Growth and Ergodicity: Has the World Converged?

John S. Landon-Lane — The University of New South Wales

J. Arnold Quinn * — University of Minnesota

January 7, 2000

Keywords: Growth, Convergence, Markov chain, Ergodicity, Mobility

JEL Classification: O5, C1

*The authors would like to thank John Geweke, Christopher Sims, Trevor Breusch, Danny Quah and seminar participants at ANU for their helpful comments. All remaining errors are our own. Contact: Dr John Landon-Lane, School of Economics, The University of New South Wales, Sydney 2052, Australia. E-mail: j.landon-lane@unsw.edu.au

Abstract

One of the core issues in the economic growth literature is the degree to which economic activity is becoming "similar" across countries. An implicit assumption within this literature is that the distribution of output per capita across countries is changing over time. This underlying assumption need not be true. This paper tests the hypothesis that the distribution of income across countries is changing where it is assumed that the income distribution evolves over time according to a simple first order Markov chain defined over a finite number of income classifications.

The underlying probability structure allows us to formally characterise the behaviour of the data with respect to growth rates and income. We are able to test whether the world is in its ergodic state. We find evidence that the world has indeed reached its ergodic or invariant state thus contradicting the implicit assumption that the income distribution of countries is changing over time. Given the definition of relative income in this paper we are able to discuss the implications of our results with respect to convergence in absolute income and convergence in growth rates. We find that the nature of the ergodic distribution that is estimated in our study precludes both convergence in absolute income and convergence in growth rates.

We also characterise the relative income distribution for our sample and find strong evidence for increased relative income mobility for "poorer" countries in the ergodic. "Richer" countries are less mobile. We also find that countries that have incomes that are close to the average are much more mobile than countries with incomes at each end of the income distribution. In fact, we find evidence of a relative income poverty trap in that countries that have a low relative income are significantly less likely to move out of their income class.

1 Introduction

One of the core issues in the economic growth literature is the degree to which economic activity is becoming "similar" across countries. This question has been addressed in a variety of manners, both theoretical and empirical.[1] For instance, the neoclassical growth model predicts that countries with similar technology and preference parameters will grow at rates inversely proportional to their output per capita. That is, poorer countries will grow at faster rates than richer countries; and as a result, countries should eventually converge in levels. On the other hand, endogenous growth models (e.g. Lucas (1988)) suggest that growth rates are independent of the initial output per capita. For example, Ventura (1997) outlines a model in which convergence in levels occurs in an endogenous growth world depending on certain preference parameters.

Ultimately, the issues addressed in the theoretical literature have empirical implications. Barro (1991) points out that there is little evidence to support absolute convergence in levels, though the data suggest convergence conditional on the stock of human capital. So that, for a given level of human capital, economic growth is inversely proportional to the initial level of physical capital. Similarly, Mankiw et al. (1992) finds that, holding population growth and capital accumulation constant, there is evidence to support the predictions of the neoclassical growth model. In the end, this line of the empirical literature addresses the question of whether we observe countries converging in levels. An implicit assumption within this literature is that the distribution of output per capita across countries is changing over time.

This underlying assumption need not be true. In fact, Parente and Prescott (1993) suggests that the disparity of wealth per capita across rich and poor countries is

[1] A small selection of papers that have discussed the idea of convergence and growth includes Lucas (1988), Barro (1991), Barro and Sala-i-Martin (1992), Mankiw, Romer and Weil (1992) and Ventura (1997). Temple (1999) and Durlauf and Quah (1999) are two papers that have reviewed the empirical contributions of the growth literature.

constant over time. While there are some countries catching up to the richer countries, there are also countries that are dropping off. These are called "growth miracles" and "growth disasters" respectively. They find that rich and poor countries tend to be growing at the same rate but that the composition of the rich and poor groups are not necessarily the same over time. There is a lot of movement, both up and down.

The finding of Parente and Prescott (1993) suggests that the distribution across countries of output per capita is constant over time. One of the aims of this paper is to test this hypothesis within an underlying probability model. In particular, we use a simple representation of the data, similar in style to the work of Parente and Prescott (1993) and Chari, Kehoe and McGratten (1995), to model the evolution of relative per capita income using a first order Markov chain. This allows us to characterise the behaviour of the data with respect to growth rates and income. We are also able to construct a test of whether the world is in its ergodic state. We find evidence that the world has indeed reached its ergodic state. Given this evidence, we then attempt to characterise the nature of the movements of countries over time. While the distribution of output per capita is invariant, we find evidence of pronounced mobility within the distribution. In addition, we find that countries cluster into three broad categories.

We will discuss evidence for or against various types convergence. We will discuss convergence of absolute incomes across countries, of relative incomes across countries, of growth rates across countries and the relative income distribution to its limiting or ergodic distribution. We hope that in all cases it is clear from context to which definition of convergence we are referring.

The remainder of the paper is organized as follows. The first order Markov chain model for relative income is defined in Section 2, which also includes a description of Bayesian methods used to estimate the model and perform inference. Section 3 provides a more detailed outline of the methods used to estimate the model and perform inferences. Section 4 contains the results and Section 5 concludes.

2 Model

A first order Markov-chain model will be used to investigate the transitions of countries over time. The use of Markov-chain models to study income dynamics has a long history with notable contributions by Champernowne (1953) and Shorrocks (1976).

One of the most appealing aspects of using a Markov-chain to model income dynamics across countries is the ability to investigate issues such as mobility and whether convergence to the ergodic state has been achieved. The ability to calculate the ergodic, or limiting, distribution allows one to test whether there is evidence of convergence or divergence of countries with respect to their income. In addition, the Markov assumption is a natural way of thinking about income dynamics while imposing only minimal theoretical structure. Ultimately, the aim is to characterise the data outside the bounds of any particular theoretical model.

The model is as follows. Let there be C classifications, where C is a finite number. In our case, the classifications will be income classes. Let $\pi_t = (\pi_{1t}, \ldots, \pi_{Ct})'$ be the distribution across the C classes, where π_{kt} is the proportion of the total population that is in class k at time t. Therefore the variable π_t defines the "state" of the world at time t. The first order Markov assumption implies that the state of the world today is dependent only on π_0. That is,

$$P(\pi_t|\pi_{t-1}, \pi_{t-2}, \ldots, \pi_{t-j}) = P(\pi_t|\pi_{t-1}) \quad \forall \; j = 2, 3, \ldots \tag{1}$$

where P(.) represents the conditional probability distribution of π. Define the probability of transiting from class i in period t-1 to class j in period t to be

$$P(\pi_t = j|\pi_{t-1} = i) \equiv p_{ij}$$

so that the Markov transition matrix, $\mathbf{P}$, can be defined as $\mathbf{P} = [p_{ij}]$. Then the first order Markov chain model is

$$\pi_t' = \pi_{t-1}'\mathbf{P}. \tag{2}$$

The initial income distribution is π_0. It is simple to show that $\pi_t' = \pi_0'\mathbf{P}^t$. An invariant or ergodic distribution, $\overline{\pi}$, is any distribution that satisfies

$$\overline{\pi}' = \overline{\pi}'\mathbf{P}. \tag{3}$$

This distribution is unique if there is only one eigenvalue of $\mathbf{P}$ with modulus one; which is simple to test.

The initial and invariant distributions are very important functions of interest if we wish to study the evolution of income distributions using a Markov chain. If the classifications are suitably defined it will be possible to test for convergence over time by comparing the initial and invariant distributions for each income class. In particular, if the classifications are defined in relation to some reference point then it will be possible to measure the degree of convergence or divergence into or from that reference point. Another very important issue is whether an observed system has converged to its invariant state or not. This clearly has important implications as to whether it is possible to find convergence or divergence at all. A world that is in its ergodic state can not be characterised as converging or diverging into or from the chosen reference point. The gap in income in such a world is constant, so there is no sense in which countries are *becoming* similar or dissimilar. Countries simply *are* similar or dissimilar.

This paper uses Bayesian inference techniques to estimate and make inferences from the Markov chain model outlined above. Bayesian methods are used as they are useful for the problem of estimating a Markov chain under the assumption that the model is in its ergodic state. In order to construct tests for ergodicity and convergence we need to first describe the sampling scheme used for this model.

We observe N countries over T time periods and place them into C classifications.

Let $i \in \{1, 2, \ldots, C\}$, $n \in \{1, 2, \ldots, N\}$, and let $t \in \{1, 2, \ldots, T\}$. For each country n, define

$$\delta_{nit} = \begin{cases} 1 & \text{if country n is in class i for time period t} \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

For each country, n, and for each time period t we observe the income class to which the country belongs, $s_{nt} \in \{1, 2, 3, \ldots, C\}$. Let $S_{NT} = \{\{s_{nt}\}_{n=1}^{N}\}_{t=1}^{T}$ be the information set at time T. Define $k_{j0} = \sum_{n=1}^{N} \delta_{nj0}$ as the number of countries that are in class j in the initial period and define $k_{ij} = \sum_{n=1}^{N} \sum_{t=1}^{T} \delta_{ni(t-1)} \delta_{njt}$ as the total number of transitions from class i in time period t-1 to class j in time period t across all time periods. The matrix $\mathbf{K} = [k_{ij}]$ will be referred to as the data transition matrix.

We implicitly assume that $\mathbf{P}$ is the same for all time periods; so that the data density, or likelihood function, for the model defined in (2) is

$$p(S_{NT}|\pi_0, \mathbf{P}) \propto \prod_{i=1}^{C} \pi_{i0}^{k_{i0}} \prod_{j=1}^{C} p_{ij}^{k_{ij}} \tag{5}$$

which is the kernel of the product of two independent multivariate Beta (Dirichlet) distributions. Natural conjugate priors for π_0 and $\mathbf{P}$ are also independent Dirichlet distributions defined as

$$p(\pi_0) = \left[\frac{\Gamma(\sum_{i=1}^{C} a_{i0})}{\prod_{i=1}^{C} \Gamma(a_{i0})} \right] \prod_{i=1}^{C} \pi_{i0}^{(a_{i0}-1)} \tag{6}$$

and

$$p(\mathbf{P}) = \prod_{i=1}^{C} \left[\frac{\Gamma(\sum_{j=1}^{C} a_{ij})}{\prod_{j=1}^{C} \Gamma(a_{ij})} \right] \prod_{j=1}^{C} \pi_{ij}^{(a_{ij}-1)}. \tag{7}$$

Here the priors are parameterised by the vector $a_0 = (a_{10}, \ldots, a_{C0})'$ and $\mathbf{A} = [a_{ij}]$. The priors have a notional sample interpretation. We can think of $a_{i0} - 1$ as the number of countries in the ith class of the initial relative income distribution of a notional sample, while $a_{ij} - 1$ can be interpreted as the total number of transitions from class i in period t-1 to class j in period t for the notional sample. Assuming that the priors are independent then the posterior distribution for (2) is

$$p(\pi_0, \mathbf{P}|S_{NT}) \propto \left[\frac{\Gamma(\sum_{i=1}^{C} a_{i0})}{\prod_{i=1}^{C}\Gamma(a_{i0})}\right] \prod_{i=1}^{C} \pi_{i0}^{(k_{i0}+a_{i0}-1)} \prod_{i=1}^{C} \left\{ \left[\frac{\Gamma(\sum_{j=1}^{C} a_{ij})}{\prod_{j=1}^{C}\Gamma(a_{ij})}\right] \prod_{j=1}^{C} \pi_{ij}^{(k_{ij}+a_{ij}-1)} \right\} \tag{8}$$

Given this framework, we are now able to make inferences about any function of the parameters π_0 and $\mathbf{P}$. In particular, posterior distributions can be analyzed for the invariant distribution, $\overline{\pi}$, and for any differences between $\overline{\pi}$ and π_0.

3 Estimation and Inferences for the First Order Markov Chain Model

The joint posterior density kernel in (8) is the kernel for the product of two Dirichlet distributions. The posterior distribution for π_0, the initial income distribution, is Dirichlet with parameters $(k_{10} + a_{10}, \ldots, k_{C0} + a_{C0})'$. The posterior distribution for $\mathbf{P}$ is the product of C independent Dirichlet distributions with parameters $(k_{i1} + a_{i1}, \ldots, k_{iC} + a_{iC})'$ for $i = 1, \ldots, C$ (Geweke 1998) .

There are two cases between which we would like to distinguish. In the first case, we do not assume that the Markov chain has reached its ergodic state. That is, there is no assumption that the initial distribution, π_0, is equal to the invariant distribution, $\overline{\pi}$. In the second case, we assume the Markov chain has reached its ergodic state, so we would need to impose the condition that the initial distribution, π_0, is equal to $\overline{\pi}$. These two cases have vastly different implications with respect to the underlying properties of the income dynamics that are being modeled.

For the first case, the posterior distribution is as reported in (8). In this case it is a simple matter to make identical and independent draws from these independent distributions using the method described in Devroye (1986). For the second case the posterior distribution is slightly different because we restrict π_0 to be equal to

$\overline{\pi}$. Geweke (1998) shows importance sampling can be used to draw from the joint distribution given in (8) under the restriction that π_0 is equal to $\overline{\pi}$.

Once we have the ability to sample from either posterior distribution it is possible to make inferences based on those draws. One such inference is to test which case is more likely. Another issue is whether there is evidence of countries converging or not. These decisions will be made through the use of Bayesian model comparison techniques outlined in Geweke (1994).

3.1 Data

The data used for this paper were obtained from the Penn World Tables version 5.6 that can be found at the Penn World Tables World Wide Web site at the University of Toronto.[2] The data consist of observations on real GDP per worker for 104 countries for the years from 1960 through 1990. A country is included in our sample if there are observations for each year 1960 through 1990. As described in Summers and Heston (1991), the real GDP per worker series is based on a measure of GDP that is weighted using a price series that is a blend of current year prices and international prices of the base year 1985. This allows us to be able to compare countries for each year as well as across years.

There are a number of ways in which we could have used this data to estimate a Markov chain model. We follow Chari et al. (1995) and calculate the geometric mean for each year and transform the data by dividing each observation by its geometric mean for each year. The goal of this technique is to minimise the effects of growth on the results. We then classify each country according to these relative income measures. As an alternative method, Quah (1993) classifies each country according to their rank in percentiles. It was noted in Durlauf and Quah (1999) that a finite state Markov chain can suffer from robustness problems when countries are classified by percentile.

[2]The address for this site is http://www.epas.utoronto.ca:8080/epas.

That is, variations in definitions of the classifications can lead to different conclusions. In addition, movement between classes for one country can cause other countries to change class as well. These problems were solved by using a stochastic kernel in place of the transition matrix (Quah 1996, Quah 1997).

Unfortunately, the issue of imposing ergodicity on such a stochastic kernel is still open. Furthermore, using relative incomes as in Chari et al. (1995) reduces much of the "phantom" transition problem described above. To test this, a reference country was left out of the sample and the analysis was performed on the rest. Relative incomes were reported with respect to the income of this reference country instead of the geometric mean of all countries. Hence the movement of one country from one class to another will not have any effect on the other countries classifications. It was found that the results for this case did not change qualitatively from the results obtained from the full sample.

We defined five classifications in terms of the countries' relative income. The classifications are defined in Table 1.

Table 1: Income Classifications

	1	2	3	4	5
lower boundary	0.0	0.35	0.7	1.4	2.8
upper boundary	0.35	0.7	1.4	2.8	5.6

This classification satisfies the criterion set out in Champernowne (1953) that the class lengths be equal in log scale. The middle class includes the relative income of 1. This was appealing in that it is possible to talk of convergence by looking at what happens to the proportion of countries in the third or middle class. For example, if all countries converge to have the same income per worker then every country would have a relative income of 1 and, in the limit, $\overline{\pi}_3$ would equal 1. This outcome is not probable but we can still compare the initial state with the ergodic state to see what is happening to the third income class. Clearly, the data are inconsistent with convergence if the proportion of countries in the third income class is lower in the ergodic state than in the initial state. It is also clear that if the world is in the ergodic state to start with then there can be no convergence. The same proportion of the countries will be in each state for each time period so that there can not be an increase in the movements of countries in to the middle income class.

3.2 Prior Definitions

We define proper natural conjugate priors for our model. All priors are based on independent Dirichlet distributions with the only difference being the manner in which they are interpreted. The general form of the prior for the initial state, π_0 and the transition matrix, $\mathbf{P}$, are defined in (6) and (7) respectively.

The prior parameters for π_0 reflect prior belief that any country in the sample has an equal probability of being in any income class. The prior parameters are found in Table 2.

Table 2: Prior parameters for π_0

a_{10}	a_{20}	a_{30}	a_{40}	a_{50}
3.375	3.375	3.375	3.375	3.375

There are three distinct priors for $\mathbf{P}$. These are found in Table 3. All priors are based on what a data transition matrix would look like for a notional sample. The priors are constructed for notional samples with different limiting behaviour. The “flat” prior for $\mathbf{P}$ is a transition matrix whose ergodic distribution is flat (see Figure 1) . The “convergence” prior for $\mathbf{P}$ is a transition matrix that has an invariant distribution that has more mass for the middle bin than in the extreme bins. The “divergence” prior for $\mathbf{P}$ is a transition matrix that has more weight in the first and fifth income classes than in the middle class. The "convergence" prior represents convergence in the sense that, in the notional sample, there are more countries in the class containing the relative income of 1 in the invariant distribution than in the notional initial distribution. The "divergence" prior represents a world where there is greater disparity in the limiting income distribution than in the notional initial distribution. The "flat" prior is the case where there is no change from the notional initial distribution to the limiting distribution. The "flat" prior is used when the ergodic assumption is imposed.

The values of a_0 and $\mathbf{A}$ were chosen to reflect a notional sample that is small relative to the number of countries that make up the data set. The row sums of $\mathbf{A}$ are all equal to 7.375 which implies that there are 2.375 countries in each class in period 0. This is consistent with the prior for a_0. This implies that there are a total of about 12 countries in the notional sample compared with the 104 countries in the data. These priors were chosen so that the data would dominate the posterior thus allowing the data to drive the results.

There are many ways in which a prior for $\mathbf{P}$ can be written so as to get invariant distributions like those reported in Figure 1. The priors chosen in this paper were constructed by altering the "flat" prior to get the appropriate invariant distributions. The prior was altered by changing the number of transitions in the first off diagonals of $\mathbf{A}$. However, in order to keep the same number of observations in each notional sample the row sum totals for $\mathbf{A}$ are kept constant. For example, for the "convergence" prior more transitions were allocated to the (1,2), (2,3), (4,3), and (5,4) elements of $\mathbf{A}$. Thus this transition matrix has a higher tendency for countries to move towards the middle bin in each time period compared to the "flat" prior. For the divergence prior more notional transitions were given to the (2,1), (3,2), and (4,5) elements of $\mathbf{A}$. This reflects the belief that almost all transitions will be between adjacent classes. The aim is to make the differences from the "flat" prior as small as possible so as to avoid biasing the tests. For example, another "convergence" prior would be an $\mathbf{A}$ where more transitions are placed in the third column of $\mathbf{A}$. This would imply a world where

Table 3: Transition Matrix Priors

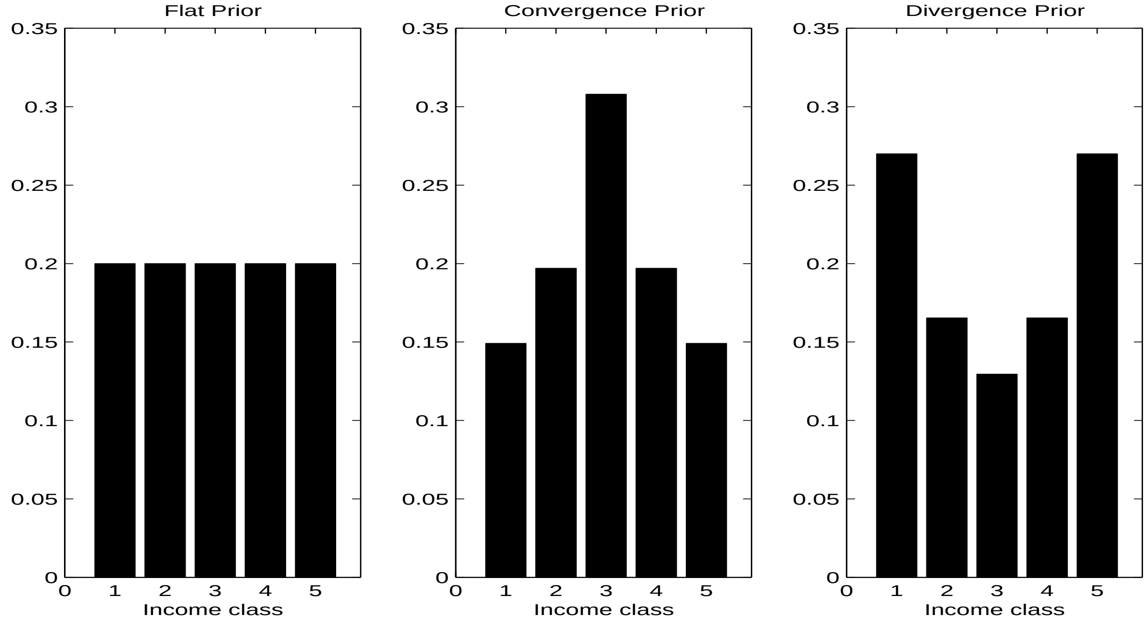
	Convergence 1	2	3	4	5	Divergence 1	2	3	4	5
1	2.31	1.70	1.30	1.04	1.01	2.68	1.37	1.19	1.08	1.04
2	1.37	2.00	1.75	1.17	1.08	1.75	2.00	1.37	1.17	1.08
3	1.21	1.41	2.12	1.41	1.21	1.28	1.59	1.62	1.59	1.28
4	1.08	1.17	1.75	2.00	1.37	1.08	1.17	1.37	2.00	1.75
5	1.01	1.04	1.30	1.70	2.31	1.04	1.08	1.19	1.37	2.68

	Flat 1	2	3	4	5
1	2.43	1.50	1.25	1.12	1.06
2	1.50	2.00	1.50	1.25	1.12
3	1.25	1.50	1.87	1.50	1.25
4	1.12	1.25	1.50	2.00	1.50
5	1.06	1.12	1.25	1.50	2.43

there are more transitions of two classes at a time, which is unlikely.

All priors implicitly assume that the bulk of transitions are to adjacent classes. In addition, all priors assume that the initial distribution has an equal proportion in each class. Clearly we would like to let the data to distinguish between the various possibilities described above. In particular we would like to let the data distinguish between a world that has reached its ergodic state versus a world that has not. We would also like to distinguish between worlds where countries are converging and worlds where they are not. The next subsection will outline how we attempt to do this.

Figure 1: Limiting Income Distributions for each Prior



3.3 Bayesian Inference for the Markov chain model

It is natural to use Bayesian model comparison (see Geweke (1994)) to test the different hypothesis described above. In general we have a number of different model specifications, M_j, $j = 1, \ldots, J$. The models all propose to describe the same set of observations, S_{NT}. However, the models are not necessarily nested. Define the marginal likelihood of a model to be

$$p(S_{NT}|M_j) = \int_{\mathcal{P}} p(S_{NT}|\pi_0, \mathbf{P}, M_j) p(\pi_0, \mathbf{P}|M_j) d(\pi_0, \mathbf{P}), \tag{9}$$

where

$$\mathcal{P} = \left\{ (\pi_{j0}, p_{ij}) : 0 \leq \pi_{j0} \leq 1, 0 \leq p_{ij} \leq 1, \sum_{j=1}^{C} \pi_{j0} = 1, \sum_{j=1}^{C} p_{ij} = 1 \quad \forall i = 1, \ldots, C \right\}.$$

Then

$$p(M_j|S_{NT}) = p(M_j) p(S_{NT}|M_j), \tag{10}$$

where $p(M_j)$ is the prior probability that model M_j is the best model that describes the data, S_{NT}. We can directly compare two models, M_i and M_j by taking the posterior odds ratio $p(M_i|S_{NT})/p(M_j|S_{NT})$. Clearly if the posterior odds ratio is greater than 1 then model i is favored over model j. If there is no prior belief that any model is better than the other candidates then the posterior odds ratio reduces to the ratio of marginal likelihoods, $p(S_{NT}|M_i)/p(S_{NT}|M_j)$, known as the Bayes factor in favor of model i over model j.

We perform two basic tests. The first aims to test whether the data supports a model where the Markov chain has reached its ergodic state versus a model where the Markov chain has not reached its ergodic state. If it is decided that there is evidence that suggests the Markov chain has not reached its ergodic state, then it is possible to test the hypothesis that the countries in the sample are growing together. In this case we have three models from which to choose, each model distinguished by the prior outlined in Section 3.2. The marginal likelihood in each case will reflect whether the data supports the prior or not. Thus the model with the highest marginal likelihood is best supported by the data. The results of these tests are reported in Section 4.

In order to calculate the Bayes factor for the first test we need to calculate the marginal likelihood for each model. Geweke (1998) describes how to do this for each model. For the Markov chain that has not reached its ergodic state the marginal likelihood can be calculated directly using the following formula;

$$p(S_{NT}|M_s) = \frac{\Gamma(\sum_{i=1}^{C} a_{i0}) \prod_{r=1}^{C} \left\{ \Gamma(\sum_{j=1}^{C} a_{ij}) \Gamma(k_{i0} + a_{i0}) \prod_{j=1}^{C} \Gamma(k_{ij} + a_{ij}) \right\}}{\Gamma\left[\sum_{i=1}^{C}(k_{i0} + a_{i0})\right] \prod_{i=1}^{C} \left\{ \Gamma(a_{i0}) \Gamma\left[\sum_{j=1}^{C}(k_{ij} + a_{ij})\right] \prod_{j=1}^{C} \Gamma(a_{ij}) \right\}}. \tag{11}$$

For the case in which we impose the condition that the Markov chain has reached its ergodic state the marginal likelihood can be consistently estimated as the average of the importance sampling weights for each draw from the posterior. The importance sampling weights for each draw are reported in Geweke (1998) to be

$$\frac{\prod_{i=1}^{C} \left\{ \Gamma(\sum_{j=1}^{C} a_{ij}) \prod_{j=1}^{C} \Gamma(a_{ij} + n_{ij}) \right\} \overline{\pi}^{k_{i0}+a_{i0}-1}}{\prod_{i=1}^{C} \left\{ \prod_{j=1}^{C} \Gamma(a_{ij}) \Gamma\left[\sum_{j=1}^{C}(a_{ij} + k_{ij})\right] \right\}}. \tag{12}$$

4 Results

Recall that our goal is to characterise the distribution of income across the countries that make up our sample. In particular, we would like to determine whether the distribution changing over time. If the distribution of income across countries is changing, then we would like to characterize the manner in which it is changing. In particular, we would like to know whether there a widening gap between the "rich" countries and the "poor" countries.

The methods described in Section 2 were used to make draws from the Markov chain model described in (2). In particular, we make 1000 independent and identically distributed (i.i.d.) draws from the posterior distribution of the parameters of a Markov chain model that has not reached its ergodic state. Once these draws were obtained it was possible to analyse the posterior distribution of any function of the parameters π_0 and $\mathbf{P}$. Possible functions of interest include the invariant income distribution, $\overline{\pi}$, and any measures of mobility that are functions of $\mathbf{P}$. The draws were made for each of the three different prior distributions of $\mathbf{P}$ defined in Section 3.2. The results from this are reported below.

Importance sampling, with weights given by (12), was used to draw from the posterior distribution of the parameters of a Markov chain in which ergodicity is imposed. The parameters of such a chain are the invariant distribution, $\overline{\pi}$ and $\mathbf{P}$ the transition matrix. Marginal likelihoods were calculated for each model and these are reported below.

As indicated above, we use data on relative per worker incomes for 104 countries from 1960 to 1990. There are a number of possible transition periods that could be used. Quah (1993) uses a transition period of 1 year. Chari et al. (1995) use a 30 year transition period. It is not uncommon in cross-section studies of growth to see authors use 5 year or 10 year transitions as well. In this paper we report results for a number of different transition periods. The results are reported for 1, 5, 10, 15, and

30 year transitions.[3] Clearly the transition matrix, $\mathbf{P}$, will have different properties as the transition period changes. Given the wide definitions of the income classifications it would be very surprising to see a country move more than one income class in one year. In fact, it would be most unlikely to see very many transitions at all over one year. You would expect $\mathbf{P}$ to have diagonal elements close to one with only the first off diagonals having non-zero elements. If the transition period were 30 years then you might expect to see some countries moving 2 or more classes.

It was found that the results of this paper were not sensitive to the transition period. In fact it was remarkable how constant the results were over all transition years. The first set of results were for the full period from 1960-1990. These are reported in Section 4.1 below. It was also found that the magnitude of the second eigenvalue of $\mathbf{P}$ was less than 1 for all cases. Hence all models had unique ergodic distributions.

4.1 Results for 1960-1990

The first set of results are for the case where we assume that $\mathbf{P}$ is constant over the whole time period, 1960-1990. We first test for ergodicity of the Markov-chain. The Bayes factors in favor of the non-ergodic model are given in Table 4.

Table 4: Bayes factors in favor of the non-ergodic model

Transition period (years)	1	5	10	15	30
Bayes factor	13.72	25.32	39.40	46.90	30.72

It is clear form the results presented in Table 4 that there is strong evidence that the Markov chain has not reached its ergodic state. The Bayes factors imply that,

[3]For example, the data transition matrix $\mathbf{K}$ for the 15 year transition was calculated by dividing the data set into two 15 year periods. Then k_{ij} is the number of countries that moved from class i at the start of either 15 year period to class j at the end of the relevant 15 year period.

conditional on the data, the non-ergodic model is at least 13 times as likely to be the correct model as the ergodic model.

Given that there is strong evidence that the Markov chain is not in its ergodic state we can now investigate any differences between the initial distribution and the ergodic or limiting income distribution of the Markov chain. The marginal likelihood for each model was calculated using (11). The results can be found in Table 5 below.

Table 5: Bayes factor in favor of “divergence” model

Transition periods (years)	1	5	10	15	30
versus “convergence” model	1.97	1.97	2.08	2.47	2.87
versus “flat” model	6.19	4.41	3.91	3.84	3.08

The Bayes factor in favor of the “divergence” model is greater than one for all cases. We can see that there is a consistent ordering of models (“divergence” $\succ$ “convergence” $\succ$ “flat”). Both models that assume π_0 is different from the prior invariant income distribution are preferred to the “flat” model where π_0 is equal to $\overline{\pi}_0$. This is consistent with the results from Table 4. We also see that the marginal likelihood of the “divergence” model is approximately twice as big as the marginal likelihood of the “convergence” model. In fact we see that as the transition period increases the “divergence” model is more likely. This is not what we would have expected if we were in a world that was converging.

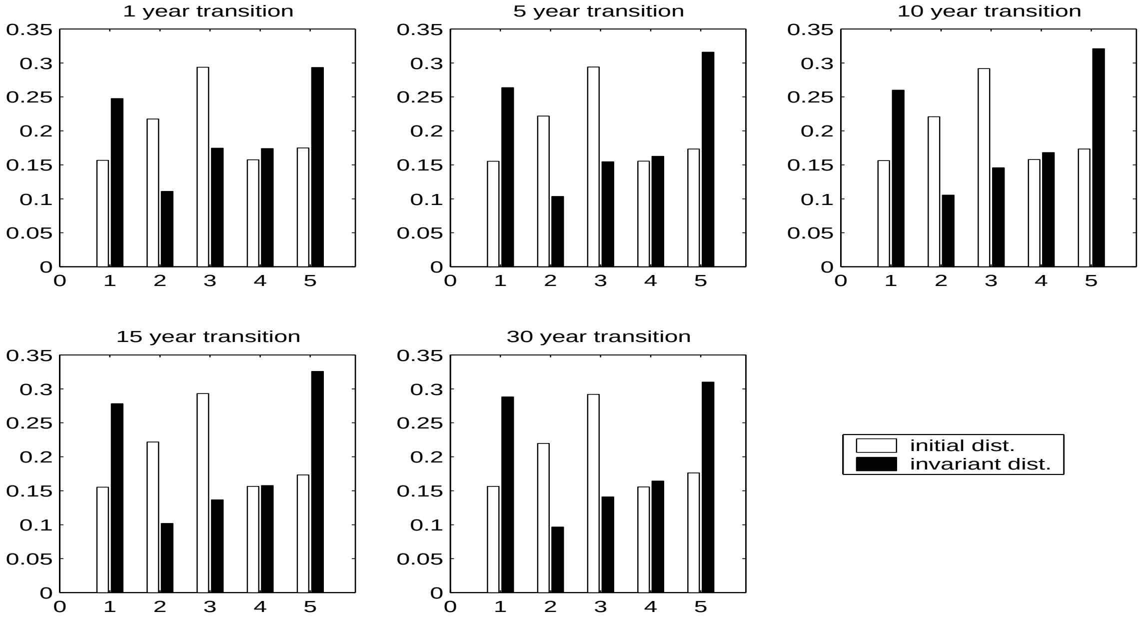
Figure 2 contains the posterior means of the π_{i0} and $\overline{\pi}_i$ for each time period used. The first observation that should be noted is the consistency of results over the different periods of transition. For all cases we see that the posterior mean of $\overline{\pi}_i$ is greater than the posterior mean of π_{i0} for $i = 1, 5$. This implies that the there is a higher proportion of countries in the first and fifth income class in the ergodic state than the initial state. There is also consistently a lower proportion of countries in the middle income class in the limit. This is further evidence that the world is not converging.

Table 6: Expected Length of Stay for each Income Class: "Divergence prior"

	Income class				
	1	2	3	4	5
1 year transition	42.36	11.88	13.89	13.67	56.39
	(10.76)	(1.78)	(1.80)	(2.04)	(18.11)
5 year transition	13.10	3.60	4.37	4.31	17.13
	(4.21)	(0.55)	(0.63)	(0.75)	(6.30)
10 year transition	7.22	2.26	3.00	2.86	8.97
	(2.35)	(0.32)	(0.47)	(0.49)	(3.21)
15 year transition	6.60	2.00	2.30	2.08	6.97
	(2.43)	(0.30)	(0.35)	(0.34)	(2.60)
30 year transition	4.80	1.45	1.89	1.79	4.25
	(2.25)	(0.18)	(0.30)	(0.33)	(1.64)

Mobility indices are another measure of the dynamics of a Markov chain. One such mobility index is the measure of expected length of stay in an income class. Table 6 contains the expected length of stay in each of the five income classes. We can see from these results that the longest expected stays are for the first and fifth income classes. We see that the middle three income classes have a much smaller expected length of stay which implies that countries with relative incomes that are close to the average are more mobile. This result together with the information from the comparison of the initial and invariant income distributions suggest that over the entire sample period countries have moved out of the middle income class towards the edges of the distribution. Therefore it is not surprising that the "divergence" model was preferred by the data.

Figure 2: Initial and Invariant Income Distributions



4.2 Results for 1970-1990

It was found in the previous section that there is substantial evidence to suggest that the world is not converging in relative incomes and that the state of the world in 1990 is different to the state of the world in 1960. This section investigates whether there is evidence that world has converged to its ergodic state when we restrict the data set to 1970 to 1990. Similar to the method of Section 4.1 we estimate (2) under the assumption of ergodicity and non-ergodicity for the period from 1970 to 1990. Marginal likelihoods and Bayes factors were calculate with the Bayes factor in favor of ergodicity reported in Table 7.

Table 7: Bayes Factor in Favor of Ergodicity

Transition period (years)	1	2	5	10
Bayes factor (1970-90)	3.31	2.50	2.85	2.41
Bayes factor (1960-70)	0.66	0.50	0.44	0.44

It is clear from the results presented in Table 7 that there is reasonably strong evidence that for the last 20 years of our data the world has reached its ergodic state. The assumption that the world has reached its ergodic state is at least as twice as likely as the alternative hypothesis that the world has not reached its ergodic state. The results reported for 1960-70 suggest that the world was not in its ergodic state from 1960-70. In order to test for consistency with the results from the previous section we also test to see whether the "divergence" assumption is preferred over the "convergence" assumption for the period 1960-70. The Bayes factors in favor of "divergence" are found in table 8.

Table 8: Bayes Factor in Favor of "Divergence": 1960-70

Transition period (years)	1	2	5	10
Bayes factor	2.00	2.08	2.50	2.72

Here we see that the "divergence" model is favored consistently for the period 1960-70. Hence the results in this section are consistent with the findings of the previous section where we found that over the period 1960-90 the world had not reached its ergodic state and the "divergence" hypothesis was preferred.

Tables 9 and 10 contain the expected length of stay in each income class for the periods 1960-70 and 1970-90 respectively. They also contain an overall mobility measure statistic defined as

$$M_P = \frac{C - \mathrm{tr}(\mathbf{P})}{C-1}. \tag{13}$$

The properties of M_P are described in detail in Geweke, Marshall and Zarkin (1986).[4] From the results reported in Tables 9 and 10 there is evidence to suggest that there is less overall mobility, as measured by M_P, in the ergodic state (1970-90) than for the

[4]In particular M_P is monotone in that $M_P(\mathbf{P}^1) > M_P(\mathbf{P}^2)$ for all transition matrices where $\mathbf{P}^1$ is more mobile than $\mathbf{P}^2$. The measure M_P is also strictly immobile in that $M_P(\mathbf{P}) > 0$. M_P is the harmonic mean of the expected length of stay in each bin.

period 1960-70. The posterior mean for M_P is uniformly lower in the ergodic state than for the transition matrix for the period 1960-70.

However, the results for the individual income classes are not so clear. The lowest income class is clearly more mobile in the ergodic state as evidenced by the lower posterior mean for the expected length of stay in class 1. However, the "richest" countries exhibit less mobility in the ergodic state as evidenced by the higher expected length of stay for the 5th income class. The results for the intermediate income classes are mixed. The second income class is more mobile in the ergodic state for three of the four cases while the middle income class in less mobile in three of the four cases. The fourth income class is uniformly less mobile. Hence we see that in the ergodic state the "poorer" countries are more mobile while the "richer" countries are less mobile. The overall effect is for less relative income mobility in the ergodic state as evidenced by the lower posterior means for M_P in Table 10.

However, as in Section 4.1 there is more mobility for countries in the middle three income classes than countries in the top and bottom income class. That is, once a country is either rich or poor, in terms of relative income, they are more likely to stay rich or poor. This result is in contrast to Parente and Prescott (1993) where they find no evidence of a poverty trap in absolute incomes. The results described above suggest that there is evidence of a poverty trap in relative incomes for the data.

Finally, Figure 3 contains the posterior means of the ergodic income distribution for the period 1970-90. We can see that the first, third, and fifth income class have the largest proportions. You can see that this distribution is substantially different from the implied ergodic distribution for $\mathbf{P}$ when data from 1960-90 is used. If we compare the initial distribution reported in Figure 2 we see that there has been movement out of the second and fourth income classes. There is evidence of the countries in the sample grouping into three distinct groups. A group of countries that have low relative incomes, a group of countries that have incomes that are close to 1, and a

Table 9: Mobility Measures: 1960-70

	Expected Length of Stay					M_P
Income class	1	2	3	4	5	
1 year transition	34.31	10.28	12.94	9.66	30.91	0.0914
	(14.90)	(2.29)	(2.54)	(2.13)	(13.14)	(0.0105)
2 year transition	21.18	6.56	7.39	5.79	17.79	0.1509
	(9.70)	(1.59)	(1.50)	(1.33)	(7.93)	(0.0166)
5 year transition	11.30	3.72	4.39	3.09	9.37	0.2747
	(6.10)	(0.94)	(0.99)	(0.74)	(4.35)	(0.0328)
10 year transition	6.18	2.42	2.70	2.15	5.28	0.4381
	(4.11)	(0.57)	(0.57)	(0.54)	(2.55)	(0.0483)

Table 10: Mobility Measures: 1970-90

	Expected Length of Stay					M_P
Income Class	1	2	3	4	5	
1 year transition	27.96	12.02	13.89	13.25	44.96	0.0742
	(5.82)	(1.94)	(2.07)	(2.01)	(11.05)	(0.0065)
2 year transition	16.24	6.46	7.69	7.83	25.70	0.1320
	(3.44)	(1.04)	(1.15)	(1.23)	(6.74)	(0.0118)
5 year transition	7.32	3.23	4.17	4.12	11.90	0.2611
	(1.62)	(0.53)	(0.63)	(0.67)	(3.25)	(0.0242)
10 year transition	4.22	2.05	3.14	2.73	6.25	0.4050
	(0.92)	(0.31)	(0.55)	(0.46)	(1.62)	(0.0363)

group of countries that have high relative incomes. Note that the makeup of each group is not necessarily the same each period.

Table 11 contains the posterior moments for $\mathbf{P}$ for each transition period. There are a number of characteristics that should be noted. The first is that, for each transition period, there is a higher probability of moving from class two to class one than from class two to class three. We therefore see a tendency for countries in relative income class two to move down rather than up. For class four we see a tendency for countries to move down rather than up in all cases except the 20 year transition case. However, countries in income class three are more likely to move up than move down. In all cases countries in the first and fifth classes are more likely to stay in their current relative income class compared to countries in the other income classes.

We should also note that as we are using the geometric mean of all the countries in the sample as the reference we can make some indirect inferences on whether the world has converged in growth rates. If all countries grow at the same rate then the geometric average will grow also at the common rate. Suppose that the ergodic distribution is such that all countries are growing at the same rate. This would imply that the relative incomes of all countries would be the same across time. If this were the case then $\mathbf{P}$ would equal the identity matrix. It is clear from Table 11 that the estimated $\mathbf{P}$ for the ergodic distribution has non-zero off diagonal elements. This would suggest that, even in the ergodic distribution, countries have different growth rates.

Table 11: Transition matrix ($\mathbf{P}$): 1970-90

1 year transition

	1	2	3	4	5
1	0.962	0.028	0.003	0.003	0.002
	(0.007)	(0.007)	(0.002)	(0.002)	(0.002)
2	0.044	0.914	0.033	0.003	0.003

Table 11: Transition matrix (**P**): 1970-90

	(0.009)	(0.013)	(0.009)	(0.003)	(0.002)
3	0.002	0.023	0.926	0.045	0.002
	(0.001)	(0.006)	(0.010)	(0.008)	(0.002)
4	0.002	0.003	0.053	0.922	0.018
	(0.002)	(0.002)	(0.009)	(0.011)	(0.005)
5	0.002	0.002	0.003	0.015	0.976
	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)

2 year transition

	1	2	3	4	5
1	0.935	0.046	0.007	0.005	0.004
	(0.013)	(0.012)	(0.005)	(0.004)	(0.004)
2	0.076	0.841	0.068	0.007	0.006
	(0.018)	(0.025)	(0.017)	(0.006)	(0.005)
3	0.004	0.046	0.867	0.077	0.004
	(0.003)	(0.011)	(0.019)	(0.015)	(0.003)
4	0.004	0.005	0.088	0.869	0.031
	(0.004)	(0.005)	(0.017)	(0.020)	(0.010)
5	0.004	0.005	0.006	0.025	0.958
	(0.003)	(0.004)	(0.005)	(0.009)	(0.010)

5 year transition

	1	2	3	4	5
1	0.857	0.099	0.017	0.013	0.012
	(0.030)	(0.028)	(0.014)	(0.011)	(0.010)
2	0.159	0.682	0.127	0.016	0.014

Table 11: Transition matrix (**P**): 1970-90

	(0.037)	(0.050)	(0.036)	(0.014)	(0.012)
3	0.010	0.090	0.754	0.134	0.010
	(0.008)	(0.024)	(0.036)	(0.028)	(0.008)
4	0.011	0.014	0.156	0.751	0.066
	(0.010)	(0.012)	(0.034)	(0.038)	(0.022)
5	0.010	0.013	0.015	0.050	0.909
	(0.009)	(0.011)	(0.013)	(0.020)	(0.023)

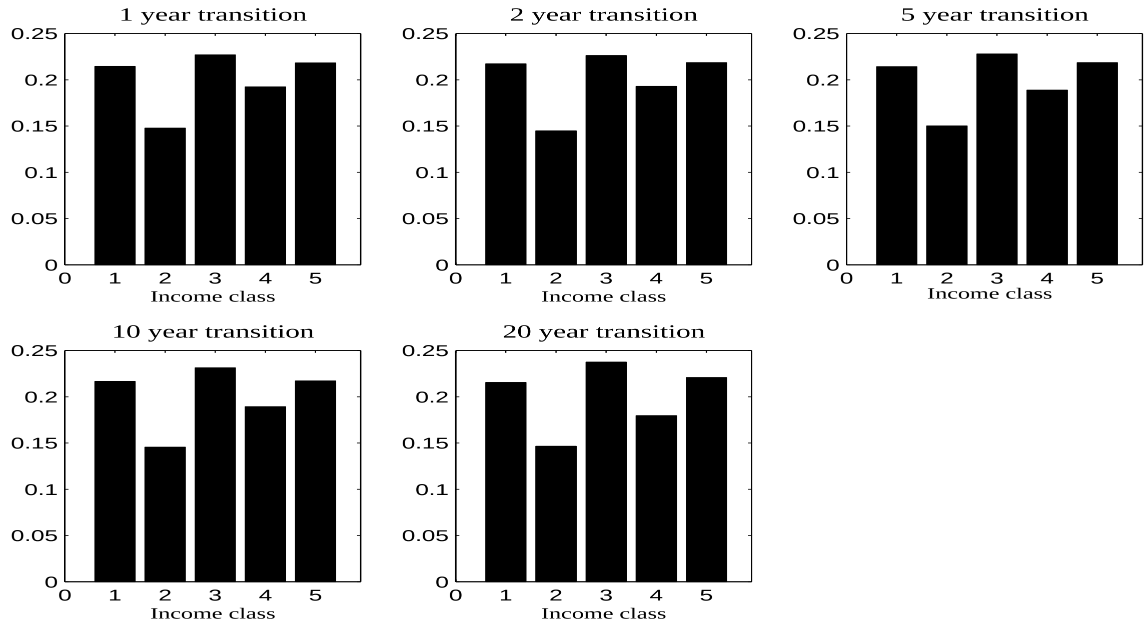
10 year transition

	1	2	3	4	5
1	0.752	0.167	0.034	0.024	0.022
	(0.052)	(0.050)	(0.026)	(0.020)	(0.019)
2	0.281	0.502	0.160	0.030	0.025
	(0.063)	(0.070)	(0.054)	(0.026)	(0.023)
3	0.0191	0.111	0.671	0.178	0.018
	(0.017)	(0.036)	(0.055)	(0.044)	(0.016)
4	0.021	0.026	0.202	0.623	0.125
	(0.019)	(0.022)	(0.053)	(0.060)	(0.040)
5	0.020	0.025	0.031	0.093	0.830
	(0.017)	(0.021)	(0.025)	(0.036)	(0.039)

4.3 Sensitivity of Results

It was noted earlier that a number of sensitivity studies were carried out. The first set of sensitivity studies involved the definition and number of income classes. Models with 3, 4, 5, 6, and 7 income classes were used. The qualitative results from all of these models were the same. The five income class model was chosen as it was a

Figure 3: Ergodic income Distribution: 1970-90



compromise between having more detail and reducing the number of zero elements in **K**. The three class model was not as sparse as the five class model but did not allow enough detail. The seven class model allowed for more detail but most of the off diagonal elements were zero, so the resulting data transition matrix was too sparse.

Once the five class model was chosen, sensitivity analysis on the bin definitions were undertaken. There clearly is an issue with respect to the bin definitions. It would be possible to get many different outcomes by choosing different bin definitions. Therefore we followed Champernowne (1953) in making the length of the income classes equal in the log scale. In that sense, the upper limit of each class was twice that of the lower limit. Having decided upon this strategy we tested the results by moving the cutoff points by 10% each way. Again, there were no qualitative changes to the results.

The final sensitivity test was a test to see whether the "phantom" transition problem described by Durlauf and Quah (1999) was significant. To do this a country that historically had a relative income close to 1 was chosen to be the reference country.

We could now see whether the observed transitions were caused by one country moving and therefore changing the relative positions of the countries. It was found again that there was no qualitative change in our results. This was not surprising as our model used relative incomes with respect to the geometric average rather than percentiles. The growth of one country is not going to have as big an impact on the geometric average as it would on the relative order of all the countries in the sample.

5 Conclusion

The intention of this paper was to investigate the statistical properties of a data set that has been used extensively in the growth literature. To that end, a first order Markov chain model was estimated using a cross-country panel for the years 1960 to 1990. The main result of this paper is that there is strong evidence that the world has converged to its ergodic distribution from 1970 to 1990. For this period we see that the relative income distribution of countries is constant over that time.

The finding that the Markov-chain has reached its ergodic state contradicts the implicit assumption of a changing income distribution used in the attempt to find evidence of (absolute) convergence of output per capita. However, this paper did not attempt to explicitly investigate the issue of conditional income convergence. This would entail grouping countries according to common attributes and studying each groups relative income distribution. This has been left for further research.

We are also able to characterise the ergodic relative income distribution for the countries in our sample. It appears that there are three clusters of countries in the ergodic distribution. It is more likely that countries in the second income class move to the lowest income class than to move to the middle income class. Once countries are in the lowest class there is a much smaller probability that they will move up to a higher class. Countries that are in the fourth class are more likely to move to the middle income class rather than the highest income class although there is positive probability that countries will move to the highest bin. Once in the highest bin there is a low chance of moving out of it. There is some evidence here for a relative income "poverty trap". Countries with low relative incomes are not likely to move out of that income class to a higher relative income.

In addition, we find that countries whose relative incomes were in the middle three income classes were more mobile than the countries in the extreme income classes. The expected length of stay for countries in the middle three income classes were approximately one half of those for the extreme income classes. We also found evidence that countries have different growth rates in the ergodic distribution thus implying that the world has not converged in growth rates.

While we have found that there is evidence that the world has converged to its ergodic distribution we have not put forward any explanation as to why the data from 1970 to 1990 supports the ergodic assumption but the data from 1960 to 1990 does not. As the model used in this paper was designed to have as little structure as possible, it is not possible for us to try to investigate why this has occurred. What we have been able to do is to provide evidence for ergodicity. The question of why has been left for further research.

References

Barro, Robert J. (1991) 'Economic growth in a cross section of countries.' *Quarterly Journal of Economics* 106(2), 407–43

Barro, Robert J., and Xavier Sala-i-Martin (1992) 'Convergence.' *Journal of Political Economy* 100(2), 223–51

Champernowne, D. G. (1953) 'A model of income distribution.' *Economic Journal* 63(250), 318–351

Chari, V. V., Patrick J. Kehoe, and Ellen R. McGratten (1995) 'The poverty of nations: A quantitative exploration.' Staff Report 204/JV, Federal Reserve Bank of Minneapolis

Devroye, L. (1986) *Non-uniform Random Variate Generation* (New York:Springer-Verlag)

Durlauf, Steven N., and Danny T. Quah (1999) 'The new empirics of economic growth.' In *Handbook of Macroeconomics,* ed. John B. Taylor and Michael Woodford (North-Holland Elsevier Science) chapter 4, pp. 231–304

Geweke, John F. (1994) 'Bayesian comparison of econometric models.' Working Paper 532, Federal Reserve Bank of Minneapolis, July

——— (1998) 'Economics 8-205, Applied Econometrics II Lecture Notes.' University of Minnesota

Geweke, John F., C. Marshall, and G. A. Zarkin (1986) 'Mobility indices in continuous time markov chains.' *Econometrica* 54(6), 1407–23

Lucas, Robert E. (1988) 'On the mechanics of economic growth.' *Journal of Monetary Economics* 22(1), 3–42

Mankiw, N. G., D. Romer, and D. N. Weil (1992) 'A contribution to the empirics of economic growth.' *Quarterly Journal of Economics* 107(2), 407–37

Parente, Stephen L., and Edward C. Prescott (1993) 'Changes in the wealth of nations.' *Federal Reserve Bank of Minneapolis Quarterly Review*

Quah, Danny T. (1993) 'Empirical cross-section dynamics in economic growth.' *European Economic Review* 37(2/3), 426–434

——— (1996) 'Convergence empirics across countries with (some) capital mobility.' *Journal of Economic Growth* 1(1), 95–124

——— (1997) 'Empirics for growth and distribution: Polarization, stratification, and convergence clubs.' *Journal of Economic Growth* 2(1), 27–59

Shorrocks, Anthony F. (1976) 'Income mobility and the markov assumption.' *Economic Journal* 86, 566–578

Summers, R., and A. Heston (1991) 'The Penn World Table (Mark 5): An expanded set of international comparisons, 1950-88.' *Quarterly Journal of Economics*

Temple, Jonathon (1999) 'The new growth evidence.' *Journal of Economic Literature* 37(1), 112–156

Ventura, J. (1997) 'Growth and interdependence.' *Quarterly Journal of Economics* 112(1), 57–84