Catching up, Forging ahead, and Falling behind:

A Panel Structure Analysis of Convergence Clubs*

Yixiao Sun
Department of Economics
Yale University

This Version: February 3, 2002

Job Market Paper II

*I am especially grateful to Peter Phillips, Donald Andrews, Chris Udry for their invaluable guidance and insightful comments on this and other projects. I would like to thank William Brainard, Eduardo Engel, Ray Fair, Michael Keane, Stefan Krieger, Giuseppe Moscarini, and Zhijie Xiao who have contributed to this paper with constructive suggestions. All errors are mine alone. *Email:* yixiao.sun@yale.edu. *Address:* Yixiao Sun, Department of Economics, Yale University, P.O. Box 208268, New Haven, CT 06520-08268, USA.

ABSTRACT

The paper advocates and implements a new panel structure model to investigate the club convergence hypothesis. The model consists of a set of linear dynamic models that characterize the behavior of growth rates within each convergence club and a logistic regression that classifies these linear models. An EM algorithm is used to estimate the system by maximum likelihood and inference is conducted using asymptotic theory derived for the model.

Our findings suggest that the world economy consists of three convergence clubs: an advanced club, an underdeveloped club and a developing club. Different convergence clubs exhibit different convergence behavior in terms of both speed of convergence and steady state growth rate. In particular, the steady state growth rates are for the three clubs are 2.09%, 0.27%, and 2.90% per year, respectively. These differences in long run growth imply that some countries will catch up and even forge ahead and some countries will fall behind.

JEL Classification: C13; C23; C51

Keywords: economic growth; convergence clubs, dynamic panel data; logistic regression.

1 Introduction

Many studies reveal persistent differences in growth rates of per capita income among countries. Some rich countries manage to sustain high growth rates over a long period while many poor countries seem to be trapped on a low growth path. These persistent differences are strikingly at odds with the neoclassical growth theory, which predicts faster growth of poorer countries. Table 1 illustrates the differences between some developed and developing countries in the last century. In general, developed countries, which started out with higher per capita income, experienced faster growth than developing countries. Differences in economic growth over a long period of time can work to stratify the world into different clubs. As suggested by Quah (1996, 1997), Paap and van Dijk (1998), the world is polarizing into clubs of rich and poor countries.

These persistent differences contradict the neoclassical growth models (e.g. Solow (1956), Swan (1956)). A common feature of these models has been the assumption of identical production functions for all countries. In consequence, a single linear dynamic model is adequate to characterize all cross-country growth behavior. That is, per capita growth converges at the same speed and to the same exogenous rate in the steady state. This is in sharp contrast with persistent differences in empirical growth rates over the last hundred years. On the other hand, there are growth models (e.g. Azariadis and Drazen (1990)) that produce multiple steady states in the growth rate of per capita income. Countries associated with the same steady state form a natural convergence club in that they grow at the same rate in the long run. The existence of a club structure is at variance with conventional studies that assume a single dynamic model for all countries. However, most studies continue to employ a single dynamic model, even though the existence of convergence clubs is supported by empirical evidence and certain theoretical reasoning. To the best of the author's knowledge, few papers have attempted to formally investigate the convergence club hypothesis.

The present paper advocates and implements a panel structure model to deal with this issue. It is a new model proposed and investigated by Sun (2001). The model assumes that countries form a number of convergence clubs in a heterogeneous world. Within each convergence club, the growth rate satisfies a dynamic linear model with the same coefficients, while these coefficients may be different across different convergence clubs. The inter-club coefficient difference implies heterogeneity in the speed of convergence and the long run growth rate across different clubs. Figure 1 illustrates the main idea. Graph (a) of the Figure describes the conditional convergence as predicted by the neoclassical growth theory. In steady states, all countries grow at the same rate, even though the levels of per capita income are heterogeneous. Graph (b) of the Figure describes the club convergence in which different clubs grow at different rates in steady states. It is our purpose to detect which countries are associated with which clubs.

We adopt a probabilistic approach to uncover the hidden club structure. Specifically, some covariates are used in a multinomial logistic regression to infer club memberships. The idea is simple. A club structure may come into being because of threshold externalities, non-convexities, and other properties related to some covariates. These covariates can thus be used to infer membership information. Potential covariates include initial human capital, initial income per capita, and ethnic diversity. This is broadly consistent with the multiplicity of steady states. Theories such as Azariadis and Drazon (1990) suggest that countries with identical structural characteristics may cluster around different steady states. According to Azariadis and Drazon, this can be attributed to differences in initial conditions.

The panel structure model thus consists of a set of linear dynamic models that characterize the behavior of growth rates and a logistic regression that classifies these linear models. From an empirical Bayes perspective, the logistic regression provides the prior probabilities that countries belong to particular clubs, whereas the linear dynamic model provides information to update the priors to get the posterior probabilities. A country is then assigned to the club to which it most likely belongs. For a given country, the assigned (ex post) club membership depends on its initial conditions, its speed of convergence, and its long run growth rate.

From the perspective of model specification, the panel structure model is also a device for controlling parameter heterogeneity. Research on convergence has accommodated cross country heterogeneity in a sequence of stages. At first, conventional cross-section analysis (Barro (1991), Mankiw, Romer and Weil (1992)) and panel data analysis (Islam (1995)) assumed complete homogeneity in steady state growth rates. Recently, Lee, Pesaran and Smith (1997, 1998) allowed complete heterogeneity in steady state growth rates. However, as pointed out by Islam (1998), extensions that allow varying growth rates run the risk of robbing the concept of convergence of any real economic meaning. Instead of assuming complete heterogeneity or homogeneity, the panel structure model sets up a structure of an intermediate form: Countries with the same initial conditions converge to the same growth rate while countries with quite different initial conditions converge to different growth rates. Through the intermediate form, the model can capture convergence patterns with more flexibility while keeping the meaning of convergence intact.

We use the Bayes Information Criterion (BIC) to determine the number of clubs. According to BIC, models with two or three clubs outperform models with four or more clubs. The three-club model considered suggests that the world economy consists of three main groups: an advanced club, an underdeveloped club, and a developing club. Within each club, countries will converge to a steady state growth rate that is positively correlated with the investment rate, initial per capita income, and human capital endowment and negatively correlated with the population growth rate and ethnic diversity. The dependence of the long run growth rate on some of the classical growth determinants casts some doubt on the neoclassical growth theory. In addition, different clubs exhibit different convergence behavior.

First, the speeds of convergence are different. The developing club converges twice as fast as the underdeveloped and advanced clubs. Even for the latter clubs, the speeds of convergence are around 20% per year, which is much larger than previously thought.

Nevertheless, our study focuses on the convergence of the growth rates of per capita income while most of the previous studies concentrated on that of the levels of per capita income.

Second, and more importantly, the long run growth rates to which each club converges are different. The long run growth rates for the three clubs are 2.09%, 0.27%, and 2.90%, respectively. The difference in the long run growth rates implies that some countries will catch up and even forge ahead. But the process of catching up will never happen automatically, as some underdeveloped countries will fall further behind in terms of both the levels and growth rates of per capita income. It is envisaged that catching up, forging ahead and falling behind happen simultaneously as individual economies interact with each other.

A number of papers have attempted to find structures in cross country growth data. For example, Durlauf and Johnson (1995) employed a regression tree approach to uncover multiple regimes. As a second example, Canova (1999) used the predictive density of the data to identify convergence clubs. The present paper differs from the existing literature in the following ways.

First, it explicitly models the club structure by allowing the club membership to depend on some country-specific characteristics. These characteristics may be interpreted as a measure of "development." For example, if one believes that democracy causally affects growth (Barro (1996)), then a democracy variable may be introduced as a "development index."

Second, historical information can be easily incorporated into the model. As argued by Durlauf (2000), historical work can provide strong priors for issues such as variable selection and grouping of countries. In the present context, if historical information strongly supports that some countries, such as some East Asian countries or the G7 countries, should belong to the same club, we can model this simply by degenerating the membership probabilities (details are provided in the next section). In this way, historical information and current data can be combined in a coherent manner.

Finally, the panel structure model is statistically appealing. Sun (2001) established the asymptotic normality of the maximum likelihood estimator of model parameters. Based on this asymptotic result, hypothesis testing and statistical inference are straightforward. In contrast, no asymptotic theory is available for the regression tree model. In addition, the panel structure approach is computationally appealing and less demanding. The widely used and convenient EM algorithm is used to estimate the parameters. In contrast, both the regression tree approach and the predictive density approach involve complicated sorting algorithms.

The rest of the paper is organized as follows. Section 2 presents a simple model that will generate multiple steady states and outlines the empirical model that we will estimate. Sections 3 and 4 describe the estimation procedure and the data employed. Section 5 presents the empirical results. Section 6 concludes.

2 Model Description

In this section, we present a simple model that produces multiple balanced growth paths. We also specify the linear models that characterize growth rate dynamics and the logistic regression that classifies these linear models.

2.1 Human Capital Accumulation with Threshold Externalities

We consider a simple overlapping generation (OLG) model with human capital accumulation[1]. Suppose a single individual is born every period. The individual lives for two periods, so that in each period there are two individuals alive, an old generation and a young generation. The young generation inherits the human capital accumulated by the old generation, i.e.

$$h_{1,t} = h_{2,t-1}. \tag{1}$$

For the young generation, human capital accumulates according to

$$h_{2,t} = h_{1,t}[1 + \gamma(\tau_{t-1})\tau_t^\theta], \tag{2}$$

where θ is a parameter that characterizes the training technology, $\tau_t \in [0, 1)$ is the fraction of time invested in education while young for an individual born at period t, and $\gamma(\tau_{t-1})$ is a step function defined as

$$\gamma(\tau_{t-1}) = \begin{cases} \underline{\gamma}, & \text{if } \tau_{t-1} \leq \tau^* \\ \overline{\gamma}, & \text{if } \tau_{t-1} > \tau^* \end{cases}, \tag{3}$$

where $0 < \tau^* < 1$ and $\underline{\gamma} << \overline{\gamma}$.

To simplify the argument to the extreme, we assume that the economy has a linear production function: $y = h\tau$, where h is the human capital and τ is the fraction of time allocated to production. An individual born at time t will choose τ_t to maximize discounted lifetime income:

$$\max(1 - \tau_t)h_{1,t} + \rho h_{1,t}[1 + \gamma(\tau_{t-1})\tau_t^\theta], \tag{4}$$

where ρ is the discount factor. The optimal allocation scheme is

$$\tau = \begin{cases} (\rho\theta\underline{\gamma})^{1/(1-\theta)}, & \text{if } \tau_{t-1} \leq \tau^*, \\ (\rho\theta\overline{\gamma})^{1/(1-\theta)}, & \text{if } \tau_{t-1} > \tau^*. \end{cases} \tag{5}$$

If the previous generation does not invest enough in education such that $\tau_{t-1} \leq \tau^*$, the next generation will not have much incentive to invest in education. In particular, if $(\rho\theta\underline{\gamma})^{1/(1-\theta)} \leq \tau^*$ and $\tau_0 \leq \tau^*$, then all the generations invest $\tau_t = (\rho\theta\underline{\gamma})^{1/(1-\theta)}$, $t \geq 1$, of their time to enhance labor quality. As a consequence, the economy is trapped on a low growth path with a growth rate given by

$$\underline{\gamma}(\rho\theta\underline{\gamma})^{\theta/(1-\theta)}. \tag{6}$$

[1] The model is based on Azariadis and Drazon (1990) and Aghion and Howitt (1998, Ch. 10).

In contrast, if the previous generation invests enough such that $\tau_{t-1} > \tau^*$, the next generation should invest $\tau = (\rho\theta\overline{\gamma})^{1/(1-\theta)}$. If $(\rho\theta\overline{\gamma})^{1/(1-\theta)} > \tau^*$ and $\tau_0 \geq \tau^*$, then all the generations will invest $\tau_t = (\rho\theta\overline{\gamma})^{1/(1-\theta)}$, for $t \geq 1$. As a consequence, the economy grows at a faster rate given by

$$\overline{\gamma}(\rho\theta\overline{\gamma})^{\theta/(1-\theta)}. \tag{7}$$

When $(\rho\theta\underline{\gamma})^{1/(1-\theta)} < \tau^* < (\rho\theta\overline{\gamma})^{1/(1-\theta)}$, both slow and fast growth paths are possible. The presence of threshold externalities in education investment thus naturally leads to multiple growth paths. The mechanism is via positive externalities in that past education investment encourages future investment.

2.2 Growth Dynamics

The simple model outlined above exemplifies the importance of human capital accumulation in determining the long run growth rate. In particular, education investment, together with some other variables such as discount factor, determines the balanced growth path that a country will follow or the club that a country will join in. We now describe the growth rate dynamics associated with each club.

Let $y_{i,t\tau}$ be the per capita income at time $t\tau$, $q_{i,t\tau} = \log y_{i,t\tau} - \log y_{i,(t-1)\tau}$. Let $\mu(i)$ be a function that maps countries into clubs, i.e. $\mu(i) = g$ if country i belongs to club g, for $g = 1, ..., G$, where G is the number of clubs. Then a partial adjustment model for the per capita growth may be given by

$$q_{i,t\tau} - q_{i,(t-1)\tau} = \kappa_{\mu(i)}[q^*_{\mu(i)} - q_{i,(t-1)\tau}] + \varepsilon_{\mu(i),t\tau}, \tag{8}$$

where $\kappa_{\mu(i)} = 1 - \exp(-\lambda_{\mu(i)}\tau)$, $\lambda_i > 0$ is the speed of adjustment (the speed of convergence); τ is the time interval; and $\varepsilon_{\mu(i),t\tau}$ is the error term. In the empirical growth literature, the time interval is usually assumed to be larger than one year in order to avoid short run fluctuations. Following this convention, we employ a time interval of five years. For ease of understanding, we can think $\tau = 1$ and the time unit is a five-year period.

Model (8) allows inter-club heterogeneity in both the speed of convergence and the steady state growth rate. The heterogeneity in the speed of convergence may be attributed to the differences in stages of development. It is reasonable to believe that the speed of convergence depends on how far a club is away from its steady state. This is supported by the studies of Evans (1997) and Lee, Pesaran and Smith (1997), who found pervasive heterogeneity in the speed of convergence.

The possibility of heterogeneous steady state growth has been demonstrated by our simple model. It may also arise from obstacles to technology transfer. The neoclassical assumption that technology is a public good so that technological progress happens at the same rate across different countries is inconsistent with the evidence on growth and development. For many less-developed countries, the process of technology transfer is hindered because of a lack of skilled labor or inferior institutional arrangements. Other obstacles include industry specialization. Less-developed countries tend to specialize in traditional production activities that involve little technological progress. These countries, as a group, tend to grow slower.

Model (8) is not tied to a specific growth model. As long as a model predicts that per capita income grows at a constant rate in the steady state, per capita growth rates should converge. The model is thus broadly consistent with any growth model that admits one or more balanced growth paths — a category that includes virtually all the growth models in the literature. To illustrate this point, consider a neoclassical growth model featuring a Cobb-Douglas production with labor-augmenting technological progress: $Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}$, where Y is output, K is capital, and L is labor. L and A are assumed to grow exogenously at rates n and q^* so that

$$L_t = L_0 e^{nt},\ A_t = A_0 e^{q^* t}. \tag{9}$$

Define output and stock of capital per unit of effective labor as $\hat{y}_t = Y_t/(A_t L_t)$ and $\hat{k}_t = K_t/(A_t L_t)$ respectively. The production function can be represented by the function $\hat{y}_t = \hat{k}_t^a$, and the evolution of the capital stock obeys

$$d\hat{k}_t/dt = s\hat{k}_t^a - (n + q^* + \delta)\hat{k}_t, \tag{10}$$

where s and δ are the savings rate and the depreciation rate, respectively. The dynamics of capital stock can be understood further by taking a first order Taylor expansion in $\log(\hat{k}_t)$ about steady state $\hat{k}^*$,

$$\frac{d}{dt}\log\hat{k}_t = -\lambda(\log(\hat{k}_t) - \log(\hat{k}^*)), \tag{11}$$

where $\lambda = (1-\alpha)(\delta + n + q^*) > 0$. This differential equation implies that

$$q_t = \frac{d}{dt}\log y_t = q^* - \lambda[\log y_0 - \left(\log A_0 + \frac{a}{a-1}\log(\frac{n+q^*+\delta}{s})\right)]\exp(-\lambda t). \tag{12}$$

So

$$\frac{dq_t}{dt} = -\lambda(q_t - q^*), \tag{13}$$

which in turn produces the empirical model (8) by discretization.

Model (8) focuses the dynamics of growth rate instead of the behavior of income level. Empirical studies in the same spirit include Jones (1995) who tested endogenous growth models using time series data, and Lee, Pesaran and Smith (1997) who studied the growth and convergence in a stochastic Solow model.

2.3 Logistic Regression

Model (8) can be rewritten as

$$q_{i,t\tau} = \alpha_{\mu(i)} + \beta_{\mu(i)} q_{i,(t-1)\tau} + \varepsilon_{\mu(i),t\tau}, \tag{14}$$

where $\alpha_{\mu(i)} = \kappa_{\mu(i)} q^*_{\mu(i)}$, $\beta_{\mu(i)} = 1 - \kappa_{\mu(i)}$ and $\varepsilon_{\mu(i),t\tau}$ is *iid* distributed with variance $\sigma^2_{\mu(i)}$.

The regression coefficients $\alpha_{\mu(i)}$ and $\beta_{\mu(i)}$, and the variance parameter $\sigma_{\mu(i)}$ take different values depending on the club membership. Let c_{ig} be the indicator of country $i's$ membership of club g, i.e. $c_{ig} = 1$ if $\mu(i) = g$. Then

$$\alpha'_{\mu(i)} = \alpha' c_i, \beta_{\mu(i)} = \beta' c_i, \sigma_{\mu(i)} = \sigma' c_i, \tag{15}$$

where $c_i = (c_{i1}, ..., c_{iG})'$, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_G)'$ $\beta = (\beta_1, \beta_2, ..., \beta_G)'$, $\sigma = (\sigma_1, \sigma_2, ..., \sigma_G)'$.

We model the distribution of c_i with a polychotomous logistic regression, which is allowed to depend on covariates specific to country i. To ensure no crossing between clubs over time, we assume these covariates are time invariant. Specifically, c_i is multinomially distributed with G categories. The probability π_{ig} that individual i belongs to club g is a function of $w_i = (w_{i1}, w_{i2}, .., w_{ik})'$ in a multinomial logistic regression:

$$\pi_{ig}(\xi) = P(c_{ig} = 1|w_i) = \frac{\exp(w_i'\xi_g)}{\sum_{j=1}^{G} \exp(w_i'\xi_j)}, \tag{16}$$

for $g = 1, 2, ..., G$. Here $\xi' = (\xi_1', ..., \xi_G')$ and ξ_g is a column vector containing the membership parameters for club g.

The first element of w_i is assumed to be one, which implies that the membership probabilities $\{\pi_{ig}\}_{i=1}^N$ are constant across countries if no covariate is included. Put another way, if we can not infer any membership information from any covariate, we assume that membership probabilities are the same for each country. Other potential variables in the vector w_i include:

(1) Human capital. The inclusion of human capital strengthens the viability of the club convergence hypothesis. As demonstrated by the model with threshold externalities, initial human capital endowment may dictate the ultimate fates of otherwise identical economies. The use of initial human capital to infer membership probabilities is also consistent with some of the ideas that underlie the endogenous growth theory, which notably emphasizes how difference in human capital can have long run growth consequences.

(2) Initial income. Durlauf and Johnson (1995) used initial income as a source of threshold in the growth process while Canova (1999) sorted countries according to initial income in order to test for breaks. The use of initial income attempts to capture the effect that initial income may have on the growth dynamics. This is consistent with various versions of the "poverty trap" theory according to which poor countries as a "club" converge towards poverty because they are all trapped. Barriers for getting out of this trap can be the lack of a "big push" (Murphy, Shleifer and Vishny (1989)).

(3) Ethnic diversity. This measures the likelihood that any two people chosen at random will be of different races or ethnicities. The justification for the use of ethnic diversity comes from Easterly and Levine (1997), who found an adverse effect of ethnolinguistic fractionalization on growth. They offered ethnic conflict as an important determinant of poor growth performance in sub-Saharan African economies. Brock and

Durlauf (2001) have recently examined the empirical evidence for their findings, using a Bayesian method that allows for model uncertainty.

Many other variables may also be relevant. For example, the incorporation of *distributions* of initial income and human capital may provide an environment in which the club convergence hypothesis is viable. This issue is left for future research.

Through the above covariates, the logistic regression (16) ensures that countries with similar characteristics belong to the same club with a large probability while countries with quite different characteristics belong to different clubs with large probabilities. From a modeling perspective, the covariates in the logistic regression control the heterogeneity not captured by the linear regression specification. The importance of the covariates lies in their capacities to determine the balanced growth path that a country will follow. We shall call these covariates development indices when there is no possibility of confusion.

If historical studies reveal that a set of countries $i \in S$ belongs to the same convergence club, we can, without loss of generality, assume that those countries belong to club g^*. In other words, we let

$$\pi_{ig} = I(g = g^*) \equiv \begin{cases} 1, & g = g^* \\ 0 & g \neq g^* \end{cases} \quad \text{for } i \in S. \tag{17}$$

If S is large enough, then all parameters specific to club g^* can be identified by using only the countries in S. Club g^* thus manifests itself through its members in S.

Equations (14), (15), (16), and (17) combine to define the panel structure model.

The panel structure model captures the common growth process through club memberships: countries in the same club follow the same growth dynamics. Conditioning on the club membership, the model assumes that the regressor errors are exchangeable in that the regression error for one country is equally likely to be the error for another country in the same club. In other words, once the club membership is realized, there exists no basis for distinguishing the regression errors in the same club. The panel structure model imposes the exchangeability assumption only within a club. In contrast, conventional empirical studies often assumed, explicitly or implicitly, that exchangeability holds for all the countries. In this case, the likelihood of a positive error for a given country needs to be the same for any other country in the sample. This is certainly a strong and restrictive assumption. To a large extent, the panel structure model avoids this untenable assumption. For more discussion on exchangeability and its implications in the context of growth convergence modeling, see Brock and Durlauf (2001).

3 Estimation Procedure

To classify countries into different clubs, we first estimate the model parameters, including the membership parameters in the logistic regression and the regression parameters in the linear regressions. Sun (2001) proposed an ML estimator and established its asymptotic normality.

3.1 ML Estimate

Let $q_i = (q_{i,\tau}, q_{i,2\tau}, ..., q_{i,T\tau})'$. Then conditioning on $\{q_{i0}\}_{i=1}^N, \{w_i\}_{i=1}^N$, the log-likelihood of the panel structure model can be expressed as

$$L(\theta, \xi|q) = \sum_{i=1}^{N} \log \sum_{g=1}^{G} \pi_{ig}(\xi) m_{ig}(q_i; \theta), \tag{18}$$

where

$$m_{ig}(q_i; \theta) = \left(2\pi\sigma_g^2\right)^{-T/2} \exp\left(-\sum_{t=1}^{T} \frac{(q_{i,t\tau} - \alpha_g - \beta_g q_{i,(t-1)\tau})^2}{2\sigma_g^2}\right), \tag{19}$$

and $\theta' = (\theta_1', \theta_2', ..., \theta_G')$, $\theta_g' = (\alpha_g', \beta_g', \sigma_g')$.

Here we assume that $\varepsilon_i = (\varepsilon_{i,\tau}, \varepsilon_{i,2\tau}, ..., \varepsilon_{i,T\tau})'$ is a normal vector after controlling for the club membership. Under this assumption, q_i follows a mixture distribution with multivariate normal components. In view of the fact that any continuous distribution can be approximated arbitrarily well by large enough normal mixtures, the normality assumption is not as restrictive as it seems. Nevertheless, it is worthwhile to relax the distributional assumption even if this may invalidate the convenient EM algorithm below.

To estimate the parameters, which include θ and ξ, we maximize the likelihood function:

$$(\hat{\theta}, \hat{\xi}) = \arg\max \sum_{i=1}^{N} \log \sum_{g=1}^{G} \pi_{ig}(\xi) m_{ig}(q_i; \theta). \tag{20}$$

Given the estimate $(\hat{\theta}, \hat{\xi})$, we can assign the memberships as follows: country i belongs to club g if

$$\pi_{ig}(\hat{\xi}) m_{ig}(y_i; \hat{\theta}) = \max_j \pi_{ij}(\hat{\xi}) m_{ij}(y_i; \hat{\theta}). \tag{21}$$

In other words, we assign a given country to the club of which it is most likely to be a member based on the posterior probabilities. We can also assign the membership on the basis of estimated prior probabilities, i.e. country i belongs to club g if

$$\pi_{ig}(\hat{\xi}) = \max_j \pi_{ij}(\hat{\xi}). \tag{22}$$

Since the posterior probabilities are close to the prior probabilities (Sun (2001)), the assigned membership is not very sensitive to which probabilities are used. We will assign the membership according to the posterior probabilities in what follows.

3.2 The EM Algorithm

To search for the maximizers of the likelihood function, we advocate the so-called Expectation Maximization (EM) algorithm (Dempster, Laird and Rubin (1977). The EM algorithm is a general technique for maximum likelihood estimation in a wide variety of situations best described as the incomplete data problem. The recent monograph by Mclachlan and Krishman (1996) provides an excellent introduction to the EM algorithm.

An application of the EM algorithm generally begins with the observation that the optimization of the likelihood function would be simplified if a set of missing variables or hidden variables were known. In our context, if the club membership c_i is observable, then the log-likelihood for $\{q_i\}$ and $\{c_i\}$ becomes

$$L(\theta, \xi | q, c) = \sum_{i=1}^{N} \sum_{g=1}^{G} c_{ig} \left\{ \log \pi_{ig}(\xi) + \log(m_{ig}(q_i; \theta)) \right\}. \tag{23}$$

The use of the indicator variable c_{ig} has allowed the logarithm to be brought inside the summation sign, substantially simplifying the maximization problem. But c_{ig} is not observable. Instead of maximizing $L(\theta, \xi|q, c)$ itself, we maximize its expectation, where the expectation is taken with respect to all the unobserved c_{ig}. Let $((\theta^{(k)})', (\xi^{(k)})')$ be the current estimate of (θ', ξ'), and $((\theta^{(k+1)})', (\xi^{(k+1)})')$ stand for the updated estimate. The EM algorithm consists of the following iterative steps (for more details, see Sun (2001)):

The E-step: The (conditional) expectation is given by

$$\begin{aligned} Q(\theta, \xi | \theta^{(k)}, \xi^{(k)}) &= E\left(L(\theta, \xi|q, c) | \theta^{(k)}, \xi^{(k)} \right) \\ &= \sum_{i=1}^{N} \sum_{g=1}^{G} p_{ig}^{(k)} \left\{ \log \pi_{ig}(\xi) + \log(m_{ig}(q_i; \theta)) \right\}, \end{aligned} \tag{24}$$

where

$$\begin{aligned} p_{ig}^{(k)} &= E(c_{ig}|\theta^{(k)}, \xi^{(k)}) = P(c_{ig} = 1|\theta^{(k)}, \xi^{(k)}) \\ &= \frac{\pi_{ig}(\xi^{(k)}) m_{ig}(q_i; \theta^{(k)})}{\sum_{j=1}^{G} \pi_{ij}(\xi^{(k)}) m_{ij}(q_i; \theta^{(k)})}. \end{aligned} \tag{25}$$

If historical information reveals that for some country $i_0, \pi_{i_0 g} = I\{g = g^*\}$, then $\pi_{i_0 g}(\xi^{(k)}) = I\{g = g^*\}$ for all k. As a consequence,

$$p_{i_0 g}^{(k)} = I\{g = g^*\} \text{ for all } k. \tag{26}$$

The M-step: To get the updated estimate $((\theta^{(k+1)})', (\xi^{(k+1)})')$, we maximize $Q(\theta, \xi|\theta^{(k)}, \xi^{(k)})$ with respect to (θ', ξ'). By inspection, the regression parameters affect Q only through the term $\sum_{i=1}^{N} \sum_{g=1}^{G} p_{ig}^{(k)} \log(m_{ig}(q_i; \theta))$ and the membership parameters affect Q only through $\sum_{i=1}^{N} \sum_{g=1}^{G} p_{ig}^{(k)} \log \pi_{ig}(\xi)$. Therefore, we can maximize the two parts in $Q(\theta, \xi|\theta^{(k)}, \xi^{(k)})$ separately. Following an argument similar to Sun (2001), we have

$$\xi^{(k+1)} = \arg\max \sum_{i=1}^{N} \sum_{g=1}^{G} p_{ig}^{(k)} \log \pi_{ig}(\xi), \tag{27}$$

$$\left(\alpha_g^{(k+1)}, \beta_g^{(k+1)}\right)' = \left(\sum_{i=1}^{N} p_{ig}^{(k)} \Gamma_i' \Gamma_i\right)^{-1} \left(\sum_{i=1}^{N} p_{ig}^{(k)} \Gamma_i' q_i\right), \tag{28}$$

and

$$(\sigma_g^2)^{(k+1)} = \frac{\sum_{i=1}^{N} p_{ig}^{(k)} \hat{u}_{ig}' \hat{u}_{ig}}{T \sum_{i=1}^{N} p_{ig}^{(k)}}, \tag{29}$$

where $\hat{u}_{ig} = (\hat{u}_{ig,\tau}, ..., \hat{u}_{ig,T\tau})'$, $\hat{u}_{ig,t\tau} = q_{i,t\tau} - \hat{\alpha}_g^{(k+1)} - \hat{\beta}_g^{(k+1)} q_{i,(t-1)\tau}$, $\Gamma_i = (1_T, x_i)$, $1_T = (1, 1, ..., 1)'$, and $x_i = (q_{i,0}, q_{i,\tau}, ..., q_{i,(T-1)\tau})'$.

The iteration calls for an initial estimate of (θ', ξ'). To initialize, we set $\xi_g^{(0)} = 0$, so that $\pi_{ig} = 1/G$ for $g = 1, 2, ..., G$, and randomly assign countries to G clubs if no prior information is available. That is, for each country i, we randomly generate an integer between 1 and G. If this random integer equals g, then we assign country i to club g. If historical analysis shows that certain countries belong to the same club, then these countries are assigned to the same club. With this assignment, we can obtain the initial estimate $((\theta_g^{(0)})', (\xi_g^{(0)})')$. The EM algorithm can then start from this initial value.

3.3 Model Selection

There are two types of uncertainties in the panel structure model. First, there is theory uncertainty. In particular, we assume that there is a set of possible development indices to include in the logistic regression, but we do not know which combination of these indices best describes the cross country growth data. Second, there is club uncertainty, i.e. there is uncertainty as to the number of clubs.

To deal with these uncertainties, we employ a Bayesian model selection technique, which is based on Bayes factors and posterior model probabilities. The basic idea is that if several model specifications $M_1, ..., M_J$ are considered, with prior probabilities $p(M_j)$, $j = 1, .., J$ (often assumed to be equal), then by Bayes theorem the posterior probability $p(M_j|D)$ of model M_j, given data D, is proportional to $p(D|M_j)p(M_j)$, namely,

$$p(M_j|D) \propto p(D|M_j)p(M_j), \tag{30}$$

where $p(D|M_j)$ is the probability of the data given model M_j. A natural model selection criterion is to choose the model with the largest posterior probability. For regular models with a large sample size, this amounts to choosing the model with the smallest BIC value.

In our context, for any given combination of indices and number of clubs, we compute the BIC value as follows

$$-2 \sum_{i=1}^{N} \log \sum_{g=1}^{G} \pi_{ig}(\xi) m_{ig}(q_i; \theta) + (2G + (G-1)k) \log(NT), \tag{31}$$

where k is the number of variables in w_i. We select the model that produces the smallest BIC value.

4 Data Description

To investigate the hypothesis of convergence clubs, we use a balanced panel data set for 85 countries averaged over five years, with beginning years 1960, 1965, 1970, 1975, 1980, and 1985. We do not use annual data because they are likely affected by short run factors. It is therefore difficult to recover long run dynamics from high frequency data. On the other hand, data availability rules out the possibility of using low frequency data. Taking these into account, we choose to employ a five-year interval, which is also the time span used by Islam (1995), among others.

The variables employed in our analysis are as follows: (i) q, the change of the log of income per capita; (ii) n, the average growth rate of the population; (iii) s, the logarithm of the ratio of real investments to real GDP; (iv) y_0, the logarithm of per capita income in 1960; (v) h, the logarithm of initial average years of secondary schooling in the total population in 1960; (vi) e, the ethnic fractionalization. Initial income, population rate, and index of ethnicity fractionalization were obtained from the Global Development Network Growth Database developed by Easterly and Yu (2000). The average proportion of real investments was constructed from the Penn World Table (Summers-Heston data set mark 5.0) while the human capital variable was obtained from Barro and Lee (2000).

Figure 2 presents kernel density estimates for the above variables. The Epanechnikov kernel, $k(x) = 3/4(1 - x^2)I(|x| < 1)$ was used and the bandwidth was based on Silverman's rule of thumb. The graphs in Figure 2 demonstrate clearly the bimodality of ethnic diversity, initial income, human capital, population growth and investment rate. However, growth rates appear to be unimodal. As is well-known, the unimodality does not rule out the existence of a club structure. The coupling of the similarity among clubs with the variation within each club tends to conceal the underlying structure.

Figure 3 presents the distribution of relative (per capita) income for each five-year interval. Relative income is income normalized by the world average. If all countries grow at more or less the same rate, the distribution of relative income should not change with time. However, Figure 3 shows a discernible pattern: the distribution exhibits twin peakedness as time passes. This twin peakedness becomes more apparent as the second mode becomes more pronounced. There is also a visible tendency for the two peaks to move apart, with the first mode moving to the left and the second one moving to the right. This tendency reveals that cross country income disparity has become larger rather than smaller as predicted by absolute convergence. A natural question is, is the tendency a permanent phenomenon? That is, will countries grow at different rates forever? We answer this question in the next section.

5 Empirical Results

Given the development indices and the number of clubs, we estimate the panel structure model using the EM algorithm. We will identify the model specification with the development indices that it uses. For example, $m(c, h, y_0)$ is the specification that uses initial human capital and income as the development indices in the logistic regression. BIC values corresponding to each model specification are listed in Table 2. The one-club model is inferior to those with more clubs. Furthermore, models with two or three clubs appear to be the best according to BIC. For specification $m(c, h, y_0)$, the three-club model achieves the smallest BIC value. For other cases, two-club models dominate the corresponding three-club models. Models with more clubs ($G = 4, 5$) yield smaller BIC values. This is because BIC penalizes models with more parameters.

For two-club models, $m(c, y_0)$ attains the minimum BIC value whereas for three-club models, $m(c, h, y_0)$ attains the minimum BIC value. We will focus on the two-club $m(c, y_0)$ model and three-club $m(c, h, y_0)$ model below.

5.1 Two-club Model

Table 3 presents the classification result for the two-club model. The first club includes most of the industrialized countries while the second club consists of the rest of the countries.

Table 4 shows that for the first club, the steady state growth rate is 1.90% while that for the second club is 1.45%. The difference in long run growth will lead to the stratification of countries. This is consistent with the hypothesis of emerging twin peaks, in which rich countries become richer while poor countries become poorer.

Table 4 reports that for the first club, the convergence rate is 14.96% per year while that for the second club is 24.85% per year. The convergence of each club happens at a rate faster than that found by conventional studies. For example, in one of the most influential studies using cross-sectional growth data, Mankiw, Romer, and Weil (1992) found that the rate of convergence is about 2% per year, which is also the rate suggested by Barro and Sala-i-Martin (1995). It should be mentioned that our study focuses on the convergence of growth rates of per capita income while most of the previous studies concentrated on that of the levels of per capita income. However, the faster rate of convergence is consistent with more recent studies. For example, Islam (1995) employed a panel data approach to control level effects and found a higher rate of conditional convergence, in which the convergence rate for the OECD countries is 10.67% per year. More recently, by allowing complete heterogeneity in both level effects and steady state growth rates, Lee, Pesaran and Smith (1997) found the speed of convergence for their full sample to be 30% per year. Nevertheless, the faster speed of convergence is conditional on the club membership. Only countries in the same club converge.

The convergence path within each club is not deterministic. The randomness of the convergence path is determined by the standard deviation of the error term. For the first club the standard deviation is 1.31% per year while that for the second club is 3.77% per year. Therefore, even if the underlying economies are in steady states, growth rates within the first club show less variation than those within the second club.

Figure 4 reports the histogram of membership probabilities. For a two club model, it suffices to know the probabilities that countries belong to one club, say the first club. The figure shows that these probabilities are close to either zero or one, meaning that in most cases we can classify countries with considerable confidence. Figure 5 presents the club structure together with the initial income. From this figure, one may argue that the club structure can be obtained by simply dividing countries into two groups based on the initial income. The growth behavior thus does not play a role in determining the club structure. But this is not true, as a similar club structure can be obtained without using any covariate. That is, a two-club $m(c)$ model produces more or less the same club structure. Therefore, the growth behavior of the first club is indeed fundamentally different from that of the second club.

Table 5 characterizes the convergence clubs through both growth determinants in a prototypical neoclassical model and covariates in the logistic regression. For a given variable v_i, the weighted averages in Table 5 are calculated using the formula $\sum \hat{\pi}_{ig} v_i / \sum \hat{\pi}_{ig}$ with estimated probabilities of club memberships $\hat{\pi}_{ig}$ as weights. The simple averages in Table 5 are arithmetic means of variables within each club. It is evident from Table 5 that the weighted averages are very close to the simple averages. This is because the estimated probabilities are close to either zero or one, as shown in Figure 4.

Table 5 reveals some interesting club features. Compared with the second club, the first club has a smaller ethnic index, a smaller population growth rate, a higher investment rate, a higher level of initial income, a larger endowment in human capital, and a faster growth rate. It seems the long run growth rate is affected by the growth determinants in a prototypical neoclassical growth model. The evidence thus casts some doubt on the neoclassical growth theory, as the theory predicts that the growth determinants affect only the long run level of per capita output but not its growth rate. This finding is consistent with that of Bernanke and Gurkaynak (2001), who found that the rates of investment and population growth are correlated with the estimated TFP growth. On the other hand, the evidence lends some support to the endogenous growth theory which emphasizes human capital as an accumulated factor that drives long run growth.

5.2 Three-club Model

We now turn to the three-club model. Tables 6 and 7 present the classification and estimation results. Figure 6 illustrates the club structure geographically.

To a great extent, the first club coincides with that in the above two-club model. The composition of the first club is thus not sensitive to model specifications. This insensitivity strengthens the belief that members of the first club do belong to the same club. The members are all advanced industrialized countries, which are committed to a market economy and a pluralistic democracy. For convenience, we shall label this club the "advanced" club.

As demonstrated in Figure 6, the second club contains mainly countries from Sub-Saharan Africa or Latin America. The geographic concentration is not surprising. On one hand, geography may affect growth, as international trade, investment, and more importantly, the spread of technology may be limited by distance. Modern advances in telecommunications and information technology may alleviate the geographical constraint. On the other hand, Gallup, Sachs and Mellinger (1998) suggested that location and climate, through their impacts on transportation costs, the burden of disease and agricultural productivity, have significant effects on income level and growth. As a result, adjacent countries or regions may grow or stagnate in tandem. We shall label the second club the "underdeveloped" club.

The third club is less homogeneous geographically. It contains countries from Asia, the Middle East, North Africa, Sub-Saharan Africa, Latin America, and Europe. Interestingly, almost all the Asian countries belong to the third club. In particular, Hong Kong and Taiwan, two of the "little tigers of Asia," belong to this club. In general, the third club countries grow very fast. We shall label this club the "developing" club.

Figure 7 is the scatterplot of membership probabilities. For a three club model, it suffices to know the probabilities that countries belong to any two clubs, say the first and second clubs. Recall that countries with membership probabilities close to (0,1), (1,0) and (0,0) are assigned to the first, second and third clubs respectively. Figure 7 shows that most countries cluster around these three points. As in the two-club model, the classification is thus unambiguous for most countries.

Figure 8 graphs initial human capital and per capita income together with the club structure. Not surprisingly, the first club has higher initial income and human capital endowment. Interestingly, a discernible line appears to divide the second and third clubs. Neither initial income nor human capital endowment is sufficient to differentiate these two clubs. It is the ratio of human capital to initial income that constitutes the distinction between these two clubs. Countries with high initial income but relatively low human capital tend to grow at unsatisfactory rate. Such countries include Argentina and Venezuela, two wealthy countries that experienced reversals of economic growth.

Table 7 shows that the steady state growth rates for the three clubs are $2.09\%, 0.27\%$, and 2.90%, respectively. The underdeveloped club will thus lag behind the advanced and developing clubs in the long run. A tendency for poor economies to catch up with rich ones has not happened and is not going to happen without serious changes in economic policies for the second club. Due to the difference in long run growth, the gap between income per capita will grow even wider, unless the future is different in important ways from the recent past.

Table 7 reveals that the long run growth of the developing club is faster than that of the advanced club. But this does not imply that all countries in the developing club will overtake the countries in the advanced club because of the randomness of the convergence path. It does imply that some countries in the developing club will catch up and probably overtake certain countries in the advanced club. Nevertheless, we do not know which countries will catch up and which countries will be overtaken. Considering the broad geographical area of the developing club, we believe that the process of catching up will not be limited to countries in any particular continent.

As in the two-club model, Table 7 reveals that the speeds of convergence differ across different clubs. The underdeveloped countries converge to their steady states twice as fast as advanced and developing countries do. The randomness of the convergence path, measured by the standard deviation of the error term is also different. For advanced countries, the path to balanced growth is much steadier than that for developing and underdeveloped countries.

Table 8 displays the characteristics of each club. The qualitative observations for the above two-club model apply. Specifically, the more developed club has a smaller ethnic index, a smaller population growth rate, a higher investment rate, a higher level of initial income, a larger endowment in human capital, and a faster growth rate than the less developed club. The degree of ethnic fractionalization, though not a critical force in determining club structure, is negatively correlated with the long run growth rate. This lends some support to the argument of Easterly and Levine (1997).

Tables 7 and 8 show that for each club the long run growth rate is less than the average growth rate. Therefore, all the clubs converge downward, which is characterized by fast growing members slowing down. It is not surprising that the growth of some advanced and developing countries will slow down eventually. What is striking is that some underdeveloped countries will follow the same pace, and that their already slow growth will become even slower in the long run. On this evidence, therefore, it appears untenable to treat the state of the world's rich nations as a condition that is potentially available to all.

5.3 Crossing-over and Misclassification

In recent years there has been renewed interest in in-depth analysis of the growth performance of individual countries. In light of this interest, it may be worthwhile to note certain "unexpected" or interesting aspects of the classification results.

First, the composition of clubs in the two and three club models reveals that Israel and Ireland are likely to have crossed over. Israel and Ireland are in the third club in the three-club model. Yet when we have two clubs with the same development indices[2], Israel and Ireland are in the first club. Although both Israel and Ireland are small nations and have been beset by generations of war and strife, they are enjoying surges of economic success, particularly in the high technology area. Both Israel and Ireland are emerging as global high technology leaders. Relative to its population, Israel produces more technology than any other country in the world. Its historically strong commitment to education has created a highly qualified work force. Similarly, a good education system has helped Ireland become one of the world's strongest producers of software and internet technology.

Second, Japan is in the second club in the two-club model while it belongs to the third club in the three-club model. Japan's economic growth showed a downtrend after the first Oil Crisis of 1973. Nevertheless, the annual average growth rates for the two periods, namely, 1960-1973 and 1973-1990, were still very high. These rates were 9.3% and 4.0% per year respectively. For the sample period, namely, 1960-1990, it is reasonable to believe that Japan belongs to the advanced club. Therefore, it is of interest to see the

[2] As in the two-club model using initial income as the development index, Israel and Ireland are in the first club in the two-club model using both initial income and human capital as development indices.

classification result when such a prior belief is imposed.

We consider two possible priors for the three-club model. First, we impose the prior that Japan and the US belong to the same group. In this case, South Africa, Israel, Greece, Ireland, and Italy join Japan in moving to the advanced club. Second, we impose the prior that G7 countries belong to the same group. In this case, Colombia, South Africa, Israel, Greece, Ireland, and Italy join Japan in moving to the first club[3]. For both cases, the club memberships remain the same for the rest of the countries. In addition, the estimation results remain more or less the same, and all the qualitative observations in the previous sub-section remain valid.

Finally, countries such as Spain are clearly misclassified. In the three-club model, Spain is in the underdeveloped club. However, all the variables indicate that Spain is closer to the first and third clubs than to the second club. Further investigation reveals some irregularity in Spain's growth dynamics. For the period from 1960 to 1975, the average growth rate of per capita GDP was very high, namely, 5.6% per year. This strong growth process stopped for the next decade. From 1975 to 1985, Spain's per capita GDP hardly improved. However, from 1985 to 1990, Spain regained its growth rate with an annual average of 4.8%. This irregularity contributes to the misclassification.

5.4 Missing data

Due to the problem of missing data, our sample consists of only 85 countries. It is certainly of interest to classify as many countries as possible. Given the estimates $\hat{\xi}$ and $\hat{\theta}$, we can classify countries that have no missing data in between. In fact, we can include these observations in the estimation stage. In this case, the time dimension varies across different countries. When the time series is too short, noises may dominate signals. As a consequence, the estimation and classification may be contaminated. In addition, we may encounter the problem of an unbounded likelihood function (see Sun (2001)), which renders the EM algorithm problematic. Taking these into account, we choose not to incorporate these observations in the estimation. On the other hand, scientific classification usually starts from grouping unequivocal subjects first and then classifies ambiguous subjects according to their similarities to the existing groups. We take this approach and expand the convergence clubs in the three-club model as much as possible.

When both the initial income and the initial human capital endowment are available, we compute the estimated prior probability according to

$$\pi_{ig} = \frac{\exp(w_i'\hat{\xi}_g)}{\sum_{j=1}^{G}\exp(w_i'\hat{\xi}_j)}, g = 1, 2, ..., G. \tag{32}$$

When either the initial income or the initial human capital endowment contains a missing value, we impose a prior probability that is proportional to the size of the convergence

[3] The membership changes for South Afria and Columbia are unexpected. This may be regarded as misclassification, as the probabilities of these two countries belonging to the first club are very close to 0.50.

club. In other words,

$$\pi_{ig} = \frac{s_g}{\sum_{j=1}^{G} s_j}, g = 1, 2, ..., G, \tag{33}$$

where s_g is the number of countries in the g-th convergence club in the previous subsection. The posterior probability is then calculated according to

$$p_{ig} = \frac{\pi_{ig} m_{ig}(q_i; \hat{\theta})}{\sum_{j=1}^{G} \pi_{ij} m_{ij}(q_i; \hat{\theta})},\ g = 1, 2, ..., G, \tag{34}$$

where $m_{ig}(q_i; \hat{\theta})$ is the likelihood function for the growth process. We classify country i into club g^* if $p_{ig^*} > 0.5$.

Table 9 presents the additional classification result, which expands the convergence clubs in Table 6. While the classification in Table 6 may be regarded as an in-sample classification, the classification in Table 9 is an out of sample classification. Figure 9 illustrates the expanded clubs geographically. As expected, Luxembourg is in the advanced club. Luxembourg is a stable, high-income economy featuring moderate growth, low inflation, and low unemployment. Asian countries such as China and Korea are in the developing club. Both China and South Korea had a low initial per capita income and enjoyed surges of economic success for most of the sample period. As in Table 6, some African countries are in the developing club while others are in the underdeveloped club. Tables 6 and 9 reveal that Venezuela, Madagascar, Mali and Chad, which experienced negative growth, are in the underdeveloped club. These countries are labeled "growth disasters" by Jones (1998, Ch.1, page 4). In contrast, newly industrialized regions or countries such as Hong Kong, Taiwan and South Korea, are in the developing club. These countries experienced astounding growth and exemplified what is meant by the term "growth miracle."

6 Conclusion

The primary purpose of this paper is to investigate the club convergence hypothesis using the panel structure approach. This novel approach provides a natural setting for detecting structures in panel data. Our findings suggest that the world economy constitutes three convergence clubs: an advanced club, an underdeveloped club and a developing club. Different convergence clubs exhibit different convergence behavior in terms of their speeds of convergence and their long run growth rates. It is envisaged from these differences that catching up, forging ahead, and falling behind happen simultaneously as individual economies interact with each other.

The present study can be extended in several ways and we briefly discuss some of the possibilities in what follows.

First, we may relax the assumption that covariates in the logistic regression are time invariant. With this relaxation, countries may switch memberships over time. A Markov Chain may be constructed to describe the membership dynamics. This dynamics may shed some light on which countries will catch up and even forge ahead and which countries will be overtaken.

Second, we may investigate the significance of other variables in determining the club structure. In particular, we may investigate the roles of income and education distributions in economic growth. This is important as the club structure may be related to the underlying economic structure. Some countries are trapped not only because of backward technology but also because of detrimental social characteristics. To a large extent, the social characteristics depend on the distributions of social and economic variables rather than the levels of these variables.

Finally, we may investigate the dynamics of per capita income using the panel structure approach. Presumedly, the dynamics of income levels is more heterogenous than that of growth rates. On one hand, we may have to relax the assumption of common fixed coefficients within each club by allowing for random coefficients (see Sun (2001)). This extension takes intra-club heterogeneity into account and may provide a more realistic model. On the other hand, we may have to allow for complete heterogeneity in level effects. A new estimation strategy is then desired to overcome the underlying problem of incidental parameters.

Figure 1: Balanced Growth Paths

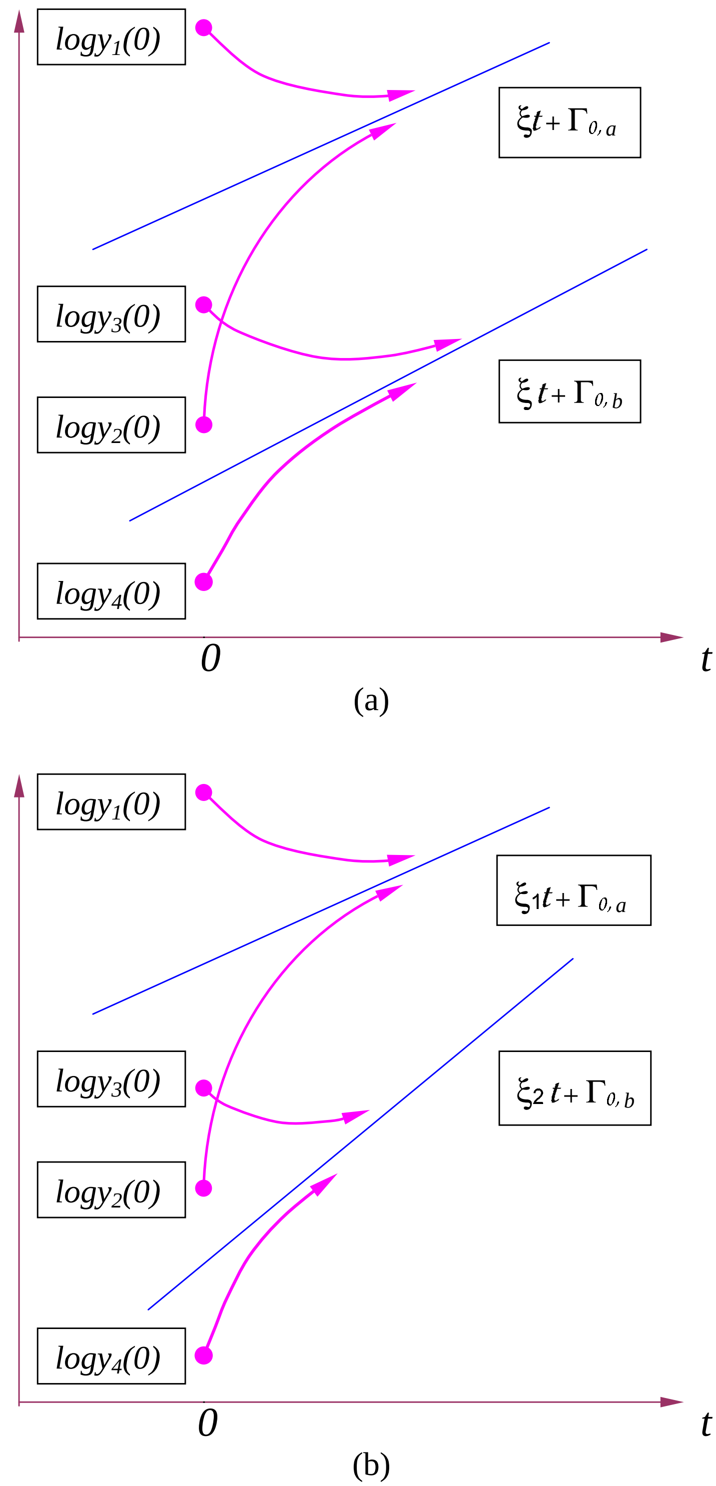


Figure 2: Density Estimates

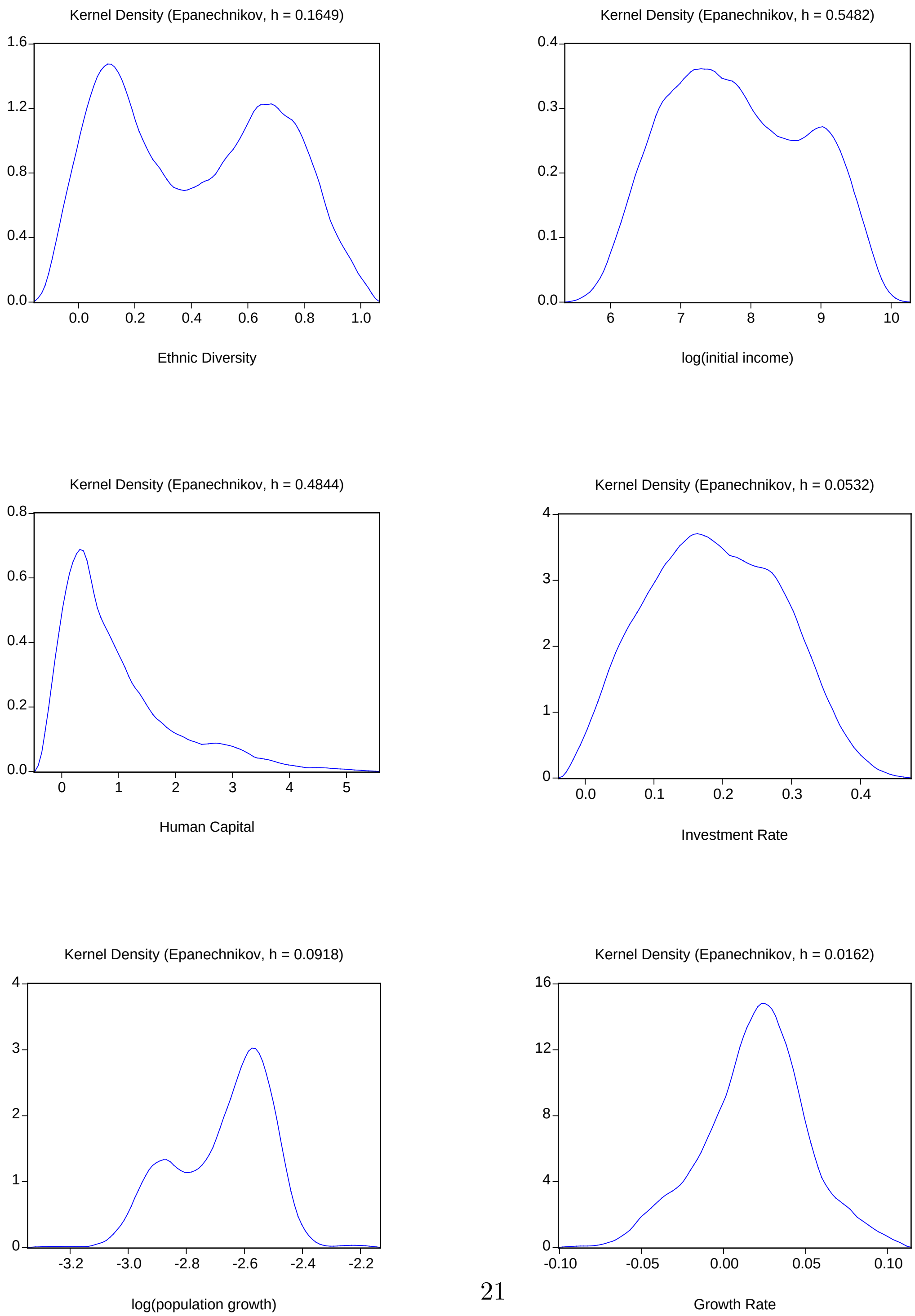


Figure 3: Comparative Statics of Relative (per capita) Income

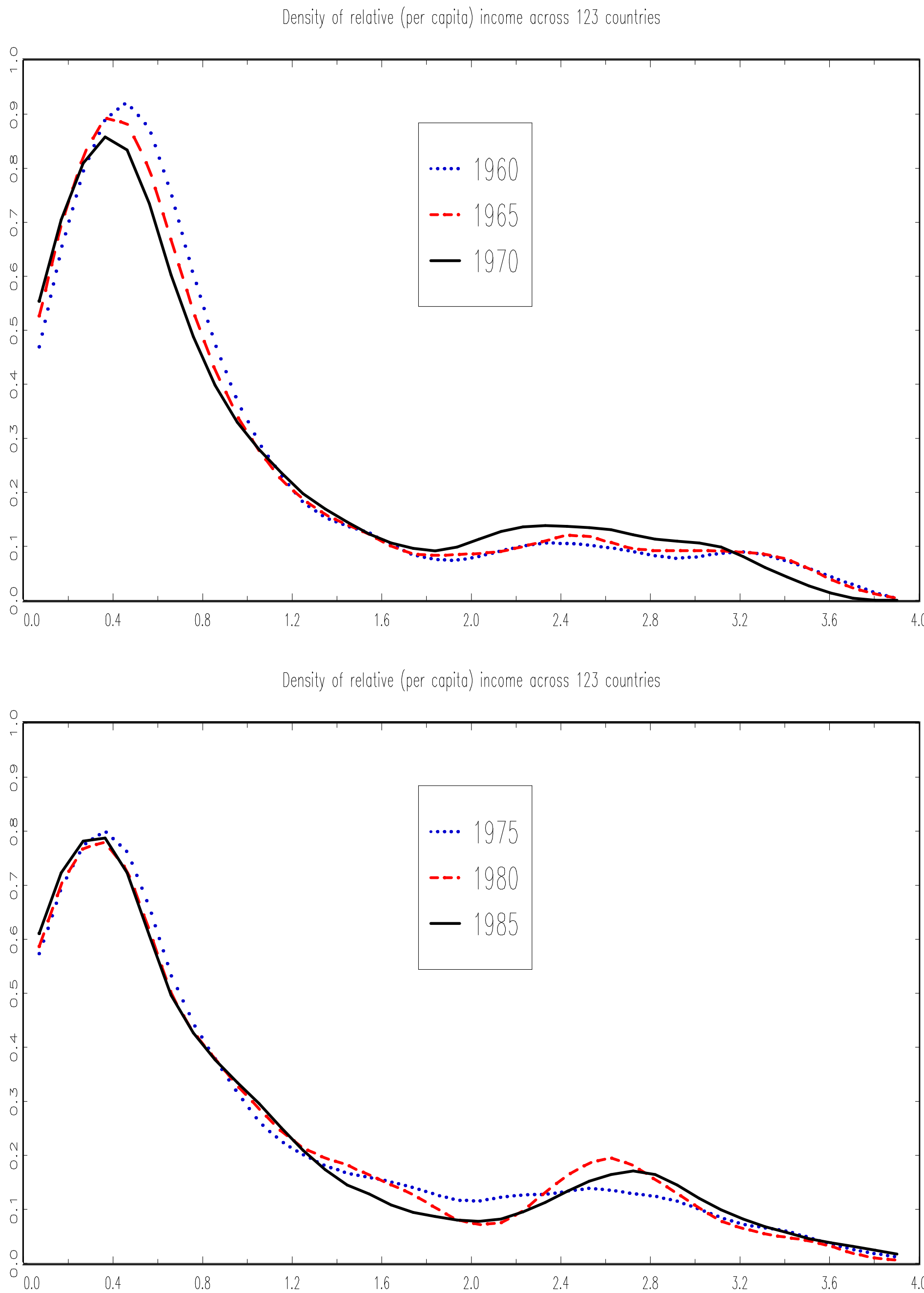
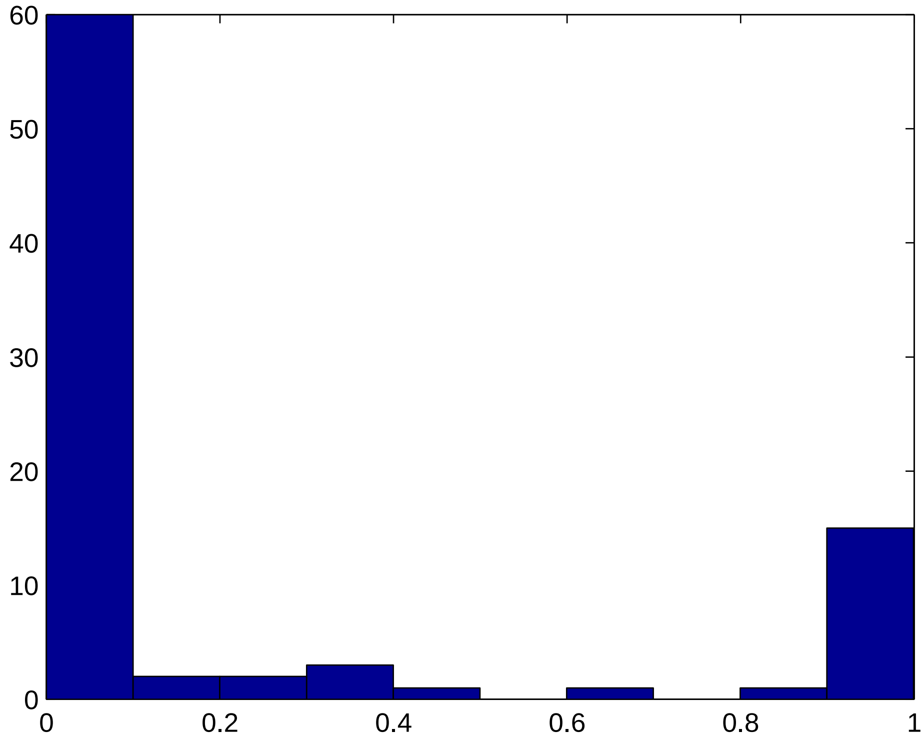


Figure 4: Histogram of Membership Probabilities: Two-club Model



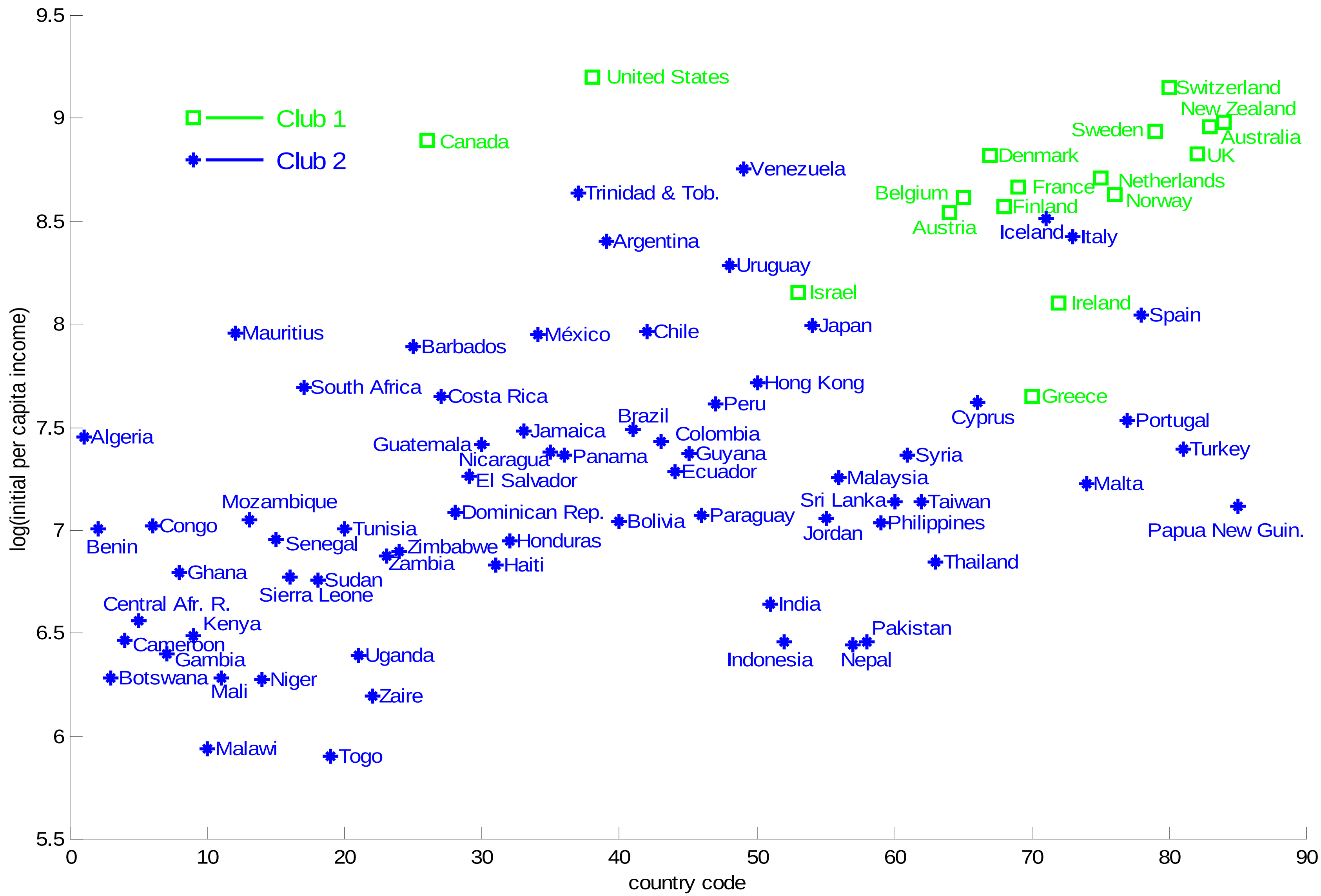


Figure 5: Two-club Structure

Figure 6: Convergence Clubs

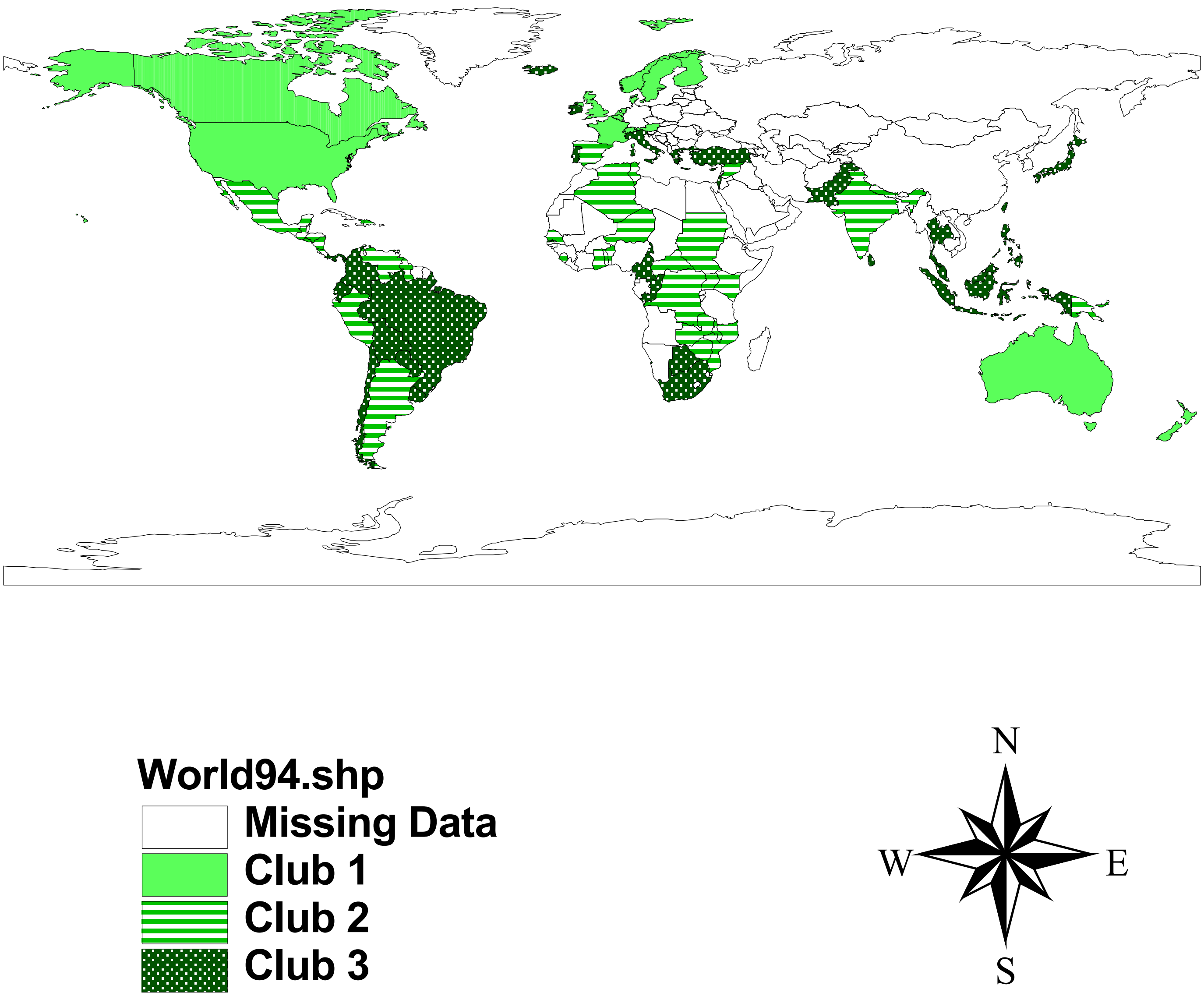
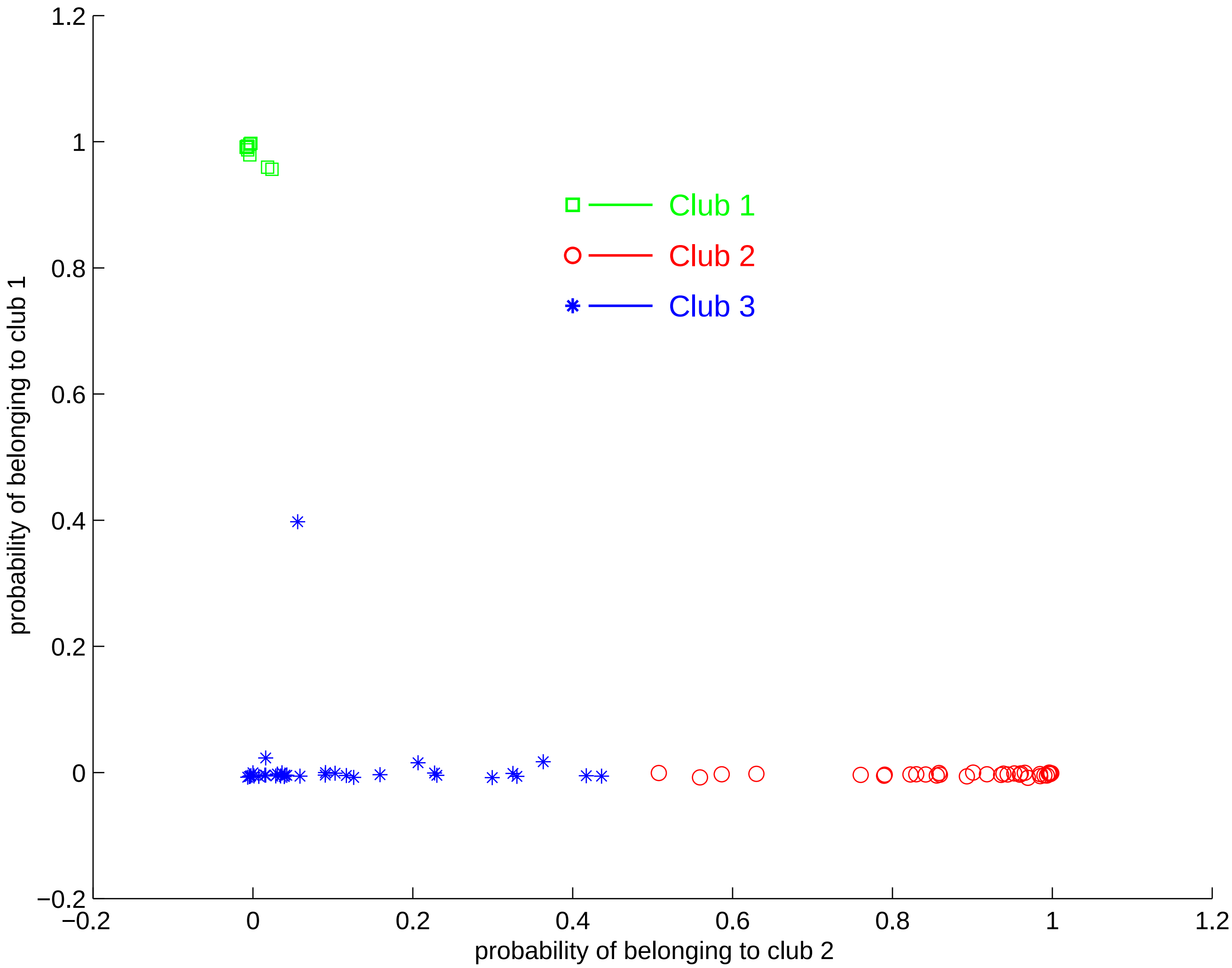


Figure 7: Scatterplot of Membership Probabilities: Three-club Model



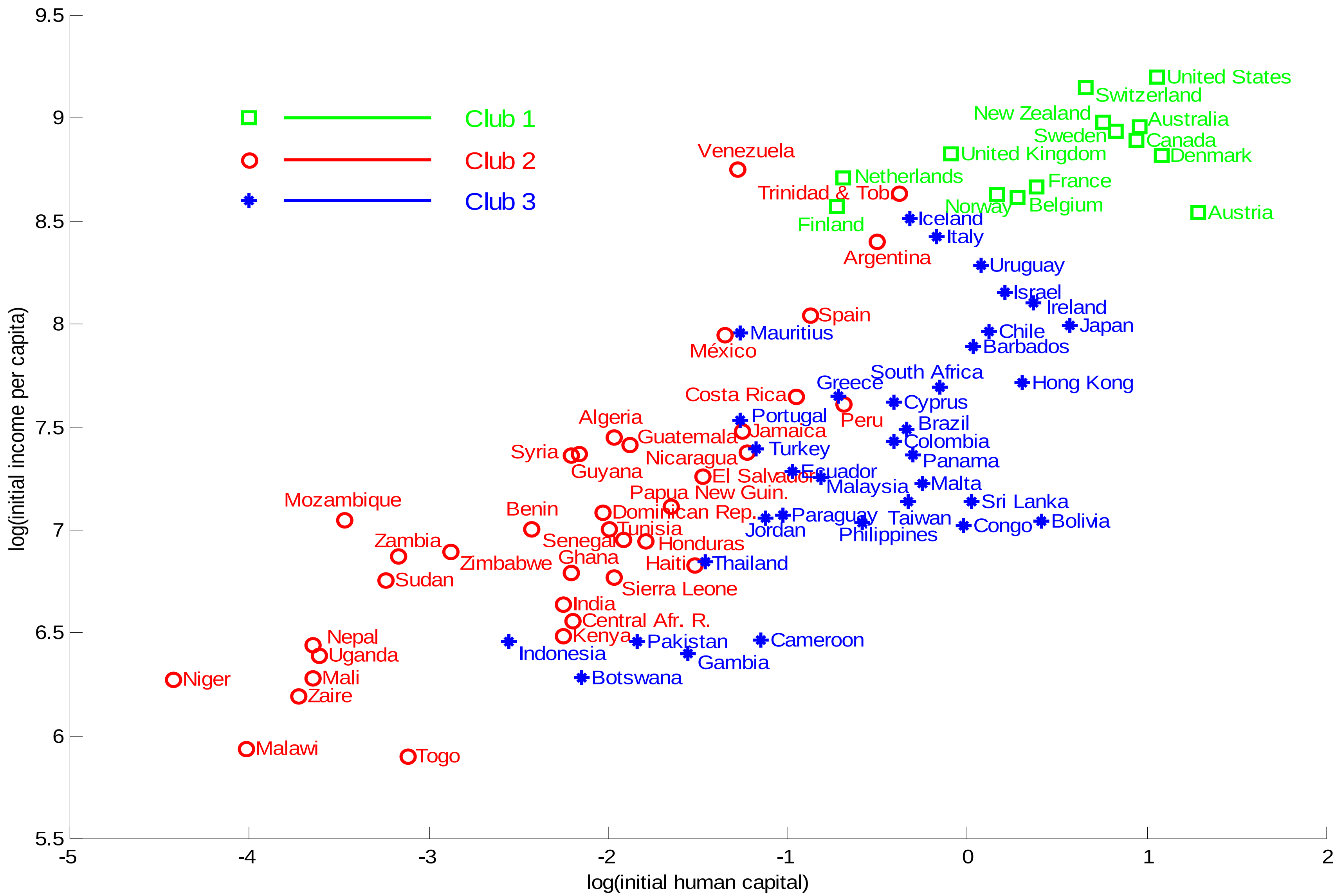


Figure 8: Three-club Structure

Figure 9: Convergence Clubs: Expanded Version

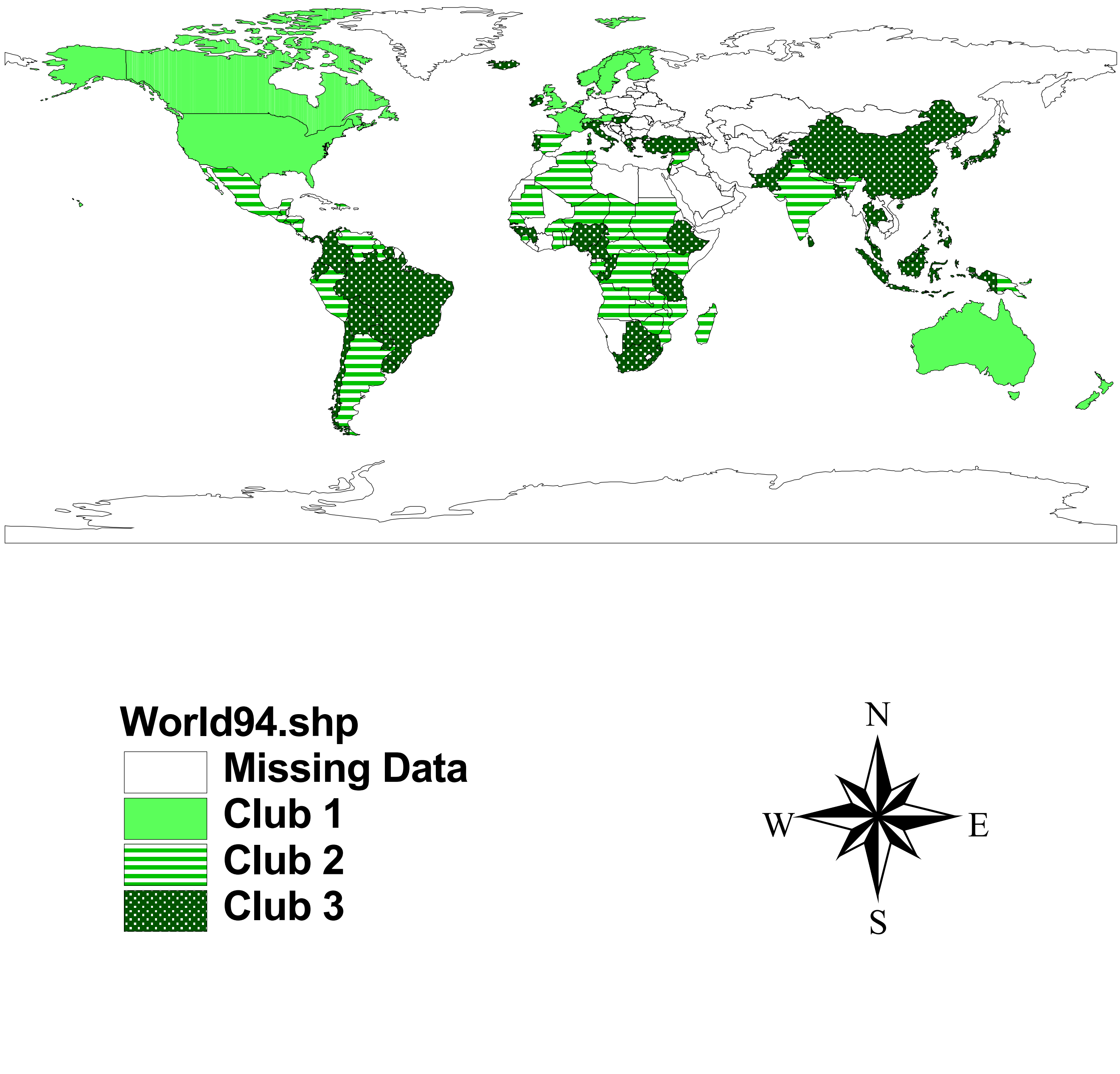


Table 1: Comparative Long-run Rates of Growth

	Time Period	GDP per capita beginning of period	GDP per capita ending of period	Annual growth (%) of GDP per capita
Austria	1870 – 1990	1442	12976	1.8478
Belgium	1870 – 1990	2009	13320	1.5888
Denmark	1870 – 1990	1543	14086	1.8600
Finland	1870 – 1990	933	14012	2.2834
France	1870 – 1990	1582	14245	1.8483
Germany (West)	1870 – 1990	1223	14288	2.0696
Italy	1870 – 1990	1216	13215	2.0081
Japan	1890 – 1990	842	16144	2.9976
Netherlands	1910 – 1990	2965	13078	1.8724
Norway	1870 – 1990	1190	15418	2.1576
Sweden	1870 – 1990	1401	14804	1.9842
Switzerland	1910 – 1990	2979	15650	2.0953
U.K.	1870 – 1990	2693	13589	1.3580
U.S.	1870 – 1990	2244	18258	1.7623
Bangladesh	1900 – 1987	349	375	0.0859
China	1900 – 1987	401	1748	1.7067
India	1900 – 1987	378	662	0.6462
Pakistan	1900 – 1987	413	885	0.8799
Philippines	1900 – 1987	718	1519	0.8650
Thailand	1900 – 1987	626	2294	1.5040
Argentina	1900 – 1987	1284	3302	1.0916
Chile	1900 – 1987	956	3393	1.4666
Colombia	1900 – 1987	610	3027	1.8583
Mexico	1900 – 1987	649	2667	1.6377
Peru	1900 – 1987	624	2380	1.5506

Data Sources: Madison (1989,1991), and Barro and Sala-i-Martin (1995).
Note: All GDP figures are in constant 1985 US dollars.

Table 2: BIC values

	$G = 1$	$G = 2$	$G = 3$	$G = 4$	$G = 5$
(c)	-389.13	-414.64^1	-406.71^2	-400.75^2	-383.31
(c, h)	-389.13	-430.48^1	-426.92^2	-409.55	-390.97
(c, e)	-389.13	-408.74^1	-397.43^2	-384.57	-360.14
(c, y_0)	-389.13	-439.47^1	-434.15^2	-413.37	-386.91
(c, e, h)	-389.13	-436.30^1	-415.61^2	-393.99	-379.19
(c, e, y_0)	-389.13	-430.84^1	-421.13^2	-361.85	-375.62
(c, h, y_0)	-389.13	-435.02^2	-435.94^1	-399.37	-371.25
(c, h, e, y_0)	-389.13	-434.04^1	-420.93^2	-393.44	-365.80

The superscripts 1 and 2 indicate the smallest and second smallest in each row, respectively.

Table 3: Convergence Clubs: Two-club Model

Club 1	Club 2			
Canada	Algeria	Tunisia	Trinidad & Tob.	Jordan
United States	Benin	Uganda	Argentina	Malaysia
Israel	Botswana	Zaire	Bolivia	Nepal
Austria	Cameroon	Zambia	Brazil	Pakistan
Belgium	Central Afr. R.	Zimbabwe	Chile	Philippines
Denmark	Congo	Barbados	Colombia	Sri Lanka
Finland	Gambia	Costa Rica	Ecuador	Syria
France	Ghana	Dominican Rep.	Guyana	Taiwan
Greece	Kenya	El Salvador	Paraguay	Thailand
Ireland	Malawi	Guatemala	Peru	Cyprus
Netherlands	Mali	Haiti	Uruguay	Iceland
Norway	Mauritius	Honduras	Venezuela	Italy
Sweden	Mozambique	Jamaica	Hong Kong	Malta
Switzerland	Niger	México	India	Portugal
United Kingdom	Senegal	Nicaragua	Indonesia	Spain
Australia	Sierra Leone	Panama	Japan	Turkey
New Zealand	South Africa	Sudan	Togo	Papua New Guin.

Table 4: Estimation of the Two-club $m(c, h, e)$ Model

	Club 1		Club 2	
	Estimate	Standard error	Estimate	Standard error
α	0.0501	0.0214	0.0516	0.0107
β	0.4734	0.1402	0.2886	0.0537
σ^2	0.0043	0.0008	0.0284	0.0022
Implied λ	0.1496	0.0585	0.2485	0.0372
Implied q^*	0.0190	0.0048	0.0145	0.0026

Table 5: Characteristics of Convergence Clubs: Two-club Model

	Club 1		Club 2	
	Weighted Average	Simple Average	Weighted Average	Simple Average
ethnic fractionalization	26.91%	26.29%	45.64%	45.19%
population growth	0.85%	0.85%	2.29%	2.28%
investment rate	28.28%	28.26%	16.36%	16.56%
initial income per capita	6332.60	6223.50	1606.60	1665.90
initial human capital	1.782	1.7625	0.391	0.398
growth rate	2.64%	2.67%	1.73%	1.74%

Table 6: Convergence Clubs: Three-club Model

Club 1	Club 2		Club 3	
Australia	Algeria	Nepal	Barbados	Italy
Austria	Argentina	Nicaragua	Bolivia	Japan
Belgium	Benin	Niger	Botswana	Jordan
Canada	Central Afr.	P. N. Guin	Brazil	Malaysia
Denmark	Costa Rica	Peru	Cameroon	Malta
Finland	Dominican	Senegal	Chile	Mauritiu
France	El Salvador	Sierra Leone	Colombia	Pakistan
Netherlands	Ghana	Spain	Congo	Panama
New Zealand	Guatemala	Sudan	Cyprus	Paraguay
Norway	Guyana	Syria	Ecuador	Philippines
Sweden	Haiti	Togo	Gambia	Portugal
Switzerland	Honduras	Trini.&Tob.	Greece	S. Africa
U.K.	India	Tunisia	Hong Kong	Sri Lanka
U.S.	Jamaica	Uganda	Iceland	Taiwan
	Kenya, Malawi	Venezuela	Indonesia	Thailand
	Mali, Mexico	Zaire, Zambia	Ireland	Turkey
	Mozambique	Zimbabwe	Israel	Uruguay

Table 7: Estimation of the Three-club $m(c, h, y_0)$ Model

	Club 1		Club 2		Club 3	
	Estimate	STD	Estimate	STD	Estimate	STD
α	0.0659	0.0396	0.0094	0.0171	0.1308	0.0263
β	0.3668	0.2631	0.3093	0.0933	0.0982	0.1040
σ^2	0.0028	0.0006	0.0213	0.0023	0.0278	0.0040
Implied λ	0.2006	0.1434	0.2347	0.0604	0.4641	0.2118
Implied q^*	0.0209	0.0044	0.0027	0.0049	0.0290	0.0036

Table 8: Characteristics of Convergence Clubs: Three-club Model

	Club 1		Club 2		Club 3	
	Weighted Average	Simple Average	Weighted Average	Simple Average	Weighted Average	Simple Average
ethnic frationalization	29.57%	29.50%	52.37%	51.05%	35.34%	35.82%
population growth	0.77%	0.75%	2.50%	2.48%	1.96%	1.97%
investment rate	28.35%	28.36%	13.33%	13.73%	20.50%	20.64%
initial income per capita	6931.10	6922.70	1504.00	1537.80	1882.20	1919.50
initial human capital	1.9227	1.9147	0.1641	0.1740	0.7214	0.7000
growth rate	2.48%	2.48%	0.52%	0.64%	3.14%	3.10%

Table 9: Convergence Clubs: Possible Members

Club 1	Club 2	Club 3	
Luxembourg	Angola	Burundi	Fiji
	Burkina Faso	Ethiopia	Hungary
	Chad	Guinea	Bahamas
	Comoros	Guinea-Bissau	Suriname
	Gabon	Nigeria	Bangladesh
	Cote d'Ivoire	Seychelles	China
	Madagascar	Swaziland	Korea
	Mauritania	Tanzania	
	Yugoslavia		

References

[1] Aghion, P. and Howitt, P. (1998), *Endogenous Growth Theory*, The MIT Press.

[2] Azariadis, C. and Drazen, A. (1990), "Threshold Externalities in Economic Development." *Quarterly Journal of Economics*, Vol. 105(2): 501-526.

[3] Bernanke, B. S. and Gurkaynak, R. S. (2001), "Is Growth Exogenous? Taking Mankiw, Romer and Weil Seriously." Forthcoming, *NBER Macroeconomics Annual.*

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

[5] Barro, R. J. (1996), "Democracy and Growth." *Journal of Economic Growth* 1(1): 1-27.

[6] Barro, R. J. and Sala-i-Martin, X. (1995), *Economic Growth*, New York: McGraw-Hill.

[7] Barro, R. and Lee, J. W. (2000), "International Data on Educational Attainment Updates and Implications." Center for International Development, Harvard University.

[8] Brock, W. A. and Durlauf, S. (2001), "Growth Empirics and Reality." Department of Economics, University of Wisconsin at Madison.

[9] Canova, F. (1999), "Testing for Convergence Clubs in Income per-capita: A Predictive Density Approach." Department of Economics and Business, Universitat Pompeu Fabra.

[10] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), "Maximum Likelihood for Incomplete Data via the EM Algorithm." *Journal of the Royal Statistics Society*, Series B, 39, 1-38.

[11] Durlauf, S. and Johnson, P. (1995), "Multiple Regimes and Cross-Country Growth Behavior." *Journal of Applied Econometrics*, 10, 365-384.

[12] Durlauf, S. (2000), "Econometric Analysis and the Study of Economic Growth: A Skeptical Perspective." Department of Economics, University of Wisconsin at Madison.

[13] Easterly, W. and Yu, H. (2000), Global Development Network Growth Database, http://www.worldbank.org/research/growth/GDNdata.htm.

[14] Easterly, W. and Levine, R. (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics,* 112(4): 1203-1250.

[15] Evans, P. (1997), "Consistent Estimation of Growth Regressions." Department of Economics, Ohio State University.

[16] Gallup, J. L., Sachs, J. D. and Mellinger, A. D. (1998), "Geography and Economic Development." NBER Working Paper, No. 6849.

[17] Islam, N. (1995), "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics*, 110, 1127-1170.

[18] Islam, N. (1998), "Growth Empirics: A Panel Data Approach–A Reply." *Quarterly Journal of Economics*, 113(1): 325-329.

[19] Jones, Charles I. (1995), "Time Series Tests of Endogenous Growth Models." *Quarterly Journal of Economics*, 110, May, 495-525.

[20] Jones, Charles I. (1998), *Introduction to Economic Growth*, W. W. Norton & Company.

[21] Lee, K., Pesaran, M. H., and Smith, R. (1997), "Growth and Convergence in a Multi-country Empirical Stochastic Solow Model." *Journal of Applied Econometrics*, 12(4), July-August, 357-392.

[22] Lee, K., Pesaran, M. H., and Smith, R. (1998), "Growth Empirics: a Panel Data Approach - a Comment." *Quarterly Journal of Economics*, 113(1), February, 319-323.

[23] Madison, A. (1989), *The World Economy in the Twentieth Century*, Paris, OECD.

[24] Madison, A. (1991), *Dynamic Forces in Capitalist Development*, Oxford, Oxford University Press.

[25] Mankiw, N. G., Romer, D. and Weil, D. N. (1992), "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics,* 107(2): 407-437.

[26] Mclachlan, G. J. and Krishman T. (1996), *The EM Algorithm and Extensions*, New York: John Wiley & Sons.

[27] Murphy, K. M., Shleifer, A. and Vishny R. W. (1989), "Industrialization and the Big Push." *Journal of Political Economy,* 97(5): 1003-1026.

[28] Paap, R. and van Dijk, H. K. (1998), "Distribution and Mobility of Wealth of Nations." *European Economic Review*, 42, 1269-1293.

[29] Quah, D. T. (1996), "Convergence Empirics across Economies with (some) Capital Mobility." *Journal of Economic Growth*, 1, 95-124.

[30] Quah, D. T. (1997), "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs." *Journal of Economic Growth,* 2(1): 27-59.

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

[32] Summers, R. and Heston, A. (1991), "The Penn World Table (mark 5): an Extended Set of International Comparisons, 1950-88." *Quarterly Journal of Economics*, 106(2): 327-336.

[33] Sun, Y. (2001), "Asymptotic Theory for Panel Structure Models." Department of Economics, Yale University.

[34] Swan, T. W. (1956), "Economic Growth and Capital Accumulation." *Economic Record*, 32, 324-361.