2. EXPLORING THE ECONOMIC CONVERGENCE IN THE EU'S NEW MEMBER STATES BY USING NON-PARAMETRIC MODELS[1]

Monica RĂILEANU SZELES[2]

Abstract

This paper analyzes the process of real economic convergence in the New Member States, by using non-parametric methods, e.g. the ash-warping test, kernel density estimation and stochastic kernel. The main findings of the paper are the bimodality of income density distribution over time and across countries, the lack of convergence at a single point in time and the presence of convergence clubs in the income distribution from 1995 to 2008.They suggest the lack of absolute convergence on long term (1995-2008), and also when looking only from 2003 onwards. The paper concludes that, in comparison with the parametrical approach, the non-parametric one gives a deeper, real and richer perspective on the process of real convergence in the NMS.

Keywords: real convergence, non-parametric models, stochastic kernel, modality

JEL Classification: O11, C14

1. Introduction

This paper applies non-parametric techniques to the analysis of real economic convergence in the New Member States (NMS) area in order to provide a broader understanding of this process and new insights than those given by the conventional parametric approach, especially when the available dataset is small. Furthermore, the

[1] *This paper was prepared during a research visit at the Halle Institute for Economic Research, in collaboration with the Institute's research staff, and presented at the Institute's seminar on post-transition issues in September 2010. The author would like to thank Hubert Gabrisch, Herbert Buscher and the participants in the research seminar entitled "Exploring the Economic Convergence in the EU New Member States by using Nonparametric Models", for their helpful comments and suggestions.*

[2] Transilvania University of Brasov and CEPS/INSTEAD Luxembourg, e-mail: monica.szeles@unitbv.ro.

non-parametric approach to economic convergence represents itself a broader analysis framework in comparison with the beta convergence, for instance. With the non-parametric techniques it is possible to derive complex insights to the convergence process, which could not be revealed by the parametric models.

The analysis of convergence relies on two fundamental approaches, i.e. the beta- and sigma-convergence models (Barro and Sala-i-Martin, 1992), which are derived from the growth theory (Sollow, 1956). Both, but especially the concept of beta-convergence, have been criticized in literature for a number of reasons, such as the assumption of linearity in the growth regressions, the Galton's fallacy problem, the impossibility to detect convergence clubs, etc. (Quah, 1993, 1996; Johnson, 2000; Rassekh, Panik and Kolluri, 2001; Linden, 2002). The use of non-parametric methods could be seen as an alternative approach to the analysis of economic convergence. The reason is that they allow modeling data without presuming that they follow a normal distribution and also allow capturing short-term divergent paths that may occur in a long convergence process.

The paper is innovative in two aspects, at both methodological and empirical levels. First, it provides a tool of analysis for the process of real convergence, when the available dataset is rather small. A small dataset would make application of regression models problematic. Second, it applies a new measurement tool, i.e. non-parametric techniques, to the analysis of real convergence in the NMS area.

The empirical part is structured as follows. First, the distribution of per capita relative income among the NMS is examined using the Kernel Gaussian density function. The graphic identification of convergence clubs within the period of analysis will also be confirmed by the ash-warping procedure. The graphic analysis is enriched when adding the stochastic kernel, which illustrates transitions from one year to another, within the NMS area. The first part of the empirical study applies non-parametric models to the analysis of economic convergence, which relaxes the assumption of linearity, specific to the parametric models. It has a strong focus on graphs and aims at identifying the number of modes in the density distribution and whether the NMS converge at a single point in time.

In the second part of the empirical analysis, random effects panel regression models are used to estimate, in a parametric framework, the beta parameter. The results of the parametric regressions will be then compared to the output of the non-parametric analysis in order to see whether the two methodologies lead to the same results and to also find whether the non-parametric models provide new information about the convergence process, as compared to the standard regression.

The non-parametric methods applied on the NMS data give insights to the convergence/divergence patterns and to the existence of convergence clubs in the process of real economic convergence, without making assumptions about the income distribution form. Even though with the non-parametric models the empirical results can be improved to a greater or lesser extent depending on the data, at a methodological level they represent one step ahead in comparison with the parametric ones.

The paper concludes on the modality of income density distribution over time and across countries, states what framework is more appropriate for the analysis of real convergence (whether the parametric or the non-parametric approach) in the NMS area, analyzes the process of real convergence on long term (1995-2008) in the NMS area and examines the short-term patterns occurring in this process.

2. Theoretical insights

The growth literature provides the basic methodological instruments for the analysis and measurement of economic convergence. Most of the theories of convergence rely on the neoclassical growth model (Solow, 1956), which implies that there is a negative relationship between the initial per capita output and its growth. According to this theory, poorer countries should advance faster than richer ones and will eventually catch up with the latter, when different countries are at different points relative to their balanced growth path, have different initial conditions, but the same steady state. This relation is referred to as absolute (unconditional) convergence. When the initial capital endowment is not the only difference among the economies, but also structural differences arise, then the convergence is referred to as being relative (conditional).

The literature of convergence is based on the seminal paper of Barro and Sala-i-Martin (1992), who introduced the concept of beta-convergence, which is the speed of convergence of an economy towards its steady state. The analysis of convergence relies on two fundamental concepts: beta- and sigma-convergence. The beta-convergence, therefore, occurs when there is a negative correlation between real per capita income growth over time and its initial level and sigma-convergence when the dispersion of real per capita income across a group of economies falls over time. The two concepts are not similar and beta-convergence is not a sufficient condition for sigma-convergence.

Despite of the standard theory which assumes that poorer countries advance faster than richer ones towards a common steady state or towards their own steady state, the empirical evidence shows the increase in inequality and income divergence over time (Pritchett, 1997). This paradox is the root of the so-called convergence clubs (Baumol, 1986), which are formed by a leader and a group of followers. According to the theory of convergence clubs, the leaders preserve their supremacy in terms of development and growth over a large period of time, while only a small number of followers converge with the leader over time. Quah (1996, 1997), followed by other economists (Galor, 1996; Kumar and Russell, 2002) observed that after 1965 the world becomes polarized into two categories, the rich and the poor, this situation being referred to as twin peaks or convergence clubs. In the context of the European integration, the concept of convergence clubs suggests that the achievement of full economic or financial convergence is problematic, and a number of countries will never completely catch up with the leaders. If the polarization phenomenon experienced at the world level will also become empirical evidence at the EU level, then the achievement of real convergence in the EU space will be problematic.

The concept of beta-convergence has been criticized in literature for several reasons (Quah, 1993, 1996; Johnson, 2000; Rassekh, Panik and Kolluri, 2001; Linden, 2002). The basic criticism of beta-convergence is the possibility of Galton's fallacy, i.e. a negative value of beta may not indicate convergence of growth rates but rather regression towards the mean (Friedman, 1992; Quah, 1993). Another criticism is that the growth regression assumes the condition of homogeneity, i.e. all economies under analysis have the same rate of convergence (Bernard and Durlauf, 1996). Therefore, the process of formation of convergence clubs[3] cannot be identified by the beta-convergence theory. Quah (1993) criticizes the concept of beta-convergence arguing that it provides no information on the way that poor economies are catching up with the richer ones. Friedman (1992) considers that the true test of convergence is a decline in the variance among individual observations. This is in fact the sigma-convergence.

3. Data

The empirical research focuses on the NMS and it is based on the data collected from the World Economic Outlook Database April 2010 (IMF). The data used here are the NMS gross domestic products at purchasing-power-parity (PPP) per capita, expressed in current USD, from 1995 to 2008. The NMS considered in the paper are Poland, Hungary, the Czech Republic, Slovenia, the Slovak Republic, Estonia, Latvia, Lithuania, Romania and Bulgaria.

In Table 1, the summary statistics show that the average of per capita GDP levels increased in the period of analysis, by a five-year growth rate of around 42% from 1995 to 2010. From 2006 to 2010, the IMF predicts the slowing-down of the five-year growth rate. Overall, the mean levels of per capita GDP in the NMS are increasing, indicating at a glance the process of catching up with the OMS[4].

Table 1

Summary statistics by sub-periods, 1995-2010

Sub-intervals	Mean	St. dev.	Min.	Max.
1995-2000	9149.36	827.41	7957.9	10435.1
2001-2005	13027.92	1486.12	11176	15292.3
2006-2010	18403.94	845.38	16977.8	19544.2

Note. For 2009 and 2010 we used the IMF predictions.

The relative income is the main indicator investigated in the empirical section, in order to ensure the comparability across countries and across years. The relative income is constructed in two ways, as to facilitate conducting both a cross-sectional and a longitudinal analysis in the empirical part. But the paper is mainly concerned with the cross-sectional representation, which requires calculating the relative income by dividing the NMS' GDP per capita levels by their mean in the same year, and then taking the natural logarithm of this value. The longitudinal approach to relative incomes is followed only in section 4c, where we explain the methodology of its construction.

---

[3] *The "convergence clubs" (Quah, 1997) denote identification of two groups of economies in the analysis of convergence: a group of convergent economies and a group of divergent economies.*

[4] *OMS stands for Old Member States.*

4. Multimodality of income distribution density

The traditional parametric models used in the analysis of income convergence are based on the assumption that data follow a certain distribution, e.g. a normal distribution. The *beta* approach, for instance, relies on another assumption that does not always hold in practice, i.e. the assumption of linearity in the relation between the economic growth and the logarithm of initial income. Due to these assumptions, the parametric models are not able to capture the process of real convergence when this process is characterized by income convergence clubs, short-term divergent paths and, in general, by non-linear dynamics.

This section examines whether the non-parametrical adjusted density is characterized by unimodality or multimodality. This could give insights regarding the existence of income convergence clubs within the NMS area in the period of analysis. All tests used in this section are applied to the logarithm of relative income per capita.

In the broad framework of the non-parametrical models and tests, several procedures have been developed to assess the modality of a univariate distribution (Cox, 1966; Good and Gaskins, 1980; Silverman, 1981; Hartigan and Hartgen, 1985). While some of the methods depend on the arbitrary choice of the scale of the effects studied (Cox, 1966; Good and Gaskins, 1980), others have incorporated automatic ways of making this choice (Silverman, 1981).

In order to test the multimodality of the relative income in the NMS, several tests have been applied, all of them relying on the Gaussian function. The intention of applying several tests to examine the multimodality of relative income density, motivated by the aim of obtaining robust results, was conditioned by the data availability and constraints. For this reason, only the results of two tests are discussed and reported here.

In the broad space of kernel density estimation the number of modes depends on the chosen bandwidth. The bandwidth is a smoothing parameter controlling for variance in the kernel probability density function, which is normally taken as a standard Gaussian function with mean zero and variance 1. For this reason, the first step in the analysis of multimodality was the selection of optimal bandwidths for each year of our analysis, using the bandwidth rules developed by Salgado-Ugarte *et al*. (1995). Silverman's Gaussian kernel bandwidths are taken as reference values in the construction of the tests described below.

a) Ash-warping

The ASH-WARPing procedure is used in this paper to estimate the non-parametric univariate density as a smoother of histograms, but also to get information about the modality in the density distribution. This procedure is derived from the general framework called WARP (weighted averaging of rounded points) developed by Hardle and Scott (1988) and it is based on the averaged shifted histogram ASH (Scott, 1985).

Despite of the theoretical aspects (Scott, 1992), the empirical evidence has shown that, when defining the histogram, the choice of origin influences the result (Silverman 1986). To solve this problem, Scott (1985) proposed averaging several histograms with different origins to produce the average shifted histogram (ASH).

In the presentation of the ash-warping method, we start by defining first the histogram[5].

If all *n* observations of a variable belong to the interval [0, Kh) and if the interval is partitioned into *K+1* bins, with *h* being the width of bins, then the *k*th bin, B_k, is defined as:

$$B_k = [kh, (k+1)h) , \qquad k = 0,\ldots,K \tag{1}$$

The histogram is defined as:

$$\hat{f}(x) = \frac{v_k}{nh} = \frac{1}{nh}\sum_i I_{(t_k,t_{k+1})}\,(x_i) \tag{2}$$

Where, v_k is the number of observations in B_k, and *I* is the indicator function, equal to one when x_i lies in the specified interval and zero otherwise.

Let be *M* a collection of hisograms $\widehat{f_1}$, $\widehat{f_2}$, ..., $\widehat{f_M}$, having the bin width *h*:

$$t_0 = 0, \frac{h}{M}, \frac{2h}{M}, \ldots, \frac{(M-1)h}{M} \tag{3}$$

The following restriction can be applied on the previous relation:

$$x_i \geq \frac{(M-1)h}{M} \tag{4}$$

With the restriction above, the unweighted ASH can be expressed as:

$$\hat{f}(\cdot) = \widehat{f_{ASH}}(\cdot) = \frac{1}{M}\sum_{i=1}^{M}\hat{f}_i(\cdot) \tag{5}$$

In a generalized form, ASH can be defined as:

$$f(x;\ M) = \frac{1}{n}\sum_{i=1-M}^{M-1}\left(1 - \frac{|i|}{M}\right)v_{k+i} \ \ for\ x\epsilon B_k \tag{6}$$

Linear interpolation schemes are sometimes used to make ASH continuous. They produce the frequency polygon of the averaged shifted histogram (FP-ASH).

ASH is a particular case of the general method WARP (weighted average of rounded points), which is defined as:

$$\hat{f}(x; M) = \frac{1}{nh}\sum_{|i|<M}^{M-1} w_{M(i)}\, v_{k+1} \qquad for\ x \in B_k \tag{7}$$

where: $w_{M(i)}$ denote the weighting operation and function, and *M* represents the number of shifted histograms to average.

In fact, the ash-warping procedure involves three steps: (1) binning the data; (2) calculating the weights, and (3) weighting the bins. Different weight functions can be used to approximate the kernel density estimator and, finally, the data are reduced to a list of bin counts along with their midpoints. The density estimate in each bin is computed as the product of the bin count and the weight.

In this paper, we have applied the ash-warping procedure to the NMS relative incomes, using the corresponding Silverman's Gaussian kernel bandwidths for each year of our analysis. The results of this procedure indicate that from 1995 to 2008 the kernel density of relative income is bimodal in 9 years and unimodal in 5 years. After 2002, the income density is bimodal each year. A detailed situation of the density modality is presented in Appendix 1 (Table 4). This is a first indication that the NMS

[5] *The presentation of ash-warping methodology is based on the Isaias Hazarmabeth Salgado-Ugarte, Makoto Shimizu, and Toru Taniuchi's paper „ASH, WARPing, and kernel density estimation for univariate data" (Stata Technical Bulletin, July 1995).*

do not tend to converge on long term at a single point, or at least that the NMS convergence cannot be seen as a gradually continuous process.

The modality of income density distribution can be also analyzed using the ash-warping procedure with the help of graphical presentation. In Figures 1, 2 and 3, the Gaussian kernel density estimation is represented for the years 1995, 2002 and 2007, by using Silverman's optimal bandwidth values. Figure 1 suggests the unimodal structure of the distribution function in 1995, while Figures 2 and 3 suggest the bimodality of density distribution[6] in 2002 and 2008.

Figure 1

Kernel density estimation, 1995

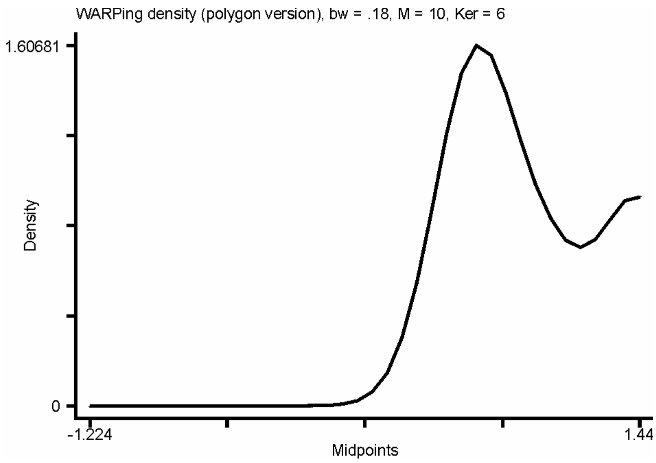
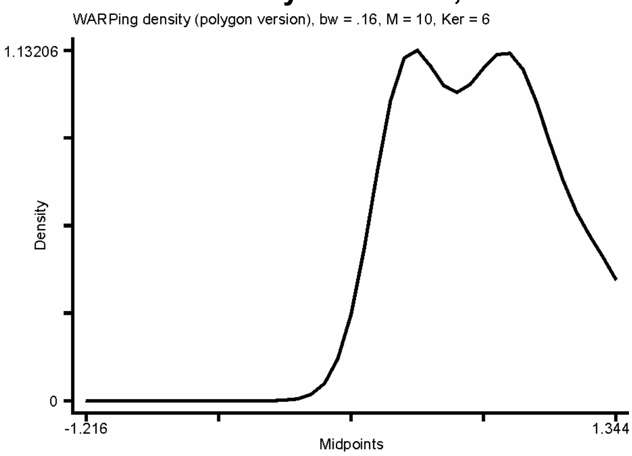


Figure 2

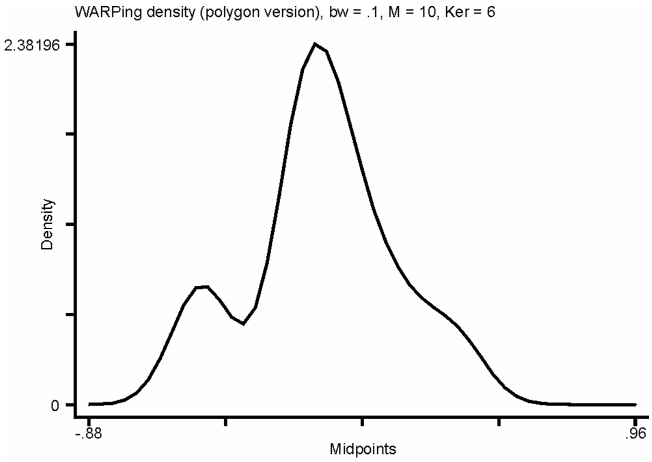
Kernel density estimation, 2002



[6] *Other years have been examined as well, but only the years signifying the beginning, the end and the switching points in the distribution function have been reported in the paper. Anyway, from 2002 to 2008, the distribution is bimodal.*

Figure 3

Kernel density estimation, 2008



b) Kernel density estimator

The kernel density estimators belong to the class of non-parametric estimators, i.e. have no fixed structure and depend upon all the data points to produce the result. In comparison with the histogram, they smooth out the contribution of each observed data point over a local neighborhood of that data point. The contribution of data point x_i to the estimate at the arbitrary point x depends upon the shape of the kernel function adopted and the width (bandwidth) accorded to it.

A typical form for the kernel density estimator is:

$$\hat{f}(x) = \frac{1}{hn}\sum_{i=1}^{n} K\left(\frac{x-X_i}{h}\right) \quad (8)$$

where: $\hat{f}(x)$ is the density estimation of the variable *x*, *n* is the number of observations, *h* is the bandwidth (smoothing parameter) and *K(·)* is the smooth and symmetric kernel function integrated to unity.

The importance of bandwidth relies on the fact that the size of bandwidth chosen for kernel density estimation determines the degree of smoothing produced. When low values are attributed to *h*, the estimated density for the data is not as smooth. The kernel density estimator uses fixed bandwidths and, thus, the estimation is sensitive to any low count interval of the distribution. Anyway, choosing the width of the bandwidth *h* is the main problem. Several procedures have been proposed in literature to find the optimal bandwidth. They range from the subjective assessment of a pleasant smooth of the result (Tarter and Kronmal, 1976) until objective methods that start with the analysis of the shape of the true density distribution. In particular, when a Gaussian kernel is used as reference function, the minimization of the mean integrated squared error (MISE) allows deriving *h* (Tukey, 1977; Scott, 1979; Silverman, 1978, 1986).

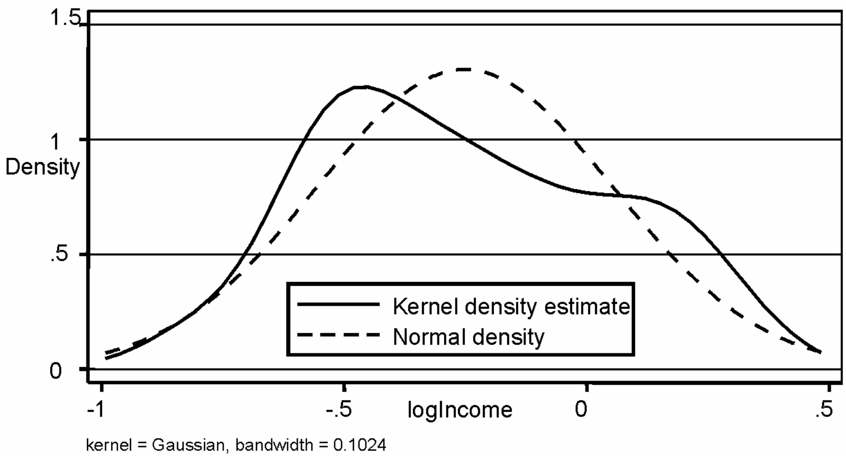
Beside the warping method, which was applied at point (*a)* of this section, the univariate kernel density is also used in the paper, with the same aim of providing a graphical representation of income density distribution. This method also allows for examining the modality of distribution. The kernel density is estimated using the Gaussian function and the results are presented in Appendix 2 (Figures 7.1-7.4). To ensure the comparability of results, the same years have been analyzed here as we did when applying the warping procedure. Both graphical representations yield the same results, with the exception of the year 2002, for which the warping method suggests a bimodal distribution, while the kernel density indicates a unimodal distribution.

In a second step, the kernel density was used to examine the modality of the relative income distribution not only among countries, but also over time, within each country. The density of the natural logarithm of relative incomes has been estimated using the Gaussian kernel. In contrast to the previous tests, this time the relative income is constructed as to reflect the longitudinal dimension of analysis, i.e. by dividing the annual values of per capita GDP by their mean, separately for each country. Due to this normalization process, a zero value on the horizontal axis indicates per capita relative income equal to the national mean of the entire period of analysis.

The results are shown in Appendix 3, where for each graph the period considered is 1995-2008. For all countries, the kernel density estimates indicate a bimodal distribution. The “twin peaks” shapes in Figures 8.1-8.10 are referred to in literature as “convergence clubs” (Quah, 1996). The density shapes give insights to the income polarization in the NMS during the transition period. Although two modes have been identified in the distribution of income densities for each country, they reflect different patterns over time. In the case of Bulgaria, Romania, the Czech Republic, the Slovak Republic and Poland, the relative income densities have two symmetric modes around the national means, while Estonia, Latvia, Lithuania and Hungary have a large mode above the national mean and a smaller one below the national mean. This reflects a more favorable income distribution for the first group of NMS from 1995 to 2008. A particular aspect regards Romania, which has a larger mode located below the national mean, and a smaller one above the mean. This suggests a higher concentration of annual incomes in the low income area.

Figure 4

Kernel density estimate, NMS area, 1995-2008



When considering a longitudinal approach not at the country level this time, but at the level of the entire NMS area, a bimodal distribution occurs again. This aspect mainly reflects the bimodality of income distribution among each country in the transition period, and, only to a lesser extent, the bimodality of income distribution across countries (see Figure 4).

c) Stochastic kernel density

The stochastic kernel density allows estimating the conditional density function, which is a transition function obtained using kernel density estimation. In contrast to other techniques specific to the measurement of convergence (beta- and sigma-convergence), it uses all information in the data, i.e. the first period, the last period and the transition process. For instance, the beta-convergence considers the transition relative to the first period, but neglects the last period, while the sigma-convergence looks at all observed periods, but only in terms of standard deviations (Weber, 2009).

In the next paragraphs of this section we introduce the stochastic kernel, starting from the density distribution. The density distribution φ_{t+1} of a variable *x* follows a first order Markov process:

$$\varphi_{t+1} = M \cdot \varphi_t \qquad (9)$$

The operator *M* maps the transition of variable *x* from its distribution in the state *t* to its distribution in the state *t+1*. It assumes either a finite number of states in φ_t distribution using the Markov Transition Matrix (Shorrocks, 1978) or using a continuous state formulation in the Stochastic Kernel (Quah, 1996). In a discrete version of the model, the operator *M* is determined by partitioning the set of possible income values into a finite number of intervals. The properties of *M* are described by a Markov chain transition matrix whose (*i*, *j*) entry is the probability that a country in state *i* transits to state *j* in the space of per capita GDP, in one time step. As the per capita GDP are stocked in a continuous variable, the transition probability matrix will be a matrix of a continuous of rows and columns. Therefore, the operator *M* can be seen as a stochastic kernel or a transition function, and real convergence can be seen as the shape of the income distribution at time $t+\tau$ over the range of incomes observed at time *t*.

According to Quah (1996), if *u* and *z* are elements of *B* and also probabilities measures in (R, R), the Stochastic Kernel is a function relating *u* and *z* by the function $M_{(u,z)}$: (R, R) -> (0, 1), such that:

(i): For each y Є R, $M_{(u,z)}(y, \cdot)$ is a probability measure in (R, R);

(ii): For each *A* Є R, $M_{(u,z)}(\cdot, A)$ is a measurable function in R;

(iii): For each *A* Є R, it is valid that $u(A) = \int M_{(u,z)}(y, A)\, dz(y)$

At an initial point in time, for given *u*, there is some fraction of economies *dz(u)* with incomes close to *u*. When being normalized as to be a fraction of the total number of economies, the number of economies in that group whose incomes fall in the subset *A* can be written as *M (y, A)*. The integral $\int M_{(u,z)}(y, A)\, dz(y)$ indicates the number of economies that end up in state *A*, regardless of their initial income levels. Stochastic Kernel *M* can be therefore seen as the description of transitions from state *y* to any other portion of the underlying state space R.

According to Arbia *et al*. (2005), the Stochastic Kernel can be also written as:

$$\varphi_{t+\tau}(y) = \int_0^\infty f_\tau(y|x)\varphi_t(x)dx \tag{10}$$

where: y is the relative per capita income in period $t+\tau$, x is the relative per capita income in period t and f_τ $(y|x)$ is the conditional density given the relative income in period t.

One of the most popular kernel functions is the standard Gaussian function with zero mean and 1 variance.

$$f(x) = \int_{-\infty}^{+\infty} f(y,x)dy = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_i)^2}{2h^2}} \tag{11}$$

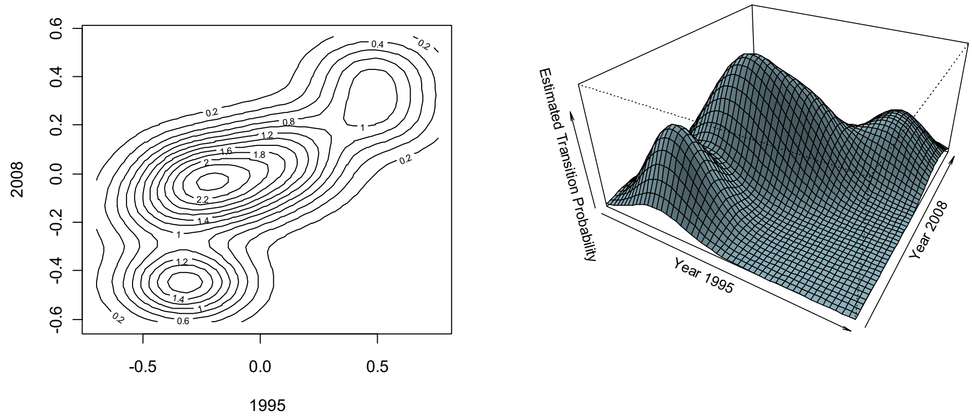
where: x is a random variable and h is the smoothing parameter called bandwidth.

The stochastic kernel represented in Figures 5 and 6 show the transition probability associated with the change in the distribution of relative incomes occurring from one period to another. For each transition considered here, two perspectives were analyzed, one being a two-dimensional representation, and another being a three-dimensional one. Anyway, they both indicate the formation of convergence clubs by highlighting "peaks" in the income distribution.

Figure 5 indicates three significant peaks in the stochastic kernel, which have occurred in the transition from 1995 to 2008. One of them is larger than the other two and reflects the transition of a sub-group of NMS countries from the poor income category to a new middle income category. This situation reflects an improvement in the relative income distribution among the NMS, since the intermediate income area, which was absent in 1995, becomes the most important category in 2008. The other two categories capture the convergence among the low income countries and the convergence towards higher incomes, respectively. This picture of stochastic kernel allows rejecting the hypothesis of relative income convergence to a single point from 1995 to 2008.

Figure 5

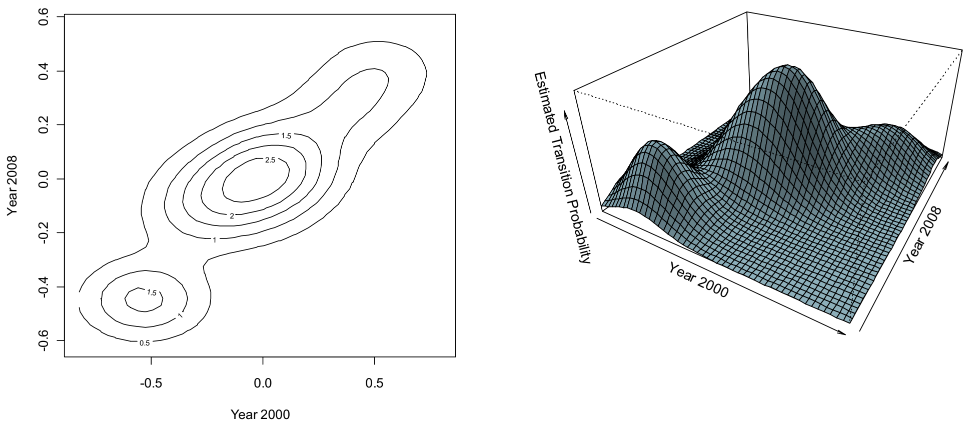
Stochastic kernel 1995-2008



Even though several sub-periods of time could be particularly examined, the paper analyzes only the sub-interval 2000-2008, in order to capture the changes in the NMS income density distribution during the global economic crisis. The stochastic kernel for this period of time is presented in Figure 6. In comparison with the period 1995-2008, from 2000 onwards the high income category became smaller, because part of the countries initially located in this category has moved into the intermediary income category by 2008. This change has reduced the number of convergence clubs to 2, with the disappearance of the high income category and the stability of the low income category over time.

Figure 6

Stochastic kernel, 2000-2008



In conclusion, the stochastic kernel analysis does not reveal, in any of the cases studied here, the convergence at a single point until 2008. The most significant patterns identified by this method in the NMS during the transition period are the emergence of a ”middle class of the NMS”, bipolarization towards the low and intermediate income categories, stability of the small, but constant poor income group and shrinking of the high income group, due to the occurrence of the global economic crisis.

The non-parametric analysis carried in Section 4 shows that the density of income distribution among the NMS cannot be considered as being unimodal from 1995 to 2008. The bimodality of income distribution arises among countries, as well as within countries. The bimodal structure of income distribution and the convergence peaks in the stochastic kernel suggest the lack of real economic convergence within the NMS area, as well as the inappropriateness of parametric models applied on our dataset.

5. A parametric approach to real convergence in the New Member States

The convergence literature is based on the seminal work of Barro and Sala-i-Martin (1991) who introduced the concept of beta-convergence. This concept implies a negative relationship between the growth rate and the initial income per capita, due to the assumption of marginal decreasing productivity. Despite of its very broad used in the empirical work on convergence, the beta-convergence approach has been criticized in the literature, one reason being the impossibility to capture the convergence clubs in the income distribution (Quah, 1996).

In this section, we apply random effects panel models to examine the unconditional beta-convergence in the NMS area. These results will be then compared with the output from the nonparametric approach, to get both empirical and methodological insights, on the basis of the IMF data and predictions about the NMS. The relevance of alternative spline regression techniques is limited here by the data availability[7].

The estimates of the first random effects panel regression, where the dependant variable is the growth rate between 1995 and 2008 and the independent variable is the logarithm of per capita GDP in 1995, are reported in Table 2.

Table 2

Random-effects panel growth regression (1995-2008)

Variable	Coefficient	St. err.
Per capita GDP 1995 (log)	-.0056	.0056
Constant	.0667	.0505
rho[8]	.17	
Nr. of obs.	140	

*Note. *** significance at 1% level, ** significance at 5% level, * significance at 10% level.*

When the entire period of time is considered, i.e. 1995-2008, the beta coefficient is very low, negative and not significant, suggesting that on long term there is no absolute convergence in the NMS area. When looking just at the period of time from 2003 onwards, the regressions still yield negative and low beta coefficients, which gradually improve in the level of significance[9]. Table 3 presents the estimates of the random effects panel regression which runs between 2003 and 2008. This time, the *beta* coefficient has a low negative value, but becomes slightly significant, indicating a slow process of convergence after 2003 in the NMS area.

[7] *The small working dataset of the NMS' GDP per capita between 1995 and 2008 makes the use of cubic splines inappropriate. If possible to be applied, this technique would allow for approaching the relationship between the dependant and independent variables on separate income ranges. By using the splines, the analysis could have revealed different patterns of convergence or divergence within this period of time. Without this transformation of this explanatory variable, the whole process of convergence will be summarized in the regression analysis by one coefficient, i.e. the beta coefficient.*

[8] *The rho statistics indicates the proportion of the total variance attributed to the panel level variance component.*

[9] *A set of regression models, starting from different years after 2003 and ending in 2008 in all cases, are tested, and all of them indicate a slow process of convergence with the gradual improvement in the level of significance when pushing ahead the first year of the regression. From this list, only one regression is reported here, in Table 3, as they all lead to the same empirical finding.*

Table 3

Random-effects panel growth regression (2003-2008)

Variable	Coefficient	St. err.
Per capita GDP 1995 (log)	-.0107*	.0066
Constant	.1224	.0596
rho	.64	
Nr. of obs.	60	

*Note. *** significance at 1% level, ** significance at 5% level, * significance at 10% level.*

In conclusion, the absolute convergence process on long term is not clearly suggested by the conventional beta approach. The evidence of convergence becomes significant, but it is still weak when running the analysis only from 2003 onwards. In comparison with the non-parametric models which suggest the absence of absolute convergence on long term and, also after 2002, the parametric regression indicates lack of absolute convergence on long term and a weak absolute convergence after 2002. In this light, the non-parametric analysis brings not only new and additional findings about the process of absolute convergence, but also different results. The presence of convergence clubs from 1995 to 2008, as well as the bimodality of income density distribution in each year from 2002 onwards, proves that the process of absolute convergence, identified in the last 6 years of our analysis by the linear regression models is not real. In addition, the empirical results obtained in this section show that the parametric analysis provides just few information about the changes occurring over time and the progress towards unconditional convergence, being unable to capture the changes in the income density distribution across countries, from one year to another.

6. Conclusions

This paper applies several non-parametric techniques to the analysis of absolute convergence in the NMS area, being oriented to provide robust conclusions, at both methodological and empirical level. Despite the fact that the methodological orientation is the primary focus of this paper, the conclusions are derived mainly from the empirical findings.

The non-parametric analysis of income density distribution in the NMS area between 1995 and 2008, as well as the parametric analysis applied in the same period of time, indicate the lack of real absolute convergence on long term, with short periods of convergence and divergence on short term. The divergence represents a yearly characteristic for the NMS area from 2002 to 2008, when using non-parametric techniques. This short-term characteristic is early signaled from 2000. Apart from the non-parametric analysis, the parametric analysis finds evidence of a weak process of absolute convergence on short term, i.e. from 2002 to 2008. This result revealed by the random effects regression should be regarded with caution, because of the presence of convergence clubs from 1995 to 2008 and also because of the bimodality of income density distribution each year after 2002.

During the transition period, the income distribution has a bimodal structure in the NMS area, which is also graphically illustrated by the convergence clubs (in the stochastic kernel analysis). This aspect is mainly driven by the bimodality that occurs in the income density distribution of each NMS across years. Even so, there are years and periods of time when the income distribution among countries also has a bimodal structure. The existence of more convergence clubs in the income density distribution, either in some years or in the transitions over time, gives insights to the convergence patterns during the period of analysis. In this light, the non-parametric analysis reveals new findings in comparison with the conventional parametric regressions, which, in turn, reduces the description of the entire convergence process at one coefficient. Even though both the parametric regressions and the nonparametric techniques suggest divergence from 1995 to 2008, the latter provides a broader framework of analysis, and becomes more credible when the number of observations in the dataset is rather small.

At the country level, the density distribution of per capita GDP is bimodal, which is not surprising as during the transition period these countries have continuously grown up and have experienced changes in the income distribution. The transitions illustrated by the stochastic kernel show that within the group of NMS there is a trend towards the mean income. This could be interpreted as the formation and consolidation of a "middle class" of NMS during the transition period, which grew up especially during the global economic crisis. This consolidation process is mainly and gradually driven by the shrinking of the high income NMS. Despite the changes produced in the upper middle income category of NMS, the category of "poor countries" remains stable over time. When looking at the entire period of transition, these changes are not sufficient to sustain the process of real convergence in the NMS area.

In conclusion, when the income density distribution is not normal, or "too non-linear", the nonparametric approach can provide complex, real and "different" information about the salient or hidden aspects of distribution or about the short-term dynamic patterns. In our paper, the non-parametric output reveals more and partially different features of the real convergence process, when being compared with the conventional beta approach.

References

Arbia, G. Basile G. and Piras, G. (2005), "Analyzing intra-distribution dynamics: A reappraisal", *REAL Working Paper Series*, 05-T-11.

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

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

Bernard, A. and Durlauf, S. (1996), „Interpreting tests of the convergence hypothesis", *Journal of Econometrics*, 71: 161--173.

Cox, D.R. (1966), "Notes on the analysis of mixed frequency distributions", *British Journal of Mathematical Statistics and Psychology*, 19: 39-47.

Friedman, M. (1992), “Do Old Fallacies Ever Die?”, *Journal of Economic Literature* 30(4): 2129–2132.

Härdle, W. and Scott, D.W. (1988), “Smoothing in low and high dimensions by weighted averaging using rounded points”, *Technical report* 88-16, Rice University.

Galor, O. (1996), "Convergence? Inferences from Theoretical Models," *Economic Journal*, Royal Economic Society, 106(437): 1056-69.

Good, I.J. and Gaskins, R.A. (1980), “Density estimation and bump-hunting by the penalized likelihood method exemplified by scattering and meteorite data”, *Journal of American Statistics Association*, 75: 42-56.

Johnson, P.A. (2000), “A non-parametric analysis of income convergence across the US”, *Economic Letters,* 69: 219-223.

Kumar, S. and Russell, R.R. (2002), “Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence”, *American Economic Review*, 92: 527-548.

Linden, M. (2002), “Trend models testing of growth convergence in 15 OECD countries, 1946-1997”, *Applied Economics,* 34: 133-142.

Pritchett, L. (1997), „Divergence”, Big Time, *Journal of Economic Perspectives*, 11(3): 3-17.

Quah, D. (1993), “Empirical cross-section dynamics in economic growth”, *European Economic Review,* 37: 426-434.

Quah, D. (1996), “Empirics for economic growth and convergence”, *European Economic Review,* 40: 1353-1375.

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

Rassekh, F., Panik, M.J. and Kolluri, B.R. (2001) "A test of the convergence hypothesis: The OECD experience, 1950-1990," *International Review of Economics & Finance*, 10(2): 147-157.

Salgado-Ugarte, I.H., Shimizu, M. and Taniuchi, T. (1995), “Practical rules for bandwidth selection in univariate density estimation”, *Stata Technical Bulletin*, 27: 5-19.

Salgado-Ugarte, I.H., Shimizu and M., Taniuchi, T. (1995) “ASH, WARPing, and kernel density estimation for univariate data”, *Stata Technical Bulletin*, 26:23-31.

Scott, D.W. (1979), “On optimal and data-based histograms”, *Biometrika*, 66: 605-610.

Scott, D.W. (1985), “Averaged shifted histograms: Effective nonparametric density estimators in several dimensions”, *Annals of Statistics*, 13: 1024-1040.

Scott, D.W. (1992), *Multivariate Density Estimation: Theory, Practice, and Visualization*, New York: John Wiley & Sons.

Silverman, B.W. (1978), “Choosing the window width when estimating a density”, *Biometrika*, 65: 1-11.

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

Silverman, B.W. (1986), *Density estimation for statistics and data analysis*, London: Chapman & Hall.

Shorrocks, A.F. (1978), “The measurement of mobility”, *Econometrica* 46(5): 1013-1102.

Solow, R. (1956), “A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70: 65-94.

Tarter, M.E. and Kronmal, R.A. (1976) “An introduction to the implementation and theory of nonparametric density estimation”, *The American Statistician*, 30: 105-112.

Tukey, J.W. (1977), *Exploratory data analysis*. Reading, MA. Addison-Wesley.

Weber, D. (2009), "European financial market integration. A closer look at government bonds in Eurozone countries”, *Working Paper* D.1.1b FINESS.

Appendix 1

Table 4

ASH Warping test

Year	Number of modes in the non-parametric density	Bandwidth *h*
1995	1	0.18
1996	2	0.17
1997	1	0.19
1998	1	0.19
1999	1	0.19
2000	2	0.19
2001	1	0.17
2002	2	0.16
2003	2	0.14
2004	2	0.13
2005	2	0.10
2006	2	0.10
2007	2	0.10
2008	2	0.10

Note. The number of modes is determined by using Silverman's Gaussian kernel bandwidth. These bandwidths are reported in the last column.

Appendix 2

Kernel density estimates by year

Figure 7.1

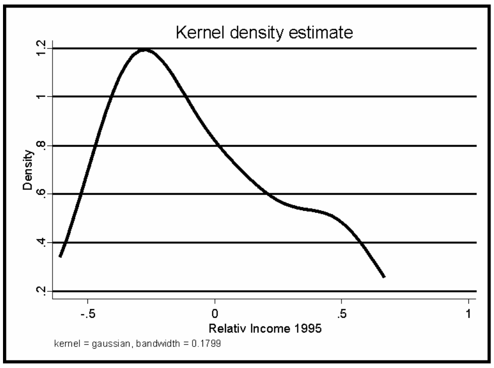


Figure 7.2

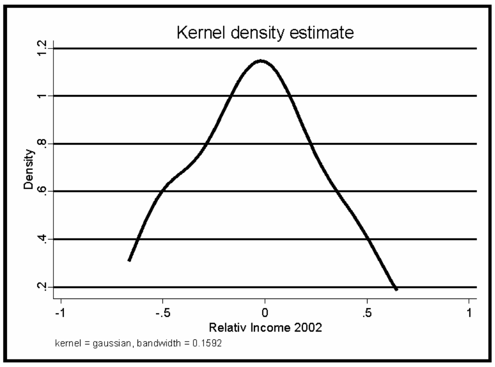


Figure 7.3

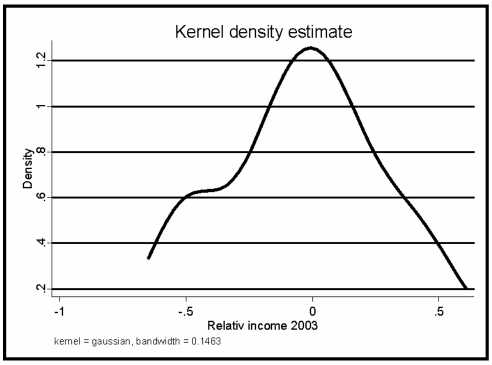
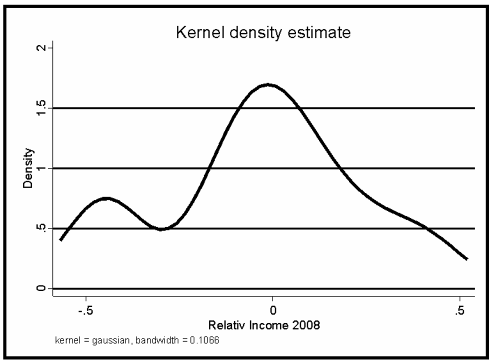


Figure 7.4



Appendix 3

Kernel density estimates by country

Figure 8.1

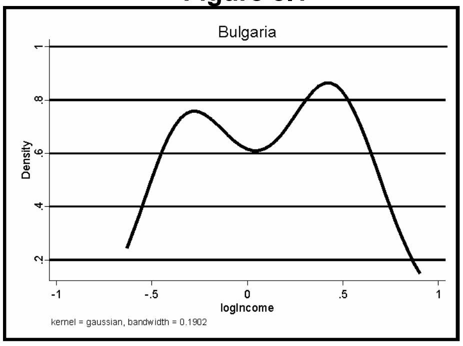


Figure 8.2

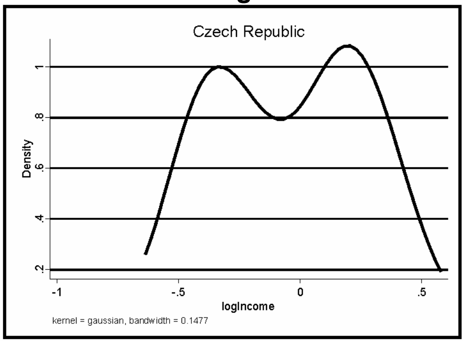


Figure 8.3.

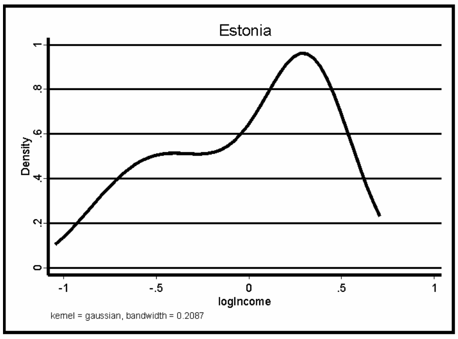


Figure 8.4.

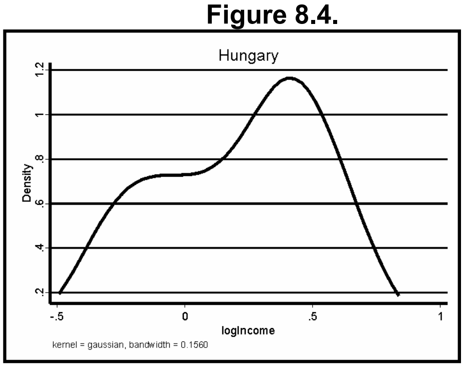


Figure 8.5

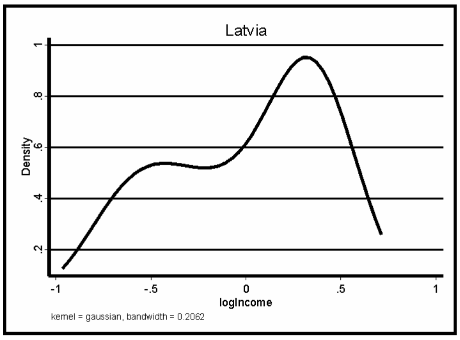


Figure 8.6

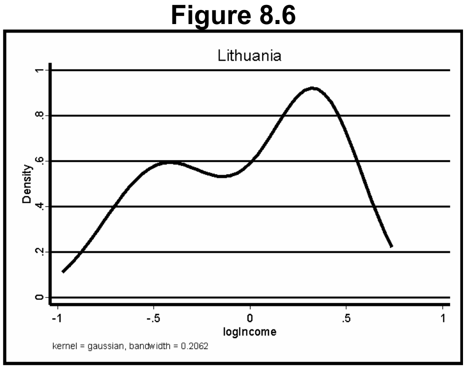


Figure 8.7

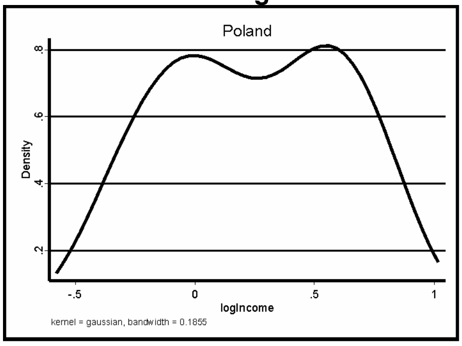


Figure 8.8

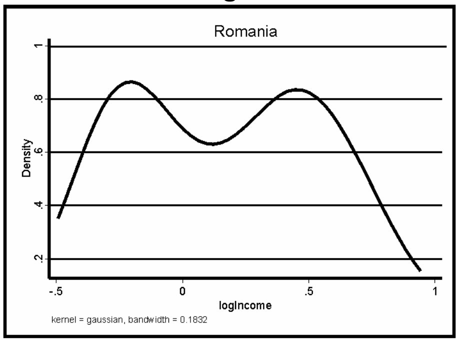


Figure 8.9

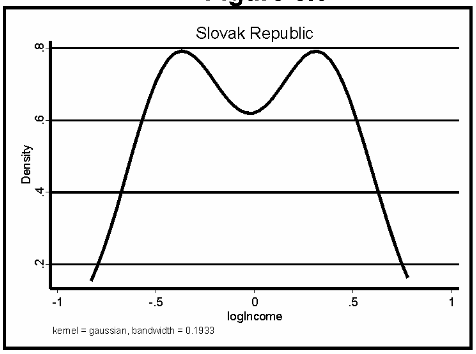


Figure 8.10