Holzmann, Hajo; Vollmer, Sebastian; Weisbrod, Julian

Working Paper

Twin peaks or three components?: analyzing the world's cross-country distribution of income

IAI Discussion Papers, No. 162

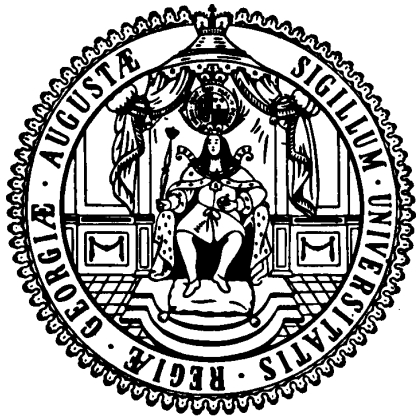
Provided in Cooperation with:
Ibero-America Institute for Economic Research, University of Goettingen

Suggested Citation: Holzmann, Hajo; Vollmer, Sebastian; Weisbrod, Julian (2007) : Twin peaks or three components?: analyzing the world's cross-country distribution of income, IAI Discussion Papers, No. 162, Georg-August-Universität Göttingen, Ibero-America Institute for Economic Research (IAI), Göttingen

This Version is available at:
https://hdl.handle.net/10419/27452

Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)



Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

Nr. 162

Twin Peaks or Three Components? – Analyzing the World's Cross-Country Distribution of Income

Hajo Holzmann, Sebastian Vollmer, Julian Weisbrod

September 2007

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173
e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

Twin Peaks or Three Components? - Analyzing the World´s Cross-Country Distribution of Income

Hajo Holzmann*
Sebastian Vollmer†
Julian Weisbrod‡

Abstract

In this paper we analyze the world´s cross-national distribution of income and its evolution from 1970 to 2003. We argue that modeling this distribution by a finite mixture and investigating its number of components has advantages over nonparametric inference concerning the number of modes. In particular, the number of components of the distribution does not depend on the scale chosen (original or logarithmic), whereas the number of modes does. Instead of so-called twin-peaks, we find that the distribution appears to have only two components in 1970-1975, but consists of three components from 1976 onwards, a low, average and high mean-income group, with group means diverging over time. Here we apply recently developed modified likelihood ratio tests for the number of components in a finite mixture. The intra distributional dynamics are investigated in detail using posterior probability estimates.

JEL classification: C12, O11, O47, F01

Keywords: cross-national income distribution; mixture models; modified likelihood ratio test; nonparametric density estimation

*Institute for Stochastics, University of Karlsruhe

†Ibero-America Institute for Economic Research and Center for Statistics, Georg-August-University Göttingen. Corresponding author: Georg-August-University Göttingen, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany. Phone: +49551398171, Fax: +49551398173, Email: Sebastian.Vollmer@wiwi.uni-goettingen.de.

‡Department of Economics, Georg-August-University Göttingen

1 Introduction

The behavior of the cross-national income distribution is for many reasons of great interest. In particular, the development of twin or even more peaks would characterize a world of growing cross-country average income polarization and suggest the existence of multiple equilibria. Numerous papers (Barro, 1991; Barro and Sala-i-Martin, 1992; Levine and Renelt, 1992; Mankiw, Romer and Weil, 1992; Jones, 1997; Quah, 1996a,b; Sala-i-Martin, 1996; Durlauf and Quah, 1999; Beaudry et al., 2005) have this debate at heart and discuss, which type of convergence governs the development of the cross-national income distribution and what is to be expected in the future. In particular, they show that a focus on β-convergence is informative on the nature of intra-distributional dynamics but cannot convey information concerning the development of the entire distribution, which appears to be polarizing. In order to overcome this traditional shortcoming of the β-convergence debate, probabilistic income mobility models are used to estimate likelihoods of convergence groups. Hence, there is a discussion on whether the twin peak phenomenon is either persistent as probabilities of switching are too low (Quah, 1996a,b) or is only a temporary occurrence due to increasing frequencies of growth miracles (Jones, 1997). Thus, existing literature either shows a descriptive picture of the cross-national income distribution by observing the development of two income per capita peaks in the cross-national income distribution, or, alternatively, is concerned with β-convergence in cross-national income growth regressions. Bianchi (1997) employed a nonparametric test for multimodality based on kernel density estimation to the cross-country income distribution. Further nonparametric approaches include those by Anderson (2004), who used stochastic dominance techniques, and by Maasoumi et al. (2007), who analyzed the cross-sectional distribution of growth rates. Regarding parametric modeling, for the cross-country income distribution, Paap and Dijk (1998) used a two-component mixture, consisting of a truncated normal distribution and a Weibull distribution.

Using nonparametric density estimates, several authors find a twin-peaked cross-country income distribution (Quah, 1996b; Bianchi, 1997). In particular, Bianchi (1997) observes

increasing evidence for bimodality in the GDP per capita distribution across countries over the period from 1970 to 1989 indicating global divergence rather than convergence. However, income data are often analyzed on a logarithmic scale, and the number of modes of the log-income distribution may differ substantially from the number of modes of the income distribution itself. In contrast, when modeling a density as a finite mixture, the number of components of this mixture is independent of the chosen scale. Furthermore, if one wishes to address convergence one can argue that it is not the number of modes in the cross-national income density which contains the most relevant information, but rather the number of convergence clubs, which correspond to the components in the finite mixture.

When modeling the cross-national income distribution by a finite mixture, determining its number of components is an essential step in the analysis. In their model, Paap and Dijk (1998) use a mixture with two distinct components, which resembles the fit of a histogram of the cross-national income distribution. Thus, the "stylized fact" of a distinction between poor and rich countries is already built into their model. We argue that via a statistical inference procedure, the data itself should determine the number of components and to this end a finite mixture with normal components of the log-income distribution (or log-normal components of the income distribution itself) is the appropriate tool. We shall apply the recently developed modified likelihood ratio test methodology (cf. Chen et al. 2001, 2004, and Chen and Kalbfleisch, 2005) for the number of components in a finite mixture to the cross-national income distribution.

In contrast to the twin-peaks literature and also to Paap and Dijk (1998), we find evidence of two components only at the beginnings of the 70s, whereas in the mid 70s, a third intermediate component emerges, which tends to separate itself ever more clearly from the poorest component. Thus, we find statistical evidence for three components rather than "twin peaks" in the cross-country income distribution.

After determining the number of components (three components from 1976 onward), we contribute to the convergence debate by extensively investigating the evolution, number of

components and the inter-distributional dynamics of the cross-country income distribution by using the posterior probability estimates from our fitted model. In particular, we find intra-distributional dynamics for Asian (upward mobility) and Latin American (downward mobility) countries. The overall picture that we obtain is that of three diverging groups in the cross-national income distribution.

The paper is structured as follows: Section 2 introduces the statistical methodology and highlights the differences (and advantages) of our approach compared with those taken in previous studies. Section 3 discusses our results, before we conclude.

2 Statistical Methodology & Data

2.1 Data

Following most other papers, our analysis is based on income data from the Penn World Tables Version 6.2 (Summer, Heston & Aten, 2006), from which we extract the real PPP GDP/per capita series of all years and countries available (chain series, base year 2000 in international $). In order to compare our observations over time, we restrict ourselves to those countries having more than half a million inhabitants, of which complete income data for the whole time period are available. This restriction leaves 124 countries for the period from 1970 to 2003 in our analysis. These countries represent about 95 percent of the world´s population.

2.2 Testing for the Number of Components in a Finite Mixture – the Modified Likelihood Ratio Test

Let f_X denote the density of the cross-country income distribution of a fixed year, and let f_Y be the corresponding density of the log-incomes, so that $f_Y(y) = f_X(e^y)e^y$. Multimodality of

f_X can arise by f_X being a finite mixture of other (unimodal) densities, so that

$$f_X(x) = p_1 g(x; \mu_1, \sigma_1) + \ldots + p_m g(x; \mu_m, \sigma_m), \qquad x > 0, \tag{1}$$

where the weights $p_i \geq 0$, $\sum_i p_i = 1$ and $g(x; \mu, \sigma)$ is a parametric family of densities, e.g. the log-normal distribution.

If f_X is a finite mixture of densities $g(\cdot; \mu, \sigma)$, there is no general simple connection between the number of modes of f_X and the number of components m. Typically, for unimodal g, the number of modes of f will be at most m, but often will be less than m. Furthermore, one is rather interested in the number of components m of the finite mixture than in the number of modes. For example, in the cross-country income distribution the components correspond to groups with different income level. Therefore, if we model the cross-country distribution of income by a finite mixture, determining its number of components statistically is a task of major importance, since this number will have essential economic consequences.

Note that the number of components is preserved if the data are transformed via a strictly monotonic transformation. In fact, if the x_i have density of form (1), the log-data $y_i = \log(x_i)$ have the transformed density

$$f_Y(y) = p_1 g(e^y; \mu_1, \sigma_1) e^y + \ldots + p_m g(e^y; \mu_m, \sigma_m) e^y.$$

Thus, the number of components is preserved, while the number of modes evidently may not. Therefore, in this paper we model f_X (and hence f_Y) as a finite mixture and then determine its number of components, mainly via hypothesis testing, but also by the use of model selection criteria. We model f_Y as a finite mixture of normal distributions, so that f_X is a finite mixture of log-normal distributions.

Estimation in finite mixture models (with a fixed number of components) typically proceeds by maximum likelihood. However, as already discussed in Pittau (2005), the likelihood function in a finite normal mixture with different variances is unbounded, thus, a global

maximizer of the likelihood function does not exist. There are some solutions around this problem. One is to look for the largest local maximum. Another is to the variances by restrictions of the form $\sigma_i^2 \leq c\,\sigma_j^2$ for all $i, j = 1, \ldots m$ and some $c > 1$ (cf. Hathaway 1985), which again leads to the existence of a global maximum and, if the true parameters satisfy the restriction, consistency. However, these solutions have practical problems, and therefore, here and regarding the analysis in section 3 we shall use finite mixtures with equal variances. For further discussion see section 4.

Testing in parametric models is often accomplished by using the likelihood ratio test (LRT). However, in order to test for the number of components in finite mixture models, it has long been known that the standard theory of the LRT does not apply. This is due to lack of identifiability under the null hypothesis: The null can be realized either by a weight zero for one of the components, or by equal parameters of two component distributions. Recently, it has been discovered that the asymptotic distribution of the LRT concerning the testing for the number of components in finite mixtures is the superior over a truncated Gaussian process (Chen and Chen, 2001; Dacunha-Castelle and Gassiat, 1999). The covariance of this process depends on the unknown true parameter, thus, the asymptotic distribution is too complicated and the LRT looses its practical appeal.

In methodologically related studies, Pittau (2005) and Pittau and Zelli (2006) fitted finite normal mixtures to the per-capita log GDP distribution across European regions in the years 1977–1996. They used a bootstrap version of the likelihood ratio test suggested by McLachlan (1987) to determine the number of components. However, this approach is computationally expensive and typically has low power properties, since under an alternative, the parameters from which resamples are obtained, are not correctly estimated (cf. Chen, Chen and Kalbfleisch, 2004, for related simulation results).

Recently, Chen et al. (2001, 2004) and Chen and Kalbfleisch (2005) suggested modified LRTs to solve these problems, which retain a comparatively simple limit theory as well as the good power properties of the LRT. We shall apply these tests to our problem concerning

the number of groups in the income distribution. At this point, we want to mention that the LRT and also the modified LRT are invariant under strictly monotonic transformation of the data (if candidate densities are correspondingly transformed). Thus, we could test on the level of the x_i as well as on the level of the y_i, the results are (in contrast to Silverman's test) completely consistent. For convenience, we shall use the log-data.

We first consider testing one against two components in a mixture. Suppose that $\phi(y;\mu,\sigma)$ is the normal distribution with mean μ and standard deviation σ, and consider the two-component mixture

$$f_Y(y;p,\mu_1,\mu_2,\sigma) = p\phi(y;\mu_1,\sigma) + (1-p)\phi(y;\mu_2,\sigma) \tag{2}$$

with equal standard deviation σ. The testing problem is

$$H_1: \quad f_Y \text{ is normally distributed} \qquad \text{against} \qquad K_1: \quad f_Y \text{ is of the form (2).}$$

The modified likelihood function is given by

$$l_n(p,\mu_1,\mu_2,\sigma) = \sum_{i=1}^{n} \log\big(p\phi(y_i;\mu_1,\sigma) + (1-p)\phi(y_i;\mu_2,\sigma)\big) \;+\; C\log\big(4p(1-p)\big),$$

where C is a fixed constant (we set $C=2$). Let $(\hat{p},\hat{\mu}_1,\hat{\mu}_2,\hat{\sigma})$ maximize $l_n(p,\mu_1,\mu_2,\sigma)$ over the full parameter space, and let $(\hat{\mu},\hat{\sigma})$ maximize $l_n(1/2,\hat{\mu},\hat{\mu},\hat{\sigma})$. The hypothesis H_1 is rejected for large values of the modified LRT statistic

$$M_n = 2\big(l_n(\hat{p},\hat{\mu}_1,\hat{\mu}_2,\hat{\sigma}) - l_n(1/2,\hat{\mu},\hat{\mu},\hat{\sigma})\big).$$

More precisely, Chen, Chen and Kalbfleisch (2001) show that for known σ, M_n asymptotically follows the distribution $1/2\chi_0^2 + 1/2\chi_1^2$, where χ_0^2 is the point mass at zero. For unknown σ, as formulated above, the precise asymptotic distribution of M_n is unknown, however, Chen and Kalbfleisch (2007) show that the χ_2^2 distribution is an upper bound to the asymptotic

distribution of M_n.

Chen, Chen and Kalbfleisch (2004) also consider the problem of testing for two against more components of a mixture distribution. More precisely, the problem is to test

$$H_2 : f_Y \text{ is of the form (2)} \qquad \text{against} \qquad K_2 : \quad f_Y \text{ has more than two components.}$$

Here, we again assume equal variances for all components, also under the alternative. Furthermore, one fixes a maximal number of components under the alternative (which can also be estimated, e.g. $m = 4$). For a mixture with m components, slightly changing the notation, the modified maximum likelihood estimators (MLEs) are defined as the maximizer of

$$l_n(\mu_1, \ldots, \mu_m, \sigma) = \sum_{i=1}^{n} \log \big(p_1\phi(y_1; \mu_1, \sigma) + \ldots + p_m\phi(y_m; \mu_m, \sigma)\big) \; + \; C \log \big(\prod_{i=1}^{m} p_i\big).$$

These estimates are then inserted into the LRT statistic. Chen, Chen and Kalbfleisch (2004) showed that given a known σ, this modified LRT is asymptotically distributed as $q\chi_0^2 + \frac{1}{2}\chi_1^2 + (1-q)\chi_2^2$, where the proportion q depends on the mixing distributions. For unknown σ, Chen and Kalbfleisch indicate that the χ_3^2 distribution is an upper bound of the asymptotic distribution.

As an illustration, we give the results of the analysis of the year 1976. The complete results are discussed in section 3.1. First we fit a single normal distribution to the log-income distribution (log to the base 10). Doing so, we obtain the following parameters: $\hat{\mu} = 3.52$ and $\hat{\sigma} = 0.46$. The modified MLEs of the two-component mixture with equal variances and penalization parameter $C = 2$ are calculated as $\hat{p} = 0.51$, $\hat{\mu}_1 = 3.15$, $\hat{\mu}_2 = 3.89$ and $\hat{\sigma} = 0.26$. The resulting value of the modified likelihood ratio function is equal to $T_n = 14.00$, which based on the upper bound of the χ_2^2-distribution yields a p-value of 0.0009. Thus, the hypothesis of a single component is clearly rejected.

Next we consider testing two against three (or more) components. Concerning the fit using three components and equal variances, the parameter estimates based on penalized maximum

likelihood are given by

$$\hat{p}_1 = 0.38, \quad \hat{p}_2 = 0.34, \quad \hat{\mu}_1 = 3.04, \quad \hat{\mu}_2 = 3.58, \quad \hat{\mu}_3 = 4.07, \quad \hat{\sigma} = 0.18. \tag{3}$$

The resulting value of the modified LR statistic is $T_n = 7.91$. Based on the upper bound by a χ^2_3-distribution, this gives a p-value of 0.048, in favor of three components. The three-component fit based on the modified MLEs, both for the y_i's as well as for the x_i's, are displayed in figure 1. Apart from testing the number of components, we also compare the mixture models via two popular model selection criteria, namely the Akaike information criterion (AIC, c.f. Akaike, 1978) and the Bayesian information criterion (BIC, Schwarz, 1978), given by $-2l + 2k$ and $-2l + k \log n$, respectively, where l is the log-likelihood, k the number of parameters and n the number of observations. The results are displayed in table 1. Here, the model selected by AIC is the model with three components, while BIC is slightly in favor of a model with only two components. Although it is theoretically known that the BIC is consistent in finite mixtures (Kerebin, 2000), in finite samples it often selects too few components. Finally, in Fig. 1 we compare the fitted three-component density with a nonparametric density estimate with bandwidth $h_c(3)$ (cf. Section 2.4). Such a comparison could also be used for a formal goodness of fit test for our mixture model, cf. e.g. Fan (1994). The nonparametric and our parametric estimate are quite close, thus, our model of the data is appropriate. The whole picture that we get from our analysis of the log cross-country income distribution in 1976, taking into account the modified likelihood ratio tests and the model selection criteria as well as the shape of nonparametric density estimates, is clearly in favor of three components.

2.3 Discriminant Analysis via Posterior Probabilities

Mixture models are routinely used for discriminant analysis, see e.g. Fraley and Raftery (2002). In our analysis of the cross-country income distribution via mixtures, once we have

a mixture fitted to the cross-country income distribution, each observation can be assigned posterior probabilities which give the probability of the observation to belong to each of the components in the mixture model.

Consider the log-income distribution in 1976. In section 2.2, we fitted a three-component normal mixture

$$f_Y(y; \hat{p}_1, \hat{p}_2, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\sigma}) = \hat{p}_1 \phi(y; \hat{\mu}_1, \hat{\sigma}) + \hat{p}_2 \phi(y; \hat{\mu}_2, \hat{\sigma}) + (1 - \hat{p}_1 - \hat{p}_2)\phi(y; \hat{\mu}_3, \hat{\sigma}), \qquad (4)$$

where the parameter estimates are given in (3). This yields three levels of income which we label poor, middle and rich, with indices $1, 2, 3$. The posterior probability of an observation y to belong to group j, $j = 1, 2$, is equal to

$$p(j; y) = \frac{\hat{p}_j \phi(y; \hat{\mu}_j, \hat{\sigma})}{f_Y(y; \hat{p}_1, \hat{p}_2, \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\sigma})},$$

and $p(3; y) = 1 - p(1; y) - p(2; y)$. Therefore, we do not merely assign an income level to each country, but rather a probability distribution, which makes transitions from one group to the other much more transparent.

If one wishes to assign a single number to each country y, one has several possibilities. One is the maximum a-posterior estimate (MPE), which assigns to observation y the j, $j \in \{1, 2, 3\}$, such that $p(j; y)$ is maximal. One can also determine the thresholds $t_{j,j+1}$, $j = 1, 2$, for the values of y at which the MPE changes between the state j and $j+1$, by solving the equations $p(j, t_{j,j+1}) = p(j + 1, t_{j,j+1})$, $j = 1, 2$, yielding the (in model (4)) unique solutions

$$t_{j,j+1} = \frac{\hat{\mu}_j + \hat{\mu}_{j+1}}{2} + \hat{\sigma}^2 \frac{\log(\hat{p}_j / \hat{p}_{j+1})}{\hat{\mu}_{j+1} - \hat{\mu}_j}, \qquad j = 1, 2.$$

If the weights $\hat{p}_1$ and $\hat{p}_2$ are sufficiently close, the values $t_{j,j+1}$ will indeed be between $\hat{\mu}_j$ and $\hat{\mu}_{j+1}$, in which case they may be properly interpreted. For example, for the year 1976 we get $t_{1,2} = 3.32$ and $t_{2,3} = 3.83$, which on the original scale correspond to the values 2089.30 and

6760.83, respectively. In 1976, by maximum a-posterior estimation there are 46 countries in the poor group, 42 countries in the middle group and 36 countries in the rich group.

Another (more informative) possibility is the posterior mean (PM) of y, which is defined as $p(1;y) + 2p(2;y) + 3p(3;y)$, a number between one and three. In our situation, since the choice of the values $1, 2, 3$ is arbitrary, this should not be interpreted as a mean but rather as a refined one-number summary of the posterior distribution. For example, if the PM of y is 1.3, then the country will belong to group 1, but will have a tendency toward group 2. A tedious but straightforward computation shows that the posterior mean in model (4) with equal $\hat{\sigma}^2$ is a monotonically increasing function of y. Thus, one can uniquely determine thresholds $s_{j,j+1}$, $j = 1, 2$, for which the PM is equal to $i + 1/2$. Solving these equations numerically for the parameters in 1976 yields the values 3.32 and 3.84.

2.4 Nonparametric Kernel Density Estimation and Twin Peaks

In this section we draw attention to possible shortcomings of previous approaches based on nonparametric kernel density estimation. Specifically, we show that the "twin peak" phenomenon in the cross-country income distribution is essentially not a structural feature, but rather an artifact which arises in nonparametric kernel density estimates of heavy-tailed distributions.

Indeed, the cross-country income distribution is concentrated in lower regions, and then has a rather long, small tail at the upper end. For example, in 2003 most values are $< 10^4$, but there is a tail up to $4 \cdot 10^4$ (cf. Fig. 2). Such a shape can lead to a very poor performance of kernel estimates with a global bandwidth (cf. Wand and Jones 1995, p. 36). As an illustration, we therefore also fitted a transformation kernel density estimate (Wand and Jones, 1995, p. 43) based on the log-transform to the data (cf. Fig. 2). Evidently, the estimates differ strongly, as the usual kernel density estimator puts too much mass to the tails. This leads to the emergence of the second peak in the "twin peaks" phenomenon of the cross-country income distribution, which is just an artifact of direct kernel density estimation with global

bandwidth of heavy-tailed data. Also note that the usual kernel estimator has a boundary problem at 0 (cf. Wand and Jones 1995, p. 46).

In summary, for the cross-country income distribution itself, simple nonparametric kernel density estimation and related inference techniques such as Silverman's (1981) test should not be used. Also note that Paap and Dijk (1998) choose their two-component mixture with distinct components to resemble the histogram of the cross-country income distribution, which is just a simple kernel estimator. Thus, their model (and in particular the number of components) is motivated by an inappropriate preliminary estimate.

The situation is different in the case of log-incomes (which are non longer heavy-tailed), for which kernel density estimates and Silverman's (1981) test are valid tools. In order to illustrate this issue, let us first briefly recall Silverman's test. Formally, a *mode* of f_X (and similarly of the kernel estimator $\hat{f}$) is a local maximum of f_X (or $\hat{f}$). Silverman (1981) showed that the number of modes of $\hat{f}$ is a right-continuous, monotonically decreasing function of the bandwidth h if the normal kernel is employed for K: $K(x) = (2\pi)^{-1} \exp(-x^2/2)$. This allowed him to define the k-critical bandwidth $h_c(k)$ as the minimal h for which $f(\cdot; h)$ still just has k modes and not yet $k+1$ modes. Based on the notion of the k-critical bandwidth, Silverman (1981) proposed a bootstrap test for the hypothesis

$$\tilde{H}_k : f \text{ has at most } k \text{ modes} \qquad \text{against} \qquad \tilde{K}_k : f \text{ has more than } k \text{ modes},$$

where in our context, $f = f_X$ (or $f = f_Y$, the density of logarithms $y_i = \log(x_i)$).

The results of Silverman's test for the year 2003 are displayed in table 3. Here one typically proceeds iteratively by testing $\tilde{H}_k$ for increasing k, starting with $k = 1$, until one finds k such that $\tilde{H}_k$ cannot be rejected with a given level α (e.g. $\alpha = 0.05$). Concerning the log-incomes and their density f_Y, the hypothesis $\tilde{H}_1$ cannot be rejected at a 5% (or even 10%) level. However, the corresponding p-value is still comparatively small. Note that this result does not mean that $\tilde{H}_1$ is true, only that there is not enough evidence to reject it on a level of 5% (or 10%). However, if one continues the analysis, one can clearly reject $\tilde{H}_2$ (p-value < 0.001),

but $\tilde{H}_3$ has a high p-value of 0.45. Thus, there is some evidence of three modes in f_Y, but none of only two modes. The associated density estimates with the critical bandwidths are displayed in Fig. 3 (right).

As an illustration we also applied Silverman's (1981) test to the original income data x_i. The hypothesis $\tilde{H}_1$ is clearly rejected with a p-value of < 0.001, and the hypothesis $\tilde{H}_2$ is not rejected with a high p-value. Thus, the procedure stops at $k = 2$, strongly indicating two modes. However, observe figure 3 (left). It shows (boundary corrected) plots of the densities with bandwidths $h_c(1)$ and $h_c(2)$. For $h_c(2)$, the third mode (which is not statistically significant according to Silverman's test) is about to occur in the two highest observations of the distribution (at about $35 \cdot 10^3$). Thus, the kernel density estimator is about to put a spurious mode into the tail, and all that Silverman's test tells us is that this mode is indeed spurious. Thus, kernel estimation and Silverman's (1981) test are inappropriate on the level of the x_i.

3 Results

3.1 Selecting the number of components

Applying the methodology above to the time range from 1970 to 2003 for the 124 countries for which we have consistent GDP data yields some surprising and telling insights into the evolution of the cross-country distribution of income. Table 4 displays the results of the modified likelihood ratio test for one vs two components and two vs three components as well as the AIC and BIC model selection criteria for the respective fitted models ranging from 1 to 4 component mixtures (all having equal variances). First of all, we note that two components are always preferable to one. In 1970 we cannot reject the hypothesis of two vs three components, however, over the first years of the 1970s the p-values are decreasing and by 1976 the modified likelihood ratio test rejects a two component model at a level of 5%. This is also supported by the values of the model selection criteria AIC and BIC, which initially

are in favor of a two component model, but over time switch toward the three component mixture model. In summary, our analysis shows that starting with a two-component (twin-peak) mixture distribution in 1970, in between the "rich" and "poor" components, in the middle of the 1970s a third component evolves in the cross-national distribution of income, thus resulting in a three-component mixture model. All subsequent distributional analysis is based on the three component mixture model from 1976 to 2003.

3.2 Evolution of the Cross-Country Distribution of Income

Table 5 summarizes the main distributional characteristics of the three component mixture model after 1976. The first three columns display the weights p_1, p_2 and p_3 of the three components in the mixture model, which can be directly interpreted as the percentage of data, i.e. the relative number of countries, ascribed to a certain component. As can be seen in figure 9, the percentage of data ascribed to the first "poor" component, despite small variation, dropped slightly over time from initially 37.9 percent in 1976 to 35.7 percent in 2003. In comparison, the second component weight gained slightly over time from 33.8 percent to 35.3 percent, leaving the third component weight largely unaltered (28.4 percent in 1970 and 29 percent in 2003). Hence, the relative number of countries ascribed to each component is rather stable over the given observational period.

Regarding the log-income data, it can be observed that the mean of the first component did not grow, but rather experienced stagnation and even slight decline. In comparison, the mean of the second and third components clearly increased over the given time period from 3.59 to 3.73 and 4.08 to 4.33 respectively. The standard deviation parameter σ of the three components remains also rather stable over the given time period. However, these model parameters are harder to interpret on the logarithmic scale. Therefore, we also computed the mean and the standard deviation of the log-normally distributed components for the original income data.

Observing table 5 and figure 6 we note that the mean GDP per capita of the countries

belonging to the first component decreased slightly over time from \$1147 to \$1128. The countries belonging to the second component saw a strongly increasing income from on average \$3998 to \$5504 which corresponds to an overall 37 percent increase between 1976 and 2003. However, over the same period the countries in the third (the richest) component experienced an increase of mean income from \$12335 to \$21938 (increase of 77 percent). Hence, from 1976 onwards the countries in the poorest component experienced stagnation in or even slightly declining average income. Moreover, despite the clear emergence of a third "transitional" component in the middle of the 1970s, the mean income gain experienced in this component is not sufficient to facilitate any catch-up to the third "rich" component, which in turn improves not only its absolute, but also its relative position. Thus, the three components of our model of the cross-national distribution of income per capita actually diverge over time. This leaves slightly over 1/3 of the poor economies in a poverty trap, whereas slightly over 1/3 of the 124 countries, "the middle group", experience growth, but not fast enough to catch-up with the rich countries club, which consists of little less than 1/3 and which improved its absolute and relative position. Thus, one may claim that the cross-national distribution of income is not converging.

3.3 Intra-distributional dynamics based on Posterior Probabilities

As mentioned in section 2.3, one major advantage of a mixture model with equal variances for the components is that it makes accessible consistent posterior analysis. In the following, we shall mainly use the posterior mean. In table 7 all countries are ranked by their change in posterior mean. The biggest winner is China which increased its posterior probability mean from 1 to 2 and is one of 14 countries which managed to move up by one component. Out of these 14 countries, half moved from the 1st component to the 2nd component, leaving the other half to move from the 2nd component to the 3rd component. Of the 12 countries which dropped by one component 5 countries dropped from the 2nd to the 1st and 7 countries dropped from the 3nd to the 2rd component. The average posterior probability mean

increased slightly from 1.91 in 1976 to 1.93 in 2003, which implies a slight increase in the size of the second and/or third components. Comparing results in 1976 and 2003, we find that 26 out of 124 countries, about 21 percent, changed the component of the cross-national income distribution (although a few further temporary changes might have taken place in between). This implies a relative low mobility of countries as, comparing 1976 and 2003, only one out of five countries changed its component position or group affiliation.

Figures 7 and 8 show a more detailed story for selected countries belonging to different regions. Most notably Asia composes half of the 14 countries which improved by one component and none of the South East Asian, East Asian or South Asian Countries experienced a deterioration of their posterior mean. Thus, the Far East is generally the most upwards mobile region, of which most countries belong to the second component and experienced on average higher growth rates than the other countries belonging to this component. In particular, China's extraordinary growth is mirrored in the jump from the very bottom to a median position of the cross-country mean income distribution. Obviously, it is these countries in particular which account for the rising mean of the 2nd component over time. In fact, the average growth rate of these countries is more than sufficient for a catch-up in mean income to the richer group of countries.

However, a second region is also very prominent in the 2nd group which lowers the average growth rate of this component, namely Latin America. Whilst only one country, Chile, managed to improve by one component, Latin America accounts for one third of the countries which moved down by one component. In particular, richer countries, like Argentina and Venezuela lost relatively and were assigned to the 2nd component in 2003. This sub-average performance of Latin America in general, of which most countries belong to the 2nd component, helps to explain why the growth rate in the mean is not sufficient to facilitate any catch-up of the entire component to the 3rd component.

The worst performing region by far is Sub-Saharan Africa, which accounts for 32 of the 46 countries, about 70 percent, belonging to the "poor" 1st component in 1976 and 32 of 44

countries, about 73 percent in 2003 respectively. It is mostly the non-existing growth record of these countries in Sub-Saharan Africa which accounts for the stagnant and even declining mean of the 1st poor component. Moreover, not only did the poorest countries remain extremely poor, but those countries which were relatively well-off in 1976, namely South Africa, and to a lesser extent Zimbabwe, belong to the group of countries, whose posterior mean decreased most. Despite the overall bleak record of Sub-Saharan Africa, there are a few examples which also show quite remarkable improvement, in particular Botswana and Lesotho. Moreover,Cameroon, Mauritius and Equatorial New Guinea even improve by one component and are the only three Sub-Saharan African Country that display upward mobility. However, these few encouraging examples are not enough to change the Sub-Saharan stagnant and very poor growth record.

Unsurprisingly, most Western countries belong firmly to the 3rd component displaying hardly any change in their posterior probability mean. It is mainly their growth record which accounts for the increase in the 3rd component mean. Eastern Europe lost in particular after the breakdown of the Iron Curtain, but had resurging growth, which lead to a rather stable position in between the 2nd and the 3rd components over time. Morocco and Egypt show the success of some Arabic countries, whilst Iraq is the extreme opposite and is the country which lost most over the time period 1976 to 2003.

Overall, the country specific data and posterior mean helps to explain the development of the cross-national distribution of income from 1976 to 2003. The following general picture emerges: First, Sub-Saharan Africa accounts mostly for the lowest component which remains stagnant and "poor". Second, the emergence of the "transition" component is mostly due to the growth spurt of the Far East and the relative decline of Latin America. The contrary growth experience accounts mainly for the relatively slow growing mean of the 2nd component. Whilst most Far Eastern countries grow fast enough to catch up with the richer countries of the 3rd component, this is not the case for most of Latin America, which experienced disappointing growth records in particular in the 1980s. Thus, the overall cross-national

income distribution does not display absolute cross-national average income convergence, but rather divergence over time, despite the fact that some countries, in particular in the Far East, are rapidly catching-up. However, in the global picture this is counteracted by the relatively poor growth record of Latin America and the average income stagnation in most parts of Sub-Saharan Africa. In particular, almost all of Sub-Saharan Africa seems to be stuck in a poverty trap in which the unit under scrutiny, the national economy, is not capable to deliver any form of sustained per capita growth. However, our data also shows that some of the most populous countries, in particular India and China, are doing extraordinarily well. Thus, the global (and not cross-national) income distribution, which takes into account the distribution of the income within countries as well as the sizes of their populations, might indeed be converging.

4 Conclusions

Previous investigations on the twin-peak phenomenon in the world´s cross-country distribution of income were mostly based on nonparametric kernel density estimates, in particular concerning the number of modes of such estimates.

In this paper we use finite mixtures in order to investigate the cross-country income distribution, since *a.* the number of modes depends on the scale (original or logarithmic) whereas the number of components in the mixture does not, *b.* finite mixtures allow for an accurate analysis of the intra-distributional dynamics by using posterior probability estimates, *c.* components in the mixture arguably correspond better to income clubs in the distribution than its modes and *d.* the heavy tail of the cross-country income distribution on the original scale can cause poor performance of nonparametric density estimates. Furthermore, we argue that, in contrast to Paap and Dijk (1998) who simply use a fixed two-component model obtained by comparison with a histogram, the number of components in the mixture model

should be determined by statistical inference.

In contrast to the twin-peaks literature, we find evidence for an emerging intermediate component in the 70s, resulting in a three-component distribution from 1976 onwards. This alone is a strong indication of divergence within the distribution and might be an indicator of convergence within groups. Diverging estimates of the three group means and very different growth rates between the groups support this conclusion. While the mean of the third (richest) component almost doubled from 1976 to 2003, the mean of the second (intermediate) component only increased by 40 percent (corresponding to a very low annual growth rate), and the first (poor) component even stagnated. One should mention that up- and downward movements of countries affect these growth rates. The growth of the third component is slowed down by countries moving up from the second component. Regarding the second component there are positive and negative effects, in which the negative effects outweigh the positive effects, since only a few countries move from the third to the second component. In the first component there should be positive effects from countries coming from the second group, which however are counterbalanced by the poor overall growth record within this component.

The regional differences are remarkable. While many Asian countries managed to catch up to the third component, the opposite is true in the case of Latin American countries. Sub-Saharan Africa seems to be stuck in the first component and looses more and more contact with the other groups. The very populous countries China and India on the other hand performed extremely well. This fact would foster convergence in a global distribution of income which takes population size and within country inequality into account (i.e. Sala-I-Martin, 2006).

A possible application of our methodology, beyond the conclusions already drawn in this paper, would be a classification of countries according to their mean income, as an alternative concept to the most prevalent "poor, middle and rich" classification of the World Bank. Indeed, the maximum posterior estimates can be used to assign countries to certain groups.

Due to its statistical nature, this approach would be less policy dependent than current approaches. The boundary points of income, separating the three groups, from our point of view currently somewhat arbitrarily obtained, could be replaced by the incomes where the maximum a-posterior estimate switches. For the year 2003 these are \$2405 and \$10859, respectively (PPP, base year 2000). However, our main aim was not to suggest a new system of classification of countries, but rather to obtain a better understanding of the cross-national distribution of income, its development, number of components and its intra-distributional dynamics over the past decades.

Appendix A: Mixtures with Distinct Variances

In Sections 2.2 and 3.1 we restricted the model class of finite normal mixtures to have equal variances. There were several reasons for this restriction.

First, maximum likelihood inference in mixtures of normal distributions with distinct variances becomes technically difficult since the likelihood function is unbounded. Moreover, the number of parameters increases drastically (up to 8 in a three-component mixture), which increases the risk of overfitting the data (in a sample of merely 124 observations).

Second, as regards the economic content, we wanted a model class which is adequate for the whole period 1976–2003 in order to draw conclusions about the evolution of the groups in the income distribution, in particular concerning their number. Furthermore, equal variances guarantee that groups are equally wide, i.e. that no components are fitted to a very small, selective group.

Third, if distinct variances are allowed, the posterior analysis is no longer consistent. In fact, higher observations can have smaller MPE or posterior mean than smaller observations, if the variance of components with smaller mean are much larger than those with higher mean. Thus, a posterior analysis does not make sense for such general models.

Nevertheless, in this section we briefly investigate the consequences if one allows the general

model class (1) for the world's cross-country income distribution, in particular if one allows for some or even all variances to be distinct. We start the analysis with the year 1970. Table 2 gives the values of the model selection criteria AIC and BIC for distinct mixture models with up to three components.

The AIC selects a model with three components and equal variances for the first two components (with small means), while the BIC selects the model with two components and equal variances (which would also be selected by the AIC3 as used in Pittau, 2005). These two competing "best" models are rather incomparable. Observing in addition the plots in figure 4, of the income distribution in 1970, one would decide for the two-component model with equal variances, since the third component in the three-component model, when compared to the nonparametric density estimate, looks much like fitting an artifact.

The picture changes over time, and in 2003 is rather in favor of the three-component model with two distinct variances, as can be observed from table 6. The resulting fit is visualized in figure 5.

However, the standard deviation σ of the third component is about twenty times smaller than the sigmas of the other two components. This makes the model inaccessible to posterior analysis. For example, posterior analysis would assign the USA to the second group. This is caused by the fact that mean income in the USA is much higher than the mean of the third component, and is thus not captured in this component. Hence, we find that the model with equal variances discussed in Section 3.1 is most adequate to describe and analyze the data at hand.

Acknowledgements

Sebastian Vollmer acknowledges financial support of the Georg Lichtenberg program "Applied Statistics & Empirical Methods". The authors would like to thank Stephan Klasen and Stefan Sperlich for helpful comments.

References

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. In: B. N. Petrov 6 F. Csaki (Eds.), *Second international symposium on information theory*, pp. 267-281

Anderson, G. (2004) Making inference about the polarization, welfare and poverty of nations: A study of 101 countries from 1970 - 1995. *J. Appl. Econ.* **19**, 537–550.

Barro, R. (1991) Economic Growth in a Cross Section of Countries, *Quarterly Journal of Economics* **106(2)**, 407-443.

Barro, R. and X. Sala-i-Martin (1992) Convergence. *Journal of Political Economy* **100(2)**, 223–251.

Beaudry, P., Collard, F. and Green, D. A. (2005) Changes in the world distribution of output per worker, 1960–1998: How a standard decomposition tells us an unorthodox story. *Review of Economics and Statistics* **87**, 741–753.

Bianchi, M. (1997) Testing for Convergence: Evidence from Non-Parametric Multimodality Tests. *Journal of Applied Econometrics* **12**, 393–409.

Chen, H., Chen, J. and Kalbfleisch, J. D. (2001) A modified likelihood ratio test for homogeneity in finite mixture models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63**, 19–29.

Chen, H., Chen, J. and Kalbfleisch, J. D. (2004) Testing for a finite mixture model with two components. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66**, 95–115.

Chen, J. and Kalbfleisch,J. D. (2005) Modified likelihood ratio test in finite mixture models with a structural parameter. *J. Statist. Plann. Inference* **129**, 93–107.

Dacunha-Castelle, D. and Gassiat, E. (1999) Testing the order of a model using locally conic parametrization: population mixtures and stationary ARMA processes. *Ann. Statist.* **27**,

1178–1209

Durlauf, S. and Quah, D. T. (1999) The new empirics of economic growth. In: Taylor, J.B. and Woodford, M. (Eds.), Handbook of Macroeconometrics I. Elsevier, Amsterdam, pp. 235–308.

Efron, B. (1979) Bootstrap Methods: Another Look at the Jackknife *Ann. Statist.* **7** (1), 1–26

Fan, Y. (1994) Testing the goodness of fit of a parametric density function by kernel method. *Econometric Theory* **10**, 316–356.

Fraley, C. and Raftery, A. E. (2002) Model-based clustering, discriminant analysis, and density estimation. *J. Amer. Statist. Assoc.* **97**, 611–631.

Hathaway, R. J. (1985) A constrained formulation of maximum-likelihood estimation for normal mixture distributions. *Ann. Statist.* **13**, 795–800.

Heston, A.; Summers, R. and Aten, B. (2006) Penn World Tables Version 6.2 *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania*

Jones, C. (1997) On the Evolution of the World Income Distribution. *The Journal of Economic Perspectives* **11**,19–36.

Keribin, C. (2000) Consistent estimation of the order of mixture models. *Sankhyā Ser. A* **62**, 49–66.

Levine, R. and D. Renelt, (1992) A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review*, **82**, 942–63.

Mankiw, G., D. Romer and D. Weil (1992) A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* **107**, 407–437.

Maasoumi, E., Racine, J. and Stengos, T. (2007) Growth and convergence: A profile of dis-

tribution dynamics and mobility. *J. Econonmetrics* **136**, 483–508.

McLachlan, G. J. (1987) On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture, *Applied Statistics*, **36**, 318–324

Pittau, M. G. (2005) Fitting regional income distributions in the European Union. *Oxford Bulletin of Economics and Statistics* **67**, 135–161.

Pittau, M. G. and Zelli, R. (2006) Empirical evidence of income dynamics across EU regions. *J. Applied Econometrics* **21**, 605–628.

Quah, D. T. (1996a) Convergence Empirics across Economies with (some) capital mobility *Journal of Economic Growth* **1**, pp. 95-124

Quah, D. T. (1996b) Twin Peaks: Growth and Convergence in Models of Distribution Dynamics. *The Economic Journal* **106**, 1045–1055.

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

Romer, P. (1990) Endogenous Technological Change. *Journal of Political Economy* **98(5)**, S71-S102.

Sala-i-Martin, X. (1996) The Classic Approach to Convergence Analysis. *The Economic Journal* **106**, 1019–1036.

Sala-i-Martin, X. (2006) The World Distribution of Income: Falling Poverty and ... Convergence, Period. *Quarterly Journal of Economics* **121**, 351–397.

Schwarz, G. (1978) Estimating the dimension of a model, *Annals of Statistics*, **6** , 461-464

Silverman, B. W. (1981) Using kernel density estimates to investigate multimodality. *J. Roy. Statist. Soc. Ser. B* **43**, 97–99.

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

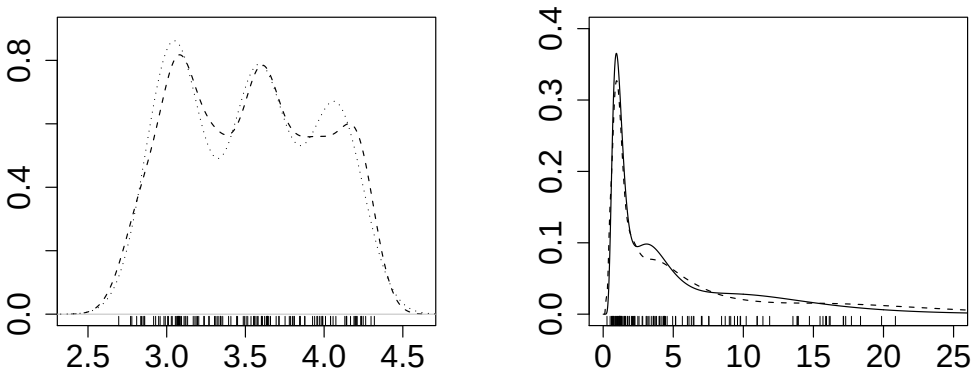


Figure 1: Left: Three-component mixture density with modified MLEs (solid line) and kernel density estimate based on $h_c(3)$ (dashed line) for the log-data (logarithm to the base 10) for 1976. Right: Corresponding three-component log-normal fit (solid line) and transformation kernel density estimate based on $h_c(3)$. Scale: x-axis 10^3, y-axis 10^{-3}.

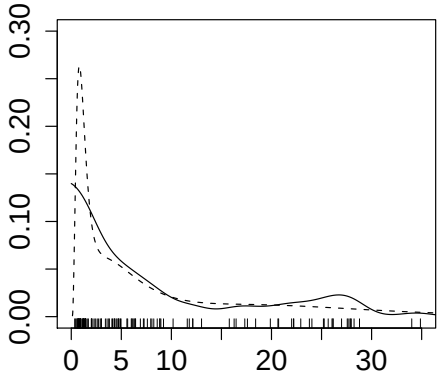


Figure 2: Kernel density estimate (solid line) and transformation kernel density estimate (dashed line) with log-transform, both with bandwidth estimated by direct plug-in, for 2003. Scale: x-axis 10^3, y-axis 10^{-3}.

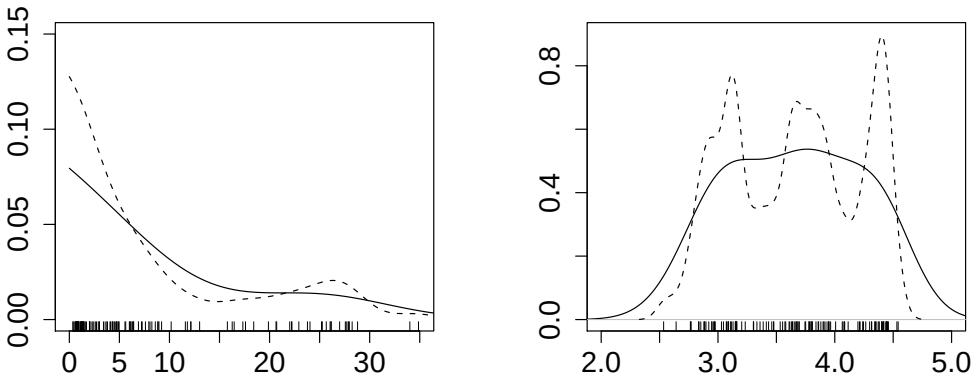


Figure 3: Left: Kernel density estimates with boundary correction at zero with bandwidths $h_c(1)$ (solid line) and $h_c(2)$ (dashed line) for the cross-country income distribution in 2003. Scale: x-axis 10^3, y-axis: 10^{-3}. Right: Kernel density estimates with bandwidths $h_c(1)$ (solid line) and $h_c(3)$ (dashed line) for the log-income distribution in 2003. For conveniently interpreting the figure, we here use the logarithm to the base 10.

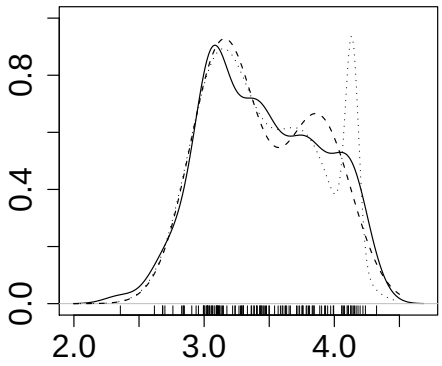


Figure 4: Density estimates for 1970: Kernel density estimate with $h_c(3)$ (solid line), two-component normal mixture with equal variances (dashed line) and three-component mixture with two distinct variances (dotted line).

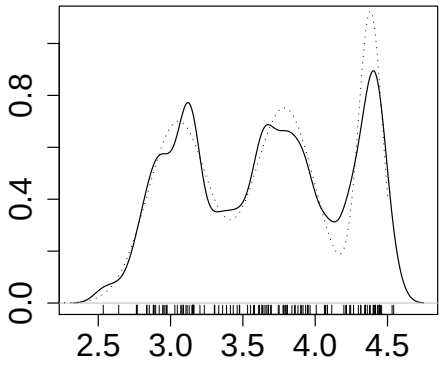


Figure 5: Density estimates for 2003: Kernel density estimate with $h_c(3)$ (solid line) and three-component mixture with two distinct variances (dotted line).

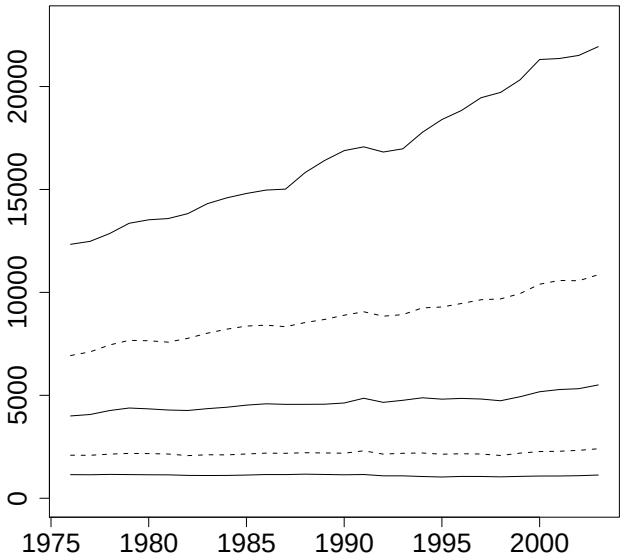


Figure 6: Means of the distinct groups (solid lines). Income levels where the maximum a-posterior estimates switch from one group to the other (dashed lines).

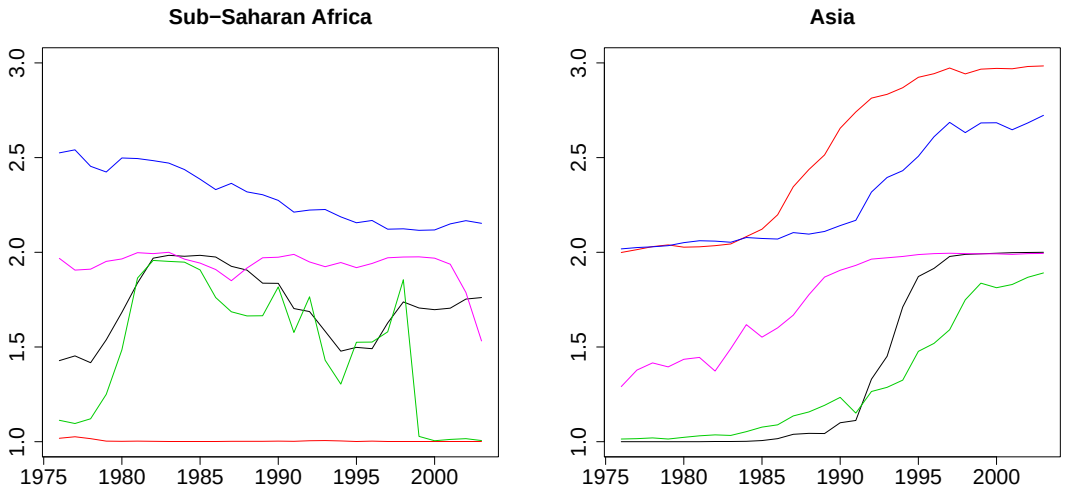


Figure 7: Posterior means of selected African (left) and Asian (right) countries. Africa: Cameroon (black), Nigeria (red), Republic of Congo (green), South Africa (blue) and Zimbabwe (pink). Asia: China (black), Republic of Korea (red), India (green), Malaysia (blue) and Indonesia (pink).

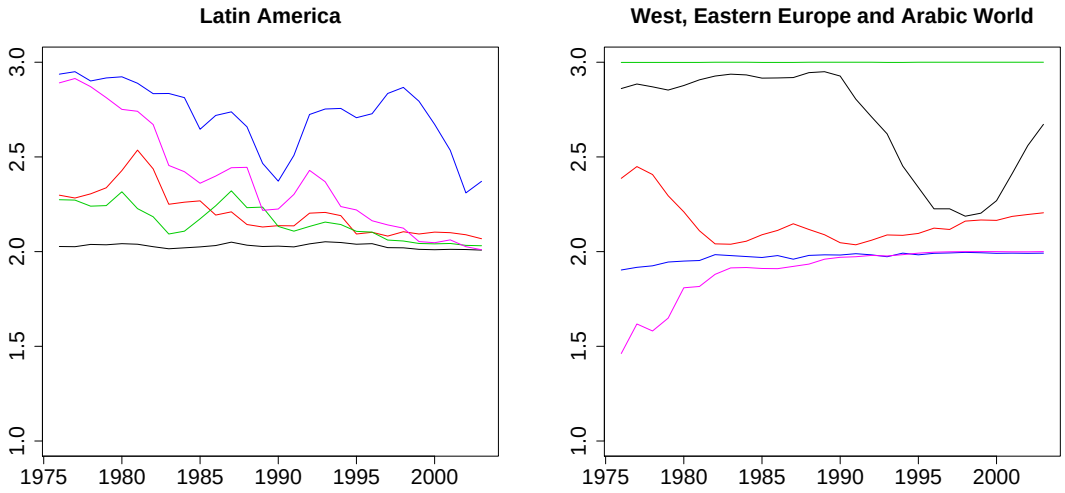


Figure 8: Posterior means of selected Latin American (left) and other (right) countries. Latin America: Colombia (black), Mexico (red), Brazil (green), Argentina (blue) and Venezuela (pink). Rest of the World: Russia (black), Poland (red), USA (green), Morocco (blue) and Egypt (pink).

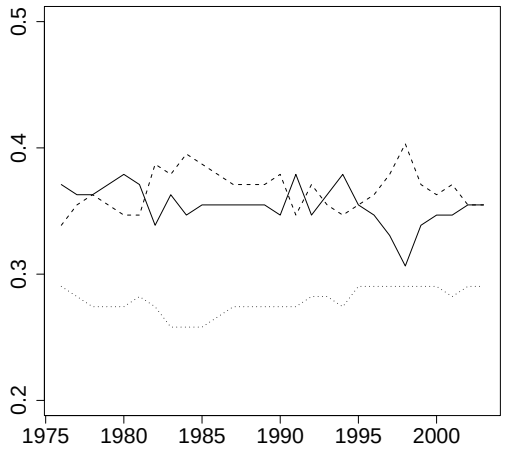


Figure 9: Weights of components. Poor (solid line), intermediate (dashed line) and rich component (dotted line).

no. components	- loglike.	no. param.	AIC	BIC
1	77.86	2	159.72	165.41
2	70.86	4	149.72	161.10
3	66.90	6	145.80	162.87

Table 1: Model selection criteria for mixtures models fitted to log cross-country income distribution 1976

no. components	no. variances	no. param.	AIC	BIC
1	1	2	152.47	158.16
2	1	4	142.64	154.01
	2	5	142.81	157.03
3	1	6	145.88	162.95
	2	7	139.73	159.64
	3	8	141.51	164.27

Table 2: Model selection criteria for mixtures models fitted to log cross-country income distribution 1970

	$h_c(1)$	p_1	$h_c(2)$	p_2	$h_c(3)$	p_3
y_i's	0.27	0.11	0.23	0.00	0.07	0.91
x_i's	5.72	0.01	2.14	0.44	1.83	0.24

Table 3: Results of Silverman's test for 2003, 124 countries

	One Component		Two Components			Three Components			Four Components	
Year	AIC	BIC	p of 1 vs 2	AIC	BIC	p of 2 vs 3	AIC	BIC	AIC	BIC
1970	152.47	158.12	0.001	142.64	153.92	0.868	145.88	162.81	145.27	167.84
1974	159.30	164.94	0.001	149.64	160.92	0.243	149.45	166.37	149.46	172.02
1975	157.47	163.11	0.001	147.49	158.77	0.171	146.47	163.39	146.40	168.97
1976	159.72	165.36	0.001	149.72	161.00	0.048	145.80	162.73	145.43	167.99
1980	167.20	172.84	0.000	155.68	166.96	0.037	151.21	168.14	155.21	177.78
1985	173.07	178.71	0.001	162.51	173.79	0.023	156.95	173.87	157.43	179.99
1990	183.20	188.84	0.001	172.40	183.68	0.006	163.84	180.77	164.47	187.03
1995	200.08	205.72	0.001	189.17	200.46	0.004	179.83	196.75	183.83	206.39
2000	204.57	210.21	0.000	192.40	203.68	0.000	173.05	189.97	173.23	195.80
2003	206.50	212.14	0.000	191.62	202.90	0.000	173.58	190.50	174.50	197.06

Table 4: Component Test and Goodness of Fit, 1970-2003

Year	p_1	p_2	p_3	μ_1	μ_2	μ_3	σ	Mean1	Mean2	Mean3	SE 1	SE 2	SE 3
1976	0.379	0.338	0.284	3.044	3.587	4.076	0.175	1147	3998	12335	310	1082	3337
1980	0.366	0.351	0.282	3.041	3.622	4.115	0.177	1141	4342	13526	313	1191	3710
1985	0.358	0.372	0.270	3.035	3.639	4.154	0.181	1126	4524	14806	316	1268	4149
1990	0.357	0.367	0.276	3.041	3.649	4.211	0.182	1140	4628	16887	320	1299	4741
1995	0.355	0.364	0.281	2.996	3.665	4.247	0.189	1032	4813	18396	303	1413	5400
2000	0.345	0.371	0.284	3.017	3.698	4.313	0.177	1079	5174	21312	295	1415	5830
2003	0.357	0.353	0.290	3.037	3.725	4.326	0.176	1128	5504	21938	307	1496	5962

Table 5: Fit of Three Component Model

no. components	no. variances	no. param.	AIC	BIC
1	1	2	206.50	212.19
2	1	4	191.62	203.00
	2	5	189.90	204.12
3	1	6	173.58	190.65
	2	7	162.13	182.03
	3	8	163.83	186.58

Table 6: Model selection criteria for mixtures models fitted to log cross-country income distribution 2003

Land	Posterior Mean 1976	Rank 1976	Posterior Mean 2003	Rank 2003	Change Posterior Mean	Group 1976 and 2003
China	1	115	2	60	1	1 to 2
Korea Republic of	1.999	60	2.984	27	0.985	2 to 3
Taiwan	2.041	47	2.994	25	0.953	2 to 3
Sri Lanka	1.087	87	1.996	66	0.909	1 to 2
India	1.014	97	1.891	76	0.877	1 to 2
Equatorial Guinea	1.787	72	2.639	36	0.852	2 to 3
Cyprus	2.195	42	2.997	22	0.802	2 to 3
Malaysia	2.018	55	2.723	34	0.705	2 to 3
Indonesia	1.291	81	1.994	69	0.703	1 to 2
Pakistan	1.035	90	1.673	79	0.638	1 to 2
Mauritius	2.339	39	2.972	29	0.633	2 to 3
Chile	2.117	43	2.724	33	0.607	2 to 3
Egypt	1.462	79	2	60	0.538	1 to 2
Botswana	1.603	74	2.077	44	0.474	2
Thailand	1.598	75	2.034	47	0.436	2
Cameroon	1.428	80	1.761	78	0.333	1 to 2
Bangladesh	1.024	93	1.255	82	0.231	1
Hungary	2.645	32	2.825	32	0.18	3
Portugal	2.832	30	2.981	28	0.149	3
Lesotho	1.001	112	1.148	85	0.147	1
Ireland	2.87	28	3	1	0.13	3
Swaziland	2.051	46	2.167	39	0.116	2
Hong Kong	2.906	26	3	1	0.094	3
Morocco	1.903	70	1.992	70	0.089	2
Singapore	2.918	24	3	1	0.082	3
Puerto Rico	2.917	25	2.998	19	0.081	3
Panama	2.036	48	2.091	43	0.055	2
Tunisia	2.003	59	2.048	46	0.045	2
Cuba	1.974	65	2.01	52	0.036	2
Dominican Republic	1.988	64	2.022	49	0.034	2
Trinidad + Tobago	2.961	21	2.989	26	0.028	3
Guinea	1.825	71	1.853	77	0.028	2
Philippines	1.951	68	1.977	73	0.026	2
Spain	2.977	19	2.996	23	0.019	3
Japan	2.991	18	2.999	11	0.008	3
Paraguay	1.993	62	2.001	59	0.008	2
Finland	2.992	15	2.999	11	0.007	3
Italy	2.992	15	2.998	19	0.006	3
Nepal	1.001	112	1.007	89	0.006	1

continued on next page

Land	Posterior Mean 1976	Rank 1976	Posterior Mean 2003	Rank 2003	Change Posterior Mean	Group 1976 and 2003
Korea Dem. Rep.	1	115	1.006	92	0.006	1
United Kingdom	2.995	14	2.999	11	0.004	3
Israel	2.992	15	2.996	23	0.004	3
Laos	1.002	111	1.006	92	0.004	1
Austria	2.997	8	3	1	0.003	3
Australia	2.997	8	3	1	0.003	3
Belgium	2.996	13	2.999	11	0.003	3
Mongolia	1.016	95	1.019	87	0.003	1
Norway	2.998	5	3	1	0.002	3
Canada	2.998	5	3	1	0.002	3
Germany	2.997	8	2.999	11	0.002	3
France	2.997	8	2.999	11	0.002	3
United States	2.999	2	3	1	0.001	3
Denmark	2.999	2	3	1	0.001	3
Netherlands	2.998	5	2.999	11	0.001	3
New Zealand	2.997	8	2.998	19	0.001	3
Ghana	1.006	105	1.007	89	0.001	1
Mali	1	115	1.001	100	0.001	1
Switzerland	3	1	3	1	0	3
Sweden	2.999	2	2.999	11	0	3
Tanzania	1	115	1	105	0	1
Malawi	1	115	1	105	0	1
Cambodia	1	115	1	105	0	1
Guinea-Bissau	1	115	1	105	0	1
Ethiopia	1	115	1	105	0	1
Bhutan	1	115	1	105	0	1
Burkina Faso	1	115	1	105	0	1
Oman	2.97	20	2.969	30	-0.001	3
Greece	2.961	21	2.96	31	-0.001	3
Burundi	1.001	112	1	105	-0.001	1
Rwanda	1.006	105	1.003	97	-0.003	1
Uganda	1.004	107	1.001	100	-0.003	1
Chad	1.003	109	1	105	-0.003	1
Gambia	1.003	109	1	105	-0.003	1
Turkey	2.008	58	2.004	57	-0.004	2
Mozambique	1.012	101	1.008	88	-0.004	1
Central African Rep.	1.004	107	1	105	-0.004	1
Guatemala	1.993	62	1.988	71	-0.005	2
Jordan	1.994	61	1.986	72	-0.008	2

continued on next page

Land	Posterior Mean 1976	Rank 1976	Posterior Mean 2003	Rank 2003	Change Posterior Mean	Group 1976 and 2003
Benin	1.013	98	1.004	96	-0.009	1
Kenya	1.01	103	1.001	100	-0.009	1
Niger	1.009	104	1	105	-0.009	1
Madagascar	1.011	102	1	105	-0.011	1
Sudan	1.013	98	1.001	100	-0.012	1
Somalia	1.013	98	1	105	-0.013	1
Congo Dem Rep.	1.015	96	1	105	-0.015	1
Papua New Guinea	2.015	57	1.999	64	-0.016	2
Fiji	2.017	56	2	60	-0.017	2
Nigeria	1.018	94	1.001	100	-0.017	1
Colombia	2.027	54	2.008	54	-0.019	2
Sierra Leone	1.027	92	1	105	-0.027	1
Algeria	2.035	50	2.007	55	-0.028	2
Romania	2.033	52	2.005	56	-0.028	2
Togo	1.03	91	1	105	-0.03	1
El Salvador	2.031	53	2	60	-0.031	2
Jamaica	2.036	48	1.999	64	-0.037	2
Ecuador	2.035	50	1.995	67	-0.04	2
Syria	1.193	83	1.153	84	-0.04	1
Bolivia	1.959	67	1.895	75	-0.064	2
Zambia	1.064	89	1	105	-0.064	1
Mauritania	1.077	88	1.007	89	-0.07	1
Namibia	2.089	45	2.004	57	-0.085	2
Senegal	1.1	86	1.006	92	-0.094	1
Congo Republic of	1.113	85	1.006	92	-0.107	1
Peru	2.114	44	1.995	67	-0.119	2
Honduras	1.503	78	1.384	81	-0.119	2 to 1
Poland	2.387	37	2.205	38	-0.182	2
Liberia	1.184	84	1	105	-0.184	1
Russia	2.861	29	2.672	35	-0.189	3
Mexico	2.298	40	2.068	45	-0.23	2
Brazil	2.274	41	2.031	48	-0.243	2
Costa Rica	2.376	38	2.124	42	-0.252	2
Comoros	1.283	82	1.002	98	-0.281	1
South Africa	2.525	35	2.153	41	-0.372	3 to 2
Uruguay	2.534	34	2.155	40	-0.379	3 to 2
Zimbabwe	1.968	66	1.532	80	-0.436	2
Cote d'Ivoire	1.598	75	1.155	83	-0.443	2 to1
Solomon Islands	1.549	77	1.036	86	-0.513	2 to 1

continued on next page

Land	Posterior Mean 1976	Rank 1976	Posterior Mean 2003	Rank 2003	Change Posterior Mean	Group 1976 and 2003
Nicaragua	2.523	36	1.965	74	-0.558	3 to 2
Argentina	2.937	23	2.371	37	-0.566	3 to 2
Ukraine	2.645	32	2.012	50	-0.633	3 to 2
Afghanistan	1.651	73	1	105	-0.651	2 to1
Iran	2.8	31	2.012	50	-0.788	3 to 2
Venezuela	2.891	27	2.01	52	-0.881	3 to 2
Iraq	1.912	69	1.002	98	-0.91	2 to 1

Table 7: Results Posterior Estimation