A Bayesian Exploration of Growth and Convergence

John S. Landon-Lane and J. Arnold Quinn*

First Version: April 24, 1996
This Version: March 12, 1999

Abstract

The topic of economic growth and convergence of countries has been an active topic in the 1990's. This paper investigates the question of whether there is any empirical evidence that countries are converging, with respect to their relative incomes, over time. To test for convergence, the evolution of the relative income for a country is modelled as a first-order Markov chain. Bayesian methods are then used investigate the posterior distributions of the parameters of the Markov chain and of other functions of interest related to the Markov chain. Issues of embeddability and mobility are discussed and Bayes factors are used to test for evidence of convergence across the countries in our sample. This is achieved through the use of carefully constructed prior distributions for the transition probabilities of the Markov chain. Contrary to existing studies we find little evidence in support of convergence of countries either across the whole data set or conditionally across subsets of countries that we believe are similar in underlying production technologies. We find that there is strong evidence that countries are diverging with respect to their relative income even for the conditional case. We also find that there is strong evidence of a structural break in the data around 1974 and that the properties of the Markov chain change significantly when this break is taken into account for all but "poor" countries.

Keywords: Growth, Convergence, Markov chain, Bayesian estimation, Mobility
JEL Classification: C1, E0, O4

*John Landon-Lane is Lecturer in the School of Economics at The University of New South Wales while J. Arnold Quinn is in the Department of Economics at the University of Minnesota. Many of the ideas in this paper were suggested to us by John Geweke. We wish to thank him for encouragement and comments on an earlier draft. We are also grateful for comments from Claustre Bajona-Xandri, Patrick Bajari, and Merrell Hora. All mistakes are our own. Contact: j.landon-lane@unsw.edu.au

1 Introduction

The topics of economic growth and convergence of countries and states have been active topics in the 1990's. Barro (1991), Barro and Sala-i-Martin (1992), Mankiw, Romer, and Weil (1992) and Ventura (1995) have all discussed the topic of growth and convergence of economies. This paper follows Chari, Kehoe, and McGratten (1995) in modeling economic growth as a first order Markov chain. Bayesian methods are used to estimate the parameters of the transition matrix along with other functions of interest for various groups of countries. Questions of convergence and structural change are addressed.

The neoclassical growth model predicts that countries with similar structural makeup, that is similar technology and preference parameters, will grow at rates inversely proportional to their size. That is, poorer countries will grow at faster rates than richer countries; and because of this countries should eventually converge in levels. Barro (1991) pointed out that the evidence does not support this contention of the model but that the data suggest that conditional on the stock of human capital, economies will converge. That is, for a given level of human capital, economic growth is inversely proportional to the initial level of physical capital. Barro called this type of convergence conditional convergence. Barro (1991) and Barro and Sala-i-Martin (1992) find strong evidence of conditional convergence across the states of the United States of America and weaker evidence for conditional convergence across countries. Barro and Sala-i-Martin (1992) look at the original members of the O.E.C.D. and report that there is some evidence of conditional convergence among these countries. Mankiw, Romer, and Weil (1992) find that, holding population growth and capital accumulation constant, there is evidence to support the predictions of the neoclassical growth model.

Using endogenous growth models, such as Lucas (1988), it would be expected that rich and poor countries would have different growth rates which would mean either rich and poor countries would be diverging or converging. Ventura (1995) outlines a model where, depending on certain preference parameters, convergence can still occur in an endogenous growth world. Parente and Prescott (1993) find that the disparity of wealth per capita across rich and poor countries is remaining 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. This is one of the facts that Chari, Kehoe, and McGratten (1995) are trying to explain. The findings of Parente and Prescott (1993) suggest that the endogenous growth models may not be completely correct in their modeling of growth.

The object of this paper is not to try to distinguish between competing models.

Rather, the aim of this paper is to ascertain, using a first order Markov chain model, whether the data support the contentions of these models. Following Chari, Kehoe, and McGratten (1995) we divide countries in our sample into discrete categories, or 'bins', according to their relative Gross Domestic Product (GDP) per worker. We calculate relative GDP per worker by comparing each country's level of GDP per worker with the geometric mean of all of the countries in the sample. Thinking of the world in this way leads us to model growth as a first order Markov chain. Modeling the world in this way also allows us to address questions of convergence and divergence. We follow the Bayesian methodology in our estimation procedures which allows us to use the Bayesian model comparison literature outlined in Geweke (1994) to test our hypotheses. We also informally discuss issues that arise from these comparisons. The paper will deal with a number of data sets comprising of groups of countries with the aim of distinguishing between absolute convergence and conditional convergence.

Before dealing with the question of convergence there needs to be a determination of whether there is a structural break in the data. As described in Eichengreen (1994), there was a breakdown of the Breton Woods agreement in 1973 and as a consequence countries moved away from fixed exchange rate regimes towards flexible exchange rate regimes. As countries moved from fixed exchange rate regimes to flexible exchange rate regimes the capital controls that were in effect for the fixed exchange rate regime were also lifted. The move to flexible exchange rates would certainly change the environment in which countries operated. A number of authors have talked about the potential importance of trade between countries as a determinant of growth (e.g. Ventura (1995) and Grossman and Helpman (1991)), so that it would be expected that a complete change in the exchange rate regimes would affect growth. Also, whether there is a structural break in the data could have an effect on results that were obtained using data that covered the period when the break occurred. For example, the data that we use for countries covers the years 1960 to 1990. We indeed find evidence, using Bayesian model comparison techniques, that the Markov chain process is different before and after 1975. In many cases we find that inference based on a one transition model (covering 1960 to 1990) differs substantively with inference based on a two transition model (from 1960 to 1974 and from 1975 to 1990).

The relative mobility of countries can also be studied. We estimate the parameters of the transition matrix of the first order Markov chain using standard Bayesian techniques that are described in Section 3 of this paper. Once we have these estimates we can test whether the first order Markov chain can be embedded into a continuous time process. If it can, then we can report continuous time mobility measures for the Markov chain. These measures will be useful in making inferences about the evolution countries over time. Discrete time mobility measures will also be reported for the instances when we cannot find evidence for embeddability. We are, however, most interested in testing for convergence or divergence, be it absolute or conditional, of countries in our sample over time. We will use the Bayesian model comparison literature to test for convergence. Convergence and divergence priors for the Markov chain will be defined and these will be used to test whether there is evidence of convergence in the data. For example, to test for convergence against a hypothesis of no convergence we would estimate the model twice. The first time we would use a convergence prior and then compare that estimated model with an estimated model with a prior that suggested no convergence. We could then form the Bayes factor of the convergence model against the no-convergence model and from that we could make inferences about whether the data supported convergence or not. As a by-product of estimating the parameters of a first order Markov chain, we are able to make inferences regarding the proximity of the final distribution to the invariant distribution implied by the model. We will also test for any evidence of conditional convergence by using a restricted data set of the O.E.C.D countries. These countries have a good chance of being economically integrated.

The layout of the paper is as follows. Section 2 will describe the model and interpret the different types of mobility measures that can be used. Section 3 will outline the econometric techniques used in the paper while Section 4 will describe the data that is used. Section 5 will present the results and Section 6 will conclude.

2 The Model

The transition of countries will be modelled as a first order discrete Markov chain. Suppose that a country can be in either of s distinct bins where the membership of a bin will be based on a country's relative GDP per worker. Define π_{it} as the probability that a country is in bin i at time t. Then for each time period t, the state of the model is described by

$$\pi_t = [\pi_{1t}, \ldots, \pi_{st}]'. \tag{1}$$

Let the probability that a country is in bin j at time t given that the country was in bin i at time $t-1$ be denoted by p_{ij}. Then the transition matrix for the first order Markov chain is

$$P = [p_{ij}]. \tag{2}$$

The transition of the model from period $t-1$ to t is then

$$\pi'_t = \pi'_{t-1}P. \tag{3}$$

That is, for each j,

$$\pi_{jt} = \sum_{i=1}^{s} p_{ij}\pi_{i,t-1}. \tag{4}$$

For a given initial state, π_0, we can express the state of the model, π_t, as

$$\pi'_t = \pi'_{t-1}P = \cdots = \pi'_0 P^t.$$

Thus, if

$$\pi^{*\prime} = \lim_{t\to\infty} \pi'_t$$

exists and is the same for all initial states, π_0, we say that π^* is the invariant distribution of the transition matrix P. The invariant distribution is of interest to us as it will give an indication as to whether transition matrix of the countries of our sample, as described by the above first order discrete Markov chain, represents a world where countries are converging, diverging or remaining roughly the same. If we observe an invariant distribution that has more mass in the middle "bins" than the current distribution then we would say that this is a world where countries are converging.

The invariant distribution satisfies

$$\pi_t^{*\prime} = \pi_t^{*\prime} P \tag{5}$$

and so the invariant distribution is just the left eigenvector of P associated with the eigenvalue equal to one. Note that P is a row stochastic matrix and so at least one of the eigenvalues of P is equal to one.

Geweke, Marshall, and Zarkin (1986b) describes a number of mobility measures for a first order discrete Markov chain. As we believe that there may be a structural break in the data we are therefore interested in how the mobility of countries changed after the structural break, if any. Because we believe the structural break was caused by the shift of exchange rate regimes to flexible exchange rates there may be evidence that there is more mobility after 1974 than before it. Geweke, Marshall, and Zarkin (1986b) describe a number of properties that mobility indices should have.

A mobility index M is a map from the space of transition matrices to the real numbers with the property that $M(I) = 0$. Geweke, Marshall, and Zarkin (1986b) outline a number of criteria that mobility indices need to satisfy. One criteria is the persistence criteria which states that mobility indices should be consistent with simple interpretations of the transition matrix. For example, if all of the off diagonal elements of a transition matrix P^1 are greater than the off diagonal elements of another transition matrix P^2 then we should see more mobility between bins for the model with transition matrix P^1. Hence we would like to see $M(P^1) > M(P^2)$. This condition is known as monotonicity (M). Other persistence criteria are immobility (I), which states that $M(P) \geq 0$; and strict immobility (SI), which states that $M(P) > 0$.

The mobility index

$$M_P(P) = \frac{s - tr(P)}{s-1} \tag{6}$$

suggested by Shorrocks (1978) satisfies (M), (I), and (SI). The above mobility index is $\frac{s}{s-1}$ times the harmonic mean of the mean length of stay in bin i, $\frac{1}{1-p_{ii}}$, over all i. Similar indices such as

$$M_B(P) = \sum_{i=1}^{s} \pi_i \sum_{j=1}^{s} p_{ij} |i-j| \quad \text{and} \tag{7}$$

$$M_U(P) = s \sum_{i=1}^{s} \pi_i \frac{(1-p_{ii})}{s-1} \tag{8}$$

satisfy (I) and (SI) but do not satisfy (M). The mobility index $M_U(P)$ is the unconditional probability of leaving the current bin.

Another criteria for mobility indices is the convergence criteria. This criteria states that the mobility index M should order transition matrices P according to how fast P^t converges to P^*, which is the limiting transition matrix. We can order the eigenvalues of P, without loss of generality, so that $1 \geq |\lambda_2| \geq \cdots \geq |\lambda_s|$. As suggested in Sommers and Conlisk (1979), π_t converges to π^* at the rate at which $|\lambda_i|^t$ converges to zero for all $i = 2, \ldots, s$; which is no faster than the rate at which $|\lambda_2|^t$ converges to zero. Thus the larger is λ_2, the slower is the rate of convergence to the invariant distribution. Note that the convergence criteria is valid only for those transition matrices for which P^* exists. It should also be noted that for P^* to exist and hence for the convergence criteria to apply, we require that $\lambda_2 < 1$. Geweke, Marshall, and Zarkin (1986b) states that any index which is a strictly decreasing function of the moduli of the eigenvalues of P will satisfy the convergence criteria. Three such indices are

$$M_E(P) = \frac{s - \sum_{i=1}^{s} |\lambda_i|}{s-1}, \tag{9}$$

$$M_D(P) = 1 - |\det(P)| \quad \text{and} \tag{10}$$

$$M_2(P) = 1 - |\lambda_2|. \tag{11}$$

In this paper we report the mobility measures defined in equations (6), (9), (10) and (11). We do not report the mobility indices defined in equations (7) and (8) as these are inconsistent within the persistence criteria. Using the mobility index described in equation (6) we will be able to compare the mobility of countries before and after 1974 to see if there is a change in the mobility of countries between bins after the change of exchange rate regimes. The last three discrete mobility measures that we report will enable us to compare the rate of transition to the invariant distribution for the implied transition matrices before and after 1974.

While we have modeled the transition of countries as a discrete time first order Markov chain, we are interested in testing whether the process can be imbedded into a continuous time process. If we can, then we are able to report more mobility statistics as described in Geweke, Marshall, and Zarkin (1986b). Suppose that we have a discrete first order Markov chain with transition matrix P. Geweke, Marshall, and Zarkin (1986a) shows that, if the discrete time process can be thought of as a process that results from a continuous time process, the rate of change of the vector of state probabilities can be expressed as

$$\dot{\pi}_t' = \pi_t' R \tag{12}$$

where the matrix R is referred to as the intensity matrix. In fact, $R = QNQ^{-1}$ where Q is the matrix of right eigenvectors of P corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_s$ and $N = \text{diag}\left(\nu_1, \ldots, \nu_s\right)$ where $\nu_i = \frac{\log(\lambda_i)}{T}$. Here, T is the number of time periods between observations. See Geweke, Marshall, and Zarkin (1986a) and Geweke, Marshall, and Zarkin (1986b) for a more detailed discussion. A discrete time process is said to be embeddable if all of the off diagonal elements of R, r_{ij} for $i \neq j$, are non-negative and if all of the diagonal elements, r_{ii}, are non-positive. That is, $r_{ij} \geq 0$ for all $i \neq j$ and $r_{ii} \leq 0$ for all $i = 1, \ldots, s$. The off diagonal elements of R can be interpreted as the instantaneous rates of transition from bin i into bin j. One problem that can arise is that $\log\left(\lambda_j\right)$ has many solutions for eigenvalues λ_j that are complex. This is known as the aliasing problem. Solutions to this problem include using $\log\left(\lambda_j\right) = \text{Log}\left(\lambda_j\right)$ where $\text{Log}\left(\lambda_j\right)$ is the value of $\log\left(\lambda_j\right)$ whose imaginary part has smallest absolute value amongst all solutions or to rule out solutions that are not embeddable.

Continuous time equivalents of the discrete time mobility measures (6) through (11) can be defined. Geweke, Marshall, and Zarkin (1986b) shows that the continuous time equivalents of (6), (7) and (10) are equivalent and can be expressed as

$$M_C^*\left(R\right) = \frac{-tr\left(R\right)}{s} = \frac{-\log\left[\det\left(P\right)\right]}{s} \tag{13}$$

Geweke, Marshall, and Zarkin (1986b) also shows that M_C^* satisfies (M), (SI) and velocity, which is defined as $M_C^*\left(kR\right) > M_C^*\left(R\right)$ for all $k > 1$; and that M_C^* is invariant to which solution of $\log\left(\lambda_j\right)$ we choose. The continuous time mobility measure corresponding to equation (11) is

$$M_2^*\left(R\right) = -\text{Re}\left(\nu_2\right) = -\text{Re}\left[\log\left(\lambda_2\right)\right]. \tag{14}$$

Embeddability is imposed and moments for continuous time mobility measures are calculated as described in Section 3. The mobility measures will enable us to determine the degree of movement of countries between bins and also the degree of convergence to the implied invariant distribution.

3 Econometric Model and Algorithms

3.1 Development of the Posterior Distribution

Let N be an $s \times s$ matrix, where n_{ij} is the observed number of governmental units that move from bin i in the first period to bin j in the last period and s is the number of bins. As before, P is an $s \times s$ transition matrix of a first order Markov chain. The likelihood function, i.e. the probability that we see N given P, is

$$L\left[N|P\right] = \prod_{i=1}^{s}\prod_{j=1}^{s} p_{ij}^{n_{ij}}. \tag{15}$$

The natural conjugate prior for this likelihood is a product of Dirichlet distributions

$$p\left[P\right] = \prod_{i=1}^{s}\prod_{j=1}^{s} p_{ij}^{\ell_{ij}}. \tag{16}$$

so that the posterior distribution of P given N is

$$p\left[P|N\right] = \prod_{i=1}^{s}\prod_{j=1}^{s} p_{ij}^{n_{ij}+\ell_{ij}} = \prod_{i=1}^{s}\prod_{j=1}^{s} p_{ij}^{a_{ij}} \propto \prod_{i=1}^{s} f_i\left[p_{i1},\ldots,p_{is}\right] \tag{17}$$

where

$$f_i\left[p_{i1},\ldots,p_{is}\right] = \frac{\Gamma\left(\sum_{j=1}^{s}\left(a_{ij}+1\right)\right)}{\prod_{j=1}^{s}\Gamma\left(a_{ij}+1\right)}\prod_{j=1}^{s} p_{ij}^{a_{ij}} \tag{18}$$

is the density of the Dirichlet $(a_{i1}+1,\ldots,a_{is}+1)$ distribution. Thus, the rows of P are independent of each other, but the elements within a row are dependent since each row must sum to one.

Since we know that the rows of P follow a Dirichlet distribution, it follows (see Berger (1985) page 561) that the mean and variance of the elements of P are

$$\begin{aligned} E\left[p_{ij}\right] &= \frac{a_{ij}+1}{\sum_{k=1}^{s}\left(a_{ik}+1\right)}; \quad \forall i=1,\ldots,s; \forall j=1,\ldots,s \quad \text{and} \\ Var\left[p_{ij}\right] &= \frac{\left[\sum_{k=1}^{s}\left(a_{ik}+1\right)-\left(a_{ij}+1\right)\right]\cdot\left(a_{ij}+1\right)}{\left[\sum_{k=1}^{s}\left(a_{ik}+1\right)\right]^2\cdot\left[1+\sum_{k=1}^{s}\left(a_{ik}+1\right)\right]}; \quad \forall i=1,\ldots,s; \quad \forall j=1,\ldots,s. \end{aligned} \tag{19}$$

Let y be an $s \times 1$ vector equal to $(y_1,\ldots,y_s)$; where y_i is distributed as a Gamma(ν_i) for all $i=1,\ldots,s$ and y_i and y_j are independent for all $i \neq j$. Now let x be an $s \times 1$ vector equal to $(x_1,\ldots,x_s)$; where x_i is defined as

$$x_i = \frac{y_i}{\sum_{j=1}^{s} y_j} \tag{20}$$

Press (1972) (page 134) shows that x has a Dirichlet$(\nu_1, \ldots, \nu_s)$ distribution. It is well known that the χ^2 distribution is a special case of the Gamma distribution; so that if z is distributed as a Gamma$\left(\frac{\nu}{2}\right)$ then $2z$ is distributed as $\chi^2(\nu)$. In this way, drawing $2y_i$ from a $\chi^2(2\nu_i)$ is equivalent to drawing y_i from a Gamma(ν_i). Thus we can draw from the $\chi^2(2\nu_i)$ distribution and divide the draw by two to get a variable distributed as a Gamma(ν_i). Note from the definition of x_i that multiplying y_i by a constant does not change the value of x_i, so we can simply draw y_i from a $\chi^2(2\nu_i)$ distribution and apply our definition of x_i to get x distributed as a Dirichlet$(\nu_i, \ldots, \nu_s)$.

Given this background, draws from the posterior of p_{ij} are made by drawing y_{ij} from a $\chi^2[2(a_{ij}+1)]$ distribution for $i = 1, \ldots, s$ and $j = 1, \ldots, s$ and then setting

$$p_{ij} = \frac{y_{ij}}{\sum_{k=1}^{s} y_{ik}}, \quad \forall i = 1, \ldots, s; \quad \forall j = 1, \ldots, s. \tag{21}$$

It then follows from the above argument that the resulting rows of P will have the appropriate Dirichlet distribution. As a result, we can make M independent and identically distributed draws of the transition matrix from the posterior distribution.

3.2 Functions of Interest

We will report a number of functions of interest, as discussed in Section 2. The proceeding discussion is meant to familiarize the reader with the algorithms that were used to calculate those functions of interest which can not be calculated analytically. It will become apparent that most of the following functions of interest are highly non-linear functions, the importance of which will be discussed at the end of the section.

Let $g\left(P^{(i)}\right)$ be any function of the i^{th} draw from the posterior distribution. Note that $g\left(P^{(j)}\right)$ and $g\left(P^{(k)}\right)$ are independent and identically distributed since $P^{(j)}$ and $P^{(k)}$ are independent and identically distributed (Lindgren 1993). Let

$$\bar{g}_M \equiv M^{-1} \sum_{i=1}^{M} g\left(P^{(i)}\right)$$

and

$$\bar{g} \equiv E\left[g(P)\,|N\right]$$

If $\bar{g} < \infty$, then the strong law of large numbers implies that $\bar{g}_M$ converges almost surely to $\bar{g}$. In addition, if we let

$$\sigma_g^2 = Var\left[g(P)\,|N\right] = E\left\{\left[g(p) - \bar{g}\right]^2 |N\right\}$$

and assume that $\sigma_g^2 < \infty$, then the central limit theorem implies that $M^{-\frac{1}{2}}\left(\bar{g}_M - \bar{g}\right)$ converges in distribution to a normal with mean 0 and variance σ_g^2. In addition, the strong law of large numbers implies that

$$s_M^2 = \frac{1}{M-1} \sum_{i=1}^{M} \left[g\left(P^{(i)}\right) - \bar{g}_M \right]^2$$

converges almost surely to σ_g^2. This was shown in Geweke (1995)

In Section 2 we noted that the invariant distribution of the transition matrix is of interest. We calculate the right eigenvectors and eigenvalues for $P^{(i)\prime}$ for each of the M draws of P from the posterior distribution. We then take $g\left(P^{(i)}\right)$ to be the vector valued function corresponding to the left eigenvector of $P^{(i)}$ corresponding to the eigenvalue equal to one. Moments for the invariant distribution are then calculated as above.

We actually calculate two separate sets of results for each model. For the first set of results, we draw M transition matrices from the posterior distribution without imposing embeddability. We do not calculate moments for continuous time mobility measures or the intensity matrix for this set of results. We impose embeddability for the second set of results so that we are able to calculate moments for the intensity matrix and continuous time equivalents of the discrete matrix mobility measures defined above. Note that imposing embeddability is achieved via acceptance sampling (see Geweke (1995)); i.e. we test whether any of the M draws from the posterior distribution are embeddable by applying the definition found in Section 2. Let K be the number (less than or equal to M) of the draws from the posterior that are embeddable. This subset of the K draws is used to calculate functions of interest and moments of those functions of interest. Of course the moment definitions above are modified to have K's rather than M's.

Note that statistical inference based on frequentist methods would lead to a very limited discussion of the above functions of interest. One could certainly calculate the maximum likelihood estimates of P,

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j=1}^{s} n_{ij}},$$

and then calculate the functions of interest of $\hat{P}$. However, the sampling distributions of the functions of interest would be very difficult to determine due to the fact that most of the functions of interest are highly non-linear functions of P. In addition, the question of embeddability (as noted in Geweke, Marshall, and Zarkin (1986a)) is almost intractable.

Frequentist methods would simply indicate that either $\hat{P}$ is embeddable or it is not embeddable. There is no sense in which we could get an idea of the probability that P is embeddable. Both the question of the distribution of functions of interest and the posterior probability of events are highly tractable when one conditions on the data and applies Bayesian methods. For instance, we can report the posterior probability that P is embeddable. We accomplish this feat using the function of interest $g\left(P^{(i)}\right)$, where $g\left(P^{(i)}\right)$ is equal to zero if $P^{(i)}$ is embeddable and one otherwise. Using the notation established above, the posterior probability of embeddability is simply $\frac{K}{M}$. Similarly, we can report the probability that $|\lambda_2|$ is between 0.20 and 0.25 using the function of interest $g\left(P^{(i)}\right)$, where $g\left(P^{(i)}\right)$ is equal to one if $|\lambda_2|$ is between 0.20 and 0.25 and zero otherwise.

3.3 Bayes Factor

It is going to be desirable to compare models. For instance, as discussed in Section 2, we have interest in testing whether the data support the hypothesis that there are actually two transition periods, one from 1960 to 1974 and another from 1975 to 1990, rather than one transition period from 1960 to 1990. Similarly, it is going to be of interest to compare different priors on P. Fortunately, we can draw from the literature on Bayesian model comparison. Following Geweke (1994), we define the marginal likelihood for model j as

$$M_{jN} = \int_{\mathcal{P}} p_j\left(P_i\right) \cdot L_j\left[N|P_i\right] dP_i \tag{22}$$

where $p_j\left[P_i\right]$ is the prior on P_i for model j, $L_j\left[N|P_i\right]$ is the likelihood of N given P_i for model j and

$$\mathcal{P} = \left\{ p_{ij} : p_{ij} \leq 1, \sum_{j=1}^{s} p_{ij} = 1 \quad \forall i = 1, \ldots, s \right\}.$$

It is important that both the prior and likelihood are proper densities. We can then define the Bayes factor of model j in favor of model k as

$$B_{jk} = \frac{M_{jN}}{M_{kN}} \tag{23}$$

The Bayes factor will give us a sense of how likely model j is relative to model k.

We can modify the notation used above to distinguish between priors and transitions. Denote the prior number of degrees of freedom under model q for p_{ij} in the first transition as ${}^{q_1}\ell_{ij}$. Denote the likelihood number of degrees of freedom for p_{ij} in the first transition as ${}^{1}n_{ij}$. Then we can write the marginal likelihood for a model q with one transition as

$$\begin{aligned} M_{qN} &= \int_{\mathcal{P}} \prod_{i=1}^{s} \frac{\Gamma\left[\sum_{j=1}^{s}\left({}^{q_1}\ell_{ij}+1\right)\right]}{\prod_{j=1}^{s}\Gamma\left[{}^{q_1}\ell_{ij}+1\right]} \prod_{i=1}^{s} p_{ij}^{{}^{q_1}\ell_{ij}} \cdot \prod_{i=1}^{s} \frac{\Gamma\left[\sum_{j=1}^{s}\left({}^{1}n_{ij}+1\right)\right]}{\prod_{j=1}^{s}\Gamma\left[{}^{1}n_{ij}+1\right]} \prod_{i=1}^{s} p_{ij}^{{}^{1}n_{ij}} dP \\ &= \frac{\prod_{i=1}^{s}\left\{\Gamma\left[\sum_{j=1}^{s}\left({}^{q_1}\ell_{ij}+1\right)\right] \cdot \Gamma\left[\sum_{j=1}^{s}\left({}^{1}n_{ij}+1\right)\right]\right\}}{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left({}^{q_1}\ell_{ij}+1\right) \cdot \prod_{j=1}^{s}\Gamma\left({}^{1}n_{ij}+1\right)} \int_{\mathcal{P}} \prod_{i=1}^{s} p_{ij}^{{}^{q_1}\ell_{ij}+{}^{1}n_{ij}} dP \end{aligned} \tag{24}$$

Notice that the integrand is simply the kernel for a Dirichlet so that the marginal likelihood is

$$M_{qN} = \frac{\prod_{i=1}^{s}\left\{\Gamma\left[\sum_{j=1}^{s}\left({}^{q_1}\ell_{ij}+1\right)\right] \cdot \Gamma\left[\sum_{j=1}^{s}\left({}^{1}n_{ij}+1\right)\right]\right\}}{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left({}^{q_1}\ell_{ij}+1\right) \cdot \prod_{j=1}^{s}\Gamma\left({}^{1}n_{ij}+1\right)} \cdot \frac{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left[{}^{q_1}\ell_{ij}+{}^{1}n_{ij}+1\right]}{\prod_{i=1}^{s}\Gamma\left[\sum_{j=1}^{s}\left({}^{q_1}\ell_{ij}+{}^{1}n_{ij}+1\right)\right]} \tag{25}$$

The marginal likelihood for a model with two transitions is simply the product of two one-transition marginal likelihoods since we are assuming that the two transition matrices are independent of each other. If we make a small modification to the notation above, we can denote the marginal likelihood for two transition model q quite simply. Let

(26)

$$M_{qN}^{(1)} = \frac{\prod_{i=1}^{s}\left\{\Gamma\left[\sum_{j=1}^{s}\left({}^{q_1}\ell_{ij}+1\right)\right]\cdot\Gamma\left[\sum_{j=1}^{s}\left({}^{1}n_{ij}+1\right)\right]\right\}}{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left({}^{q_1}\ell_{ij}+1\right)\cdot\prod_{j=1}^{s}\Gamma\left({}^{1}n_{ij}+1\right)}\cdot\frac{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left[{}^{q_1}\ell_{ij}+{}^{1}n_{ij}+1\right]}{\prod_{i=1}^{s}\Gamma\left[\sum_{j=1}^{s}\left({}^{q_1}\ell_{ij}+{}^{1}n_{ij}+1\right)\right]}$$

and

(27)

$$M_{qN}^{(2)} = \frac{\prod_{i=1}^{s}\left\{\Gamma\left[\sum_{j=1}^{s}\left({}^{q_2}\ell_{ij}+1\right)\right]\cdot\Gamma\left[\sum_{j=1}^{s}\left({}^{2}n_{ij}+1\right)\right]\right\}}{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left({}^{q_2}\ell_{ij}+1\right)\cdot\prod_{j=1}^{s}\Gamma\left({}^{2}n_{ij}+1\right)}\cdot\frac{\prod_{i=1}^{s}\prod_{j=1}^{s}\Gamma\left[{}^{q_2}\ell_{ij}+{}^{2}n_{ij}+1\right]}{\prod_{i=1}^{s}\Gamma\left[\sum_{j=1}^{s}\left({}^{q_2}\ell_{ij}+{}^{2}n_{ij}+1\right)\right]}$$

Then the marginal likelihood, M_{qN}, for two transition model q is

$$M_{qN} = M_{qN}^{(1)}\cdot M_{qN}^{(2)}. \tag{28}$$

The intuition behind the marginal likelihood is relatively straight forward. The marginal likelihood will be large when the prior number of degrees of freedom "resemble" the likelihood number of degrees of freedom.

4 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.[1]

The data consist of observations on real GDP per worker for 104 countries for the years 1960 through 1990. The criterion for a country to be included in our sample is that 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. Version 5.6 of the Penn World Tables is essentially the same as version 5 that is described in Summers and Heston (1991). The changes in version 5.6 that affect the data we used are the corrections to some errors in population and labor force participation rate estimates. For a complete description of the changes to versions 5 and 5.5 that were made in version 5.6 see the World Wide Web page at the University of Toronto. For a complete description of the Penn World Tables version 5 see Summers and Heston (1991). We then followed Chari,

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

Kehoe, and McGratten (1995) and expressed our data relative to the geometric mean of the whole sample for each year.

5 Results

5.1 Evidence for convergence

The results in this section are based on the data set containing all of the countries. A number of different models are considered. For the purposes of this section, we use two different classification systems. We present results for a model with five bins (relative real GDP per worker from 0.000 to 0.350, 0.350 to 0.700, 0.700 to 1.400, 1.400 to 2.800 and 2.800 to 5.600) and for a model with four bins (relative real GDP per worker from 0.000 to 0.300, 0.300 to 1.000, 1.000 to 3.000 and 3.000 to 6.000). Both classification systems have advantages and disadvantages. The five bin classification has greater detail and allows us to think about "middle" countries as belonging to only one bin. However, the five bin classification thinly spreads the data. This artificially hinders the possibility of finding embeddable transition matrices (Geweke, Marshall, and Zarkin 1986b). The four bin classification allows for fewer "empty cells", but also decreases the detail of the inference.

We consider four general classes of priors. The priors have been created to be consistent with various theories of growth. Flat priors give equal prior probability to moving from any bin i to any bin j. Diagonal priors have been created so that the number of degrees of freedom declines exponential away from the main diagonal. The invariant distribution implied by these priors can be characterized as having essentially equal measures of countries in each bin. Convergence priors use the diagonal priors as a base, and then give additional weight to the middle bin. The invariant distribution implied by convergence priors is characterized by having more weight in the middle bins relative to the end bins. Divergence priors also use the diagonal priors as a base, but put additional weight in the top and bottom bins. This causes the implied invariant distribution to have more weight at the end bins relative to the middle bins. One can consider the prior number of degrees of freedom as a notional data set, where the number of degrees of freedom, ℓ_{ij}, is interpreted as the number of countries in the artificial data that move from bin i to bin j. Each prior has been set up so that the number of notional data points in a given bin is small relative to the number of actual data points in that bin. These priors are combined with assumptions regarding the number of transitions to arrive at the priors used. Tables 1 through 3 list the prior number of degrees of freedom and the implied transition matrix mean and invariant distribution for each 5 bin prior considered. Four bin priors are similar.

Bayes factors for all the models considered can be found in Tables 4 and 5 . Several patterns are apparent. There is strong evidence in favor of two transition models. For instance, model D5 is a two transition generalization of model D3. The Bayes factor in favor of model D5 versus model D3 is 147.1089 and 22.9768 in the five bin and four bin models respectively. This implies that one needs to give model D3 a prior probability 147 times the prior probability given to model D5 in order to have equal posterior probability for the two models. In fact, model F1 is the only one transition model that has a Bayes factor greater than one versus its two transition counterpart. In addition, priors which are consistent with the divergence of nations are strongly favored to priors which are consistent with the convergence of nations. The Bayes factor for model D5 versus C5 is 35.9977 and 51.3290 in the five bin and four bin models respectively. Notice that "divergence" priors like D5 are slightly favored to simple "diagonal" priors like F5. Nonetheless, results are robust across priors given the number of bins. In the spirit of brevity, we will present results for model D5, the model which has a Bayes factor greater than one versus all other priors considered, and its one transition counterpart D3.

We will start by presenting the results for the version of model D5 which has five bins. The posterior probability of embeddability is 0.00072 in the first transition and 0.00022 in the second transition. We have 50,000 draws from the posterior distribution for P, which implies that only 36 of the first transition draws were embeddable and only 11 of the second transition draws were embeddable. Due to the small sample size, we do not present results for the subset of draws which are embeddable.

Table 6 presents the first and second (analytical) moments for the transition matrices for model D5. In both the first and the second transition, the probability of staying in either the top or the bottom bin is much greater than the probability of staying in one of the middle bins. This is exactly the opposite of what one would expect if countries were converging. In spite of this similarity across transitions, the convergence story is different in the two transitions. If we use the probability of leaving bins as a measure of mobility, it is apparent that nations are more mobile in the second transition period than they are in the first transition period. The probability of staying in bins one, three and five are similar across the two transitions. However, the probability of staying in bin two is more than one and a half standard deviations smaller in transition two than in transition one. The probability of staying in bin four is almost a standard deviation smaller in transition two than it is in transition one. In general, we see that there is much more movement from the top bins to the middle bins in transition two than there is in transition one. While we can not say that there is evidence for convergence of nations in transition two, we should notice that transition two is more consistent with the convergence of nations than is transition one. Finally, note that the mean transition matrix for model D3, presented in Table 8, masks the difference between the two transition periods. The mean transition matrix is generally more similar to transition one of D5 than it is to transition two of D5.

Inference using the transition matrix is somewhat difficult because there are a large number of parameters at which to look. Inference using the invariant distribution is somewhat more straight-forward. Table 7 presents the posterior moments for the invariant distribution of the transition matrix for model D5 along with the distribution of countries in 1960, 1974, 1975 and 1990. Table 9 presents similar information for model D3. The data suggest that countries are becoming increasingly different over time. Comparing the distribution of nations in 1960 and 1990, it appears that countries are diverging. There were more many countries in the top and bottom bins in 1990 than there were in 1960. The invariant distribution for model D3 tells the same story.

There is a high posterior probability (above 0.90) that the invariant distribution for model D3 has more nations in either of the two extreme bins than does the initial distribution of nations in 1960. Furthermore, the invariant distribution for model D3 is more "extreme" than the distribution of nations in 1990. That is, model D3 suggests continued divergence of nations. The first transition of model D5 provides the same message. Comparisons of the 1960 distribution of nations with both the 1974 distribution of nations and the mean invariant distribution indicate a divergence of nations. Transition two of model D5 suggests weaker evidence that the divergence of nations will continue. The invariant distribution for the top three bins is somewhat different in the second transition than it is in the first transition. We see a movement of countries from the top bin to the third and fourth bin in transition two relative to transition one. In fact, the invariant distribution for the top three bins in transition two is not much more extreme than the distribution of nations in 1990. There is only a 0.4993 posterior probability that the proportion of countries in the top bin in the invariant is greater than the proportion of countries in the top bin in 1990.

We can interpret the posterior probability that the invariant distribution is greater than the final distribution as a measure as how similar the two distributions are. That is, if the posterior probability is either close to one or close to zero we can infer that the invariant distribution and the final distribution are not close. For the one transition model D3 we find that there is little evidence to suggest that the final distribution and the invariant distribution are close. However, for the second transition for model D5 we see relatively strong evidence that the two distributions are close. We can also see this by noting that the final distribution is within one half of a standard deviation of the invariant distribution for each bin.

As a final note, the speed of convergence to the invariant distribution, as measured by the modulus of the second eigenvalue of the transition matrix, is essentially the same across the two transitions. The mean of $|\lambda_2|$ is 0.7886 with standard deviation of 0.0694 in the first transition of model D5 and mean of $|\lambda_2|$ is 0.7982 with standard deviation of 0.0563 in the second transition of model D5. The posterior probability that the $|\lambda_2|$ is bigger in transition two than in transition one is 0.5343.

Matrix mobility measures are reported in Table 10. Consistent with the discussion of the transition matrix moments, $M_p(P)$ is probably larger in transition one than in transition two. The posterior probability that nations are more mobile in transition two than in transition one, using $M_p(P)$ as the mobility measure, is 0.8185. The question of speed of convergence to the invariant distribution is ambiguous. There is very high posterior probability that either $M_D(P)$ or $M_E(P)$ is greater in the second transition than it is in the first transition. However, there is only a 0.4657 posterior probability that $M_2(P)$ is greater in the second transition as compared to the first transition. Given these results, there seems to be reasonable evidence that nations are more mobile in transition two than they are in transition one.

The four bin first order Markov chain model results are similar to the five bin model. The posterior probability of embeddability is non-trivial for both of the transition matrices in model D5. The posterior probability of embeddability is 0.1688 for the first transition and 0.0477 for the second transition.

The transition matrix moments for model D5 and model D3 can be found in Tables 11 and 12 respectively. Relative to transition one, there is more movement out of bins in the second transition in model D5. Most of the additional movement is to lower bins. In general, we see the pattern that divergence of countries in the bottom two bins is prevalent in both transitions, while the divergence of countries in the top two bins is less extreme in the second transition than it is in the first transition. The same pattern can be seen in the intensity matrices reported in Table 13 and 14. Finally, one should also note that inference made from a one transition model (model D3) is more consistent with the inference made for the first transition of model D5 than for the second transition of model D5.

The evidence from the invariant distributions again points to a divergence of nations. This pattern is also evident in the change in of the actual distribution of nations from 1960 to 1990. However, it is instructive to analyze this pattern in the two transition periods we have been discussing. It is especially interesting to note that the single transition model D3 (invariant distributions are listed in Tables 15 and 16) indicates general divergence. The invariant distribution for first transition in model D5 also suggests divergence; but with more countries moving into the top bin than into the bottom bin. This pattern reverses itself in the second transition period. We see no evidence that the final distribution of countries in 1990 is close to the implied invariant distribution in the case of the single transition. However, as was the case for the five bin model, there is relatively strong evidence that the final distribution in 1990 is close to implied invariant distribution for the second transition.

Measures of mobility are found in Table 17. We see that there is overwhelming evidence that nations are more mobile in the second transition period than in the first, both in the sense of transition across states and in the sense of speed of convergence to the invariant distribution. For a subset of matrix mobility measures, the posterior probability that the second period transition matrix displays more mobility than the first period mobility matrix is greater than 0.90.

As a brief summary, both the four and five bin first order Markov chain models of nations suggests that nations are diverging rather than converging. However, it is also apparent that this divergence has not been as marked since 1975. In addition, the type of divergence is different across the two transition periods discussed. From 1960 to 1974, more countries were becoming richer than were becoming poorer. From 1975 to 1990, we see exactly the opposite phenomenon. It is important to note that the one transition models mask this pattern. Finally, the five bin Markov chain model suggests that countries that are in the top three bins are not diverging in the second transition period.

5.2 Evidence for Conditional Convergence

Given results from Barro (1991) and Barro and Sala-i-Martin (1992), it is not surprising that countries do not appear to be converging. However we have yet to discuss *conditional* convergence. We will now consider subsets of the 104 countries discussed above. In particular, we have attempted to group countries so that the national stock of human capital within the group is roughly the same; so that we can apply Barro's definition of conditional convergence. We follow Barro (1991) and Barro and Sala-i-Martin (1992) by looking at the countries in the O.E.C.D.. In addition, we naively group countries based on their relative real GDP per worker in 1960. We group all countries with a relative real GDP per worker greater than 2.00 ("rich countries") and we group all countries with a relative real GDP per worker less than 0.50 ("poor countries"). Table 18 lists the countries in each group.

For each group, the relative (with respect to other countries in the group) real GDP per worker was obtained by calculating the geometric mean of real GDP per worker within the group and then dividing actual real GDP per worker by the group's geometric mean. As is expected, we see much less variation in relative real GDP per worker within groups than we see in the all country models. Due to the lack of variation, we use a three bin classification system. The bins for the "poor" countries are: 0.00 to 0.80, 0.80 to 1.20, 1.20 to 3.60; the bins for the "rich" countries and the O.E.C.D. countries are: 0.00 to 0.80, 0.80 to 1.20, 1.20 to 2.40. The three bin priors are similar to the five bin priors listed in Tables 1 to 3. In the interest of brevity, the full set of results will not be presented for each group, though they are available upon request. The results are robust across priors. Model D3 has a Bayes factor greater than one against all one transition models for all three groupings, while model D5 has a Bayes factor greater than one against all two transition models for the three groupings. Therefore we will only present results for models D3 and D5. Tables 19 through 21 contains the Bayes factors for all three groupings.

Inference for the "poor" countries does not change substantially when going from a one transition model to a two transition model. The invariant distributions for models D3 and D5 can be found in Tables 22 and 23. The data and both model D3 and model D5 indicate that the distribution for "poor" countries is shifting down. This is weak evidence that the "poor" countries are diverging. For the "poor" countries, we see little evidence that the final distribution is close to the implied invariant distribution for either the single transition model or the two transition model. The mobility measures reported in Table 28 indicate little difference between the mobility of countries across the two transitions in Model D5; further evidence that no structural break exists for the "poor" countries.

Quite to the contrary, results for the "rich" countries are very different across the two transitions in model D5. The first transition in model D5 strongly suggests that the "rich" countries are converging. The posterior probability that the proportion of countries in the middle bin of the invariant distribution is greater than the proportion of countries actually in the middle bin in 1960 is over 0.97 (see Table 24). The second transition on model D5 presents exactly the opposite picture. There is less than 0.01 posterior probability that the proportion of countries in the middle bin of the invariant distribution is greater than the proportion of countries actually in the middle bin in 1975. Notice that model D3 does not suggest strong evidence for either divergence or convergence. It is less clear for the "rich" countries whether the final distribution is close to the implied invariant distribution. The middle bin is not very close to its invariant distribution while the other bins are somewhat close.

The evidence for convergence of the O.E.C.D. countries depends on the model one considers. Evidence from the actual distributions of countries in a one transition model like model D3 points to convergence of the O.E.C.D. countries; the proportion of countries in the middle bins goes from 0.2222 in 1960 to 0.4444 in 1990 (see Table 26). The invariant distribution for model D3 suggests weaker evidence for convergence; movement from the top bin to the middle is apparent, but the proportion of countries in the bottom bin stays roughly the same. Evidence from the two transition model D5 suggests even weaker evidence of convergence. The results from the first transition of model D5 are very similar to the results from model D3. However, there is additional movement into the bottom bin in the second transition of model D5. Thus, it appears that the O.E.C.D. countries are converging if one considers a one transition model, while the evidence for convergence is much weaker when a two transition model is entertained. Note that the mobility measures for the O.E.C.D. countries (presented in Table 29) suggest that the O.E.C.D. countries are more mobile in transition two than in transition one of model D5. The evidence from the invariant distributions suggest that the O.E.C.D. countries have not reached their implied invariant distributions for either of the models considered.

In summary, we can note several interesting patterns. First, "poor" countries do not seem to be affected by the phenomenon that indicates a structural break in 1974. Second, both the "rich" countries and O.E.C.D. countries do seem to behave differently after 1974.

We find some evidence for conditional convergence of the O.E.C.D. countries especially if we consider a one transition model. However, we do not have strong evidence for conditional convergence of countries that have been naively grouped together as "rich" in 1960.

6 Conclusion

A first order Markov chain model was estimated using Bayesian techniques. We used Bayesian model comparison techniques to check for a structural break in the data, and given those results, we looked for evidence of convergence. We then selected the best model according to the Bayes Factor criteria and reported a number of function of interest regarding the posterior distribution of the model in an attempt to address whether the distribution of countries has converged to its implied invariant distribution.

A number of patterns emerged from the analysis. There is strong evidence for a structural break, in the early 1970's, for the data and the model that we used. This has implications for researchers who wish to make inferences based on data from the Penn World Tables. In particular, using a one transition Markov chain model to make inferences regarding the transition of countries across time, on the weight of the evidence, is not appropriate. It was found that the first transition of a two transition model was consistent with a one transition model but that the second transition was quite different. It is interesting to note that the evidence for a structural break is not strong for "poor" countries. If the structural break was indeed caused by an increased movement of capital, then this would suggest that the "poor" countries have yet to benefit from increased capital mobility.

For the sample as a whole, there is little evidence supporting convergence of countries in levels. In fact the Bayes factors consistently favour divergence priors. There is strong evidence in favor of divergence of countries across the full data set. However we do see evidence of convergence when we restrict our attention to O.E.C.D. countries. This evidence is stronger for the first transition of the two transition model than the second transition in regard to the O.E.C.D. countries. An interesting result is that, when restricting attention to "rich" countries rather than O.E.C.D. countries, we see evidence for divergence after 1974 rather than convergence. The reason for this could be the mobility of capital after the change in exchange rate regime. Countries that offer the best return for capital would benefit from increased capital flows while countries that have net capital outflow could perform worse. We also see evidence of countries being more mobile after 1974 than before. The evidence for "poor" countries is significantly different from that of the "rich" countries in our sample. We do not see evidence of a structural break around 1974 which suggests that the change in exchange rate regimes has had little effect on the economic growth for these countries. "Poor" countries have not benefited from capital flows as much as the 'rich" countries. This could be due to owners of capital preferring investments that have a lower risk of failure.

We addressed the question of whether countries had converged to their implied invariant distribution. We saw that the conclusions we drew were vastly different if we were looking at a single transition model or a two transition model. By looking at the posterior probability that the implied invariant distribution was greater than the actual distribution we could infer whether the two were close. Posterior probabilities that were either close to one or close to zero led us to infer that the invariant distribution and the actual distribution were not close. We found that, using the complete sample, the distribution of countries in 1990 was close to the implied invariant distribution for the second transition of the two transition model. We could not make the same inference for the one transition model. This is further evidence that the issue of whether there is a structural break in the data is indeed an important one. The question of whether an economic process has settled down to its stationary or equilibrium state is important to researchers wanting to use the data to contrast results obtained from models assumed to be in equilibrium. Our results suggest that if we think of the growth of countries across time as a first order Markov chain, with a structural break at 1974, then the data suggests that countries are close to the invariant or stationary distribution implied by the second period's transition matrix.

References

Barro, R. J. (1991): "Economic Growth in a Cross Section of Countries," *Quarterly Review of Economics*, 106.

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

Berger, J. O. (1985): *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag, New York, 2 edn.

Chari, V. V., R. J. Kehoe, and E. R. McGratten (1995): "The Poverty of Nations: A Quantitative Exploration," Staff Report 204, Federal Reserve Bank of Minneapolis.

Eichengreen, B. (1994): "History of International Monetary Systems: Implications for Research in International Macroeconomics and Finance," in *Handbook of International Macroeconomics*, ed. by F. van der Ploeg. Blackwell, Oxford, United Kingdom.

Geweke, J. F. (1994): "Bayesian Comparison of Econometric Models," Working Paper 532, Federal Reserve Bank of Minneapolis.

——— (1995): "Monte Carlo Simulation and Numerical Integration," Staff Report 192, Federal Reserve Bank of Minneapolis.

Geweke, J. F., C. Marshall, and G. A. Zarkin (1986a): "Exact Inference for Continuous Time Markov Chain Models," *Review of Economic Studies*, 53.

——— (1986b): "Mobility Indices in Continuous Time Markov Chains," *Econometrica*, 54.

Grossman, G. M., and E. Helpman (1991): *Innovation and Growth in the Global Economy*. MIT Press.

Lindgren, B. (1993): *Statistical Theory*. Chapman and Hall, Inc., 4 edn.

Lucas, R. E. (1988): "On the Mechanics of Economic Growth," *Journal of Monetary Economics*, 22.

Mankiw, N. G., D. Romer, and D. N. Weil (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Review of Economics*, 107.

Parente, S. L., and E. C. Prescott (1993): "Changes in the Wealth of Nations," *Federal Reserve Bank of Minneapolis Quarterly Review*, 17.

Press, S. J. (1972): *Applied Multivariate Analysis*. Holt,Rinehart and Winston, Inc.

Shorrocks, A. F. (1978): "The Measurement of Mobility," *Econometrica*, 46.

Sommers, P. M., and J. Conlisk (1979): "Eigenvalue Immobility Measures for Markov Chains," *Journal of Mathematical Sociology*, 6.

Summers, R., and A. Heston (1991): "The Penn World Table(Mark 5): An Expanded Set of International Comparisons, 1950-88," *Quarterly Review of Economics*, 106.

Ventura, J. (1995): "Growth and Interdependence," mimeo, Harvard University.

Table 1: "Flat and diagonal" priors

Priors					Expected Value					Inv. Dist.
F1-One Transition										
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
F2-Two Transition/Transition 1										
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
F2-Two Transition/Transition 2										
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
0.500	0.500	0.500	0.500	0.500	0.200	0.200	0.200	0.200	0.200	0.200
F3-One Transition										
1.333	0.490	0.180	0.066	0.024	0.329	0.210	0.166	0.150	0.144	0.193
0.490	1.333	0.490	0.180	0.066	0.197	0.309	0.197	0.156	0.141	0.203
0.180	0.490	1.333	0.490	0.180	0.154	0.194	0.304	0.194	0.154	0.207
0.066	0.180	0.490	1.333	0.490	0.141	0.156	0.197	0.309	0.197	0.205
0.024	0.066	0.180	0.490	1.333	0.144	0.150	0.166	0.210	0.329	0.193
F4-Two Transition/Transition 1										
2.000	0.736	0.271	0.100	0.037	0.368	0.213	0.156	0.135	0.127	0.191

Table 1: "Flat and diagonal" priors

0.736	2.000	0.736	0.271	0.100	0.196	0.339	0.196	0.144	0.124	0.205
0.271	0.736	2.000	0.736	0.271	0.141	0.193	0.333	0.193	0.147	0.207
0.100	0.271	0.736	2.000	0.736	0.124	0.144	0.196	0.339	0.196	0.205
0.037	0.100	0.271	0.736	2.000	0.127	0.135	0.156	0.213	0.368	0.191

F4-Two Transition/Transition 2

1.666	0.613	0.226	0.083	0.031	0.350	0.212	0.161	0.142	0.135	0.192
0.613	1.666	0.613	0.226	0.083	0.197	0.325	0.197	0.149	0.132	0.205
0.226	0.613	1.666	0.613	0.226	0.147	0.193	0.320	0.193	0.147	0.207
0.083	0.226	0.613	1.666	0.613	0.132	0.149	0.197	0.325	0.197	0.204
0.031	0.083	0.226	0.613	1.666	0.135	0.142	0.161	0.212	0.350	0.192

F5-Two Transition/Transition 1

2.000	0.736	0.271	0.100	0.037	0.368	0.213	0.156	0.135	0.127	0.191
0.736	2.000	0.736	0.271	0.100	0.196	0.339	0.196	0.144	0.124	0.205
0.271	0.736	2.000	0.736	0.271	0.141	0.193	0.333	0.193	0.147	0.207
0.100	0.271	0.736	2.000	0.736	0.124	0.144	0.196	0.339	0.196	0.205
0.037	0.100	0.271	0.736	2.000	0.127	0.135	0.156	0.213	0.368	0.191

F5-Two Transition/Transition 2

2.000	0.736	0.271	0.100	0.037	0.368	0.213	0.156	0.135	0.127	0.191
0.736	2.000	0.736	0.271	0.100	0.196	0.339	0.196	0.144	0.124	0.205
0.271	0.736	2.000	0.736	0.271	0.141	0.193	0.333	0.193	0.147	0.207
0.100	0.271	0.736	2.000	0.736	0.124	0.144	0.196	0.339	0.196	0.205
0.037	0.100	0.271	0.736	2.000	0.127	0.135	0.156	0.213	0.368	0.191

Table 2: "Convergence" priors

Priors					Expected Value					Inv. Dist.
C3-One Transition										
0.297	0.232	0.180	0.031	0.005	0.226	0.205	0.214	0.179	0.175	0.174
0.109	0.630	0.490	0.085	0.015	0.175	0.257	0.235	0.171	0.160	0.199
0.040	0.232	1.333	0.232	0.040	0.151	0.179	0.339	0.179	0.151	0.253
0.015	0.085	0.490	0.630	0.109	0.160	0.171	0.235	0.257	0.175	0.201
0.005	0.031	0.180	0.232	0.297	0.175	0.179	0.205	0.214	0.226	0.173
C4-Two Transition/Transition 1										
0.446	0.347	0.271	0.047	0.008	0.236	0.220	0.208	0.171	0.165	0.165
0.164	0.944	0.736	0.128	0.022	0.166	0.278	0.248	0.161	0.146	0.198
0.060	0.347	2.000	0.347	0.060	0.136	0.172	0.384	0.172	0.136	0.274
0.022	0.128	0.736	0.944	0.164	0.146	0.161	0.248	0.278	0.166	0.199
0.008	0.047	0.271	0.347	0.446	0.165	0.171	0.208	0.220	0.236	0.163
C4-Two Transition/Transition 2										
0.372	0.290	0.226	0.039	0.007	0.231	0.217	0.207	0.175	0.170	0.169
0.137	0.787	0.613	0.107	0.019	0.171	0.268	0.242	0.166	0.153	0.199
0.050	0.290	1.666	0.290	0.050	0.143	0.176	0.363	0.176	0.143	0.264
0.019	0.107	0.613	0.787	0.137	0.153	0.166	0.242	0.268	0.171	0.201
0.007	0.039	0.226	0.290	0.372	0.170	0.175	0.207	0.217	0.231	0.168
C5-Two Transition/Transition 1										
0.446	0.347	0.271	0.047	0.008	0.236	0.220	0.208	0.171	0.165	0.165
0.164	0.944	0.736	0.128	0.022	0.166	0.278	0.248	0.161	0.146	0.198
0.060	0.347	2.000	0.347	0.060	0.136	0.172	0.384	0.172	0.136	0.274
0.022	0.128	0.736	0.944	0.164	0.146	0.161	0.248	0.278	0.166	0.199
0.008	0.047	0.271	0.347	0.446	0.165	0.171	0.208	0.220	0.236	0.163
C5-Two Transition/Transition 2										
0.446	0.347	0.271	0.047	0.008	0.236	0.220	0.208	0.171	0.165	0.165
0.164	0.944	0.736	0.128	0.022	0.166	0.278	0.248	0.161	0.146	0.198
0.060	0.347	2.000	0.347	0.060	0.136	0.172	0.384	0.172	0.136	0.274
0.022	0.128	0.736	0.944	0.164	0.146	0.161	0.248	0.278	0.166	0.199
0.008	0.047	0.271	0.347	0.446	0.165	0.171	0.208	0.220	0.236	0.163

Table 3: “Divergence” priors

Priors					Expected Value					Inv. Dist.
D3-One Transition										
1.333	0.297	0.066	0.015	0.003	0.347	0.193	0.159	0.151	0.149	0.219
0.297	0.490	0.109	0.024	0.015	0.219	0.251	0.187	0.173	0.171	0.193
0.066	0.109	0.180	0.109	0.066	0.193	0.201	0.213	0.201	0.193	0.178
0.015	0.024	0.109	0.490	0.297	0.171	0.173	0.187	0.251	0.219	0.190
0.003	0.015	0.066	0.297	1.333	0.149	0.151	0.159	0.193	0.347	0.219
D4-Two Transition/Transition 1										
2.000	0.446	0.100	0.022	0.005	0.396	0.191	0.145	0.135	0.133	0.227
0.446	0.736	0.164	0.037	0.022	0.226	0.271	0.182	0.162	0.160	0.190
0.100	0.164	0.271	0.164	0.100	0.190	0.201	0.219	0.201	0.190	0.171
0.022	0.037	0.164	0.736	0.446	0.160	0.162	0.182	0.271	0.226	0.188
0.005	0.022	0.100	0.446	2.000	0.133	0.135	0.145	0.191	0.396	0.225
D4-Two Transition/Transition 2										
1.666	0.372	0.083	0.019	0.004	0.373	0.192	0.152	0.143	0.141	0.222
0.372	0.613	0.137	0.031	0.019	0.222	0.261	0.184	0.167	0.165	0.190
0.083	0.137	0.226	0.137	0.083	0.191	0.201	0.216	0.201	0.191	0.174
0.019	0.031	0.137	0.613	0.371	0.165	0.167	0.184	0.261	0.222	0.191
0.004	0.019	0.083	0.372	1.666	0.141	0.143	0.152	0.192	0.373	0.223
D5-Two Transition/Transition 1										
2.000	0.446	0.100	0.022	0.005	0.396	0.191	0.145	0.135	0.133	0.227
0.446	0.736	0.164	0.037	0.022	0.226	0.271	0.182	0.162	0.160	0.190
0.100	0.164	0.271	0.164	0.100	0.190	0.201	0.219	0.201	0.190	0.171
0.022	0.037	0.164	0.736	0.446	0.160	0.162	0.182	0.271	0.226	0.188
0.005	0.022	0.100	0.446	2.000	0.133	0.135	0.145	0.191	0.396	0.225
D5-Two Transition/Transition 2										
2.000	0.446	0.100	0.022	0.005	0.396	0.191	0.145	0.135	0.133	0.227
0.446	0.736	0.164	0.037	0.022	0.226	0.271	0.182	0.162	0.160	0.190
0.100	0.164	0.271	0.164	0.100	0.190	0.201	0.219	0.201	0.190	0.171
0.022	0.037	0.164	0.736	0.446	0.160	0.162	0.182	0.271	0.226	0.188
0.005	0.022	0.100	0.446	2.000	0.133	0.135	0.145	0.191	0.396	0.225

Table 4: Bayes Factors: Five State Models

	F1	F2	F3	F4	F5	C3	C4	C5	D3	D4	D5
F1	1.0000	1.3943	0.0301	0.0006	0.0006	0.0659	0.0042	0.0045	0.0184	0.0001	0.0001
F2	0.7172	1.0000	0.0216	0.0004	0.0004	0.0473	0.0030	0.0032	0.0132	0.0001	0.0001
F3	33.2362	46.3414	1.0000	0.0187	0.0185	2.1913	0.1390	0.1498	0.6121	0.0047	0.0042
F4	1778.8722	2480.2904	53.5221	1.0000	0.9903	117.2835	7.4406	8.0165	32.7597	0.2515	0.2227
F5	1796.3299	2504.6318	54.0474	1.0098	1.0000	118.4340	7.5137	8.0952	33.0812	0.2539	0.2249
C3	15.1662	21.1469	0.4564	0.0085	0.0084	1.0000	0.0634	0.0684	0.2793	0.0021	0.0019
C4	239.0568	333.3281	7.1932	0.1344	0.1331	15.7626	1.0000	1.0774	4.4028	0.0338	0.0299
C5	221.8843	309.3837	6.6765	0.1247	0.1235	14.6303	0.9282	1.0000	4.0865	0.0314	0.0278
D3	54.2964	75.7081	1.6338	0.0305	0.0302	3.5801	0.2271	0.2447	1.0000	0.0077	0.0068
D4	7073.2519	9862.5649	212.8342	3.9766	3.9379	466.3836	29.5882	31.8781	130.2711	1.0000	0.8856
D5	7987.3207	11137.0936	240.3385	4.4905	4.4405	526.6538	33.9977	35.9977	147.1059	1.1292	1.0000

Table 5: Bayes Factors: Four State Models

	F1	F2	F3	F4	F5	C3	C4	C5	D3	D4	D5
F1	1.0000	1.1312	0.1440	0.0206	0.0222	0.3084	0.1392	0.1692	0.0738	0.0034	0.0032
F2	0.8840	1.0000	0.1273	0.0182	0.0196	0.2726	0.1231	0.1457	0.0652	0.0030	0.0028
F3	6.9447	7.8557	1.0000	0.1431	0.1543	2.1415	0.9669	1.1449	0.5125	0.0236	0.0223
F4	48.5446	54.9132	6.9902	1.0000	1.0783	14.9696	6.7585	8.0032	3.5825	0.1647	0.1559
F5	45.5446	50.9281	6.4829	0.9274	1.0000	13.8833	6.2680	7.4224	3.3225	0.1527	0.1446
C3	3.2428	3.6682	0.4670	0.0668	0.0720	1.0000	0.4515	0.5346	0.2393	0.0110	0.0104
C4	7.1826	8.1248	1.0343	0.1480	0.1595	2.2149	1.0000	1.1842	0.5301	0.0244	0.0231
C5	6.0655	6.6812	0.8734	0.1250	0.1347	1.8705	0.8445	1.0000	0.4476	0.0206	0.0195
D3	13.5500	15.3276	1.9512	0.2791	0.3010	4.1785	1.8865	2.2340	1.0000	0.0460	0.0435
D4	294.8030	333.4771	42.4516	6.0730	6.5482	90.9106	41.0442	48.6034	21.7567	1.0000	0.9469
D5	311.3353	352.1782	44.8323	6.4136	6.9154	96.0088	43.3460	51.3290	22.8768	1.0561	1.0000

Table 6: Transition matrix moments for model D5: 5 states

Transition 1

From 1960 to 1974

0.7636 (0.0857)	0.1038 (0.0615)	0.0467 (0.0425)	0.0434 (0.0411)	0.0426 (0.0408)
0.2532 (0.0789)	0.5692 (0.0898)	0.1076 (0.0562)	0.0353 (0.0335)	0.0348 (0.0332)
0.0299 (0.0277)	0.1132 (0.0515)	0.5237 (0.0812)	0.3034 (0.0748)	0.0299 (0.0277)
0.0456 (0.0431)	0.0463 (0.0434)	0.0966 (0.0611)	0.5238 (0.1032)	0.2877 (0.0936)
0.0409 (0.0392)	0.0416 (0.0395)	0.0448 (0.0409)	0.0588 (0.0465)	0.8139 (0.0770)

Transition 2

From 1975 to 1990

0.7777 (0.0752)	0.1165 (0.0580)	0.0372 (0.0342)	0.0346 (0.0330)	0.0340 (0.0328)
0.3181 (0.0943)	0.4160 (0.0998)	0.1779 (0.0774)	0.0433 (0.0417)	0.0437 (0.0414)
0.0396 (0.0363)	0.1138 (0.0592)	0.5493 (0.0927)	0.2577 (0.0815)	0.0396 (0.0363)
0.0387 (0.0368)	0.0393 (0.0371)	0.2713 (0.0849)	0.4445 (0.0949)	0.2062 (0.0773)
0.0329 (0.0317)	0.0334 (0.0320)	0.0360 (0.0331)	0.1454 (0.0627)	0.7523 (0.0768)

Table 7: Invariant distributions for Model D5: 5 states

Transition 1

State	1960	1974	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1974
1	0.1538	0.2019	0.2588	0.8416	0.6697
			(0.1074)		
2	0.2212	0.1827	0.1364	0.0788	0.1897
			(0.0412)		
3	0.2981	0.2019	0.1128	0.0009	0.0403
			(0.0454)		
4	0.1538	0.1923	0.1462	0.3941	0.1850
			(0.0546)		
5	0.1635	0.2115	0.3458	0.9511	0.8618
			(0.1262)		

Transition 2

State	1975	1990	Invariant Distribution	Post. Prob. ID > 1975	Post. Prob ID > 1990	Post. Prob. ID2 >ID1
1	0.2115	0.2500	0.2763	0.7059	0.5525	0.5514
			(0.1042)			
2	0.1635	0.1154	0.1146	0.1198	0.4448	0.3852
			(0.0412)			
3	0.2115	0.2212	0.1926	0.3431	0.2905	0.8601
			(0.0612)			
4	0.1923	0.1827	0.1748	0.3398	0.4078	0.6583
			(0.0526)			
5	0.2212	0.2308	0.2417	0.5415	0.4993	0.2521
			(0.0916)			

Table 8: Transition matrix moments for model D3: 5 states

0.7631 (0.0873)	0.1011 (0.0619)	0.0469 (0.0434)	0.0447 (0.0424)	0.0442 (0.0422)
0.4250 (0.0904)	0.2934 (0.0832)	0.1766 (0.0697)	0.0699 (0.0466)	0.0251 (0.0336)
0.0292 (0.0275)	0.1399 (0.0566)	0.4703 (0.0815)	0.2494 (0.0706)	0.1113 (0.0513)
0.0463 (0.0439)	0.0467 (0.0441)	0.1873 (0.0815)	0.4326 (0.1035)	0.2871 (0.0945)
0.0423 (0.0405)	0.0428 (0.0407)	0.0450 (0.0417)	0.1390 (0.0696)	0.7309 (0.0892)

Table 9: Invariant distributions for model D3: 5 states

State	1960	1990	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1990
1	0.1538	0.2500	0.2928	0.9145	0.6065
			(0.1121)		
2	0.2212	0.1154	0.0948	0.0024	0.2561
			(0.0351)		
3	0.2981	0.2212	0.1388	0.0033	0.0551
			(0.0475)		
4	0.1538	0.1827	0.1673	0.5613	0.3552
			(0.0542)		
5	0.1635	0.2308	0.3063	0.9334	0.7845
			(0.1044)		

Table 10: Discrete mobility measures

	M_p	M_e	M_d	M_2
Transition 1	0.4513	0.4496	0.9241	0.2114
	(0.0493)	(0.0486)	(0.0338)	(0.0694)
Transition 2	0.5150	0.5134	0.9698	0.2018
	(0.0495)	(0.0480)	(0.0233)	(0.0563)
PP	0.8185	0.8233	0.8766	0.4657

PP = Posterior probability that the transition two measure is greater than the transition one measure

Table 11: Transition matrix moments for model D5: 4 states

Transition 1

From 1960 to 1974

0.8157 (0.0834)	0.0790 (0.0580)	0.0553 (0.0492)	0.0500 (0.0469)
0.1075 (0.0467)	0.7444 (0.0657)	0.1216 (0.0492)	0.0265 (0.0242)
0.0265 (0.0242)	0.0751 (0.0397)	0.7212 (0.0676)	0.1773 (0.0575)
0.0419 (0.0396)	0.0463 (0.0415)	0.1068 (0.0610)	0.8050 (0.0783)

Transition 2

From 1975 to 1990

0.7544 (0.0868)	0.1537 (0.0727)	0.0483 (0.0432)	0.0437 (0.0412)
0.2672 (0.0727)	0.5838 (0.0810)	0.1174 (0.0529)	0.0316 (0.0288)
0.0271 (0.0248)	0.1482 (0.0542)	0.6670 (0.0718)	0.1577 (0.0556)
0.0348 (0.0331)	0.0385 (0.0348)	0.1563 (0.0656)	0.7704 (0.0760)

Table 12: Transition matrix moments for model D3: 4 states

0.7665 (0.0937)	0.1246 (0.0731)	0.0563 (0.0510)	0.0526 (0.0494)
0.2696 (0.0674)	0.4885 (0.0759)	0.2161 (0.0625)	0.0258 (0.0241)
0.0258 (0.0241)	0.1689 (0.0569)	0.5593 (0.0754)	0.2460 (0.0654)
0.0436 (0.0413)	0.0467 (0.0427)	0.1460 (0.0715)	0.7637 (0.0860)

Table 13: Intensity matrix moments for model D5: 4 states

Transition 1

From 1960 to 1974

-0.1980 (0.1079)	0.0926 (0.0756)	0.0568 (0.0528)	0.0486 (0.0489)
0.1282 (0.0625)	-0.3037 (0.0957)	0.1430 (0.0659)	0.0325 (0.0315)
0.0321 (0.0313)	0.0928 (0.0550)	-0.3403 (0.1023)	0.2154 (0.0836)
0.0403 (0.0399)	0.0446 (0.0414)	0.1357 (0.0842)	-0.2206 (0.1068)

Transition 2

From 1975 to 1990

-0.2994 (0.1328)	0.2041 (0.1135)	0.0523 (0.0476)	0.0430 (0.0405)
0.3519 (0.1246)	-0.5480 (0.1514)	0.1537 (0.0832)	0.0424 (0.0425)
0.0336 (0.0330)	0.1802 (0.0758)	-0.4138 (0.1135)	0.2000 (0.0797)
0.0361 (0.0340)	0.0442 (0.0420)	0.1911 (0.0936)	-0.2715 (0.1083)

Table 14: Intensity matrix moments for model D3: 4 states

-0.2691 (0.1318)	0.1470 (0.1062)	0.0722 (0.0654)	0.0500 (0.0485)
0.3674 (0.1189)	-0.7316 (0.1637)	0.3262 (0.1191)	0.0380 (0.0364)
0.0373 (0.0359)	0.2381 (0.1064)	-0.5924 (0.1537)	0.3169 (0.1077)
0.0427 (0.0427)	0.0572 (0.0525)	0.1761 (0.0996)	-0.2760 (0.1218)

Table 15: Invariant distributions for Model D5: 4states

Transition 1

State	1960	1974	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1974
1	0.1250	0.1538	0.2408	0.8451	0.7426
			(0.1210)		
2	0.3558	0.3077	0.1936	0.0363	0.0883
			(0.0799)		
3	0.3558	0.3269	0.2436	0.0992	0.1593
			(0.0840)		
4	0.1635	0.2115	0.3220	0.9268	0.8161
			(0.1215)		

Transition 2

State	1975	1990	Invariant Distribution	Post. Prob. ID > 1975	Post. Prob ID > 1990	Post. Prob. ID2 >ID1
1	0.1538	0.2115	0.3068	0.9471	0.8038	0.6796
			(0.1087)			
2	0.2885	0.2500	0.2118	0.1233	0.2601	0.5832
			(0.0649)			
3	0.3462	0.3077	0.2291	0.0637	0.1405	0.4537
			(0.0724)			
4	0.2115	0.2308	0.2523	0.6192	0.5378	0.3334
			(0.1001)			

Table 16: Invariant distributions for model D3: 4 states

State	1960	1990	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1990
1	0.1250	0.2115	0.2972	0.9617	0.7479
			(0.1184)		
2	0.3558	0.2500	0.1650	0.0013	0.0686
			(0.0537)		
3	0.3558	0.3077	0.2158	0.0242	0.0877
			(0.0651)		
4	0.1635	0.2308	0.3221	0.9384	0.7751
			(0.1146)		

Table 17: Mobility measures

	M_p	M_e	M_d	M_2	M_c^*	M_2^*
Transition 1	0.3047	0.3043	0.6706	0.1858	0.2657	0.1851
	(0.0494)	(0.0492)	(0.0727)	(0.0611)	(0.0538)	(0.0622)
Transition 2	0.4084	0.4082	0.8136	0.2053	0.3832	0.2253
	(0.0529)	(0.0528)	(0.0592)	(0.0610)	(0.0685)	(0.0672)
PP	0.9230	0.9240	0.9360	0.5855	0.9266	0.6766

PP = Posterior probability that the transition two measure is greater than the transition one measure

Table 18: Group classifications

O.E.C.D Countries	"Rich" Countries	"Poor" Countries
Australia	Argentina	Benin
Austria	Australia	Burkina Faso
Belgium	Austria	Burundi
Canada	Belgium	Cameroon
Czechoslovakia	Canada	Cape Verde Is.
Denmark	Chile	Central African Rep.
Finland	Denmark	Chad
France	Finland	China
Greece	France	Comoros
Iceland	Iceland	Gambia
Ireland	Iran	Ghana
Italy	Ireland	Guinea
Japan	Israel	Guinea-Biss.
Luxemborg	Italy	India
Mexico	Luxemborg	Indonesia
Netherlands	Mexico	Ivory Coast
New Zealand	Netherlands	Kenya
Norway	New Zealand	Lesotho
Portugal	Norway	Malawi
Spain	Sweden	Mali
Sweden	Switzerland	Mozambique
Switzerland	Trinidad & Tobago	Nigeria
Turkey	United Kingdom	Pakistan
United Kingdom	Uraguay	Rwanda
U.S.A	U.S.A.	Thailand
West Germany	Venezuela	Togo
	West Germany	Uganda

Table 19: Bayes Factors: "Poor" countries

	F1	F2	F3	F4	F5	C3	C4	C5	D3	D4	D5
F1	1.0000	0.7927	0.6007	0.2612	0.2701	0.6634	0.3938	0.4110	0.4885	0.1558	0.1553
F2	1.2615	1.0000	0.7578	0.3295	0.3407	0.8369	0.4967	0.5185	0.6163	0.1965	0.1959
F3	1.6648	1.3197	1.0000	0.4348	0.4497	1.1044	0.6555	0.6842	0.8133	0.2593	0.2585
F4	3.8289	3.0352	2.3000	1.0000	1.0342	2.5401	1.5077	1.5737	1.8705	0.5964	0.5945
F5	3.7023	2.9349	2.2239	0.9670	1.0000	2.4561	1.4578	1.5217	1.8087	0.5767	0.5749
C3	1.5074	1.1949	0.9055	0.3937	0.4071	1.0000	0.5936	0.6195	0.7364	0.2348	0.2341
C4	2.5396	2.0132	1.5255	0.6633	0.6859	1.6848	1.0000	1.0438	1.2407	0.3956	0.3943
C5	2.4331	1.9287	1.4615	0.6355	0.6572	1.6141	0.9581	1.0000	1.1886	0.3790	0.3778
D3	2.0470	1.6226	1.2296	0.5346	0.5529	1.3580	0.8060	0.8413	1.0000	0.3188	0.3178
D4	6.4201	5.0892	3.8564	1.6767	1.7341	4.2591	2.5280	2.6387	3.1364	1.0000	0.9969
D5	6.4401	5.1051	3.8685	1.6820	1.7395	4.2724	2.5359	2.6469	3.1462	1.0031	1.0000

Table 20: Bayes Factors: "Rich" countries

	F1	F2	F3	F4	F5	C3	C4	C5	D3	D4	D5
F1	1.0000	0.7273	0.5715	0.2226	0.2194	0.6465	0.3598	0.3654	0.4590	0.1328	0.1265
F2	1.3750	1.0000	0.7859	0.3061	0.3017	0.8890	0.4948	0.5025	0.6312	0.1826	0.1740
F3	1.7497	1.2725	1.0000	0.3895	0.3839	1.1313	0.6296	0.6394	0.8032	0.2323	0.2214
F4	4.4917	3.2666	2.5671	1.0000	0.9856	2.9041	1.6163	1.6414	2.0619	0.5964	0.5684
F5	4.5573	3.3144	2.6046	1.0146	1.0000	2.9465	1.6399	1.6654	2.0920	0.6051	0.5767
C3	1.5467	1.1248	0.8840	0.3443	0.3394	1.0000	0.5566	0.5652	0.7100	0.2054	0.1950
C4	2.7791	2.0211	1.5883	0.6187	0.6098	1.7968	1.0000	1.0156	1.2757	0.3690	0.3517
C5	2.7364	1.9901	1.5639	0.6092	0.6004	1.7692	0.9847	1.0000	1.2561	0.3633	0.3463
D3	2.1784	1.5843	1.2450	0.4850	0.4780	1.4085	0.7839	0.7961	1.0000	0.2892	0.2757
D4	7.5319	5.4776	4.3047	1.6768	1.6527	4.8697	2.7102	2.7525	3.4575	1.0000	0.9531
D5	7.9025	5.7471	4.5164	1.7594	1.7340	5.1093	2.8436	2.8879	3.6276	1.0492	1.0000

Table 21: Bayes Factors: "O.E.C.D." countries

	F1	F2	F3	F4	F5	C3	C4	C5	D3	D4	D5
F1	1.0000	0.8604	0.5588	0.2358	0.2420	0.6098	0.3785	0.3963	0.4344	0.1259	0.1235
F2	1.1623	1.0000	0.6494	0.2741	0.2813	0.7088	0.4399	0.4606	0.5049	0.1463	0.1435
F3	1.7897	1.5399	1.0000	0.4221	0.4331	1.0915	0.6773	0.7092	0.7775	0.2253	0.2210
F4	4.2405	3.6485	2.3694	1.0000	1.0262	2.5860	1.6048	1.6804	1.8421	0.5338	0.5236
F5	4.1322	3.5553	2.3089	0.9745	1.0000	2.5200	1.5638	1.6375	1.7951	0.5202	0.5103
C3	1.6398	1.4108	0.9162	0.3867	0.3968	1.0000	0.6206	0.6498	0.7123	0.2064	0.2025
C4	2.6423	2.2735	1.4764	0.6231	0.6395	1.6114	1.0000	1.0471	1.1479	0.3326	0.3263
C5	2.5235	2.1712	1.4100	0.5951	0.6107	1.5389	0.9550	1.0000	1.0962	0.3177	0.3116
D3	2.3020	1.9806	1.2862	0.5429	0.5571	1.4038	0.8712	0.9122	1.0000	0.2898	0.2843
D4	7.9440	6.8350	4.4387	1.8734	1.9225	4.8446	3.0064	3.1480	3.4510	1.0000	0.9810
D5	8.0980	6.9675	4.5248	1.9097	1.9598	4.9385	3.0647	3.2091	3.5179	1.0194	1.0000

Table 22: Invariant distributions for Model D5: "Poor" countries

Transition 1

State	1960	1974	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1974
1	0.2593	0.2963	0.3236	0.6915	0.5634
			(0.1142)		
2	0.4074	0.3704	0.3399	0.2240	0.3586
			(0.0904)		
3	0.3333	0.3333	0.3365	0.4714	0.4714
			(0.1355)		

Transition 2

State	1975	1990	Invariant Distribution	Post. Prob. ID > 1975	Post. Prob ID > 1990	Post. Prob. ID2 >ID1
1	0.2593	0.3333	0.3863	0.8401	0.6351	0.6407
			(0.1275)			
2	0.4444	0.4074	0.3560	0.1704	0.2857	0.5497
			(0.0937)			
3	0.2963	0.2593	0.2577	0.3306	0.4520	0.3307
			(0.1112)			

Table 23: Invariant distributions for Model D3: "Poor" countries

State	1960	1990	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1990
1	0.2593	0.3333	0.3631	0.8601	0.6028
			(0.0974)		
2	0.4074	0.4074	0.3867	0.3877	0.3877
			(0.0799)		
3	0.3333	0.2593	0.2502	0.1975	0.4248
			(0.1021)		

Table 24: Invariant distributions for Model D5: "Rich" countries

Transition 1

State	1960	1974	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1974
1	0.2963	0.1852	0.2222	0.2260	0.5957
			(0.1048)		
2	0.3333	0.5556	0.5781	0.9729	0.5605
			(0.1306)		
3	0.3704	0.2593	0.1996	0.0903	0.2738
			(0.1175)		

Transition 2

State	1975	1990	Invariant Distribution	Post. Prob. ID > 1975	Post. Prob ID > 1990	Post. Prob. ID2 >ID1
1	0.1852	0.2222	0.2716	0.7024	0.5860	0.5997
			(0.1377)			
2	0.5556	0.3704	0.2808	0.0057	0.1890	0.0366
			(0.1005)			
3	0.2593	0.4074	0.4476	0.9039	0.5798	0.9051
			(0.1490)			

Table 25: Invariant distributions for Model D3: "Rich" countries

State	1960	1990	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1990
1	0.2963	0.2222	0.2330	0.2344	0.5049
			(0.0950)		
2	0.3333	0.3704	0.3504	0.5584	0.3917
			(0.0887)		
3	0.3704	0.4074	0.4166	0.6558	0.5176
			(0.1068)		

Table 26: Invariant distributions for Model D5: O.E.C.D. countries

Transition 1

State	1960	1974	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1974
1	0.2593	0.2222	0.2537	0.4208	0.5135
			(0.1495)		
2	0.2222	0.3333	0.3915	0.9040	0.6323
			(0.1376)		
3	0.4815	0.4074	0.3548	0.1976	0.3482
			(0.1449)		

State	1975	1990	Invariant Distribution	Post. Prob. ID > 1975	Post. Prob ID > 1990	Post. Prob. ID2 >ID1
1	0.1852	0.1852	0.2716	0.7023	0.7023	0.5446
			(0.1378)			
2	0.3704	0.4444	0.4397	0.6836	0.4688	0.6075
			(0.1319)			
3	0.4074	0.3333	0.2887	0.1769	0.3395	0.3671
			(0.1282)			

Table 27: Invariant distributions for Model D3: O.E.C.D. countries

State	1960	1990	Invariant Distribution	Post. Prob. ID > 1960	Post. Prob ID > 1990
1	0.2593	0.1852	0.2464	0.4122	0.6633
			(0.1160)		
2	0.2222	0.4444	0.4220	0.9912	0.3750
			(0.0952)		
3	0.4815	0.3333	0.3316	0.0998	0.4829
			(0.1139)		

Table 28: Mobility measures: “Poor” countries

	M_p	M_e	M_d	M_2	M_c^*	M_2^*
Transition 1	0.6791	0.6449	0.9507	0.4397	0.7312	0.5996
	(0.1128)	(0.0887)	(0.0978)	(0.1318)	(0.2443)	(0.2331)
Transition 2	0.6699	0.6516	0.9197	0.4831	0.7035	0.6665
	(0.1142)	(0.0966)	(0.0905)	(0.1348)	(0.2283)	(0.2513)
PP	0.4783	0.5255	0.4068	0.5929	0.4667	0.5725

PP = Posterior probability that the transition two measure is greater than the transition one measure

Table 29: Mobility measures: O.E.C.D. countries

	M_p	M_e	M_d	M_2	M_c^*	M_2^*
Transition 1	0.4734	0.4717	0.7351	0.3391	0.4232	0.3983
	(0.1114)	(0.1091)	(0.1273)	(0.1188)	(0.1528)	(0.1690)
Transition 2	0.5096	0.5071	0.7641	0.3970	0.4740	0.4796
	(0.1126)	(0.1101)	(0.1159)	(0.1240)	(0.1637)	(0.1909)
PP	0.5897	0.5888	0.5666	0.6332	0.5939	0.6275

PP = Posterior probability that the transition two measure is greater than the transition one measure