basile and

Gianfranco Piras

REAL 05-T-11 December, 2005

Analyzing Intra-distribution Dynamics: a Reappraisal

Giuseppe Arbia

University of Pescara and LUISS

Roberto Basile[♣]

ISAE and University of Macerata

Gianfranco Piras

University of Pescara and University of Illinois at Urbana Champaign

Abstract: In this paper we suggest an alternative estimator and an alternative graphical analysis, both developed by Hyndman *et al.* (1996), to describe the law of motion of cross-sectional distributions of per-capita income and its components in Europe. This estimator has better properties than the kernel density estimator generally used in the literature on intra-distribution dynamics (cf. Quah, 1997). By using the new estimator, we obtain evidence of a very strong persistent behavior of the regions considered in the study, that is poor regions tend to remain poorer and rich regions tend to remain richer. These results are also in line with the most recent literature available on the distribution dynamic approach to regional convergence (Pittau and Zelli, 2006).

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 developed by Barro and Sala–i-Martin (1995), known as the β-convergence approach. This method has been discussed from different points of view (see Durlauf and Quah, 1999, 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

[♣] Corresponding author: 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.

The present paper has been written while Roberto Basile and Gianfranco Piras were visiting the Regional Economics Applications Laboratory at UIUC. A particular thank is due to Geoffrey J.D.Hewings both for providing hospitality during that period and for all the comments and suggestions helpul to improve the paper. Moreover, the authors wish to acknowledge the help of Robert Hyndman for having read previous version of the paper and for having contributed to the final version with very interesting comments. All the suggestions we obtained from the participants at the REAL seminar series were also appreciated.

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. Generally speaking, 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 overcome such a problem. This method, known as the *intra-distribution dynamics* approach (Quah, 1996, 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, 1996; López-Bazo *et al*., 1999; Fingleton, 1997, 1999) or, more recently, through the estimation of conditional densities using stochastic kernel estimators (Quah, 1997; Magrini, 2004; Cheshire and Magrini, 2005). 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. However, it has to been noticed that the conditional density function is a "sequence of univariate functions" (Hyndman *et al.*, 1996).

The aim of this paper is to suggest an alternative technique to describe the law of motion of cross-regional distributions of per-capita income, labor productivity and unemployment rates in Europe. To achieve this aim we will make use of an alternative estimator and of alternative graphical evidences (based on the so-called *stacked density plot and highest density plot*), both developed by Hyndman *et al*. (1996). In particular, they notice that the conditional mean function obtained from the kernel estimation of the conditional density is equivalent to a kernel smoother. Starting from the observation that the kernel smoothers present some undesirable bias properties, 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. Furthermore, they show that their modified estimator has a smaller integrated mean square error than the standard kernel estimator.

The layout of the paper is the following. In Section 2, we recall the intra-distribution dynamics approach and describe the conditional density estimator developed by Hyndman *et al*. (1996). In Section 3, we report the estimation results obtained applying this estimator to data on per-capita GDP, labor productivity and unemployment rates for European regions over the period 1980-2003. Section 4 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, 1994, 1996, 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 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 say h different income classes and T time periods and define F_t as the 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 autoregressive process:

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

The matrix ***M*** is usually defined as the transition probability of a Markov process. Each element of ***M*** describes the probability that an economy belonging to class *i* in time period *t* will move to class *j* in the next time period (Quah, 1993). Iterations of (1) yield a predictor for future cross-section distributions

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

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

Implications for the convergence debate are then drawn from the study of the functions $\phi_{t+\tau}$'s: if they display a tendency towards a single point mass, then we can conclude that there is convergence towards equality. Conversely, if $\phi_{t+\tau}$ displays a tendency towards a two-point (or bimodal measure), this might be interpreted as a sign of income polarization.

Even if intuitively appealing, the Markov Chain approach is not free of criticisms. In fact, the Markov property assumes that in each moment of time the temporal process is dependent on only the previous period in time (a process is said to be a Markov chain if the random variable at time *t+1* depends exclusively on the information set at time *t* and not on any other previous period in time). For this reason, Bickenbach and Bode (2003) pointed out that Markov chain theory imposes restrictions on the data-generating process if applied to analyze the regional convergence process. In particular, the assumption that per-capita income follows a stationary first order Markov process is often unrealistic and needs to be verified in each empirical application by using appropriate statistical testing procedures. By using data on the U.S., Bickenbach and Bode (2003) proved that the per-capita income over a fairly long period of time did not follow a common stationary fist order Markov process.

One way to overcome this problem is to avoid any arbitrary discretization and to allow the number of cells of the Markov transition probability matrix to tend to infinity (Quah, 1997). In this case, 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.

2.2 The kernel conditional density estimator

Operationally, the *transition dynamics approach* consists of the estimation and visualization of 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 $\{(X_1,Y_1),\dots,(X_n,Y_n)\}$ and the observations by $\{(x_1,y_1),\dots,(x_n,y_n)\}$; 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), $h_\tau(x)$ be the marginal density of X and $f_\tau(y \mid x) = g_\tau(x,y)/h_\tau(x)$ 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:

$$\hat{f}_\tau(y \mid x) = \frac{\hat{g}_\tau(x,y)}{\hat{h}_\tau(x)} \qquad (4)$$

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 multiplicative joint density of *(X,Y)* and

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

the estimated marginal density.[1]

Equation (4) can also be written as:

$$\hat{f}_\tau(y\,|\,x)=\frac{1}{b}\sum_{i=1}^{n}w_i(x)K\left(\frac{\|y-Y_i\|_y}{b}\right) \qquad (5)$$

where

$$w_i(x)=K\left(\frac{\|x-X_i\|_x}{a}\right)\Bigg/\sum_{j=1}^{n}K\left(\frac{\|x-X_j\|_x}{a}\right).$$

Equation (5) suggests that the conditional density estimate at $X=x_0$ can be obtained by summing the n kernel functions in the Y-space, weighted by $\{w_i(x)\}$ in the X-space. In other words, equation (5) can be interpreted as the Nadaraya-Watson kernel regression (or locally weighted averaging) of $K\left(\frac{\|y-Y_i\|_y}{b}\right)$ on X_i. (see Hyndman and Yao, 2002). 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.[2] As usual, small bandwidths produce small bias and large variance whereas large bandwidths give large bias and small variance. The optimal bandwidth might be derived by differentiating the integrated mean square error function (IMSE) with respect to a and b and setting the derivatives to 0 (Bashtannyk and Hyndman, 2001). However, this requires additional

---

[1] $\|.\|_x$ and $\|.\|_y$ 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.

[2] In the original Rosenblatt's estimator $a=b$.

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 *x*. When the data are not homogenously distributed over all the sample space (i.e. 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* are included in the window. The fraction *m/n* is called the span of the kernel smoother.

2.3 A nonparametric conditional density estimator with mean-bias correction

Hyndman *et al*. (1996) observe that the estimation of the conditional mean function obtained from the kernel (Equation 5) is equivalent to the Nadaraya (1964) and Watson (1964) kernel regression function:

$$\hat{m}(x) = \int y\hat{f}_\tau(y \mid x)\,dy = \sum_i^n w_i(x) Y_i \qquad (6)$$

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 (Bowman and Azzalini, 1997).

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

$$\hat{f}_\tau^*(y \mid x) = \frac{1}{b}\sum_{i=1}^n w_i(x) K\left(\frac{\left\| y - Y_i^*(x)\right\|_y}{b}\right) \qquad (7)$$

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 $e \mid (X = x)$.

Hyndman *et al.* (1996) observe that the error term (e_i) has the same distribution of y_i except for a shift in the conditional mean. Thus, one may start by applying the standard kernel density estimator to the points $\{x_i, e_i\}$, and then add 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, 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 $\hat{r}(x) = \hat{m}(x) = \sum_{i=1}^{n} w_i(x) Y_i$ (i.e. 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, $\hat{m}(x)$. In other words, using the method developed by Hyndman *et al.* (1996), the mean function of $\hat{f}_\tau^*(y \mid x)$ 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. Moreover, Hyndman *et al.* (1996) show that the modified estimator has a smaller integrated mean square error than the standard kernel estimator.

3. Some evidences on regional convergence in Europe

3.1 Data

In this paper, we use the intra-distribution dynamics approach described in the previous section to explain the law of motion of cross-regional distributions of per-capita income and its components (i.e., labor productivity and unemployment rate) in Europe. Following de la Fuente (2002), the per-capita income can be expressed as the product of two main components, income per-worker and workers per population unit. The employment component of income per-capita, in particular, depends crucially on labor force participation and unemployment rates. Thus, income per-capita can be written in the form

$$\frac{Y}{P} = \frac{Y}{L} \times \frac{L}{P} \approx \frac{Y}{L} \times (1-u) \qquad (8)$$

where $\frac{Y}{P}$ is regional gross value added measured at constant prices, $\frac{Y}{L}$ is income per worker, also expressed at constant prices, and u is the unemployment rate (defined as the ratio between unemployment and total labour force). All variables are normalized with respect to the EU average. Figure 1 presents all the variables considered in the analysis, where the value at time t is plotted against the value at $t+\tau$. The scatter of points is very close both for the per-capita GDP and for the labor productivity (except for the higher part of the distribution and for a middle region in the case of labor productivity), while the distribution in the case of the unemployment rate is much more dispersed. Working with relative values helps to remove co-movements due to the European wide business cycle and trends in the average values. The period considered extends from 1980 to 2003. The number of NUTS2 regions included in the sample is 184 (see Appendix 1 for full details of the regional coverage). All series are drawn from the Cambridge Econometrics Dataset.[3] In the first step, we set $\tau = 15$ and we estimate $f_{15}(y \mid x)$ by using a kernel estimator with a constant bandwidth parameter a (Equation 5). Then, we estimate a conditional density using the modified estimator with mean bias correction (Equation 7). In particular, in the second step the conditional densities were estimated using a *lowess* (locally quadratic)[4] mean with a span c=0.2 (see Cleveland, 1979; Cleveland and Devlin, 1988). Smoothing parameters for the conditional density estimation are a = 0.15 and b=0.10 in the case of per capita GDP, a = 0.13 and b=0.08 in the case of labour productivity and a = 0.22 and b=0.19 in the case of unemployment rates.[5]

3.2 New graphical methods for visualizing intra-distribution dynamics

All the studies on intra-distribution dynamics which make use of non-parametric stochastic kernel estimators provide three-dimensional perspective plots (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 function as a bivariate density function, while this

[3] Groningen and Luxemburg were excluded from the sample since they appeared to be outliers.

[4] 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. The combination of three features - nearest neighbours, smoothed weight function (the tri-cube kernel), and local expected value formed via locally weighted regressions - helps local regression outperform many other scatter-plot smoothers (such as moving averages and overlapping regressions).

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

function must be interpreted as "a sequence of univariate densities" (Hyndman *et al*. 1996) of per-capita income levels (or of its components) conditional on certain initial levels.

In the present paper, new graphical methods for visualizing conditional density estimators developed by Hyndman *et al*. (1996) and Hyndman (1996) are used. The first graphical method, called the "*stacked conditional density plot*" (see figures 2-7, left hand side panel), displays a number of conditional densities plotted side by side in a perspective plot. It facilitates viewing the changes in the shape of the distributions of the variables observed at time $t+\tau$ over the range of the same variable observed at time *t*. In other terms, each univariate density plot describes transitions over 15 years from a given income value in period *t*. Such a representation is equivalent to a transition probability matrix with a continuum of rows and columns. 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. 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 uni-modal distributions, the *HDRs* are exactly the usual probabilities around the mean value; however, in the case of multi-modal distributions, the *HDR* displays different disjointed sub-regions. For each variable, Figures 2-7 (right hand side panel) show a plot of the 25% (the darker shaded region), 50%, 75% and 90% (the lighter shaded region) *HDR*s computed from the density estimates shown in panel (a). If the 25% or the 50% *HDRs* cross the 45-degree diagonal, it means that most of the elements in the distribution remain where they began. Thus, it is quite clear that this method is particularly informative for the analysis of regional growth behavior, since it highlights the dynamics of the entire cross-section distribution. Clearly, it remains important to analyze any other moments of the distribution (such as mean and variance) and any other central points. In particular, one may wish to analyze the modes, the values of *y* where the density function takes on its maximum values. Indeed, 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 (Scott, 1992). In each figure, the highest modes for each conditional density estimate are superimposed on the *HDR* plots and they are shown as a bullet.

3.3 Empirical evidence

For the case of per-capita incomes, figure 2 shows the stacked density plot and the *HDR* plot of conditional density for transitions of 15 years based on the kernel estimator with a fixed bandwidth parameter *a*. The results obtained are consistent with those discussed in previous work (Magrini and Cheshire, 2005; Brasili and Gutierez, 2004). In particular, the two plots provide some evidence of convergence even if with a rate that appears to be very slow. In particular, regions that at the beginning of the period had a per-capita income level much lower than the EU average appear more likely to improve their relative position over the next 15 years; the first three modes of the lower tail of the distribution are above the main diagonal. Conversely, regions that at the beginning of the period had a per-capita income level higher than the EU average appear more likely to worsen their relative position over the next 15 years; the modes of the upper tail of the distribution are always below the main diagonal. This means that the poorer economies are catching up with the richer ones but this process appears to be very slow because the most of the mass of the probability distribution is still close to the 45-degree diagonal. Finally, it is quite revealing to note that there are no signs of bimodality in the distribution at any level of per capita income at time *t*.

Figure 3 reports the results based on the modified conditional density estimator with mean function specified by a *loess* smoother. The two plots provide strong evidence of persistence; in most of the cases, regions remain where they started. In other words, almost all the modes appear to lie on the 45-degree diagonal, and also the mass of the probability is very concentrated around the diagonal. However, there is still evidence of some changes in the relative positions for the very high and very low part of the distribution. In particular, regions that at the beginning of the period had a very low per-capita income level with respect to the EU average appear more likely to improve their relative position over the next 15 years (the lower mode in figure 3). Conversely, regions that at the beginning of the period had a per-capita income level particularly high with respect to the EU average appear more likely to worsen their relative position over the next 15 years (the two upper modes in figure 3). In other words, there are some regions (those between 0 and 0.5 in the distribution of the per-capita GDP at time *t*) that were so poor at the beginning that their relative position could only have improved at the end of the time period. On the other hand, regions belonging to the part of the distribution greater than 2.5 were so rich in relative terms at the beginning of the period that it would have been difficult to believe that the growth rate all over the time span would have not slowed down in relative terms. In fact, they tend to worsen their relative position in the considered time interval and most of the distribution for those regions stands below the main diagonal. These results are perfectly in line with those presented in Pittau and Zelli (2006).

In the case of labor productivity estimated with a kernel estimator with fixed bandwidth *a* (figure 4), there are also signs of convergence, even if the picture is much more complicated. The modes of the distributions estimated in correspondence with relatively low levels of labor productivity at time *t* lie below the 45-degree diagonal, while the modes of the distributions estimated in relation to the relatively high levels of labor productivity at time *t* lie above the same line. Some exceptions are found in the middle part of the distribution. Figure 5 shows the evidence obtained using the modified conditional density estimator. The evidence of persistence in this case is even stronger than in the case of the per-capita GDP since, in addition, the modes of the very high values of the distribution at time *t* appear to lie on the 45-degree diagonal, the only exception remaining in the left tail of the distribution.

The evidence shown in figure 6 (referring to data on the rate of unemployment using the kernel estimation with fixed bandwidth) appears consistent with the findings reported in Overman and Puga (1999). Regions that at the beginning of the period were characterized by a very low relative unemployment rate with respect to the European average have a propensity to worsen their relative position over the next 15 years (in other words, the unemployment rate is rising). Generally, regions with unemployment rate in line with the mean of the distribution at time *t* exhibit a lower growth rate. Furthermore, some evidence of bimodality occurs for regions with relatively high initial unemployment rates (higher than 2.5), while regions belonging to the very upper tail of the distribution (higher than 2.8) are likely to show either a divergent path (i.e. the unemployment rate increases) or a converging path (i.e. decreasing unemployment rate). Thus, this evidence is strongly influenced by the fact that the distribution of the unemployment rate is highly positively skewed as it can be seen from the scatterplot in figure 1. Finally, the results obtained using the modified conditional density estimator (figure 7) again generate evidence of a strong persistent pattern. Even if there is still support for the finding that lower regions are increasing the unemployment rate (and the increase seems higher in this case), it is no longer true that the very upper regions are experiencing a dual outcome. The number of modes lying on the 45-degree diagonal appear to be very high, again giving the impression that nothing has been changing in the distribution of the unemployment rate over the 15-years period considered.

4. 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 income. In particular, the empirical applications of the kernel stochastic approach to the case of European regions report evidence of some degree of convergence (see, in particular, Brasili and Gutierez, 2004); 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 (for example, the "twin peaks" distribution has been identified for example by Magrini and Cheshire, 2005); some rich regions are converging to an 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 can be criticized from two different point of view. First, the kernel density estimator is usually implemented applying the same 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 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 alternative graphical visualization of the conditional density estimations, both developed by Hyndman *et al.* (1996), to describe the law of motion of cross-sectional distributions of per-capita income and of its components (labor productivity and unemployment rates) in Europe. This estimator has better properties than the kernel density estimator with one common (constant) bandwidth parameter generally used in the literature on intra-distribution dynamics.

Applying the Hyndman *et al*. (1996) method to European data, we obtain interesting evidence that enriches the debate on the distribution dynamics. In particular, for all the variables under analysis, even if with small differences, we observe that the most of the modes are lying on the 45-degrees diagonal. From an economic point of view, this means that there is a strong persistent behavior of the European regions considered in the present study. Alternatively, it may be stated that the picture of the disparities is not changing over the 15-years interval considered, and almost all the regions appear to remain where they were at the beginning.

In this paper, we have suggested some technical improvements to the study of the intra-distribution dynamics approach, but many questions still remain open. In future work, we will investigate the determinants of the patterns of cross-sectional growth, by combining the new methodology proposed here with the conditioning schemes for cross-sectional distributions proposed in the literature (see, e.g., Quah, 1997). This analysis will be helpful in producing suggestions for a set of regional policies intended to reduce disparities. Also the study of the ergodic distribution and the further development of models incorporating spatial dependence can be supportive for this aim.

Bibliography

Barro R.J. and X. Sala-i-Martin, (1995), *Economic growth*, McGraw Hill, New York.

Bashtannyk D. and Hyndman R. (2001), Bandwidth Selection for Kernel

Conditional Density Estimation, *Computational Statistics and Data Analysis*, 36, 279-98.

Birckenbach F. and Bode E. (2003), Evaluating the Markov Property in Studies of Economic Convergence. *International Regional Science Review*, Vol. 26, No. 3, 363-392.

Bowman A.W. and Azzalini A. (1997), *Applied Smoothing Techniques for Data Analysis. The kernel approach with S-Plus illustrations*, Clarendon Press, Oxford.

Brasili C. and Gutierrez L. (2004), Regional convergence across European Union, Development and Comp Systems 0402002, Economics Working Paper Archive EconWPA

Cheshire P. and Magrini S. (2005), Analyzing growth distribution dynamics: isolating divergence factors, Paper presented at the 45th European Regional Science Association Congress, Amsterdam.

Cleveland W.S., (1979), *Robust locally-weighted regression and scatterplot smoothing*, "Journal of the American Statistical Association", n.74.

Cleveland W.S. and S.J. Devlin, (1988), *Locally-weighted regression: an approach to regression analysis by local fitting*, "Journal of the American Statistical Association", n.83.

Copas, J. B. (1995) ‘‘Local likelihood based on kernel censoring’’, J. Roy. Statist. Soc. Ser. B 57, 221–235.

de la Fuente, A. (2002), "On the Sources of Convergence : A Close Look at the Spanish Regions". *European Economic Review*, 46 (3), 569-599.

Durlauf S.N. and D. Quah (1999), The new empirics of economic growth. In *Handbook of macroeconomics. Volume 1A*. Amsterdam, New York and Oxford: Elsevier Science, North-Holland, 235-308.

Fan, J., Yao, Q. and 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 (3): 385-.

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

Hall, P., Wolff, R. and Yao, Q. (1999) ''Methods for estimating a conditional distribution function'', J. Amer. Statist. Assoc. 94, 154–163.

Hjort, N. L. and Jones, M. C. (1996) ''Locally parametric nonparametric density estimation'', Ann. Statist. 24, 1619–1647.

Hyndman R.J. (1996), "Computing and Graphing Highest Density Regions", *The American Statistician*, 50, 2, 120-126.

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

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

Loader, C. R. (1996) ''Local likelihood density estimation'', Ann. Statist. 24, 1602–1618.

López-Bazo E, Vayá E, Mora AJ and Suriñach J (1999) Regional economic dynamics and convergence in the European Union, *The Annals of Regional Science* 33: 343-370.

Magrini S. (2004), *Regional (Di)Convergence*, in V. Henderson, J.F. Thisse (eds.), *Handbook of Regional and Urban Economics*, Volume 4.

Nadaraya E.A. (1964), On Estimatine Regression, *Theory Probab. Applic.*, 15, pp. 134-137.

Overman H.G. and D. Puga (1999), Unemployment clusters across Europe's regions and countries, *Economic Policy* 34, April 2002: 115-147

Pittau M.G. and R. Zelli (2006), Income dynamics across EU regions: empirical evidence from kernel estimator, *Journal of Applied Econometrics (forthcoming)*.

Quah D. (1993), Galton's fallacy and tests of the convergence hypothesis. *Scandinavian J. Economics* 95, 427–443.

Quah D. (1994), One business cycle and one trend from (many,) many disaggregates. *Eur. Economic Review* 38, 605–613.

Quah D. (1996), Empirics for economic growth and convergence. Eur. Economic Rev. 40, 1353–1375.

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

Rosenblat, M. (1969), Conditional Probability Density and Regression Estimators, in Krishnaiah, P.R. (a cura di), *Multivariate Analysis II*, Academic Press NY, pp. 25-31.

Scott D.W. (1992), Multivariate density estimation: theory, practice, and visualization, Wiley: New York.

Watson G.S. (1964), Smooth Regression Analysis, *Sankhya*, 26, pp. 359-72.

APPENDIX 1: THE SAMPLE

BE1	Bruxelles-Brussels
BE21	Antwerpen
BE22	Limburg
BE23	Oost-Vlaanderen
BE24	Vlaams Brabant
BE25	West-Vlaanderen
BE31	Brabant Wallon
BE32	Hainaut
BE33	Liège
BE34	Luxembourg
BE35	Namur
DK	Denmark
DE11	Stuttgart
DE12	Karlsruhe
DE13	Freiburg
DE14	Tübingen
DE21	Oberbayern
DE22	Niederbayern
DE23	Oberpfalz
DE24	Oberfranken
DE25	Mittelfranken
DE26	Unterfranken
DE27	Schwaben
DE5	Bremen
DE6	Hamburg
DE71	Darmstadt
DE72	Gießen
DE73	Kassel
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
DEC	Saarland
DEF	Schleswig-Holstein
GR11	Anatoliki Makedonia
GR12	Kentriki Makedonia
GR13	Dytiki Makedonia
GR14	Thessalia
GR21	Ipeiros
GR22	Ionia Nisia
GR23	Dytiki Ellada
GR24	Sterea Ellada
GR25	Peloponnisos
GR3	Attiki
GR41	Voreio Aigaio
GR42	Notio Aigaio
GR43	Kriti
ES11	Galicia
ES12	Principado de Asturias
ES13	Cantabria
ES21	Pais Vasco
ES22	Navarra
ES23	La Rioja
ES24	Aragón
ES3	Comunidad de Madrid
ES41	Castilla y León
ES42	Castilla-la Mancha
ES43	Extremadura
ES51	Cataluña
ES52	Comunidad Valenciana
ES53	Illes Balears
ES61	Andalucia
ES62	Región de Murcia
FR1	Île de France
FR21	Champagne-Ardenne
FR22	Picardie
FR23	Haute-Normandie
FR24	Centre
FR25	Basse-Normandie
FR26	Bourgogne
FR3	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	Prov.-Alpes-Côte d'Azur
FR83	Corse
IE01	Border, Midl. and Western
IE02	Southern and Eastern
ITC1	Piemonte
ITC2	Valle d'Aosta
ITC3	Liguria
ITC4	Lombardia
ITD1+2	Trentino-Alto Adige
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
NL12	Friesland
NL13	Drenthe
NL21	Overijssel
NL22	Gelderland
NL31	Utrecht
NL32	Noord-Holland
NL33	Zuid-Holland
NL34	Zeeland
NL41	Noord-Brabant
NL42	Limburg
AT11	Burgenland
AT12	Niederösterreich
AT13	Wien
AT21	Kärnten
AT22	Steiermark
AT31	Oberösterreich
AT32	Salzburg
AT33	Tirol
AT34	Vorarlberg
PT11	Norte
PT15	Algarve
PT16	Centro
PT17	Lisboa
PT18	Alentejo
SE01	Stockholm
SE02	Östra Mellansverige
SE04	Sydsverige
SE06	Norra Mellansverige
SE07	Mellersta Norrland
SE08	Övre Norrland
SE09	Småland med öarna
SE0A	Västsverige
UKC1	Tees Valley and Durham
UKC2	Northumberland et al.
UKD1	Cumbria
UKD2	Cheshire
UKD3	Greater Manchester
UKD4	Lancashire
UKD5	Merseyside
UKE1	East Riding et al.
UKE2	North Yorkshire
UKE3	South Yorkshire
UKE4	West Yorkshire
UKF1	Derbyshire et al.
UKF2	Leicestershire et al.
UKF3	Lincolnshire
UKG1	Herefordshire et al.
UKG2	Shropshire et al.
UKG3	West Midlands
UKH1	East Anglia
UKH2	Bedfordshire, Hertfordshire
UKH3	Essex
UKI1	Inner London
UKI2	Outer London
UKJ1	Berkshire, Bucks and Oxon
UKJ2	Surrey et al.
UKJ3	Hampshire et al.
UKJ4	Kent
UKK1	Gloucestershire et al.
UKK2	Dorset and Somerset
UKK3	Cornwall et al.
UKK4	Devon
UKL1	West Wales et al.
UKL2	East Wales
UKM1	North Eastern Scotland
UKM2	Eastern Scotland
UKM3	South Western Scotland
UKM4	Highlands and Islands
UKN	Northern Ireland

Figure 1: Scatter plots of Regional Per-Capita Income, Labor Productivity and Unemployment Rates at time t and t+15

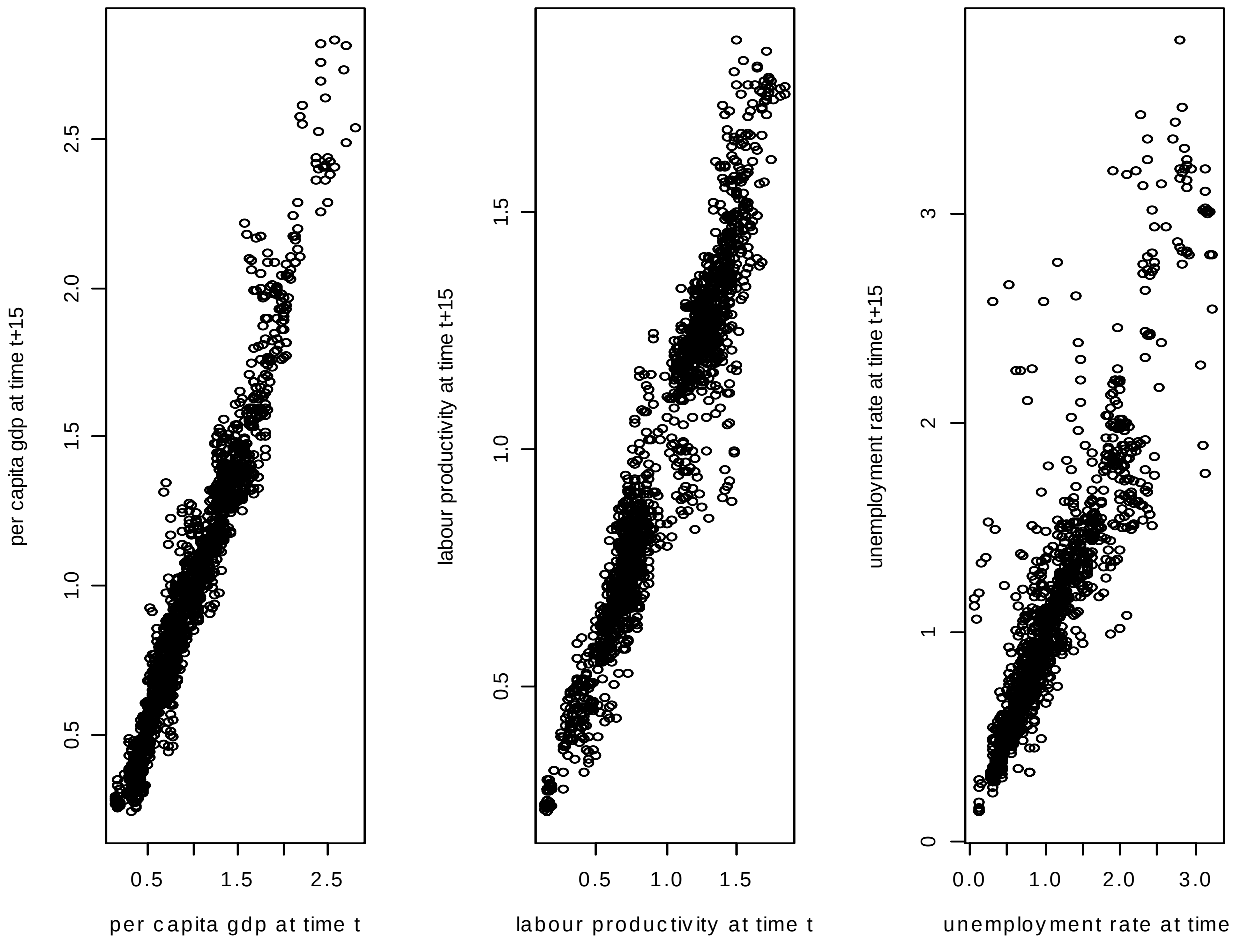


Figure 2: Intra-Distribution Dynamics of Regional Per-Capita Incomes

Stacked density plot (left hand side panel) and HDR plot (right hand side panel) of conditional density for transitions of 15 years based on the kernel estimator with a fixed bandwidth

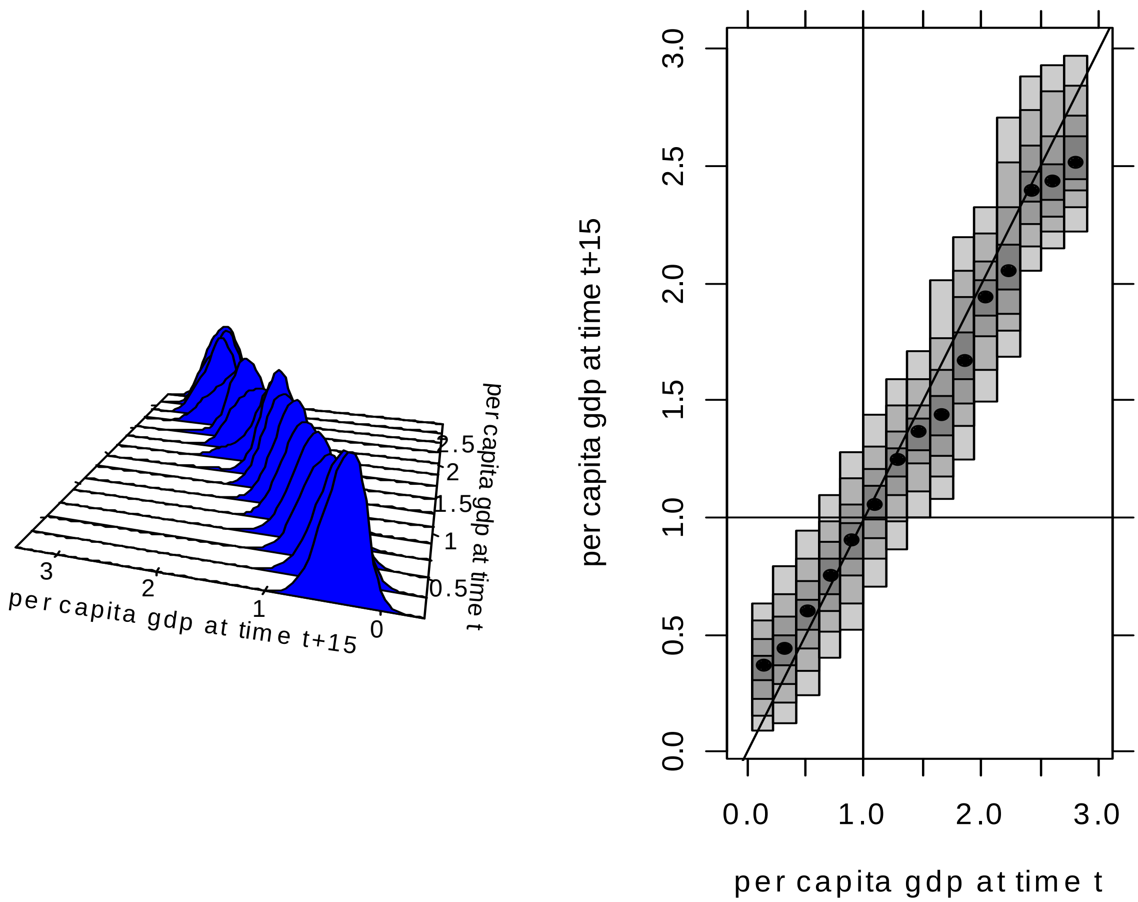


Figure 3: Intra-Distribution Dynamics of Regional Per-Capita Incomes

Stacked density plot (left hand side panel) and HDR plot (right hand side panel) of conditional density for transitions of 15 years based on the modified conditional density estimator with mean function specified by a loess smoother

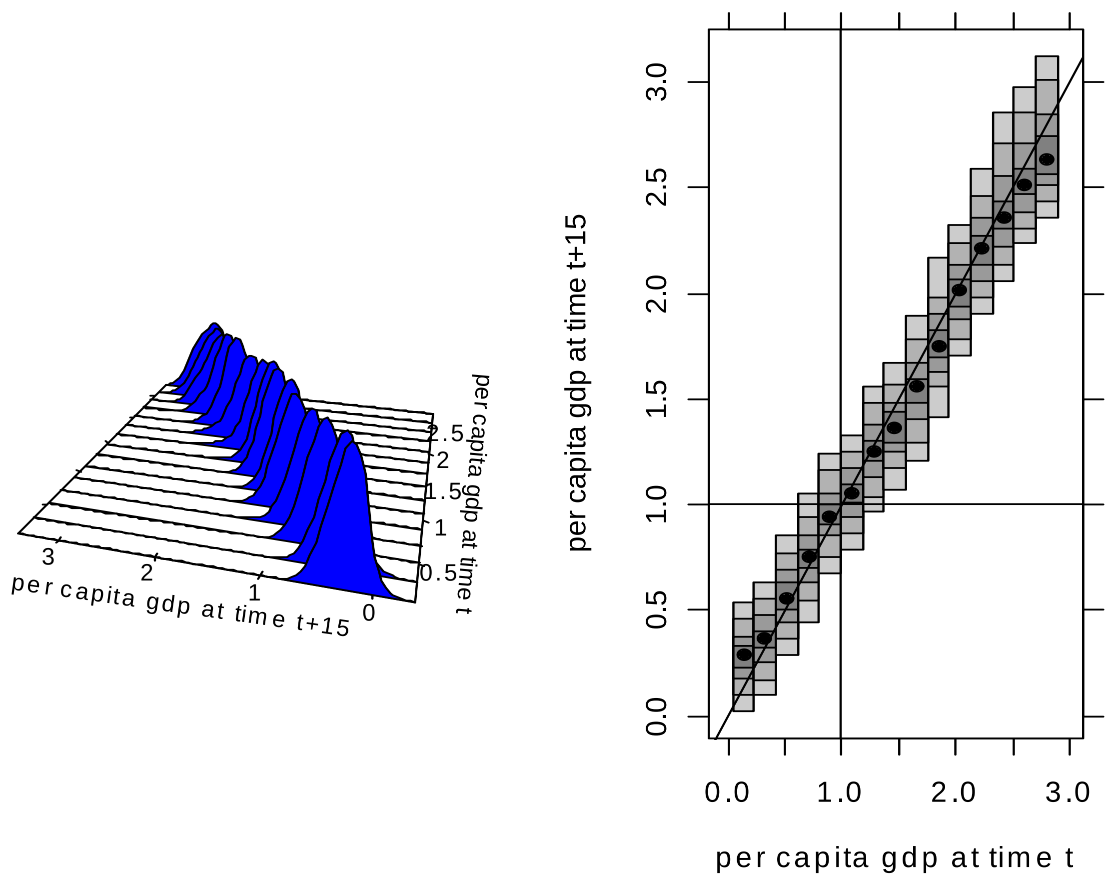


Figure 4: Intra-Distribution Dynamics of Regional Labor Productivity

Stacked density plot (left hand side panel) and HDR plot (right hand side panel) of conditional density for transitions of 15 years based on the kernel estimator with a fixed bandwidth

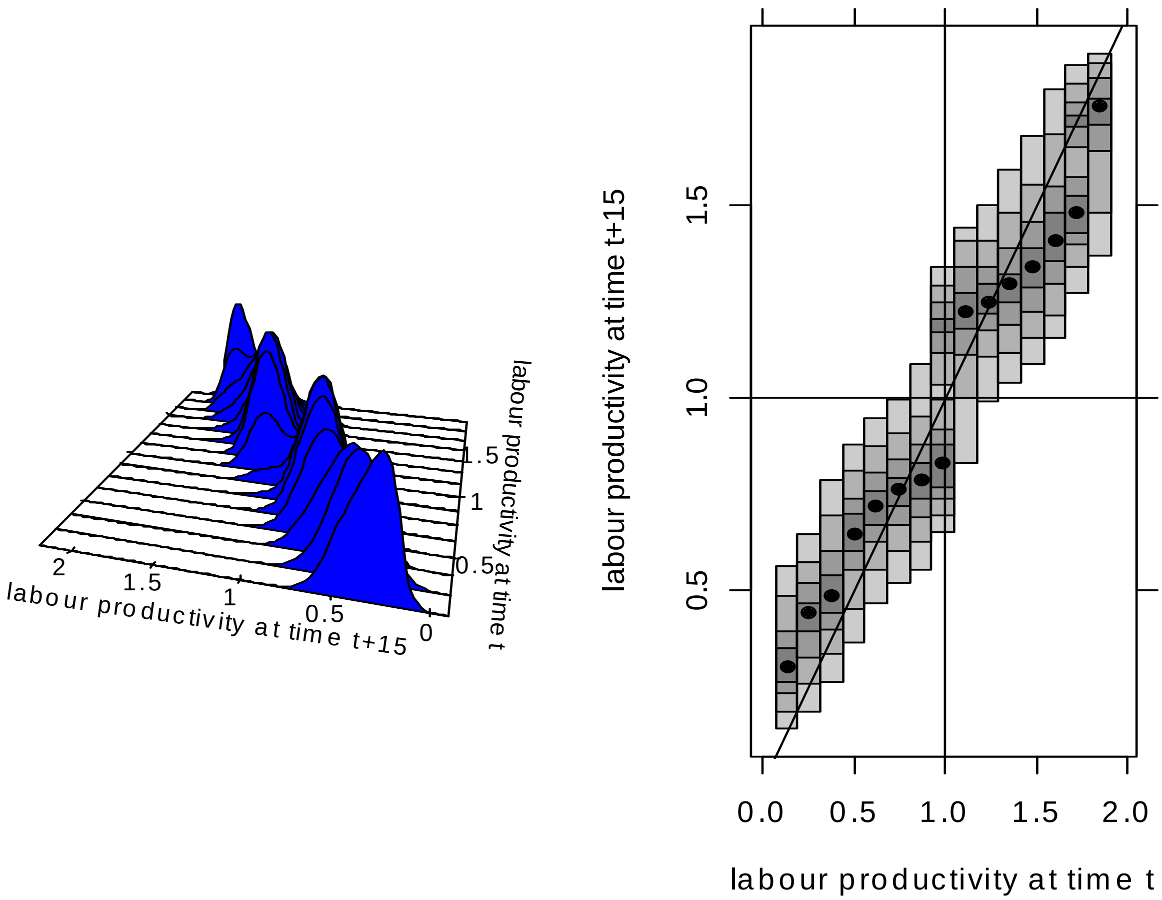
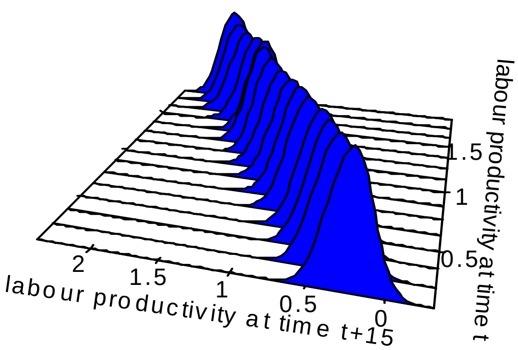


Figure 5: Intra-Distribution Dynamics of Regional Labor Productivity

Stacked density plot (left hand side panel) and HDR plot (right hand side panel) of conditional density for transitions of 15 years based on the modified conditional density estimator with mean function specified by a loess smoother



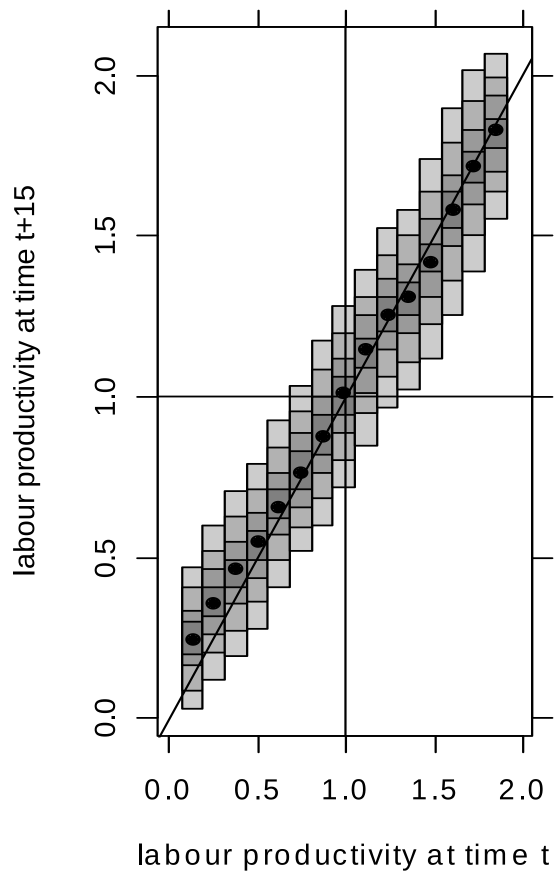


Figure 6: Intra-Distribution Dynamics of Regional Unemployment Rates

Stacked density plot (left hand side panel) and HDR plot (right hand side panel) of conditional density for transitions of 15 years based on the kernel estimator with a fixed bandwidth

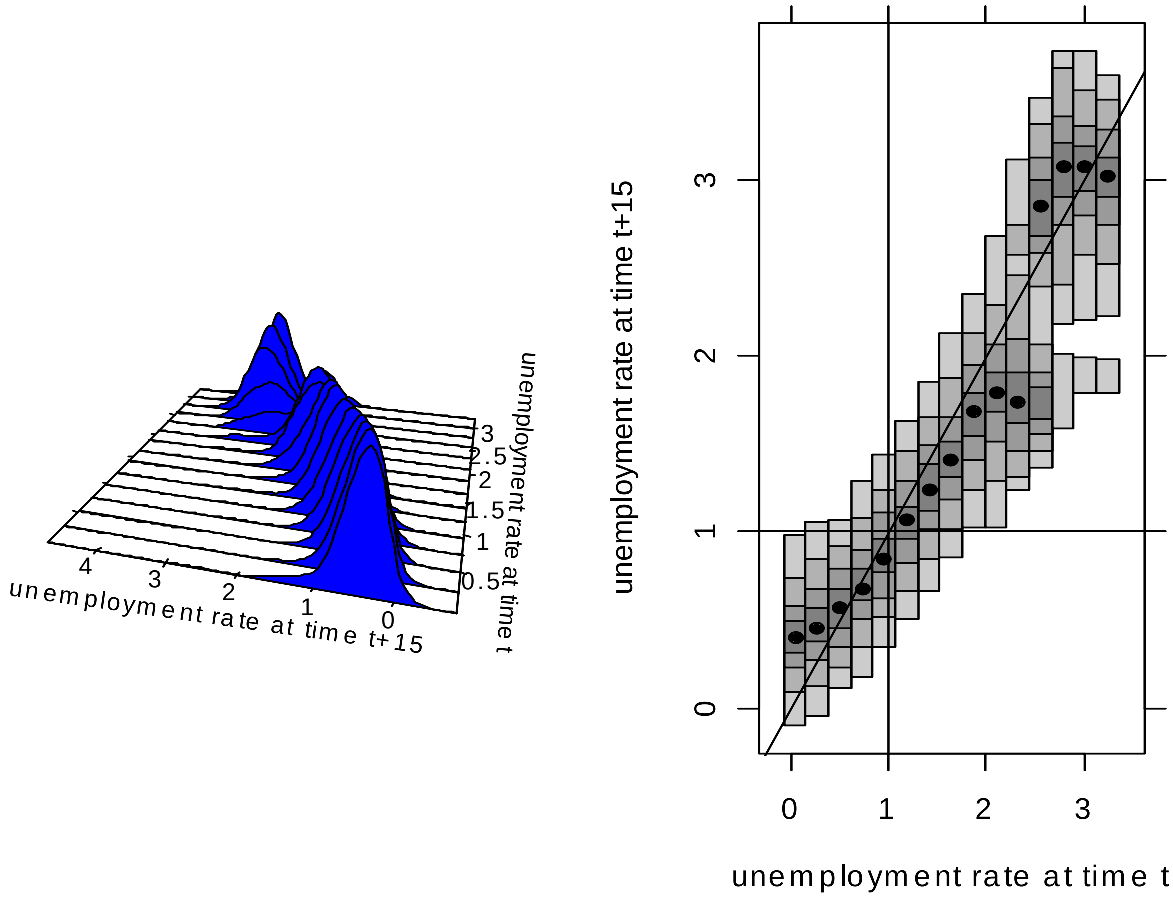


Figure 7: Intra-Distribution Dynamics of Regional Unemployment Rates

Stacked density plot (left hand side panel) and HDR plot (right hand side panel) of conditional density for transitions of 15 years based on the modified conditional density estimator with mean function specified by a loess smoother