CHAPTER 5

Distribution and Mobility of Wealth of Nations*

Richard Paap† and Herman K. van Dijk‡

Abstract

We estimate the empirical bimodal cross-section distribution of real Gross Domestic Product per capita of 120 countries over the period 1960–1989 by a mixture of a Weibull and a truncated normal density. The components of the mixture represent a group of poor and a group of rich countries, while the mixing proportion describes the distribution over poor and rich. This enables us to analyse the development of the mean and variance of both groups separately and the switches of countries between the two groups over time. Empirical evidence indicates that the means of the two groups are diverging in terms of levels, but that the growth rates of the means of the two groups over the period 1960–1989 are the same.

1 Introduction

Empirical evidence on convergence of national economies has usually been investigated by regressing growth rates of real Gross Domestic Product [GDP] on initial levels, sometimes after correcting for exogenous variables (conditional convergence), see among others, Baumol (1986), Barro (1991), Mankiw, Romer and Weil (1992) and Sala-i-Martin (1994). A negative regression coefficient, usually labelled the β-coefficient, is interpreted as an indication of so called β-convergence. It implies that countries with a relatively low level of GDP grow faster than countries with a high level of GDP, indicating catching-up, compare also Abramowitz (1986).

---

* Reprint of Richard Paap and Herman K. van Dijk (1998) Distribution and Mobility of Wealth of Nations, *European Economic Review*, 42, 1269-1293. Copyright (1998), with permission from Elsevier.

† Econometric and Tinbergen Institutes, Erasmus University Rotterdam.

‡ Econometric and Tinbergen Institutes, Erasmus University Rotterdam.
Email: hkvandijk@few.eur.nl

A different concept of convergence, called σ-convergence, refers to a reduction in cross-sectional variance or dispersion over time, see Barro and Sala-i-Martin (1992). Friedman (1992) and especially Quah (1993a) show, using Galton's fallacy, that a negative β-regression coefficient can be perfectly consistent with the absence of σ-convergence, even when conditioning on exogenous variables. Furthermore, Levine and Renelt (1991, 1992) discuss the robustness of the regression approach with respect to the conditioning variables and the consistency of the results, see also Durlauf and Johnson (1995). Another limitation of the regression approach is that the dynamics of the economic process is summarised in a growth rate and an initial level, neglecting the short run dynamics of the variable investigated.

This paper deals with the analysis of convergence in terms of several characteristics of the distribution of real GDP per capita and is related to the work by Quah (1993a,b), Desdoigts (1994), Jones (1997), Quah (1996a,b) and Bianchi (1997). In these studies nonparametric methods are usually applied to analyse convergence. In the present paper we take a parametric approach. More, generally, we analyse the *development of the distribution and mobility of wealth* of 120 countries from 1960 until 1989. As measure for wealth we take the real Gross Domestic Product per capita, which can be interpreted as a rough approximation of the basic idea about wealth, see Parente and Prescott (1993)[10]. We start with presenting some stylized facts on the observed real GDP per capita over the period. This leads to the conclusion that the data may be described by a bimodal distribution. Next, we divide the further analysis into two parts[11]. In the first part, the empirical cross sectional bimodal distribution of the real GDP per capita in each year is described by a finite mixture density. Efficient estimation of the parameters of several classes of finite mixtures results in a partitioning of the countries into two groups in each year, a group with a relatively high level of real GDP per capita and a group with a low level of real GDP per capita and two estimated conditional density functions for the two groups. The use of mixtures enables us to analyse the distribution of countries over poor and rich as well as the development of the distribution of each group.

In the second part, the results of the estimated mixture distributions are used to consider the intra-distribution dynamics. By examining the movements of countries between the poor group and the rich group, we obtain insight into the extent of catching-up of poor countries with rich countries.

The outline of the paper is as follows. In section 2 we describe the data and present some stylized facts. In section 3 we briefly discuss the interpretation, representation and estimation of finite mixture distributions. Section 4 considers the estimation results of the mixture distribution for the cross-section real GDP per capita distribution including the development of the mean, the variance and the mixing parameter through time. The mobility in wealth between and within both groups is investigated in section 5. The final section contains our conclusions.

---

[10] Of course, the real GDP per capita of a country is a measure which neglects information about the spread of wealth among people living in this country. There can be a small group of persons living in a country with a high level of income, while the majority has low income.

[11] Here we differ from Quah who considers the year by year distribution and intra-distribution dynamics simultaneously.

2 Stylized Facts

In order to analyse the distribution and mobility of wealth of nations empirically, one needs a suitable data set containing per capita data over a long period for a large number of countries. Usually, one has data over several periods (years) but only a limited number of (industrialized) countries or one has many countries over a small number of years. In this paper we analyse the distribution and mobility of wealth using a reasonably large collection of countries over 30 years. The obvious data set for our analysis is the Penn World Table version 5.6 of Summers and Heston (1991). This table contains a set of economic time series, based on national accounts covering 152 countries for the period 1950–1992. Because observations are not available for each country over the whole period, we focus on the period 1960–1989. By restricting ourselves to this period, there remain observations for 120 countries. The variable we analyse in this paper is the real Gross Domestic Product [GDP] per capita, which is constructed by dividing nominal GDP per capita by a special price index made up of the weighted averages across countries of relative prices of all goods in a particular basket of final goods and services. This is intended to make real GDP per capita comparable across time and countries. For a discussion of the construction of the special price index and the data in general, we refer to Summers and Heston (1991).

Figure 5.1 shows smoothed versions of histograms for real GDP per capita of 120 countries in each year[12]. Several features of the data are shown in this figure. First, the cross-section distribution of the real GDP per capita is bimodal. There is a group of countries with a relative small real GDP per capita (poor countries) and a smaller group of countries with a relative large real GDP per capita (rich countries). Second, the gap between these groups seems to become larger over time, as the peak of the real GDP per capita of the rich countries shifts more to the right than the peak of the poor countries, leaving very few countries in a middle group.

In order to obtain better insight into the stylized facts of our data set we divide our sample into six subperiods of five years and compute the average real GDP per capita for all 120 countries over these subperiods, *i.e.* for 1960–1964, 1965–1969, 1970–1974, 1975–1979, 1980–1984 and 1985–1989. Figure 5.2 displays the histograms for the mean real GDP per capita in each subperiod in a 3-dimensional space, similar to figure 5.1. This figure shows the data features mentioned before even more clearly. In addition, we notice that the variance of the poor group in the early sixties seems to be smaller than in the early eighties. For the rich countries this seems to be the opposite. The same features of the data can be detected from figure 5.3 which shows the histograms of the real GDP per capita in the six subperiods in a one-dimensional setting. The six histograms give good insight in the development of the cross-section distribution of the real GDP per capita. From the stylized facts

[12] This figure is constructed by making a histogram for real GDP per capita in each year and putting these histograms in a 3-dimensional space. For visual convenience we use small ribbons, which connect the midpoints of the bars, instead of 3-dimensional bars. Furthermore, the real GDP per capita data, are divided by 1000 for the convenience of representation, like in the remainder of this paper.

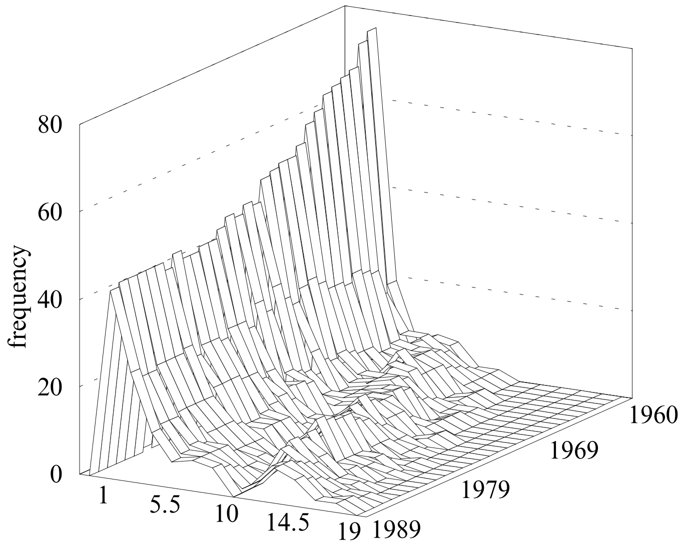

Fig. 5.1: Histograms of real GDP per capita divided by 1000 (1960–1984).

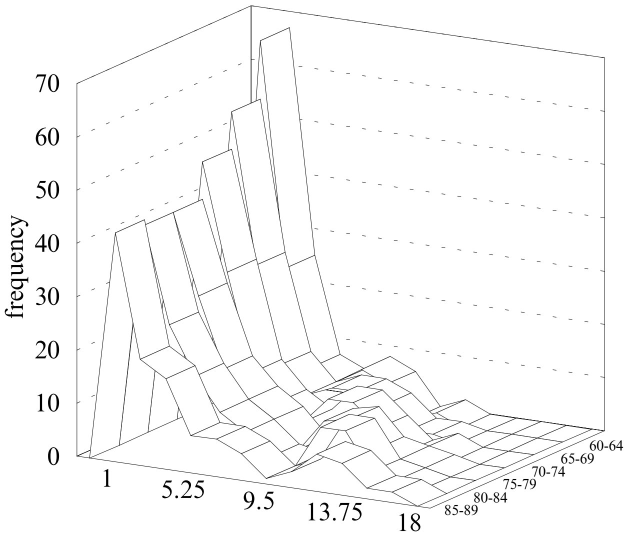

Fig. 5.2: Histograms of the average real GDP per capita divided by 1000 in six subperiods.

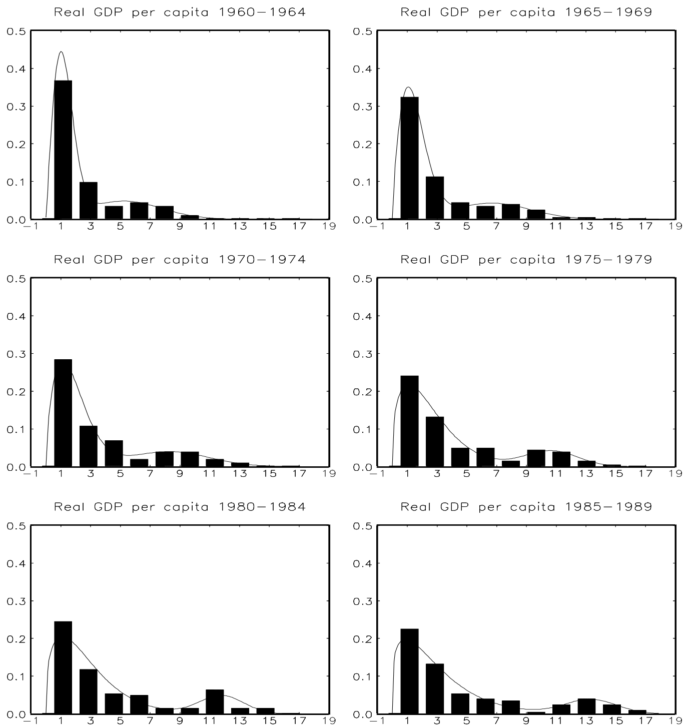

Fig. 5.3: Histograms of the average real GDP per capita divided by 1000 with the fitted density functions.

we conclude that: the distribution in each period is bimodal; a gap arises between the poor and rich group, which increases over time; the number of countries with an extremely low real GDP per capita decreases, but the spread of wealth within the poor group seems to rise. Similar findings are reported in *e.g.* Quah (1993a,b), Bianchi (1997) and Quah (1996a,b).

We end this section with three remarks. First on the loss of individual information through our histogram analysis, we note that a data summarization of 3600 individual observations into 30 yearly histograms - with only a relatively small number of cells - involves some loss of individual information. The optimal level of aggregation of information depends on the purpose of the empirical analysis. We are interested in describing and estimating efficiently such stylized facts as the behaviour of the poor and rich countries and their relative position through the post-war period. From the data summarization presented in this section we conclude as main stylized fact the bimodality of the empirical distribution of real GDP per capita.

Second on the relative merits of parametric and non-parametric analysis of income distributions, we note that we estimate the bimodal cross-section distribution of real GDP per capita per year by means of a mixture of two densities using individual observations per country. A mixture density belongs to a parametric class of densities which are defined as a convex combination of two or more densities. In our case these densities describe the distribution of the poor and the distribution of the rich countries, with a mixing distribution, representing the distribution over poor and rich. The separate analysis of the components of the mixture and of the relative importance of these components over time are the main advantages over a non-parametric approach as performed by, for instance, Desdoigt (1994). A clear choice between a parametric or a non-parametric approach depends on the availability of large data sets and on the purpose of the analysis. If there are many data over a long period then the asymptotically valid non-parametric approach is attractive in the sense that one can let the data ‘speak for themselves’. Often in economics there are not enough data to have a reliable non-parametric analysis. The parametric analysis is attractive in case there are no overly restrictive assumptions. In the next section we perform a sensitivity analysis with respect to the chosen functional form of the components of the mixture. One might also discuss the proper number of components in the mixture. Our choice of bandwidth and therefore the number of classes in the histograms are to some extent arbitrary. Using a different bandwidth in the histograms may result in the conjecture of more than two modes in the cross-section distribution. It is difficult to estimates a component of a mixture if the number of observations belonging to the components is very small, see also section 3 for a discussion about singularities in the likelihood function. Furthermore, the extra modes which occur using a smaller bandwidth, may also be due to noise. Bianchi (1997) rejects the hypothesis of more than two modes using a non-parametric approach based on the choice of the bandwidth. This supports our choice of two components in the mixture.

Third, on the choice between level, log of the level and relative level of real GDP per capita we note that in this paper we are interested in investigating convergence in the level of real GDP per capita. That is, that convergence implies that the differences in the level of real GDP per capita between countries disappear. As a byproduct we test in section 4 whether the growth rates of the the rich and the poor group of countries are the same. Another option is to scale the data by the sum of the real GDP per capita in each year as suggested by Canova and Marcet (1995) or to analyse log transformed data to test for convergence in relative welfare. In section 3 we show that our analysis is not sensitive to scaling the data in each year by a constant. A log transformation makes the data more homogenous and the evidence of bimodality in the data is considerably reduced, see Bianchi (1997). Homogeneity of the data is an attractive feature if one has to meet the assumptions of classical regression models, *e.g.* when testing for β-convergence. Also, one may use data on real GDP per worker instead of real GDP per capita in order to analyse convergence in productivity. In the present paper, we have chosen to focus on testing for convergence in the level of real GDP per capita.

3 Finite Mixture Distributions

We briefly discuss the representation, interpretation and estimation of mixtures distributions. For a good introductory survey of finite mixture distribution reference is made to Everitt and Hand (1981) or Titterington, Smith and Makov (1985). For our purpose it suffices to restrict ourselves to finite mixtures with a multinomial mixing distribution. In this case, the mixture density function g is defined as

$$g(y;\theta_1,...\theta_S,\lambda_1,...\lambda_{S-1}) = \sum_{s=1}^{S}\lambda_s f(y;\theta_s) \text{ with } \lambda_S = 1-\sum_{s=1}^{S-1}\lambda_s, \tag{5.1}$$

where S denotes the number of components in the mixture; $f(y;\theta_s)$, $s = 1,...,S$ are probability density functions evaluated at y depending on a parameter vector θ_s; and λ_s, $s = 1,...,S-1$ represent the mixing proportions. An example of a finite mixture distribution is a mixture of two normal distributions. The density function g evaluated at y_i is given by

$$g(y_i;\theta_1,\theta_2,\lambda) = \frac{\lambda}{\sigma_1\sqrt{2\pi}}\exp\left(-\frac{(y_i-\mu_1)^2}{2\sigma_1^2}\right) + \frac{(1-\lambda)}{\sigma_2\sqrt{2\pi}}\exp\left(-\frac{(y_i-\mu_2)^2}{2\sigma_2^2}\right), \tag{5.2}$$

where $\theta_1 = \{\mu_1,\sigma_1^2\}$ and $\theta_2 = \{\mu_2,\sigma_2^2\}$ denote the mean and the variance of the normal distribution of each component and λ represents the mixing proportion. For suitable chosen parameters, this mixture distribution is bimodal[13].

Interpretation

Representing the bimodal distribution of the data by a mixture of two densities is a convenient and interpretable way of describing the real GDP per capita. The distribution of the real GDP per capita of the poor countries is described by the first component of the mixture and the distribution of the rich countries by the second component. The mixing parameter λ gives the ex-ante probability that a country belongs to the first component of the mixture. Formally, the probability density function for the real GDP per capita for country i, denoted by y_i for $i = 1,..,N$ can be written as

$$g(y_i;\theta_1,\theta_2,\lambda) = \lambda f(y_i \mid s_i = 1;\theta_{s_i}) + (1-\lambda) f(y_i \mid s_i = 2;\theta_{s_i}), \tag{5.3}$$

[13] A sufficient condition that a value λ exists such that the mixture of two normal distributions is bimodal is $(\mu_2-\mu_1)^2 < (8\sigma_1^2\sigma_2^2)/(\sigma_1^2+\sigma_2^2)$.

where $\lambda = P[s_i = 1]$ and $1 - \lambda = P[s_i = 2]$ are the ex-ante probabilities that country i is poor or rich and where $f(y_i \mid s_i = 1; \theta_{s_i})$ and $f(y_i \mid s_i = 2; \theta_{s_i})$ are conditional probability density functions given that country i is poor or rich. The mean and variance of the conditional distribution of component s can be interpreted as the mean and the variance of the real GDP per capita of countries belonging to component s.

An attractive feature of our approach is that the mixing parameter λ is an *endogenous* parameter which determines the relative importance of each component in the mixture distribution.[14] So, *a priori* we do not impose an absolute borderline between the rich and the poor countries but let the data determine the relative importance of each group. One may interpret a mixture model as an unobserved component model in the following sense. To generate an observation y_i from a mixture, a country is selected to be poor with probability λ or to be rich with probability $(1-\lambda)$, or in other words the value of s_i is determined. Given that the country is poor the value of the real GDP per capita, y_i is generated by the conditional density function $f(y_i \mid s_i = 1; \theta_{s_i})$ (or $f(y_i \mid s_i = 2; \theta_{s_i})$ in case the country is rich). However, we only observe the value of the real GDP per capita y_i and not the value of s_i. Given the realized value of the real GDP per capita y_i and given the values of the parameters θ_1, θ_2 and λ, we can make inference about the value of s_i. The conditional probability that observation y_i is generated by the first component ($s_i = 1$) for the mixture defined in (5.3) is defined as

$$\Pr[s_i = 1 \mid y_i; \theta_1, \theta_2, \lambda] = \frac{\lambda f(y_i \mid s_i = 1; \theta_{s_i})}{\lambda f(y_i \mid s_i = 1; \theta_{s_i}) + (1-\lambda) f(y_i \mid s_i = 2; \theta_{s_i})} \tag{5.4}$$

This conditional probability denotes the ex-post probability that a country is poor and is used for the investigation of mobility in wealth in section 5. Note that the ex-post probability of being rich $\Pr[s_i = 2 \mid y_i; \theta_1, \theta_2, \lambda]$ equals $1 - \Pr[s_i = 1 \mid y_i; \theta_1, \theta_2, \lambda]$ by definition.

In practice, we do not know the true values of the parameters θ_1, θ_2, λ and we have to replace them by their estimates. The estimated λ can be interpreted as the proportion of countries belonging to the first component, *i.e.* the percentage of poor countries, while the probability in (5.4) can be seen as the relative ex-post contribution of country i to the first component. (Note that in case of a mixture of normal densities the estimated mean $\hat{\mu}_1 = \frac{1}{\lambda N} \sum_{i=1}^{N} \Pr[s_1 = 1 \mid y_i; \hat{\theta}_1, \hat{\theta}_2 \hat{\lambda}]\, y_i$, *i.e.* a weighted average of the observations.)

Since countries can switch over time from being poor to being rich and *vice versa* the mixing proportion λ can change through time. The growth in real GDP per capita causes changes in the means of the mixture components through time. Further, countries belonging to a group do not need to have the same growth rates, which implies that the variance does not have to be the same over time. Note that a change in the mean and/or the variance of a component can also be caused by movement of countries between the rich and the poor group.

[14] Durlauf and Johnson (1995) use the regression tree technique to endogenously split the data in multiple regimes.

Estimation

Several methods have been proposed to estimate the parameters of a mixture, *e.g.* maximum likelihood and the methods of moments, see Everitt and Hand (1981). We follow the maximum likelihood approach, which implies maximising the following criterion function

$$\mathscr{L}(Y;\theta_1,\theta_2,\lambda) = \prod_{i=1}^{N} g(y_i;\theta_1,\theta_2,\lambda), \tag{5.5}$$

where the density function g is given by (5.2) and $Y = \{y_1,...,y_N\}$. From the first order conditions, it is easy to see that maximising the likelihood implies a non-linear optimisation problem. Standard numerical algorithms can be use to maximise the likelihood function. Note that the likelihood function (5.5) for estimation of a mixture of normal densities (5.2) has not a global maximum, since a singularity in the likelihood function arises, whenever one of the components is imputed to have a mean equal to one of the observations (μ_1 equals y_i) with zero variance ($\sigma_1^2 \to 0$). At that point the value of the likelihood function becomes infinite. Kiefer (1978) shows that if there exists a local maximum in the interior of the parameter region then this maximum yields consistent, asymptotically normal estimators of the parameters. In this case, the ML estimators are not values of the parameters which maximize the likelihood function globally, but are those solutions of the likelihood equations, which yields asymptotically the largest value of the likelihood function. In practice, if a numerical optimisation algorithm gets "stuck" at a singularity, the easiest strategy is to try a different starting value. Another solution is to use a quasi Bayesian approach by multiplying the likelihood function by a prior density to cancel out the singularity problem (see Hamilton, 1991).

A mixture of two normal densities does not suffice to describe our bimodal distributions. It is clear from figure 5.3 that the first component of the mixture distribution is skew. Another point is that real GDP per capita can never be negative, so a mixture of normal densities is, strictly speaking, not appropriate. Possible candidates to describe the distribution of the poor countries (first component) are *e.g.* the Weibull distribution, the gamma distribution and the lognormal distribution. For the distribution of the rich countries a normal distribution (truncated at 0) seems appropriate.

We have estimated several combinations of the proposed distributions and compared the fit to select the best candidates. To analyse the fit of these distribution we divide the data in each of the six subperiods in equally-sized intervals. In each subperiod we compare the number of observations in each interval with the expected number of observations in the interval based on the estimated mixture distribution using a χ^2 goodness of fit test. We note that this strategy is dependent on the number of intervals. We choose 8 through 15 equally spaced intervals to evaluate the estimated mixtures. This means that we perform $(15-7)\times 6 = 48$ goodness of fit tests for each candidate mixture density. Table 5.1 shows the number of rejections for different mixtures in each subperiod using a 5% level of significance. We note that three cases result in four rejections: the mixture of a Weibull or a gamma with a truncated normal density and the mixture of two truncated normal distributions. The other mixtures including the mixtures containing the lognormal distribution perform worse. To choose between the three best fitting mixtures, we look at the number of rejections at the 1% and 10% level. In that case the mixture of a Weibull and a truncated normal distribution produces the best fit.

Table 5.1: The outcomes of χ^2 goodness of fit test for different mixture distributions[1].

components[2]		subperiod					
first	second	1960–64	1965–69	1970–74	1975–79	1980–84	1984–89
normal	normal	0	1	1	1	1	0
gamma	normal	0	0	0	1	3	0
gamma	gamma	0	0	0	2	3	0
lognormal	normal	0	0	0	1	4	0
lognormal	lognormal	0	0	0	1	5	2
Weibull	normal	0	0	0	1	3	0
Weibull	Weibull	0	0	0	1	4	0

[1] The cell denotes the number of rejections at a 5% level out of eight χ^2 goodness of fit test in each subperiod. The data in each subperiod are divided in 8 through 15 equally-sized interval. The χ^2 test compares the number of observations in each interval with the expected number of observations in the interval based on the estimated mixture distribution.
[2] Normal means truncated normal with 0 as point of truncation.

Figure 5.3 shows the fitted density of a mixture of a Weibull and a truncated normal together with the histograms of figure 5.3. The histograms have been normalised such that the area under the bars is equal to one in order to compare them with the density functions. The estimated mixtures fit the histograms reasonably well. Therefore, we decide to consider in this paper a mixture of a Weibull and a truncated normal density. Since a gamma and a truncated normal distribution are also good candidates to describe the first component, we discuss the robustness of our results with respect to the other two mixtures at the end of each section. The density function h of a mixture of a Weibull and a truncated normal evaluated at y_i is given by

$$h(y_i;\beta_1,\alpha_1,\lambda,\mu_2,\sigma_2) = \lambda\frac{\beta_1}{\alpha_1^{\beta_1}}y_i^{\beta_1-1}\exp\left(\left(-\frac{y_1}{\alpha_1}\right)^{\beta_1}\right) + (1-\lambda)\frac{\phi(y_i;\mu_2,\sigma_2^2)}{\Phi(\mu_2/\sigma_2)}, \tag{5.6}$$

where $\phi(y;\mu_2,\sigma_2^2)$ represents the probability density function of a normal distribution with mean μ_2 and variance σ_2^2 and Φ the cumulative density function of a standard normal distribution. The parameters α_1 and β_1 are the scale and location parameters of the Weibull component. The parameters of the mixture $\{\alpha_1,\beta_1,\mu_2,\sigma_2,\lambda\}$ are estimated by maximizing the likelihood function

$$\mathscr{L}(Y;\beta_1,\alpha_1,\mu_2,\sigma_2,\lambda)=\prod_{i=1}^{N}h(y_i;\beta_1,\alpha_1,\mu_2,\sigma_2,\lambda), \tag{5.7}$$

where the density function h is given in (5.6). Here we face, of course, the same problem with the singularity in the likelihood function as in the case of a mixture of two normal densities and we opt for the same solution as before. The numerical algorithm to maximise the likelihood functions (5.7) is Newton-Raphson. A range of starting values is used to find the maximum. In case two or more maxima are found the maximum with the largest value of the likelihood function is chosen. Finally, it can easily be shown that scaling of the data via multiplying by a constant k does not influence the estimated value of the mixing parameter and changes the other parameters in the corresponding way, $k\alpha_1$, $k\mu_2$ and $k\sigma$. Therefore, scaling the data by the sum of the real GDP per capita in a year does not alter the conclusions, since the means and the variances of the components change accordingly.

4 Distribution of Wealth

To describe the cross-section distribution of real GDP per capita over the 120 countries in each year, we estimate a mixture of a Weibull and a truncated normal density. First, we focus on the six subperiods. The first five columns of table 5.2 show the parameter estimates of the fitted mixture distributions in every subperiod. Apart from the mixing proportion λ it is difficult to interpret the estimated scale and location parameters directly, since they do not represent the means and variances of the components. Therefore, the second panel of the table shows the means and the variances of the poor and the rich group based on the parameters estimates together with the mean and variance of all countries. Note that the truncation of the normal component becomes less important in the end of the sample.

From the sixth column of table 5.2 we notice that the mixing proportions indicate an almost constant percentage of poor countries in the first three subperiods followed by a substantially increase after the subperiod 1970–1974. There are 14% more poor countries in the final subperiod than in the first subperiod. A Likelihood Ratio [LR] test for equal mixing proportions in the first and final subperiod equals, however, 2.56 which is not significant at a 5% level (the 95% percentile of the χ^2 distribution with one degree of freedom equals 3.84). The LR test is computed by comparing the sum of the maximum likelihoods of the two unrestricted densities with the maximum likelihood of the mixture densities in the first and final period estimated under the restriction of equal mixing parameters.

The seventh column of table 5.2 shows the mean of all countries in every subperiod. The mean has increased monotonically over time. The same is true for the means of the poor and the rich group. Notice that the mean real GDP per capita of both groups has grown faster than the overall mean. This is possible because the relative number of poor countries has increased over time. The difference between the mean of the poor and rich group is about 4.1 in the first subperiod, while in the

Table 5.2: Estimates of mixture parameters, means and variances of real GDP per capita of the poor and the rich component and of all countries in the six subperiods.[1]

sub-period	$\hat{\beta}_1$	$\hat{\mu}_2$	$\hat{\alpha}_1$	$\hat{\sigma}_2^2$	$\hat{\lambda}$	mean			variance		
						all	poor	rich	all	poor	rich
60–64	2.02	5.24	1.40	6.27	0.70	2.46	1.24	5.35	5.55	0.41	5.66
65–69	1.79	6.68	1.68	6.50	0.73	2.91	1.49	6.71	7.75	0.74	6.27
70–74	1.63	8.29	2.03	7.05	0.73	3.53	1.81	8.30	11.20	1.31	6.98
75–79	1.38	10.66	2.76	3.13	0.82	4.01	2.52	10.66	13.71	3.41	3.13
80–84	1.35	11.77	2.95	2.18	0.82	4.32	2.70	11.78	16.29	4.08	2.18
84–89	1.25	13.32	3.15	3.02	0.84	4.62	2.93	13.33	20.60	5.57	3.02

[1] The mean and variance of the truncated normal component are computed using the formulae in appendix A of Maddala (1986). The mean and the variance of a Weibull distribution are $\alpha_1\Gamma(1+1/\beta_1)$ and $\alpha_1^2\Gamma(1+2/\beta_1)-(\alpha_1\Gamma(1+1/\beta_1))^2$ respectively.

final subperiod this difference is 10.4. This indicates that the means of the real GDP per capita of the two groups are diverging. However, the growth rates in the mean of both groups are roughly the same. The mean of real GDP per capita of the rich countries in the final subperiod is two and a half times larger than in the first subperiod. For the poor group this factor is about 2.4. A LR test for equal growth rates equals 0.12, which is not significant at a 5% percent level. This means that although the difference in the mean between the poor and the rich group gets larger over the last 25 years, the growth rates of the means of both groups over this period are not significantly different. To compute the LR test we estimate the mixture distribution in the first period and the final period jointly under the restriction of an equal growth rate.

The final three columns of table 5.2 display the variance of the poor, the variance of rich and the variance of all countries. The total variance has increased monotonically over the last 25 years. The same conclusion can be drawn for the spread of wealth within the poor group, which indicates the absence of convergence within the poor group. For the group of rich countries an increase in the spread of wealth is followed by a decrease after the subperiod 1970–1974.

We have to interpret the results of the diverging means with care. Changes in the mean of each component over time can be caused by two forces. First, the real GDP per capita of countries in a group can increase over time. Second, countries can switch from the poor to the rich group and *vice versa*, which can lead to a change in the ratio of the means of the rich and the poor group. A typical example of the latter occurs when only the very rich countries stay in the rich group. The same kind of reasoning counts for the variances of each component. Changes in the variances of the components can also be caused by changes in the mixing parameter.

To correct for the effect of the decrease in the number of rich countries on the development of the means and variances of the components, we estimate in each period a mixture of a Weibull and a truncated normal with equal mixing proportions. We analyse three different scenarios. First, we determine an optimal mixing parameter for the six subperiods by jointly estimating the mixture densities under the restriction of equal mixing parameters. Next, we set the mixing parameter equal to the estimated mixing parameter in the final subperiod (=0.84) and equal to the estimated parameter in the first subperiod (=0.70). Notice that we theoretically still allow for switches of countries between the poor and the rich group. Using the same techniques that we apply in the next section, we can show that the number of switches between the two groups is low. This means that the rich and the poor group contain almost the same countries in every subperiod.

Table 5.3 shows the means and the variances of each component under the different restrictions on λ. Several conclusions emerge from the results of this table. Not surprisingly, fixing the mixing parameter results in different values for the means of both groups. However, for all three scenarios, the means of the poor and the rich group still diverge, which implies that the change in the number of rich countries is not the driving force in the diverging process. Note that the growth rates in real GDP per capita over the last 25 years of the rich and the poor group are still about the same.

The variances of the components are more sensitive to the value of the mixing parameter. Under equal mixing parameters, the variance of the poor group still increases over time. For the rich group the situation is different. From the lower left panel of table 5.3 we observe that the variance of the countries, which were rich in the beginning of the sample, is increasing over time. This indicates that the decrease in variance, when we allow for a changing mixing parameter, is mainly due to the decrease in the number of rich countries. Hence, a number of countries, which originally were located in a middle group, was not capable of catching-up with the remaining rich countries. The lower right panel of table 5.3 shows the development of the variance of the countries, who ended up rich in the last subperiod. We still notice the decrease in the variance after the period 1970–1974 and the increase after 1980–1984 but the changes in the variances are less pronounced.

The results in tables 5.2 and 5.3 are not suitable to notice short run patterns, since we have considered the average real GDP of five consecutive years. In the remainder of this section we analyse the distribution of the real GDP per capita using a mixture of a Weibull and a truncated normal density for each year from 1960 until 1989. Instead of using tables with parameter estimates, we report the main results in several graphs, which show the interesting aspects of the estimated distributions[15].

Figure 5.4 shows the estimated values of the mixing proportions λ. In 1960 the percentage of poor countries was about 71%. In the first part of our sample there is an overall effect of a decrease in the number of poor countries to 67% in 1973, but after 1973 the number of poor countries has risen especially during the period 1975–1977. At the end of the sample the percentage of poor countries seems to stabilise around 83%. These results match the outcomes of table 5.2.

---

[15] A detailed outline of the parameter estimates can be obtained from the authors.

Table 5.3: The means and the variances of the poor and the rich component in the six subperiods for different values of the mixing parameters[1].

	unrestricted value of λ				constant value of $\lambda = 0.79$			
	mean		variance		mean		variance	
subperiod	poor	rich	poor	rich	poor	rich	poor	rich
1960-1964	1.24	5.35	0.41	5.66	1.35	6.11	0.57	4.23
1965-1969	1.49	6.71	0.74	6.27	1.59	7.32	0.93	4.70
1970-1974	1.81	8.30	1.31	6.98	1.97	8.99	1.70	4.90
1975-1979	2.52	10.66	3.41	3.13	2.44	10.44	3.09	3.80
1980-1984	2.70	11.78	4.08	2.18	2.68	11.73	3.92	2.30
1984-1989	2.93	13.33	5.57	3.02	2.82	13.08	4.94	3.98

	initial period value of $\lambda = 0.70$				final period value of $\lambda = 0.84$			
	mean		variance		mean		variance	
subperiod	poor	rich	poor	rich	poor	rich	poor	rich
1960-1964	1.24	5.35	0.41	5.66	1.42	6.52	0.70	3.34
1965-1969	1.46	6.44	0.69	6.90	1.66	7.64	1.08	3.94
1970-1974	1.73	7.89	1.14	8.15	2.08	9.34	2.03	4.00
1975-1979	2.04	8.95	1.82	8.40	2.58	10.77	3.67	2.81
1980-1984	2.25	10.18	2.33	7.99	2.72	11.80	4.19	2.13
1985-1989	2.28	10.69	2.34	13.16	2.93	13.33	5.57	3.02

[1] The results in the upper right corner are ased on a joint estimate of the six mixture densities with equal λ parameter. In the lower panel of the table the λ is equal to the estimated λ in the first and the final subperiod respectively, see the sixth column of table 5.2.

Figure 5.5 shows means and variances of the real GDP per capita in each year for the period 1960–1989, which are based on the parameter estimates of the mixtures. The left panel of the figure shows the overall means and the means of each component. The mean of the real GDP per capita of all countries has increased almost monotonically during the whole period. There are small decreases in the periods 1974–1975 and 1980–1983 reflecting the oil crisis and the crisis in the beginning of the eighties. These periods of decrease can also be detected in the mean of the poor group and the mean of the rich group. In 1960 the difference in the means is about 3.8, while in 1989 this difference is 11. The means of both groups are diverging, which leads to a gap between the poor and the rich group. If we however look at the growth rates of both groups we see that for the poor group the real GDP per capita in 1989 is about 2.5 times larger than in 1960, while for the rich group the factor is about 2.8. A LR test for equal growth rates equals 0.37, which is not significant at 5% level of significance. Therefore, this implies again that although the means

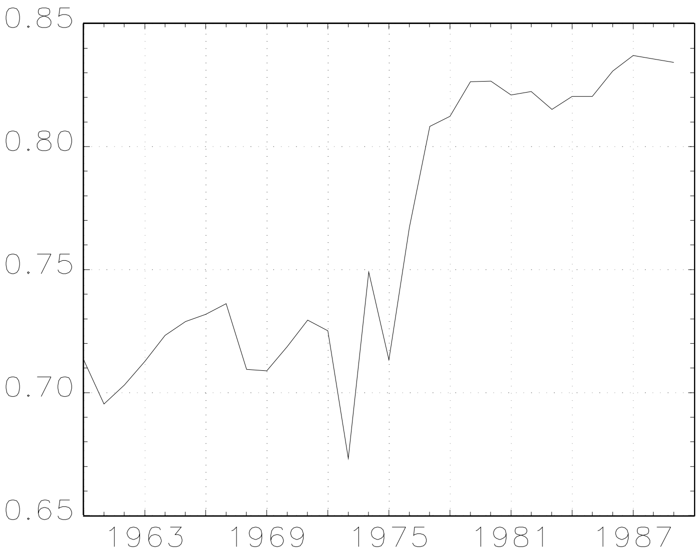

Fig. 5.4: Estimated mixing proportions in each year (1960–1989).

of the poor and the rich group are diverging, the growth rates of the means of both groups are the same over the period 1960–1989. To investigate whether changes in the mixing parameter are responsible for the effects on the means, we estimate the mixture densities under the restriction of equal mixing parameters like in table 5.3. Unreported results show that although we find slightly different values for the means of the poor and the rich group, the means of the two groups are still diverging and the growth rates of the two groups are still about the same.

The right panel of figure 5.5 shows the variance for all countries and for the poor and the rich group in every year. The variance of the real GDP per capita of all countries has risen during 1960–1989 indicating an increase in the spread of wealth between all countries. There are two short periods with a decrease in the variance, *i.e.* 1974–1975 and 1980–1982. The same periods can be found in the variance of the poor group. Unreported estimation results show that the increase in the variance of the pour group remains if we fix the mixing proportion λ. The sharp increase in the variance of the poor group after 1975 is due to the increase in the number of poor countries.

Figure 5.5 shows an increase in the spread of wealth within the rich group until 1973. After the oil crisis the variance has decreased strongly until 1982. In the period

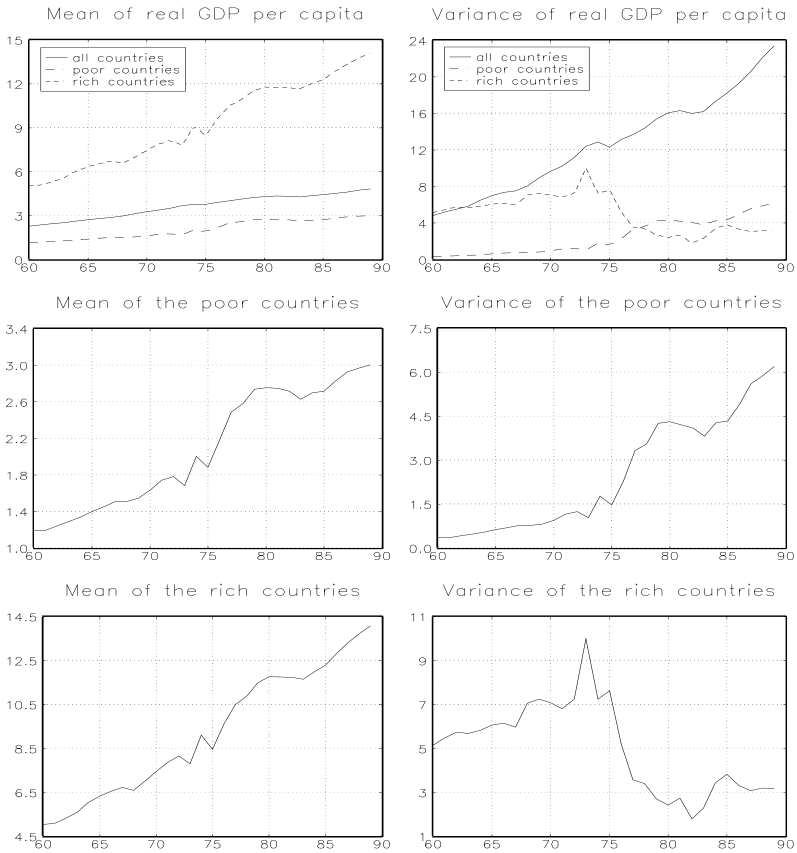

Fig. 5.5: Means and variances of the poor and the rich component and all countries.

1982-1986 there is an increase in the variance. The same analysis as in table 5.3 shows that the variance of the countries, which are rich in the beginning of the sample, is increasing over time and that the variance of the countries, which are rich at the end of the sample, does not decrease. Furthermore, the decrease in the variance of the rich component during the two crises still remains if we fix the mixing parameter, but the decreases are much smaller. In summary, the analysis shows that especially in the middle of the seventies a number of countries was not capable of catching-up with the rich countries and became poor. This has caused a gap between the poor and the rich group. The movement of the poorest rich countries from the rich to the poor group leads to an increase in the variance of the poor countries and a decrease in the variance of the very rich countries.

In this section we have analysed the development of the real GDP per capita over time using a mixture of a Weibull and a truncated normal density. We have seen that the number of poor countries has increased over the last 30 years. The difference in the mean of the real GDP per capita of the poor and the rich group is increasing, indicating no convergence in the level. However, there is no significant difference in the growth rates of both groups, which suggests convergence in growth rates. The spread of wealth within the poor group increases. This is partially caused by the increase in the number of poor countries. For the rich group there is some indication for convergence as the spread of wealth of the rich group has decreased during the two crises in our sample. The largest part of these decreases is however due to the decrease in the number of rich countries. These rich countries were not capable of catching-up with the very rich countries.

In order to investigate the sensitivity of the results with respect with our choice of mixture, we performed the same analysis of cross-section distribution of the real GDP per capita using a mixture of a gamma and a truncated normal density and a mixture of two truncated normal densities, which also produce a reasonable fit according to table 5.1. The results coming out of these analyses are roughly the same. The main difference lies in the estimated mixing proportions before 1974, using a mixture of two truncated normals. The estimated mixing proportions are about 0.10 smaller compared to the mixtures of a Weibull or gamma and a truncated normal. In the next section we analyse the intra-distribution movement of countries within the estimated mixtures. We also consider in more detail the switches of countries between the poor and the rich group.

5 Mobility in Wealth

So far our analysis was limited to describing the development of the distribution of real GDP per capita in each year. In this section we consider the intra-distribution mobility of wealth. The obvious strategy is to look at switches of countries and/or groups of countries from the poor to the rich group and *vice versa*. From figure 5.4 we observe that the mixing proportion has risen during the period 1960–1989 indicating an increase in the number of poor countries. One might conclude that the main mobility between the two groups consists of countries moving from the rich to the poor group. However, even when the mixing parameter is rising over time, there can be switches from poor to rich, when the number of rich countries that become poor is larger than the number of poor countries that become rich. We start analysing mobility in wealth by considering the individual switches of countries between the two groups.

To analyse the mobility between groups, we need to decide whether a country is rich or poor. We can do inference about this question based on the ex-post conditional probability that an observation is generated by one of the components of the mixture, see (5.4). We declare a country poor, if the ex-post conditional probability that a country belongs to the first component of the mixture is larger than 50%, *i.e.* $\Pr[s_i = 1 \mid y_i; \hat{\theta}_1, \hat{\theta}_2, \hat{\lambda}] > 0.5$, otherwise the country is labelled as rich. Note that this means that a rich country can become poor even if the level of real GDP per capita of this country does not change or even increases. Such a situation can for instance arise when the other rich and the poor countries grow faster than this country. In summary, switching from rich to poor depends on the relative movement of a country in the distribution with respect to the other countries.

Table 5.4 displays the number of countries that belong to each group based on the ex-post conditional probability. We see that the number of poor countries has risen from 87 in first subperiod to 100 in the last subperiod. The movements from the poor group to the rich group can be summarised as follows. After the first subperiod only Hong Kong moves from the poor group to the rich group and stays in the rich group for two subperiods. However, after 1974 Hong Kong moves back to the rich group. Furthermore, Barbados moves from the poor to the rich group after the second subperiod and stays in the rich group for only one subperiod. The number of movements from the rich group to the poor group is much larger. Especially after the subperiod 1970–1974 many countries have moved from the rich to the poor group including Argentina, Puerto Rico, Iran and Israel, Spain and Ireland. These countries were not able to catch-up with the very rich countries. After 1979 only Venezuela, Trinidad and Saudi Arabia have moved from the rich group to the poor group. Before 1970, Martinique, Barbados, Mexico and Chile have moved from rich to the poor group. There are 19 countries that are rich in every period, *i.e.* Canada, the USA, Japan, Australia, New Zealand, Iceland, Switzerland, Sweden and all countries of the European Union except for Greece, Ireland, Portugal and Spain. There are 86 countries including most of the African and Asian countries that are poor in every period.

The same analysis can be performed using the estimation results in each year. Figure 5.6 shows the number of rich and poor countries in each year based on the ex-post probabilities of the estimated mixtures. In the period 1960–1973 the number of poor countries drops from 88 to 83. After 1975 we see an increase in the number

Table 5.4: The number of rich and poor countries in each subperiod based on ex-post probabilities[1].

subperiod	60–64	65–69	70–74	75–79	80–84	85–89
# poor	87 (→ 86)	90 (→ 89)	90 (→ 90)	98 (→ 97)	98 (→ 98)	100
# rich	33 (→ 29)	30 (→ 29)	30 (→ 22)	22 (→ 21)	22 (→ 20)	20

[1] In parentheses the number of countries that are in the same group the following period.

of poor countries resulting in 99 poor and 21 rich countries in the final year of our sample. The majority of the switches is from the rich to the poor group.

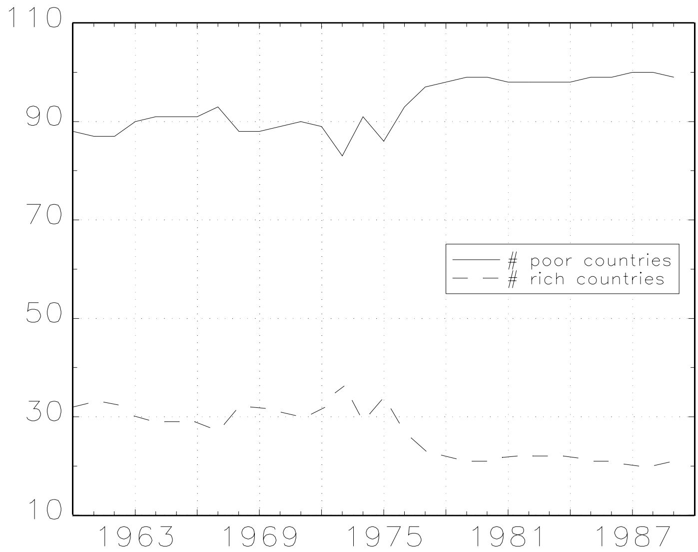

Fig. 5.6: Number of rich and poor countries in each year based on ex-post probabilities.

To investigate the intra-distribution movements of countries we follow the strategy proposed by Quah (1993a). He analyses the intra-distribution dynamics of real GDP per worker over time by a so called fractile Markov Chain. Formally, let F_t denote the distribution of real GDP per worker at time t and suppose that the distribution at time $t+1$ can be written as

$$F_{t+1} = MF_t, \tag{5.8}$$

where M is an operator which maps the distribution F at time t into the distribution at time $t+1$. Iteration of (5.8) gives a prediction for future distributions of the ex-post probabilities

$$F_{t+k} = \underbrace{M...M}_{k} F_t = M^k F_t. \tag{5.9}$$

Quah (1993a) approximates the operator M by an transition matrix by discretising the distribution F_t into intervals. Then M becomes a transition matrix of a

Markov chain. The ergodic probabilities of the Markov chain give insight in the limiting distribution over the states[16]. The transition matrix M is estimated by averaging the total number of switches between the predefined intervals on F. A more technical description of analysing mobility using Markov chains can be found in Shorrocks (1978) and Geweke, Marshall and Zarkin (1986).

In this paper we use the simple framework of Quah (1993a) to analyse the movements of countries between rich and poor. For the distribution F_t we choose the cross-section distribution of the *ex-post probabilities of being poor* in year t (Since the ex-post probability of being poor is equal to one minus the ex-post probability of being rich, we can limit ourselves to analysing the first probabilities.). To estimate the M matrix we divide the the cross-section distribution of ex-post probabilities of being poor at time t, F_t into equally-sized intervals, which is in the line of Quah (1993b). The $[0,1]$ interval on which F_t is defined, is divided into 2, 3 and 4 equally-sized intervals. In the case of 2 equally-sized intervals, we consider movements from the rich to the poor group and *vice versa*. The division into 3 intervals is useful to analyse whether countries who initially belong to a "middle" group can catch up with the rich countries or fall behind. Movements within the rich and the poor group can be analysed if we use 4 subdivision. The transition matrix M is estimated by averaging the total number of switching between the states over 30 years.

Table 5.5 shows the estimated values of M for the three proposed subdivisions. The transition matrix of the 2-state Markov process shows that the probability of staying poor is larger than the probability of staying rich. The ergodic probabilities of being poor is 0.83, which matches the estimates of the mixing proportions in the last years of our sample period. The transition matrix of the 3-state Markov chain shows the probability of moving from the middle group to the poor group is larger than *vice versa*, which indicates that the probability of catching up is smaller than the probability of falling behind. The ergodic probability of being in the middle group shows that the middle group is vanishing. This matches our earlier findings on the divergence of the levels of the means of the poor and the rich group in section 4 and corresponds with the stylized facts, discussed in section 2. We note that the inconsistency in the ergodic probabilities (0.83 for 2-state, 0.86 for 3-state) is due to the relatively small sample size. The transition matrix of the 4-state Markov process show that if a country is very poor there is almost no chance of becoming rich anymore. The probability to catch up is larger for countries who are in the middle rich group than for countries in the middle poor group. The diagonal elements of the transition matrices are always larger than 0.5, except for the state enabled middle rich in the 4-state Markov process. Further, only sub- and superdiagonal elements differ substantially from zero except for the transition from middle rich to rich, indicating that there are almost no major movements in relative wealth. This implies that the rate at which convergence proceeds, is not large enough for the poorest countries to escape from a poverty trap. Similar findings are reported in Quah (1993a,b).

[16] As Quah (1993b) indicates, this framework is much too simple for forecasting. The limiting distribution should be interpreted as an indication for the long-run tendencies in the data rather than a forecast.

There is no need that the transition matrix M is time invariant or that the law of motion for F_t is first order. The former statement is not straightforward to analyse in the present framework. The latter however, can be analysed by considering, for instance, second and higher order Markov chains and compare the estimates of the second order transition matrix with estimates of M from table 5.5 to the power two or to compare the ergodic probabilities. In table 5.6 we show the ergodic probabilities based on a first, second and a third order Markov process. We see that if we increase the order of the chain the ergodic probability of being poor increases. However, the conclusions about the long-run tendencies in the data stay the same.

In this section we have analysed the mobility in wealth using the outcomes of the estimated mixtures of a Weibull and a truncated normal density. The main mobility

Table 5.5: Intra-distribution movements in real GDP per capita analysed using a first order Markov Chain on the ex-post probabilities.

first order Markov process (2-states)

	poor	rich
poor	0.99	0.05
rich	0.01	0.95
ergodic[1]	0.83	0.17

first order Markov process (3-states)

	poor	middle	rich
poor	0.99	0.24	0.01
middle	0.01	0.60	0.03
rich	0.00	0.17	0.96
ergodic[1]	0.86	0.02	0.12

first order Markov process (4-states)

	very poor	middle poor	middle rich	very rich
very poor	0.99	0.24	0.10	0.00
middle poor	0.01	0.52	0.31	0.01
middle rich	0.00	0.22	0.38	0.02
very rich	0.00	0.02	0.21	0.97
ergodic[1]	0.88	0.02	0.02	0.08

[1] Ergodic probabilities of the Markov Chain.

Table 5.6: Ergodic probabilities of a first, second and a third order Markov process with 2, 3 and 4 states, see table 5.5.

	order	poor	rich		
	1	0.83	0.17		
2 subdivisions	2	0.87	0.13		
	3	0.89	0.11		
	order	poor	middle	rich	
	1	0.86	0.02	0.12	
3 subdivisions	2	0.90	0.02	0.08	
	3	0.92	0.02	0.06	
	order	very poor	middle poor	middle rich	very rich
	1	0.88	0.02	0.02	0.08
4 subdivisions	2	0.93	0.01	0.01	0.05
	3	0.94	0.01	0.01	0.04

we have detected is movements of countries from the rich group to the poor group, which have caused the increase in the number of poor countries. The middle group has vanished into the poor group because of the inability of poor countries to catch up with the rich countries. The main results stay the same if we use a Gamma instead of a Weibull distribution to describe the distribution of the poor countries. If we however take a mixture of two truncated normal distributions, we observe a bit more mobility in the beginning of the sample, but after 1975 the results are the same.

6 Conclusion

In this paper we have analysed the distribution of real GDP per capita over 120 countries during the period 1960–1989. The cross-section distribution of the real per capita GDP turns out to be bimodal, displaying a relative large group of poor countries and a small group of rich countries. The analysis is split up in two parts. In the first part we describe the bimodal distributions in each year by a mixture of a Weibull and a truncated normal density and analyse the mixing proportions, the means and variances of the components of the mixture. In the second part we use the estimated mixture distributions for analysing intra-distribution mobility.

The analysis of the cross section distributions shows that the means of the real GDP per capita of the poor and the rich group are diverging, resulting in an increasing gap between the poor and the rich group in terms of levels. However, there is indication of convergence in growth rates between the two groups. The analysis of the mixing proportions shows a large increase in the number of poor countries in the middle of the seventies, which results in an increase in the spread of wealth within the poor group and a decline in the spread of wealth within the rich group. The analysis of the mobility of wealth shows that the main mobility is from rich to poor and the "middle" group between poor and rich disappears. The probability to catch up for the poor countries is smaller than the probability of falling behind. The rate at which convergence proceeds, is not large enough for the poorest countries to escape from a poverty trap.

The results have to be interpreted with care and further research is needed. Specific further research topics are to consider conditioning variables and to link up with endogenous growth models.

Acknowledgements

This research was sponsored by the Economic Research Foundation, part of the Netherlands Organization for Scientific Research (N.W.O.). We thank Suzanne Bijl for preliminary data analysis, Gordon Anderson, John Geweke, Andrew Lewbell, Gary Chamberlain, David Hendry, Mars Cramer and participants of several seminars and of the seventh World Congress of the Econometric Society in Tokyo. We are, in particular, indebted to three anonymous referees for very helpful comments on an earlier version of this paper.

References

Abramowitz, M. (1986) Catching Up, Forging Ahead, and Falling Behind, *Journal of Economic History*, **46**, 385–406.

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

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

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

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

Canova, F. and A. Marcet (1995) The Poor Stay Poor: Nonconvergence Across Countries and Regions, cEPR Discussion Paper Series, No. 1265.

Desdoigts, A. (1994) Changes in the World Income Distribution: A Non-Parametric Approach to Challenge the Neo-Classical Convergence Argument, Working Paper, Université Paris Dauphine.

Durlauf, S.N. and P. Johnson (1995) Multiple Regimes and Cross-Country Growth Behaviour, *Journal of Applied Econometrics*, **10**, 365–384.

Everitt, B. S. and D. J. Hand (1981) *Finite Mixture Distributions*, Monographs on Applied Probability and Statistics, Chapman and Hall, London.

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

Geweke, J., R.C. Marshall and G.A. Zarkin (1986) Mobility Indices in Continuous Time Markov Chains, *Econometrica*, **54**, 1407–1423.

Hamilton, J.D. (1991) A Quasi-Bayesian Approach to Estimating Parameters of Mixtures of Normal Distributions, *Journal of Business and Economic Statistics*, **9**, 27–39.

Jones, C.I. (1997) Convergence Revisited, *Journal of Economic Growth*, **2**, 131–153.

Kiefer, N.M. (1978) Discrete Parameter Variation: Efficient Estimation of a Switching Regression Model, *Econometrica*, **46**, 427–434.

Levine, R. and D. Renelt (1991) Cross Country Studies of Growth and Policy: Some Methodological, Conceptual, and Statistical Problems, World Bank Working Paper Series, No. 608.

Levine, R. and D. Renelt (1992) A Sensitivity Analysis of Cross Country Regressions, *American Economic Review*, **82**, 942–963.

Mankiw, N.G., D. Romer and D.N. Weil (1992) A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics*, **429**, 407–437.

Parente, S.L. and E.C. Prescott (1993) Changes in the Wealth of Nations, *Quarterly Review, Federal Reserve Bank of Minneapolis*, **Spring 1993**, 3–16.

Quah, D. (1993a) Galton's Fallacy and Tests of the Convergence Hypothesis, *Scandinavian Journal of Economics*, **95**, 427–443.

Quah, D. (1993b) Empirical Cross-section Dynamics in Economic Growth, *European Economic Review*, **37**, 426–434.

Quah, D. (1996a) Convergence Empirics Across Economies with (some) Capital Mobility, *Journal of Economic Growth*, **1**, 95–124.

Quah, D. (1996b) Empirics for Economic Growth and Convergence, *European Economic Review*, **40**, 1353–1376.

Sala-i-Martin, X. (1994) Cross-sectional Regression and the Empirics of Economic Growth, *European Economic Review*, **38**, 739–747.

Shorrocks, A.F. (1978) The Measurement of Mobility, *Econometrica*, **46**, 1013–1024.

Summers, R. and A. Heston (1991) The Penn World Table (mark 5): An Expanded Set of International Comparisons, 1950–1988, *Quarterly Journal of Economics*, **106**, 327–368.

Titterington, D.M., A.F.M. Smith and U.E. Makov (1985) *Statistical Analysis of Finite Mixture Distributions*, Wiley and Sons, New York.