

Economic Transition and Growth*

Peter C. B. Phillips

*Cowles Foundation, Yale University
University of Auckland & University of York*

Donggyu Sul

*Department of Economics
University of Auckland*

August 16, 2007

Abstract

Some extensions of neoclassical growth models are discussed that allow for cross section heterogeneity among economies and evolution in rates of technological progress over time. The models offer a spectrum of transitional behavior among economies that includes convergence to a common steady state path as well as various forms of transitional divergence and convergence. Mechanisms for modeling such transitions, measuring them econometrically, assessing group behavior and selecting subgroups are developed in the paper. Some econometric issues with the commonly used augmented Solow regressions are pointed out, including problems of endogeneity and omitted variable bias which arise under conditions of transitional heterogeneity. Alternative regression methods for analyzing economic transition are given which lead to a new test of the convergence hypothesis and a new procedure for detecting club convergence clusters. Transition curves for individual economies and subgroups of economies are estimated in a series of empirical applications of the methods to regional US data, OECD data and Penn World Table data.

Keywords: Augmented Solow regression, Convergence club, Decay model, Economic growth, Growth convergence, Heterogeneity, Neoclassical growth, Relative transition, Transition curve, Transitional divergence.

JEL Classification Numbers: O30; O40; C33.

First Completed Draft: January 2005

*Some of the results given here were contained in an earlier working paper circulated under the title “Elusive Empirical Shadow of Growth Convergence”. Thanks go to Steven Durlauf, Kenneth West, Mario Gruchini, Paul Evans, and Feng Zhu for comments on the original version of the paper. Phillips gratefully acknowledges research support from a Kelly Fellowship and the NSF under Grant Nos. SES 04-142254 and SES 06-47086.

“The legacy of economic growth that we have inherited from the industrial revolution is an irreversible gain to humanity, of a magnitude that is still unknown....The legacy of inequality, the concomitant of this gain, is a historical transient”. Lucas (2002, pp.174-175).

1 Introduction

In his study of the growth of nations in the world economy over the last 250 years, Lucas (2002) argues that the enormous income inequality across countries that followed in the swath of the industrial revolution has now peaked. Instead, in the twenty first century, as countries increasingly participate in the economic benefits of industrialization, this income inequality will prove to be a historical transient. Building on a model of Becker, Murphy and Tamura (1990), Lucas develops a theory that seeks to explain the transition that has occurred in the world economy from the stagnant steady state economies that persisted until around 1800 to modern economies that experience sustained income growth. Human capital accumulation is posited as the engine of this growth and the mechanism by which it is accomplished comes by way of a demographic transition that emerges from the inclusion of fertility decision making into the theory of growth. These arguments involve two forms of transition: a primary economic transition involving the move toward sustained economic growth and a secondary, facilitating demographic transition associated with declining fertility. Lucas supports the arguments by some descriptive data analysis that document the transitions and suggest the emergent transience in income inequality mentioned in the headnote quotation.

The present paper looks at the phenomenon of ‘economic transition’ from an econometric perspective. We ask two main questions and then proceed to develop an econometric methodology for studying issues of economic transition empirically. The first question concerns neoclassical economic growth and asks if the model has the capacity to generate transitional heterogeneity of economic growth patterns across countries that are consistent with historical income inequality while still allowing for some form of ultimate growth convergence. Such behavior would have to accommodate transient divergence in growth patterns. So, a subsidiary question relates to the conditions under which such transitional economic divergence could occur and how it might be parameterized and evaluated empirically.

In seeking to address the first question, we use a neoclassical growth model under heterogeneous technological progress, as suggested by Barro and Sala-i-Martin (1997), Howitt and Mayer-Foulkes (2005) and Parente and Prescott (1994). Within this framework we develop a nonlinear dynamic factor model for log per capita real income of the form

$$\log y_{it} = a_{it} + x_{it}t = b_{it}\mu_t, \tag{1}$$

where the component a_{it} embodies transitional dynamics for real effective capital and the com-

ponent $x_{it}t$ captures the idiosyncratic time paths of technological progress. Both components are permitted to be heterogeneous across individuals and over time. The dynamic factor formulation $b_{it}\mu_t$ involves a growth component, μ_t , that is common across individuals (for instance, μ_t may represent a proxy for commonly available world technology), and individual transition factors (b_{it}) that measure how individual economic performance relates over time to μ_t .

In contrast to (1), transitional dynamics for log per capita real income are often expressed in the form

$$\log y_{it} = a_i + \exp(-\beta t) + xt, \quad (2)$$

whose first component, a_i , incorporates initial conditions and steady state levels. In this model, where the transition parameter β and the growth rate x are homogeneous, neoclassical theory does not naturally accommodate such enormous differences in observed income growth as the world economy has witnessed in the success of the Asian Dragons or the growth disasters of Sub Saharan Africa in relation to other developing countries. However, when we permit cross sectional and temporal heterogeneity in these parameters – replacing β and x in (2) by β_{it} and x_{it} , neoclassical growth can provide for such forms of transitional cross sectional divergence. With these extensions, the model may also allow for ultimate growth convergence, thereby making cross country income inequality a transient phenomena, as argued by Lucas.

In such an extended model, the speed of convergence parameter β_{it} may reasonably be regarded as an increasing function of technological progress x_{it} . Accordingly, poor economies with a low level of technological accumulation may begin with a low β_{it} and a correspondingly slow speed of convergence. As such countries learn faster (e.g., from improvements in education and the diffusion of technology), their x_{it} rises and may exceed the rate of technological creation in rich nations. So, β_{it} rises and the speed of convergence of these economies begins to accelerate. Conversely, if a poor country responds slowly to the diffusion of technology by learning slowly or through suffering a major economic disaster which inhibits its capacity to adopt new technology, its speed of convergence is correspondingly slower in relation to other countries (including rich countries), thereby producing the phenomenon of transitionally divergent behavior in relation to other countries. In other words, heterogeneous neoclassical economic growth may accommodate a family of potential growth paths in which some divergence may be manifest. If over time the speed of learning in the divergent economies becomes faster than the speed of technology creation in convergent rich economies, there is recovery and catch-up. In this event, the inequality that was initially generated by the divergence becomes transient, and ultimate convergence in world economic growth can be achieved.

Transitional economic behavior of the type described in the last paragraph leads to another major question: what variables govern the behavior of x_{it} and influence its transitional heterogeneity. While this question is not directly addressed in the present paper, the methods developed here for studying empirical economic transitions in growth performance are suited to

address similar issues regarding the transition behavior of the factors that influence economic growth.

To accommodate the time series and cross sectional heterogeneity of technological progress in growth empirics, this paper proposes a new econometric approach based on the analysis of an economy's transition path in conjunction with its growth performance. The transition path can be measured by considering the relative share of per capita log real income of country i in total income, or $h_{it} = \log y_{it} / \overline{\log y_t}$, where $\overline{\log y_t}$ denotes the cross section average of log per capita real income in the panel or a suitable subgroup of the panel¹. Under certain regularity conditions on the growth paths, the quantity h_{it} eliminates the common growth components (at least to the first order), and provides a measure of each individual country's share in common growth and technological progress. Moreover, since h_{it} is time dependent, it describes how this share evolves over time, thereby providing a measure of economic transition. In effect, h_{it} is a time dependent parameter that traces out a *transition curve* for economy i , indicating that economy's share of total income in period t . If there is a common source of sustained economic growth μ_t , then with the diffusion of technology and learning across countries, learning through formal education, and on the job learning (Lucas, 2002), we may reasonably suppose that all countries ultimately come to share (to a greater or lesser extent) in this growth experience. In this context, the parameter h_{it} captures individual economic transitions as individual countries experience this phenomenon to varying extents. As with the Galton fallacy, we do not expect all countries to converge. There will always be an empirical distribution of growth and per capita income among nations, as indeed there is between individuals within a country. However, there can still be convergence in the sense of an elimination of divergent behavior (as even the poorest countries begin to catch up) and an ultimate narrowing of the differences. Transitional growth empirics of the type considered in this paper seek to map these differences over time in an orderly manner that provides information about the transition behavior of countries in a world economy as they evolve toward a limit distribution in which all countries share in the common component in economic growth

The paper is organized as follows. Section 2 outlines some of the stylized facts that have emerged in the economic growth literature and provides some new ways of looking at these regularities. Section 3 studies some of the issues that arise in allowing for heterogeneity in neoclassical growth models and examines links between temporal heterogeneity in the speed of convergence and transitional divergence. Section 4 formalizes the concept of an economy's transition curve, which reveals the extent to which an individual economy shares at each point in time in the common growth component of a group of economies. Also, we develop an econometric formulation of this concept, which provides the time profile of transition for one

¹The idea of measuring transitions by means of a transition parameter was first suggested in the working paper Phillips and Sul (2003).

economy relative to a group average. This relative transition curve is identified and can be fitted using various smoothing methods. Use of the transition curve concept is demonstrated, some of its properties are discussed, and a new regression method is given for studying convergence, transition and divergence among economies. Empirical applications of these methods are reported in Section 5, where we study regional transitions in the US, national economic performance in the OECD nations, and growth and transitional divergence in the world economy using the Penn World Tables (PWT). Some conclusions and prospects for further research are given in Section 6. Supporting technical arguments, a demonstration of bias and inconsistency in augmented Solow regressions under heterogeneous technological progress, and information on the data are given in the Appendix. Further aspects of the econometric methodology used here are discussed in Phillips and Sul (2007).

2 Heterogeneous Technology and Growth

We start from the neoclassical theory of growth convergence and seek to build some connections between the theoretical formulations and observed empirical regularities. Write the production function in the neoclassical theory of growth with labor augmented technological progress as $Y = F(K, LHA)$ and define

$$\tilde{y} = f(\tilde{k}), \quad \tilde{y} = Y/LHA, \quad \tilde{k} = K/LHA, \quad y = \tilde{y}HA = \tilde{y}A \quad (3)$$

where Y is total output, L is the quantity of labor input, H is the stock of human capital, A is the state of technology, K is physical capital, and \tilde{y} is output per effective labor unit. In the last part of (3), H is normalized to unity so that technology A is defined broadly to encompass the effects of human capital.

It is commonly assumed that technological progress follows a simple exponential path of the form

$$A_{it} = A_{i0}e^{xt}, \quad (4)$$

where the growth rate of technology is common across countries. After imposing homogeneity restrictions on $A_i(0)$ and assuming a Cobb-Douglas production function, the transition dynamics of real per capita income have the form

$$\log y_{it} = \log \tilde{y}^* + [\log \tilde{y}_{i0} - \log \tilde{y}^*] e^{-\beta t} + \log A_0 + xt \quad (5)$$

where \tilde{y}^* is the steady state level of real effective per capita income and β is the speed of convergence given by $(1 - \alpha)(n + x + \delta)$. Note that α is the capital share, n is the population growth rate and δ is the discount rate.

The assumption of homogeneous technological progress is obviously restrictive and presumes that all economies experience technological improvements at the same rate $x_{it} = x$ over

time, while operating from different initial levels (A_{i0}). A more plausible assumption is that the technology growth rates x_{it} differ across countries and over time but may possibly converge to the same rate x as $t \rightarrow \infty$. In such a case, the evolution of A_{it} is inevitably more complex than (4), thereby accommodating a wider range of possible growth behavior. This motivation underlies the framework for empirical analysis that we develop later in this paper. A number of papers have started to assume and analyze heterogeneity in technological progress. For example, Parente and Prescott (1994) introduce an ‘adoption barrier’ to explain the cross sectional income heterogeneity, while Howitt and Mayer-Foulkes (2005) cite the ‘cost of learning’ as a possible explanation for heterogeneous technology.

We assume that the evolution of A_{it} follows

$$A_{it} = A_{i0}e^{x_{it}t}, \quad (6)$$

and thus the growth rate of A_{it} is time varying and may differ across countries. In the long run, technological progress is assumed to converge to the same constant rate x , that is, $\lim_{t \rightarrow \infty} x_{it} = x$.² The source of cross sectional heterogeneity for given time t can be explained as follows: Technology can be thought of as a public good and thus to be widely available. Let C_t denote the common technology available at time t and let $C_0 = e^c > 0$ be origination technology at time $t = 0$. The evolution of C_t is assumed to follow

$$C_t = C_0e^{\xi t} = e^{\xi t + c}. \quad (7)$$

Developed countries create the common technology C_t at the constant proportional rate ξ and they have access to the full extent of C_t in production. Followers, like the developing nations, generally have to learn earlier technology first to digest the more advanced technology, and develop an infrastructure to absorb and utilize it. As a result, it may be assumed that such countries cannot fully share in the present level of C_t . Depending on the speed of learning in these countries and the time from of their initial exposure to the common technology, the actual technological progress of developing countries is likely to differ across i over time. To model such cross section and temporal heterogeneity, we may treat C_t as a factor of production which different countries share in at their own idiosyncratic rate. More specifically, we set

$$A_{it} = C_t^{\gamma_{it}} = e^{\gamma_{it}(\xi t + c)} = e^{x_{it}t}, \quad \text{for } x_{it} = \gamma_{it}\xi + c\frac{\gamma_{it}}{t} \quad (8)$$

which yields (6) with the initialization absorbed into the expression for x_{it} . Note that technological convergence

$$\lim_{t \rightarrow \infty} x_{it} = x \quad (9)$$

²Barro and Sala-i-Martin (1997) consider the case where an initially poor country learns faster by copying advanced technology which was created by an initially rich country. However, they do not require overall convergence of x_{it} to x but rather local convergence of the x_{it} to some constants, x_g for $g < N$. This assumption generates conditional convergence.

occurs iff $\lim_{t \rightarrow \infty} \gamma_{it} = \gamma$ for some constant γ , so that in this event individual shares in common technology equalize over time. The rate of technological progress is $x_{it} + \dot{x}_{it}t = x_{it} \left(1 + \frac{\dot{x}_{it}}{x_{it}}t\right)$, which differs from x_{it} in general but is approximately x_{it} for large t if x_{it} converges fast enough³, although this will not be assumed in what follows.

To obtain an explicit expression of the transitional path of log per capita real income, we make the following assumptions.

Assumption 1 *Labor grows exogenously at the rate n , so that $L_{it} = L_{i0}e^{nt}$, while technology is assumed to grow exogenously according to the path $A_{it} = e^{x_{it}t}$.*

Assumption 2 *Savings grow exogenously and homogeneously at rate s .*

The homogenous growth rates for population and saving may be relaxed at the cost of greater complexity, but the main results of the paper do not change. Here, we stress the importance of temporal heterogeneity in technological progress, measured by x_{it} , and it is convenient to simply assume that saving and population growth rates are constant across countries. The two assumptions, combined with (9), imply that all economies ultimately converge.

As long as $x_{it} \rightarrow x$, the steady state per capita real income and capital stocks remain the same. However, the dynamic evolution of the variables does change. Capital stock is determined by

$$\dot{k}_{it} = sk_{it}^\alpha - (n + x_{it} + \delta)k_{it}. \quad (10)$$

As shown in the Appendix (see (34)), this system has a local linear approximation about the steady state path whose solution takes the following form

$$\log k_{it} = \log k^* + (\log k_{i0} - \log k^*)e^{-\beta_{it}t}. \quad (11)$$

In this system, β_{it} is a time varying speed of convergence parameter whose value is given by (see (37) and the Appendix for the derivation)

$$\beta_{it} = \beta - \frac{1}{t} \log \left\{ 1 - d_{i1} \int_0^t e^{\beta s} (x_{is} - x) ds \right\}, \quad (12)$$

where $d_{i1} = 1/(\log k_{i0} - \log k^*)$. Clearly, β_{it} depends on the whole time profile of the rate of technological progress $\{x_{is}\}_{s \leq t}$ since initialization at $t = 0$. Note that in cases where technology is convergent and (9) holds, the deviations $x_{is} - x$ are bounded and so $\beta_{it} \rightarrow \beta > 0$ as $t \rightarrow \infty$.⁴

In a similar way to (11), log per capita real income, $\log y_{it}$, evolves under cross sectional heterogeneous technology progress as

$$\log y_{it} = \log \tilde{y}^* + [\log \tilde{y}_{i0} - \log \tilde{y}^*]e^{-\beta_{it}t} + x_{it}t. \quad (13)$$

³That is if $\frac{\dot{x}_{it}}{x_{it}} = o\left(\frac{1}{t}\right)$ as $t \rightarrow \infty$.

⁴Note that β_{it} can be expressed as $\beta + O(t^{-1} \log t)$.

Thus, the transition path of log per capita real income depends on the rate of technological progress, x_{it} , which enters through the final term of (13) and through β_{it} in view of (12). In view of the latter, $\log y_{it}$ depends on the whole past trajectory $\{x_{is}\}_{s \leq t}$.

As is apparent from (13) and (5), when $x_{it} = x$, the relative income differential between economies, $(\log y_{it} - \log y_{jt})$, is explained only by the initial real effective per capita income. However, when $x_{it} \neq x$ during transition periods, the relative technological differential between x_{it} and x_{jt} (and the historical trajectory of this differential) also contributes to the income difference. Note that $e^{-\beta_{it}t} \rightarrow 0$ as $t \rightarrow \infty$, and if the convergence rate of this exponential term in (13) is fast relative to the convergence rate of x_{it} , then the main long run determinant of the relative income difference is the difference in the rates of technological accumulation. In this case, the relative income difference between two economies may be well explained by the relative difference in technology accumulation. For large t , $\log y_{it}$ eventually follows a long run path determined by the term $x_{it}t$ in (13). Hence analyzing the dynamics of $\log A_{it}$ and the past history of x_{it} are key elements in understanding transitional income dynamics.

3 Economic Transition Curves

This section consists of two subsections. We start by developing an econometric formulation of the neoclassical model given earlier that allows for heterogeneity in the speed of convergence and transition effects over time. The new formulation is a nonlinear factor model and involves the product of a time varying idiosyncratic element (b_{it}), which measures individual transition effects, and a common (stochastic) trend factor (μ_t) which captures the effects of common technology. Next, we suggest a mechanism for measuring the transition effects by means of a transition curve, which may be interpreted as approximating the trajectory over time of the idiosyncratic elements b_{it} . The second subsection applies this model and approach to three panel data sets, which display various forms of transition effects, including some cases of transitional divergence.

3.1 Transition and Relative Transition Curves

It is helpful in this development to use some general specification of the trending mechanism. It is sufficient for our purpose that there be some underlying trend mechanism, which may have both deterministic and stochastic components, and that this trend mechanism be a common element (for instance arising from knowledge, technology and industry in developed countries) in which individual economies can share. We denote this common trend element by μ_t . The extent to which economies do share in the common trend depends on their individual characteristics and this will ultimately be manifest in their growth performance and the shape of any economic transitions that occur.

From (13), the actual transition path of log per capita real income can be written as follows

$$\log y_{it} = \log \tilde{y}_i^* + \log A_{i0} + [\log \tilde{y}_{i0} - \log \tilde{y}^*] e^{-\beta_{it}} + x_{it}t = a_{it} + x_{it}t, \quad (14)$$

where

$$a_{it} = \log \tilde{y}_i^* + \log A_{i0} + [\log \tilde{y}_{i0} - \log \tilde{y}^*] e^{-\beta_{it}}. \quad (15)$$

As $t \rightarrow \infty$, (15) is a decay model for a_{it} which captures the evolution $a_{it} \rightarrow \log \tilde{y}_i^* + \log A_{i0}$. Correspondingly for large t , $\log y_{it}$ eventually follows a long run path that is determined by the term $x_{it}t$ in (14).

Following the discussion above, the growth path $x_{it}t$ is presumed to have some elements (and sources) that are common across economies. We use μ_t to represent this common growth component and can think of μ_t as being dependent on a common technology variable like C_t in (7), which enters as a factor of production for each individual economy. According to this view, all economies share to a greater or lesser extent in certain elements that promote growth – such as the industrial and scientific revolutions and internet technology. We may then write (14) in the following form

$$\log y_{it} = \left(\frac{a_{it} + x_{it}t}{\mu_t} \right) \mu_t = b_{it}\mu_t, \quad (16)$$

where b_{it} explicitly measures the share of the common trend μ_t that economy i experiences. In general, the coefficient b_{it} measures the transition path of an economy to the common steady state growth path determined by μ_t . During transition, b_{it} depends on the speed of convergence parameter β_{it} , the rate of technical progress parameter x_{it} and the initial technical endowment and steady state levels through the parameter a_{it} .

Note that growth convergence requires the following condition.

$$x_{it} \rightarrow x, \quad \text{for all } i \text{ as } t \rightarrow \infty. \quad (17)$$

Condition (17) is sufficient for the convergence of the growth rate of $\log y_{it}$, but only necessary for the level convergence of $\log y_{it}$.

In a neoclassical growth framework, steady state common growth for $\log y_{it}$ may be represented in terms of a simple linear deterministic trend $\mu_t = t$. Such a formulation is explicit in (7), for example. Then, according to (16), $b_{it} = x_{it} + a_{it}/t$ and, further, under the growth convergence condition (17) we have the convergence $b_{it} \rightarrow x$ as $t \rightarrow \infty$. Further, when the economies have heterogeneous technology and x_{it} converges to x_i , we have

$$b_{it} = x_{it} + \frac{a_{it}}{t} \rightarrow x_i, \quad \text{as } t \rightarrow \infty, \quad (18)$$

so that x_i determines the growth rate of economy i in the steady state.

In more general models and in empirical applications, the common growth component μ_t may be expected to have both deterministic and stochastic elements, such as a unit root

stochastic trend with drift. In the latter example, μ_t is still dominated by a linear trend asymptotically and conditions like (18) then hold as limits in probability. While this case covers most practical applications, we may sometimes want to allow for formulations of the common growth path μ_t that differ from a linear trend even asymptotically. Furthermore, a general specification allows for the possibility that some economies may diverge from the growth path μ_t , while others may converge to it. These extensions involve some technical complications that can be accommodated by allowing the functions to be regularly varying at infinity (that is, they behave asymptotically like power functions). We may also allow for individual country standardizations for log per capita income, so that expansion rates may differ, as well as imposing a common standardization for μ_t . Details of such extensions of the present set up will be reported elsewhere.

The estimation of b_{it} is not possible without imposing some smoothness or structural restrictions since the total number of unknowns is the same as the number of observations. Parametric assumptions enable the time profile of b_{it} to be fitted by filtering methods such as the Kalman filter. Smoothness conditions and deterministic assumptions allow for nonparametric estimation by kernel methods or sieve techniques. However, such methods are presently not well developed for fitting stochastic processes, rather than deterministic functions.

An alternate approach to modeling the transition elements b_{it} that is convenient in the present context is to construct the following relative transition coefficient

$$h_{it} = \frac{\log y_{it}}{N^{-1} \sum_{i=1}^N \log y_{it}} = \frac{b_{it}}{N^{-1} \sum_{i=1}^N b_{it}}, \quad (19)$$

which eliminates the common growth component by scaling and measures the transition element for economy i relative to the cross section average. The variable h_{it} traces out an individual trajectory for each i relative to the average, so we call h_{it} the ‘relative transition path’. At the same time, h_{it} measures economy i ’s relative departure from the common steady state growth path μ_t . Thus, any divergences from μ_t are reflected in the transition paths h_{it} . While many paths are possible, a case of particular interest and empirical importance occurs when an economy slips behind in the growth tables and diverges from others in the group. We may then use the transition path to measure the extent of the divergent behavior and to assess whether or not the divergence is transient.

When there is a common (limiting) transition behavior across economies, we have $h_{it} = h_t$ across i , and when there is ultimate growth convergence we have

$$h_{it} \rightarrow 1, \quad \text{for all } i, \text{ as } t \rightarrow \infty \quad (20)$$

This framework of growth convergence admits a family of relative transitions, where the curves traced out by h_{it} may differ across i in the short run, while allowing for ultimate convergence when (20) holds in the long run. Removing the common (steady state) trend function μ_t , Fig.

1 shows some examples of relative transition paths, each satisfying the growth convergence condition (20).

While the criterion for the ultimate convergence of economy i to the steady state is given by (20), the manner of economic transition and convergence can be very different across economies. Fig. 1 shows three different stylized paths. Economies 2 and 3 have quite different initializations and their transitions also differ. While both relative transition parameters converge monotonically to unity, path 3 involves transition from a high initial state, typical of an already advanced industrial economy, whereas path 2 involves transition from a low initial state that is typical of a newly industrialized and fast growing economy. Economy 1, on the other hand, has the same initialization as 2 but its relative transition involves an initial phase of divergence from the group, followed by a catch up period, and later convergence. Such a transition is typical of a developing country that grows slowly in an initial phase (transition phase A), begins to turn its economic performance around (phase B) and then catches up and converges (phase C).

Also, as Fig. 1 illustrates in a stylized way, when there is temporal and cross section heterogeneity, there exists an infinite number of possible transition paths some showing periods of transitional divergence (such as economy 1) even in cases where there is ultimate convergence.

3.2 Transition Phases and Divergence: Some Graphical Illustrations

As suggested in this stylized diagram, diverse patterns of economic transition are possible when we allow for cross sectional and time series heterogeneity in the parameters of a neoclassical growth model. This potential for diversity in transition is illustrated in the following empirical examples involving regional and national economic growth. Similar panel data sets to those used here have been extensively analyzed in the growth convergence literature in the past. Our application now focuses attention on the phenomenon of economic transition as part of a larger empirical story regarding convergence and divergence issues. We start by providing some graphical illustrations of the various phases of transition in the empirical data and then proceed to conduct some formal statistical tests based on empirical (decay model) regressions.

The first illustration is based on regional economic growth among the 48 contiguous U.S. states⁵. In this example, there is reasonable prior support for a common rate of technological progress and ultimate growth convergence but we may well expect appreciable heterogeneity across states in the transition paths. Fig. 2 displays the relative transition parameters calculated for log per capita income in the 48 states over the period from 1929 to 1998 after eliminating business cycle components.⁶ Evidently, there is heterogeneity across states, but

⁵The data source for U.S. state per capita real income is the Bureau of Economic Analysis.

⁶Both HP and trend regression methods were used to eliminate business cycle components and the results showed little sensitivity to the method employed.

also a marked reduction in dispersion of the transition curves over this period, together with some clear evidence that the relative transition curves narrow towards unity, as indicated in the convergence criterion (20).

The second illustration involves a panel of real per capita income for 18 western OECD countries taken from the OECD historical data set. Panel A in Fig. 3 displays the relative transition parameters for log per capita income in these 18 OECD countries between 1929 and 2001. The countries were selected on the basis of data availability and are listed in the Data Appendix. The observed time profiles of transition for these OECD nations are quite different from those of Fig. 2, even though the time frame is similar. For the OECD nations, the relative transition parameters initially seem to display no coherent pattern and, in some cases, even appear to diverge before World War II. Around 1950, however, the pattern of transition appears to change and subsequently becomes similar to that of Fig. 2. Over the latter part of the period, there is a noticeable narrowing in the transition curves towards unity, indicating a clear tendency to converge towards the end of the period.

Panels B and C in Fig. 3 show the relative transition curves for certain subgroups of countries against the benchmark of the U.S.⁷. We have created five economic subgroups in this exercise. Except for the former U.K. colonies, all subgroups show clear evidence of some transitional divergence with a turn-around by the end of WWII. After that, all of these subgroups reveal a strong tendency towards convergence with the U.S.. Evidently, panel B in Fig. 3 provides an empirical illustration of the stylized patterns of economic performance characterized as phases B and C of Fig. 1. Extending the panel back to 1870 and through to 1930, panel C in Fig. 3 shows transition curves that are similar in form to phase A (transitional divergence) in the stylized patterns of Fig. 1, with evidence of the phase B turn-around coming towards the end of this period.

The final illustration is based on log per capita income in 98 PWT countries in the world economy over various periods. The country selection is mainly based on data availability. Given the large number of countries and the wide variation in the data, it is helpful to take subgroup averages to reduce the number of transition curves, which we show against the benchmark of the 19 OECD countries.⁸ The subgroups are based on total population and geographical region. Phase A transitions are found in two of these subgroups – the countries of Sub-Saharan

⁷For all subgroups, the U.S. enters as the numeraire country. The relative transition curves for subgroup k are calculated as follows: Let N_k be the total number of countries in the subgroup. Since the U.S. is always included in the subgroup, the total number of countries excluding the U.S. is $N_k - 1$. First we calculate $h_{it} = \ln y_{it} \left(\left[N_k^{-1} \sum_{i=1}^{N_k} \ln y_{it} \right] \right)^{-1}$ for $i = 1, \dots, N_k$. Next, we take the cross sectional average of the h_{it} excluding the USA. That is, $h_{kt} = \frac{1}{N_k - 1} \sum_{i \neq USA}^{N_k} h_{it}$.

⁸The relative transition curves are calculated in a similar fashion to that described in footnote 11. The only difference is that instead of U.S. as numeraire, the 19 OECD countries are used as numeraire countries. That is, the cross sectional average of h_{it} is given by $h_{kt} = \frac{1}{N_k - 19} \sum_{i \neq OECD}^{N_k} h_{it}$.

Africa, and the Latin American & Caribbean economies from 1960 to 2003 (Panel 3 in Fig. 4). Phases B and C occurred in three cases – India, China and Korea – over the period from 1953 to 2003 (Panel 1 in Fig. 4). Finally, phase C transitions are evident in two subgroups – the Asian dragons and the newly industrialized economies (NIEs) from 1960 to 2003 shown in Panel 2 in Fig. 4.

From these findings about the present standing of these economic groups and assuming that the world economies are in transition to ultimate convergence on a path that is related to long run historical OECD growth, then we can expect that China, India will continue to grow faster over the next decade than the OECD nations as they experience phase C transition; and, sooner or later, we might expect to witness the Sub-Saharan and Latin American countries entering phase B transition when they begin to turn around their economic performance and start to catch up with the 19 OECD countries. However, from the evidence to date in these figures, we cannot distinguish for the Sub-Saharan Africa and Latin American countries whether or when such changes may occur.

We now provide some formal econometric procedures for evaluating transition curves to shed light on growth convergence and convergence clustering issues.

4 Testing Growth Convergence

The conventional conditional β -convergence test is designed to assess whether the speed of convergence parameter (β) in (5) is positive or negative. This simple and rather intuitive test works under homogeneity of technology progress and has been widely applied. When technology is heterogeneous across countries allowing transition periods of the type just described, then the speed of convergence β_{it} is time varying and depends on the time profile of the rate of technological progress $\{x_{is}\}_{s \leq t}$. This dependence is a considerable complication that affects the properties of the regression equation that is traditionally used in tests that focus attention on β . Under heterogeneity, as we have seen, growth convergence depends on whether $x_{it} \rightarrow x$, for all i , as in (17), or whether $b_{it} \rightarrow b$ and $h_{it} \rightarrow 1$, for all i , as in (20). In such cases, it seems appropriate to direct attention specifically at these conditions, and we do so below.

4.1 Pitfalls of Existing Convergence Tests

First, we briefly discuss some issues that arise with other approaches when there is heterogeneous transition. Perhaps the most popular approach in empirical work is the augmented Solow regression (Barro and Sala-i-Martin, 1992; Mankiw, Romer and Weil, 1992), which takes the general form

$$\log \left(\frac{y_{it+1}}{y_{it}} \right) = a_0 + a_1 \log y_{it} + \mathbf{z}'_i a_2 + e_i \quad (21)$$

where z_i is a vector of proxy (determining) variables for the steady state log levels $\log \tilde{y}_i^*$. When there is heterogeneity in technological progress, the parameter β_{it} is time and country varying, and this materially affects the specification (21) in important ways. As shown in the Appendix, the coefficient a_1 in (21) is then a functional of β_{it} and x_{it} . In particular, if we assume that the β_{it} are independently drawn from some distribution across economies, then a_1 can be written in terms of the moment generating function of that distribution, so that for some given t we have

$$a_1 = - \left(1 - E \left[e^{-\beta_i^+} \right] \right), \quad \text{where } \beta_i^+ = \beta_{it+1} + (\beta_{it+1} - \beta_{it})t, \quad (22)$$

and where β_{it} is defined in (12). The residual is absorbed into the error e_i in (21). Moreover, as shown in the Appendix, under heterogeneous technological progress, the regression error e_i includes terms that involve $\log A_{it}$, $\log A_{it+1}$ and variables that are correlated with $\log y_{it}$ and z_i . Hence, least squares regression on (21) is inconsistent and the bias arises from two separate sources – omitted variables and endogeneity. The Appendix provides a more complete analysis. The main issue regarding the use of augmented Solow regression such as (21) in practice is that, under conditions of transitional heterogeneity, estimation of the coefficient a_1 is biased and inconsistent, so negative estimates of a_1 can not be directly interpreted as evidence of growth convergence. Moreover, $a_1 = - \left(1 - E \left[e^{-\beta_i^+} \right] \right) < 0$ does not imply that $x_{it} = x$ or that $x_{it} \rightarrow x$ as $t \rightarrow \infty$, so that it is possible to get negative estimates of a_1 even under conditions of divergence. A further issue in the use of augmented Solow regressions is the inconsistency of estimates of the coefficient vector a_2 of the proxy variables in (21). Again, due to omitted variable and endogeneity bias, estimates of a_2 may suffer a reversal of sign, leading to false inferences concerning the direction of influence certain growth determinants. This issue is being investigated more fully in other work.

Another common approach to growth convergence analysis involves the use of cointegration and unit root tests. The conventional cointegration test approach is based on examining the time series properties of log level inter-economy differentials such as $\log y_{it} - \log y_{jt}$. In particular, such series can be tested for the presence of a stochastic trend by means of an empirical regression of the form

$$\log y_{it} - \log y_{jt} = c_1 + c_2 t + \rho (\log y_{it-1} - \log y_{jt-1}) + error \quad (23)$$

Bernard and Durlauf (1995,96) and Evans and Karras (1996) omit the linear trend term in this regression and test whether $\rho < 1$ (convergence) or $\rho = 1$ (divergence). If $c_1 \neq 0$ but $\rho < 1$, Evans and Karras (1996) interpret this as conditional convergence. Carlino and Mills (1993) include the linear trend term in (23), and test whether $c_2 = 0$ and $\rho < 1$. When the model involves heterogeneous transition, we may write the inter-economy differentials as

$$\log y_{it} - \log y_{jt} = (a_{it} - a_{jt}) + (x_{it} - x_{jt})t = (b_{it} - b_{jt})\mu_t,$$

using (14) and (16). Suppose the common growth component μ_t follows a stochastic trend with drift, so that $\mu_t = \mu_0 + \mu t + \sum_{s=1}^t \zeta_s$ for some stationary process ζ_s and initialization μ_0 . Then

$$\log y_{it} - \log y_{jt} = (b_{it} - b_{jt})(\mu_0 + \mu t) + (b_{it} - b_{jt}) \sum_{s=1}^t \zeta_s,$$

and $\log y_{it} - \log y_{jt}$ has a unit root component for all t . Existence of a unit root in the differential $\log y_{it} - \log y_{jt}$ does not necessarily imply growth divergence because it is possible that the convergence condition $b_{it} - b_{jt} \rightarrow_p 0$ still holds.⁹ Hence, depending on the stochastic order of the idiosyncratic factor differential $b_{it} - b_{jt}$, cointegration type tests may or may not reveal whether growth convergence applies.

4.2 The $\log t$ Convergence Test

Since growth convergence under transitional heterogeneity depends explicitly on the condition $b_{it} - b_{jt} \rightarrow_p 0$, it seems appropriate to develop an approach which focuses attention on this condition or some convenient version of it. A particularly simple approach is to work with the relative transition coefficients h_{it} , rather than the coefficients b_{it} , because h_{it} may be directly measured from the data in some cases, as in (19), or is easily computed in others where some prefiltering (e.g., business cycle removal) is performed. In both cases, the common growth component is eliminated. Also, when there is ultimate growth convergence, we have the limit $h_{it} \rightarrow_p 1$ for all i as $t \rightarrow \infty$, and then the mean square transition differential

$$H_t = N^{-1} \sum_{i=1}^N (h_{it} - 1)^2$$

provides a quadratic distance measure for the panel from the common limit. Under convergence, the distance $H_t \rightarrow 0$ as $t \rightarrow \infty$. When convergence does not hold, the distance remains positive as $t \rightarrow \infty$. There are various possibilities: H_t may converge to a non zero constant, it may remain bounded above zero but not converge, or it may diverge. In the case where there is club convergence, H_t typically converges to a positive constant, as shown in the Appendix. With limited time series evidence, it is naturally difficult to distinguish whether H_t converges to zero or a positive constant. To sharpen the distinction and assist in empirical testing it is helpful to use a specific model of transition that is conducive to econometric testing.

In related work (Phillips and Sul, 2007) we have developed such a model and testing procedure. The test is based on a simple time series regression and involves a one-sided t test of the null hypothesis of convergence against alternatives which include no convergence and partial convergence among subgroups. The test is called the ‘ $\log t$ ’ convergence test because

⁹In that event, we have $\log y_{it} - \log y_{jt} = O_p\left(\mu(b_{it} - b_{jt})t + (b_{it} - b_{jt})t^{1/2}\right) = O_p\left((b_{it} - b_{jt})t^{1/2}\right)$, with the second equality holding when $\mu = 0$ and there is no drift.

the t statistic refers to the coefficient of the $\log t$ regressor in the regression equation - see (27) below. What follows is a brief outline of the model and the econometric testing procedure.

To formulate a null hypothesis of growth convergence, we use a semiparametric model for the transition coefficients that allows for heterogeneity over time and across individuals as

$$b_{it} = b_i + \frac{\sigma_i \xi_{it}}{L(t) t^\alpha}, \quad (24)$$

where b_i is fixed, ξ_{it} is $iid(0, 1)$ across i but may be weakly dependent over t , and $L(t)$ is a slowly varying function (like $\log t$) for which $L(t) \rightarrow \infty$ as $t \rightarrow \infty$. The parameter α governs the rate at which the cross section variation over the transitions decays to zero over time. This formulation ensures that b_{it} converges to b_i for all $\alpha \geq 0$, which therefore becomes a null hypothesis of interest. If this null holds and $b_i = b_j$ for $i \neq j$, the model allows for transitional periods in which $b_{it} \neq b_{jt}$, thereby incorporating the possibility of transitional heterogeneity or even transitional divergence across i . As shown in Phillips and Sul (2007), further heterogeneity may be introduced by allowing the decay rate α and slowly varying function $L(t)$ to be individual specific without affecting the manner in which the test is conducted. Some regularity conditions on the idiosyncratic scale parameters σ_i and the random variables ξ_{it} are required in order to develop rigorous asymptotics for the regression and these are detailed in Phillips and Sul (2007).

The null hypothesis of convergence may be written as

$$H_0 : b_i = b \ \& \ \alpha \geq 0, \quad (25)$$

which involves the weak inequality $\alpha \geq 0$, since

$$\begin{aligned} \lim_{t \rightarrow \infty} b_{it} = b & \quad \text{iff} \quad b_i = b \ \text{and} \ \alpha \geq 0 \\ \lim_{t \rightarrow \infty} b_{it} \neq b & \quad \text{iff} \quad b_i \neq b \ \text{and/or} \ \alpha < 0 \end{aligned} .$$

The alternative hypothesis is given by

$$H_A : \{b_i = b \ \text{for all } i \ \text{with } \alpha < 0\} \ \text{or} \ \{b_i \neq b \ \text{for some } i \ \text{with } \alpha \geq 0, \ \text{or } \alpha < 0\} .$$

One role of the slowly varying component $L(t)$ in (24) is to ensure that convergence holds even when $\alpha = 0$, although possibly at a very slow rate. This formulation accommodates some interesting empirical possibilities where there is slow transition and slow convergence. The alternative hypothesis includes straightforward divergence but more importantly also includes the possibility of club convergence. For example, if there are two convergent clubs, then the alternative can include such a case in which

$$H_A : b_{it} \rightarrow \begin{cases} b_1 \ \text{and} \ \alpha \geq 0 & \text{if } i \in G_1 \\ b_2 \ \text{and} \ \alpha \geq 0 & \text{if } i \in G_2 \end{cases}$$

where the number of individuals in G_1 and G_2 aggregates to N .

Corresponding to the decay model (24) and under growth convergence the transition distance H_t is shown in Phillips and Sul (2007) to have the limiting form

$$H_t \sim \frac{A}{L(t)^2 t^{2\alpha}} \quad \text{as } t \rightarrow \infty, \quad (26)$$

for some constant $A > 0$. Setting $L(t) = \log t$, this formulation leads to the following ‘log t ’ regression model

$$\log \frac{H_1}{H_t} - 2 \log(\log t) = \beta_0 + \beta_1 \log t + u_t, \quad \text{for } t = T_0, \dots, T. \quad (27)$$

In (27), the initial observation in the regression is $T_0 = [rT]$ for some $r > 0$, so that empirical log t regressions are based on time series data in which the first $r\%$ of the data is discarded¹⁰. The second term on the left hand side in (27), $-2 \log(\log t)$, plays the role of a penalty function and improves test performance particularly under the alternative. For instance, under the alternative of club convergence, the transition distance H_t converges to a positive quantity as $t \rightarrow \infty$ and the inclusion of the penalty term in the regression gives the test discriminatory power between overall convergence and club convergence.

Under the null of growth convergence, the point estimate of the parameter β_1 converges in probability to the scaled speed of convergence parameter 2α . The corresponding t -statistic in the regression is constructed in the usual way using HAC standard errors. This t -statistic diverges to positive infinity when $\alpha > 0$ and converges weakly to a standard normal distribution when $\alpha = 0$. The convergence test then proceeds as a one sided t test of $\alpha \geq 0$. Under the alternative of growth divergence or club convergence, the point estimate of β_1 converges to zero regardless of the true value of α , but its t -statistic diverges to negative infinity, thereby giving the one sided t -test discriminatory power against these alternatives.

In conducting the regression (27) we may employ either h_{it} computed directly from $\log y_{it}$ as in (19) or, as discussed above, use filtered data \hat{h}_{it} based on fitted values $\widehat{\log y_{it}}$ from a coordinate trend regression (Phillips, 2005) or a smoothing filter that removes business cycles.

We are interested not only in the sign of the coefficient $\beta_1 = 2\alpha$ of $\log t$ in (27) but also its magnitude, which measures the speed of convergence of b_{it} .¹¹ Hence if $\beta_1 \geq 2$ (i.e., $\alpha \geq 1$) and the common growth component μ_t follows a random walk with drift or a trend stationary process, then values of β_1 that are this large will imply convergence in level per capita incomes. Meanwhile, if $2 > \beta_1 \geq 0$, then this speed of convergence corresponds with conditional convergence, i.e., income growth rates converge over time.

¹⁰Phillips and Sul (2007) suggest r values in the interval $[0.2, 0.3]$. This data trimming focuses attention on the latter part of the sample data, validates the regression equation in terms of the asymptotic representation of the transition distance (26), and ensures test consistency in growth convergence applications.

¹¹In particular, under (24) we have $b_{it} = b_i + O_p\left(\frac{1}{L(t)t^\alpha}\right)$

4.3 Empirical Evidence

We employ the $\log t$ test with four panels. The first panel consists of income data for the 48 contiguous U.S. states from 1929 to 1998, the second panel consists of 18 Western OECD countries from 1870 to 2001. The third panel consists of 152 PWT countries from 1970 to 2003.

Table 1 reports the results of applying the $\log t$ test with three panel data sets. The last two columns in the table report the point estimates β and their standard errors, which are computed using an automated HAC procedure. For the state income data, the estimate $\hat{\beta}_1 = 0.46$ is significantly positive, so that there is strong evidence in support of H_0 in (25) and for convergence as defined in (20). If the common stochastic trend component follows either a random walk with a drift or a trend stationary process, then the speed of convergence parameter is significantly below 2, so that absolute level convergence can be rejected.

The second panel involves log per capita GDP for 18 Western OECD countries. The point estimate of $\hat{\beta}_1 = 1.71$ is large for the period 1870-2001, so that the null hypothesis of convergence is strongly accepted, but $\hat{\beta}_1$ is also significantly below 2, so that the hypothesis of absolute level convergence is rejected. In order to analyze the impact of transition on the $\log t$ test, we did some subsample analyses. The first subsample covers the period 1870 to 1929 and can be characterized as phase A transition as discussed in the context of Fig. 7. During this period, we find that $\hat{\beta}_1 = -0.42$, which implies divergence. The second subsample covers 1911 to 1970 and this period has phase B (or transitional divergence and turn around) characteristics. For this period, the point estimate is $\hat{\beta}_1 = -0.11$ and is not significantly different from zero. The last subsample covers 1940-2001 and for this period the point estimate is $\hat{\beta}_1 = 1.14$, confirming convergence. The point estimate of the convergence parameter increases as the initialization period is moved forward and as later observations are included.

For the third panel, there is no evidence of convergence: the point estimates, $\hat{\beta}_1$, are significantly less than zero for all case and the estimated standard errors are so small for these large cross section groups that the null hypothesis of convergence is rejected even at the 0.1% level. This confirms earlier findings.

5 Growth Convergence Clubs and Economic Transition

The $\log t$ regression test has power against cases of club convergence, so we can expect the null hypothesis of convergence to be rejected for data in which there is evidence of club convergence. Accordingly, we investigate the possibility of a club convergence pattern among the 152 PWT countries where overall convergence is firmly rejected. To do so, we utilize a clustering mechanism test procedure which relies on the following stepwise and cross section recursive application of $\log t$ regression tests. The procedure facilitates the empirical discovery

of club convergence clusters, using the fact that the $\log t$ regression test has power against club convergence alternatives.

5.1 Clustering Algorithm and Convergence Club Empirics

A detailed analysis of the clustering procedure is given in Phillips and Sul (2007). The constructive steps for implementing the procedure are briefly summarized as follows.

Step 1 (Cross section ordering by final observation) Order the countries according to the amount of final period income.

Step 2 (Form a core primary group of k^* countries) Selecting the first k highest individuals in the panel to form the subgroup G_k for some $2 \leq k < N$, run the $\log t$ regression and calculate the convergence test statistic $t_k = t(G_k)$ for this subgroup. Choose the core group size k^* by maximizing t_k over k according to the criterion:¹²

$$k^* = \arg \max_k \{t_k\} \quad \text{subject to } \min \{t_k\} > -1.65, \quad (28)$$

If the condition $\min \{t_k\} > -1.65$ does not hold for $k = 2$, then the highest individual in G_k can be dropped from each subgroup and new subgroups $G_{2j} = \{2, \dots, j\}$ formed for $3 \leq j \leq N$. The step can be repeated with test statistics $t_j = t(G_{2j})$.¹³

Step 3 (Sieve the data for new club members) Add one country at a time to the core primary group with k^* members and run the $\log t$ test again. Include the new country in the convergence club if the associated t -statistic is greater than the criterion c^* . For small T ($T \leq 30$, say) the sieve criterion c^* can be set to zero to ensure selection is conservative, whereas for large T , c^* can be set to the asymptotic 5% critical value -1.65. Repeat this procedure for all the remaining countries. Form the first sub-convergence group from the core primary group supplemented by those countries satisfying the sign criterion. As c^* approaches to zero, the sieve condition becomes more conservative.

Step 4 (Recursion and stopping rule) Form a second group from those countries for which the sieve condition fails in Step 3. Run the $\log t$ test to see if $t_{\beta} > -1.65$ on this group, i.e. if this group satisfies the convergence test. If so, conclude that there are two convergence club groups: the core primary group and the second group. If not, repeat step 1 through step 3 to see if this second group can itself be subdivided into convergence

¹²The condition $\min \{t_k\} > -1.65$ plays a key role in ensuring that the null hypothesis of convergence is supported for each k .

¹³If the condition $\min \{t_j\} > -1.65$ is not satisfied for the first $j = 2$, the step may be repeated again, dropping the highest individuals in G_j and proceeding as before. If the condition does not hold for all such sequential pairs, then we conclude that there are no convergence subgroups in the panel.

clusters. If there is no k in Step 2 for which $t_k > -1.65$, conclude that the remaining countries do not contain a convergence subgroup and so the remaining countries have divergent behavior.

The sieve criterion can be set to $c^* = 0$ whenever we want to make particularly conservative selections into subgroups according to convergence characteristics (that is, selections with higher confidence of accuracy with respect to inclusion into subgroups). Subsequently, we may run $\log t$ regression tests across these subgroups to determine the evidence in support of merging clubs into larger clubs.

Table 2 shows the results of applying these clustering procedures to the PWT data involving 152 countries over the period 1970 - 2003. The table has three panels. The left panel (headed “Initial Classification”) shows the empirical results obtained from a direct application of the clustering algorithm described above. The second column under this heading reports the respective fitted regression coefficients $\hat{\beta}$ and HAC standard errors in parentheses. The algorithm classifies the country data into 6 subgroups. The leading 5 of these subgroups form convergence clubs. In each case the fitted regression coefficient is significantly positive, revealing strong empirical support for the club classification. The final group has a fitted coefficient that is significantly negative, thereby rejecting convergence and revealing evidence of divergence. For clubs 1 through 5, while the point estimates of β are all significantly positive, they are also significantly less than 2.0. So there is strong evidence of conditional convergence but little evidence of level convergence within each of these clubs.

The middle panel (headed “Tests of Club Merging”) reports the tests conducted to determine whether any of the original subgroups can be merged to form larger convergence clubs. We consider adjacent subgroups in the original classification and each cell in the panel reports the fitted regression coefficient and corresponding HAC standard error. Except for clubs 4 and 5, there is no evidence to support mergers of the original groupings. Hence, the first three subgroups are taken to form separate convergence clubs, while the aggregate of subgroups 4 and 5 constitute a fourth but somewhat weaker convergence club. Note that the point estimate $\hat{\beta} = -0.044$ of β is negative but is not statistically significant. In last subgroup (6), there appears to be no evidence of any convergent clubs and no evidence of merging behavior with club 5.

The right panel of Table 2 gives the final empirical classification from this clustering analysis into 4 growth convergence clubs and one divergent subgroup of countries. Below we report evidence in support of transitioning between these clubs and possible transitioning over time.

Figure 5 displays the country names arranged according to each of the 4 convergence clubs and the divergent group. The first convergence club (Club 1) consists of a large group of 50 countries. The main distinguishing feature of these countries is that they can be classified as either developed countries or as very fast growing countries. We also note that none of the

countries in Club 1 has experienced any lengthy or serious political disturbance since 1970. The second convergence club (Club 2) consists of 30 countries. There seems to be no single distinguishing characteristic of these countries. Some of the countries such as India, Brazil and Turkey are fast growing economies, especially during early 2000s. Meanwhile, several of the other countries have experienced episodes of financial and political disturbance since 1970 – for example, Mexico, Argentina, and some of the Eastern and Central European countries. Thus, while there is heterogeneity in country type and experience within this club, the evidence in support of economic growth convergence is substantial ($\hat{\beta} = 0.24$, $s.e.(\hat{\beta}) = 0.035$). The third convergence club (Club 3) consists of 21 countries. Compared with Club 2, all the economies in Club 3 are generally slowly growing, and many have experienced political disturbances. The last convergence club (Club 4) consists of 38 poor countries. All of them have experienced at least one of following problems: (i) long external wars; (ii) internal regional, ethnic and/or religious conflicts; (iii) corrupt and inefficient political systems, coups and military dictatorships. Group 5 consists of 13 depressed economies with serious ongoing ethnic or religious conflicts. Given recent historical experience, several nations in Club 4 might be expected to transition to group 5 once data is extended to 2007. For example, Iraq and Sudan continue to be ravaged by war, and Zimbabwe and North Korea continue to suffer from serious political corruption and inefficient economic management under dictators.

In sum, about one third of the 152 economies (those in Club 1) are developed and experiencing strong growth, one third (those in Club 4 and Group 5) are underdeveloped or seriously depressed with internal or external conflicts, and the remaining one third (those in Clubs 2 and 3) form a mixed collection of economies, some of which are growing slowly and others of which have experienced episodes of financial and political instability.

5.2 Economic Growth Transition

As the world economy and its constitution of nation states change, we may anticipate some evolution in the group clusters over time. The presence of 5 distinct groups indicates substantial diversity in economic performance among countries and raises the possibility of some transitioning between the groups. Many paths of transition are possible, depending on the individual circumstances of each country, its exposure to common world technology, and the extent to which it may have experienced or be undergoing political upheaval and ethnic or social conflict. The present section explores empirical evidence in the world economy for such transitioning in economic performance over time.

The clustering procedure outlined in the previous subsection allows us to examine evidence for transitioning between groups, or the possibility of sequential club convergence where part of one group moves towards another group. In particular, we can use the $\log t$ test to assess evidence of convergence between neighboring members of different club clusters. This test is

performed by running the $\log t$ test regression using data that includes a fraction (λ_1) of the lowest (in terms of final income) members in the upper club together with a fraction (λ_2) of the highest (in terms of final income) members in the lower club. In our present application, we set $\lambda_1 = \lambda_2 = 0.5$.

Table 3 shows the results of this empirical analysis. Panel A reports results of $\log t$ convergence tests across groups. The findings indicate strong evidence of transitioning across the top club clusters. In particular, the data show support for conditional convergence between the 15 highest countries in Club 2 and the 25 lowest countries in Club 1 ($\hat{\beta} = 0.465$, $s.e.(\hat{\beta}) = 0.049$). Similarly, there is support for conditional convergence between the 15 lowest countries in Club 2 and the 10 highest countries in Club 3 ($\hat{\beta} = 0.554$, $s.e.(\hat{\beta}) = 0.055$). Thus, these groups may be understood to be in a state of transition. There is no empirical evidence of transitioning behavior between Club 4 and Club 3. This may be explained by the fact that most of the economies in Club 4 have suffered from various internal and external conflicts and political and social instability, as discussed above.

Some comments on the interpretation of transition behavior are in order. One possible interpretation is that, although certain club clusters may have been identified and overall convergence among countries may have been rejected with given sample data, there remains the possibility that certain clubs may be slowly converging to one another. Another interpretation, and one that may be more realistic in the presence of substantial income diversity across countries, is that some countries within a certain convergence group may exhibit a tendency to be in transition towards a higher or lower club, leaving open the possibility of joining the new club in the future.

To examine evidence of the latter form of transitioning, we repeat the exercise above by resetting the data initialization from 1970 to 1960 so as to achieve a longer time series sample. In making this re-initialization, one third of the 152 PWT countries drop from the sample, giving a new country count in the subgroups as follows: Club 1 (17), Club 2 (12), Club 3 (7), Club 4 (12), and Group 5 (6). The results are shown in the left of Panel B in Table 3. Using this new panel of 98 countries over the period from 1960 to 1985, we find evidence to support the conclusion of four convergence clubs and a fifth divergent group, giving results that broadly correspond to those for the period 1970 – 2003. In reaching this determination, however, we find that the boundaries between the groups is not as clear for this earlier subperiod. As before, the algorithm originally classifies the country data into 6 subgroups and then some reduction is achieved on a further pass through the data, leading to 4 clubs and a residual divergent group. Taking the first two clubs together (Club 1 and Club 2) the regression test does not reject the null of convergence, but the test does reject the inclusion of Club 3 into this larger group. Similar patterns apply with subgroups 2 and 3, subgroups 3 and 4, and subgroups 4 and 5. Thus, there is greater ambiguity about the Club boundaries for the period 1960-1985.

In making the final determination, we have followed the previous rule and classified subgroups 1,2,3 as Clubs 1,2 and 3, while classifying the sum of subgroups 4 and 5 as Club 4, and leaving the residual group as a divergent subgroup.

The right part of Panel B in Table 2 reports the results of applying the same clustering algorithm to the same 98 PWT countries over the period 1970 - 2003. We now find evidence of 5 Clubs overall. Over time, the number count in each of the first 4 Clubs has risen and the number of countries in group 5 has fallen. Moreover, the 7 countries in group 5 satisfy the growth convergence test ($\hat{\beta} = 0.170$, $s.e.(\hat{\beta}) = 0.055$), so that this group is now classified as Club 5.

Figures 6 and 7 summarizes the transitional movements across country groups over time. These figures list the countries in each club based on the clustering results obtained for the period between 1960 and 1985. The figures further display the changes in club membership that have taken place by 2003. Figure 6 focuses on the transitions to Club 1, while Figure 7 tracks the remaining transitions. There are 10 new members in Club 1. About 40% of the countries in Club 2 move to Club 1 over the later time period. Thailand moved from Club 3 to Club 1, and three countries (Cape Verde, Chile and China) jumped to Club 1 from Club 4. Finally, Equatorial Guinea jumped from Group 5 to Club 1. There are 12 new members in Club 2 comprising 7 countries from Club 3, 4 countries from Club 4, and India, which moved from Group 5 to Club 2. Similarly, there are significant changes in the membership of Clubs 3 and 4. One country, the Congo Republic, slipped to Club 4 from Club 3. Meanwhile, 7 countries from Club 4 and 3 countries from Group 5 joined Club 3. Except for the 7 countries that remained in Group 5, 21 countries in this group joined Club 4.

In sum, therefore, there is some strong empirical support for economic transitioning and evolving membership of convergence clubs. Only for those particularly poor countries where overall growth rates are negative is there no evidence of economic transition from Group 5 to a convergence club. These countries seem to be caught in a “phase A” pattern of transitional divergence and seem not to be sharing in the global common growth component. With the short time series sample used here, of course, it is difficult to assess whether or not the negative growths sustained by some Sub-Saharan countries is permanent or transitory. By contrast, there is ample empirical evidence that emerging countries such as the Eastern European countries and parts of the former Soviet Union, have shown dramatic *U*-shape transition patterns of real income dynamics during the sample period, undergoing all three phases of transition in this short period.

6 Conclusion

As authors such as Durlauf and Quah (1999) have noted, the study of cross country economic growth often reveals more about heterogeneity in economic performance than it does about

convergence. Indeed, just as the distribution of income within nations displays inequality that evolves over time, the distribution of income across nations moves over time, often in ways that cannot be anticipated. Nonetheless, it is also evident that the benefits of modern technology are spreading across national borders and influencing economic performance. Of course, this diffusion occurs more quickly in some cases and for some countries than it does for others. Thus, while there are good reasons to expect some convergence in economic performance, especially with the growth of regional economic unions, there are also reasons to expect that the paths of transition in economic performance may be very different across nations. Indeed, in the process of observing nations over time, we observe many different forms of transitional behavior. Some groups of countries or economic regions behave in a similar way over time and appear to moving on a path towards some steady state growth pattern. Others appear to diverge over certain periods of time, fall behind and then turn around and show evidence of catching up.

This paper provides some mechanisms for thinking about such transitions, modeling them in a manner that is compatible with a neoclassical framework, and measuring them econometrically. To do so, we focus not on individual economic growth but on economic growth relative to the average performance in a subgroup of economies or an individual benchmark like that of the US economy which is relevant because it provides for an underlying growth pattern based on the latest common technology. This process enables us to identify the relative transitions that occur within these subgroups and to measure these transitions against the correlative of a common growth trend. Thus, in measuring a country's economic transition curve, we are able to assess its path over time relative to a useful benchmark. In this approach, the transition curve of an economy is an individual characteristic, allowing for many ways in which a neoclassical steady state can be approached, including the possibility of growth convergence clusters or even transitional divergence from the steady state.

The reality of economic transition raises questions about the relevant factors that influence transition. Just as a host of variables have been considered in analyzing the determinants of growth (Barro, 1997; Sala-i-Martin, 1997), there are similarly a large number of factors potentially influencing transition. These factors range over the many economic, social, cultural and political facets that characterize individual countries. In ongoing work the authors are using the methods of this paper to explore the manner in which such factors may influence economic transition. Just as there is transition in economic growth performance over time, we may also expect transition behavior (relative to some benchmark) in many of the factors that influence growth, such as human capital and educational attainment. Linkages between these two forms of transitional behavior contribute to our understanding of the time-forms of long run economic performance and the various transitions that individual economies experience. The methodology of the present paper is intended to advance this fundamental enquiry.

References

- [1] Barro, R. J. (1997). *Determinants of Economic Growth*. MIT Press.
- [2] Barro, R.J., and Sala-i-Martin (1992), “Convergence”, *Journal of Political Economy* 100 (2): 223-251.
- [3] Barro, R.J., and X. Sala-i-Martin (1995), *Economic Growth*. New York: McGraw-Hill.
- [4] Barro, R.J., and X. Sala-i-Martin (1997), “Technological Diffusion, Convergence, and Growth,” *Journal of Economic Growth*, 2, 1–26.
- [5] Becker, G.S., K.M. Murphy, and R. Tamura, (1990) “Human Capital, Fertility, and Economic Growth,” *Journal of Political Economy* 98 (5) 12–37.
- [6] Bernard, A.B, and S.N. Durlauf (1995), “Convergence in International Output”, *Journal of Applied Econometrics* 10(2): 97-108.
- [7] Bernard, A.B, and S.N. Durlauf (1996), “Interpreting Tests of the Convergence Hypothesis”, *Journal of Econometrics*, 71(1-2): 161-174.
- [8] Bianchi, M. (1997). “Testing for convergence: Evidence from Non-Parametric Multimodality Tests”. *Journal of Applied Econometrics*, 12, 393-409.
- [9] Carlino, G.A. and L.O. Mills (1993), “Are U.S. Regional Incomes Converging?: A Time Series Analysis,” *Journal of Monetary Economics*, 32, 335–346.
- [10] DeLong, J.B (1988), “Productivity Growth, Convergence, and Welfare: A Comment”, *American Economic Review* 78(5):1138-55.
- [11] Den Haan, W.J. (1995), “Convergence in Stochastic Growth Models: the Importance of Understanding Why Income Levels Differ”, *Journal of Monetary Economics* 35(1): 65-82.
- [12] Durlauf, S. N., P. A. Johnson and J. Temple (2004). *Growth Econometrics*. University of Wisconsin, mimeographed.
- [13] Durlauf, S.N. and D.T. Quah (1999), “The New Empirics of Economic Growth”, Ch. 4 in *Handbook of Macroeconomics 1A*, (Elsevier)
- [14] Easterly, W. (2001) “The Middle Class Consensus and Economic Development,” *Journal of Economic Growth*, Vol. 6 (4), 317–336.
- [15] Evans, P. and G. Karras (1996), “Convergence Revisited,” *Journal of Monetary Economics*, 37, 249–265.

- [16] Galton, F. (1886). "Regression toward mediocrity in hereditary stature." *Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246-263.
- [17] Galton, F. (1989) *National Inheritance*. London: Macmillan.
- [18] Guerre, E. (2004). "Design-Adaptive Pointwise Nonparametric Regression Estimation For Recurrent Markov Time Series," mimeo, Université Pierre et Marie
- [19] Hart, P. E. (1995) "Galtonian regression across countries and the convergence of productivity." *Oxford Bulletin of Economics and Statistics*, 57, 287-293.
- [20] Hodrick, R. and E. Prescott (1997). "Post war business cycles: An empirical investigation". *Journal of Money Credit and Banking*, 29, 1-16.
- [21] Howitt, P. and D. Mayer-Foulkes, (2005). "R&D, Implementation and Stagnation: A Schumpeterian Theory of Convergence Clubs," *Journal of Money, Credit and Banking*, 37, 147-177.
- [22] Kitagawa, G. and W. Gersch (1996). *Smoothness Priors Analysis of Time Series*. New York: Springer Verlag
- [23] Lucas, R.E. Jr. (2002) "The Industrial Revolution: Past and Future," in Lucas, R. E. Jr *Lectures on Economic Growth*. Cambridge, Massachusetts: Harvard University Press
- [24] Mankiw, N.G., D. Romer, and D.N. Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107, 407-38.
- [25] Parente, S.L. and E.C. Prescott (1994), "Barriers to Technology Adoption and Development," *Journal of Political Economy*, 102, 298-321.
- [26] Phillips, P. C. B. (2005). "Challenges of Trending Time Series Econometrics", *Mathematics and Computers in Simulation*, 68, 401-416.
- [27] Phillips, P.C.B. and H.R. Moon (1999) : Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, 67, 1057-1111.
- [28] Phillips, P. C. B. and S. Jin, (2002). "Limit Behavior of the Whittaker (HP) Filter". Yale University, mimeographed.
- [29] Phillips, P. C. B. and D. Sul (2003). "The elusive empirical shadow of growth convergence", Cowles Foundation Discussion Paper #1398, Yale University.
- [30] Phillips, P. C. B. and D. Sul (2007). "Transition modeling and econometric convergence tests", *forthcoming in Econometrica*

- [31] Quah, D. (1993), “Galton’s Fallacy and Tests of the Convergence Hypothesis,” *Scandinavian Journal of Economics* 95, 427-443.
- [32] Sala-i-Martin, X. (1997). “I just ran two million regressions” *American Economic Review*, 87, 178-183.
- [33] Sala-i-Martin, X. (2002), “The Disturbing “Rise” of Global Income Inequality,” NBER working paper, No. 8904.
- [34] Sun, Y. (2001), “Catching up, Forging ahead, and Falling Behind: A Panel Structure Analysis of Convergence Clubs”. UCSD mimeographed.
- [35] Whittaker, E. T. (1923). On a new method of graduation. *Proceedings of the Edinburgh Mathematical Association*, 78, 81-89.
- [36] Zhu, F. (2004). *Three Essays in Macroeconomic Empirics and Monetary Theory*. Ph.D Dissertation, Yale University.

Technical Appendix

1. Transition Paths under Heterogeneous Technology

We start by giving the transition path for $\log k_{it}$ under homogenous technological progress and then generalize to the case of heterogenous technological progress. Under homogeneous technological progress, the evolution of $k = K/AL$ under Cobb-Douglas production is governed by

$$\dot{k}_{it} = sk_{it}^\alpha - (n + x + \delta)k_{it}, \quad (29)$$

where s is the saving rate. The steady-state value k^* is obtained by setting $\dot{k}_{it} = 0$, giving

$$k^* = \left(\frac{s}{n + x + \delta} \right)^{1/(1-\alpha)}. \quad (30)$$

In the steady state, A_t grows at the constant rate x , so steady state income per capita (denoted by the affix s) is given by

$$\left(\frac{Y_t}{L_t} \right)_s = \left(\frac{K_t}{A_t L_t} \right)_s^\alpha A_t = \left(\frac{s}{n + x + \delta} \right)^{\alpha/(1-\alpha)} A_t, \quad (31)$$

and then

$$\log \left(\frac{Y_t}{L_t} \right)_s = \log A_0 + xt + \frac{\alpha}{1-\alpha} \log s - \frac{\alpha}{1-\alpha} \log (n + x + \delta). \quad (32)$$

The local transition path for $\log k_t$ in the neighborhood of the steady state is obtained by taking the first order Taylor expansion of (29), or equivalently

$$\frac{\dot{k}_{it}}{k_{it}} = sk_{it}^{\alpha-1} - (n + x + \delta),$$

about $\log k^*$. We have

$$\begin{aligned} sk_{it}^{\alpha-1} - sk^{*\alpha-1} &\simeq s(\alpha-1)k^{*\alpha-1}(\log k_{it} - \log k^*) \\ &= -\beta(\log k_{it} - \log k^*), \quad \text{say,} \end{aligned}$$

with $\beta = (1-\alpha)sk^{*\alpha-1} = (1-\alpha)(n+x+\delta)$. Hence, the local transition path is governed by the linear system

$$\dot{k}_{it}/k_{it} = -\beta(\log k_{it} - \log k^*). \quad (33)$$

When technological progress is heterogenous, the steady state value k^* does not change but the transition path for $\log k_{it}$ now depends on the transition path for x_{it}

$$\dot{k}_{it}/k_{it} = sk_{it}^{\alpha-1} - (n + x_{it} + \delta) = sk_{it}^{\alpha-1} - (n + x + \delta) - (x_{it} - x)$$

As above, we have the expansion

$$sk_{it}^{\alpha-1} - (n + x + \delta) \simeq -\beta(\log k_{it} - \log k^*),$$

which leads to the following dynamic system governing the local transition path about the steady state

$$\dot{k}_{it}/k_{it} = -\beta (\log k_{it} - \log k^*) - (x_{it} - x). \quad (34)$$

For an advanced economy where the technological growth rate has stabilized we have $x_{it} = x$, in which case the transitional path dynamics (34) is the same as (33). For a less developed economy, however, the general solution of (34) gives the following transitional path

$$\log k_{it} = \log k^* + d_{i0}e^{-\beta t} - e^{-\beta t} \int_0^t e^{\beta s} (x_{is} - x) ds,$$

which depends on the full past trajectory $\{x_{is}\}_{s \leq t}$ of the technology parameter and its deviation from x . For $t = 0$, we have the initialization $\log k_{i0} = \log k^* + d_0$, so that $d_{i0} = (\log k_{i0} - \log k^*)$, giving the explicit transition path

$$\log k_{it} = \log k^* + (\log k_{i0} - \log k^*) e^{-\beta t} - e^{-\beta t} \int_0^t e^{\beta s} (x_{is} - x) ds. \quad (35)$$

Setting $d_{i1} = 1/(\log k_{i0} - \log k^*)$, the transitional path (35) has the form

$$\begin{aligned} \log k_{it} &= \log k^* + (\log k_{i0} - \log k^*) \exp \left[-\beta t + \log \left\{ 1 - d_{i1} \int_0^t e^{\beta s} (x_{is} - x) ds \right\} \right] \\ &= \log k^* + (\log k_{i0} - \log k^*) \exp(-\beta_{it} t), \end{aligned} \quad (36)$$

where

$$\beta_{it} = \beta - \frac{1}{t} \log \left\{ 1 - d_{i1} \int_0^t e^{\beta s} (x_{is} - x) ds \right\}. \quad (37)$$

Using this solution, re-express $\log k_{it+q}$ as a function of the initial observation at q , viz., $\log k_{iq}$, so that

$$\log k_{iq} - \log k^* = (\log k_{i0} - \log k^*) \exp(-\beta_{iq} q),$$

and then

$$\begin{aligned} \log k_{it+q} &= \log k^* + (\log k_{i0} - \log k^*) \exp(-\beta_{it+q}(t+q)) \\ &= \log k^* + (\log k_{i0} - \log k^*) \exp(-\beta_{iq} q) \exp(\beta_{iq} q - \beta_{it+q}(t+q)) \\ &= \log k^* + (\log k_{iq} - \log k^*) \exp(-\beta_{it}^+ t), \end{aligned} \quad (38)$$

where $\beta_{it}^+ = \beta_{it+q}(t+q)/t - \beta_{iq} q/t$. Since

$$\begin{aligned} \beta_{it+q} &= \beta - (t+q)^{-1} \log \left\{ 1 - d_{i1} \int_0^{t+q} e^{\beta s} (x_{is} - x) ds \right\}, \\ \beta_{iq} &= \beta - q^{-1} \log \left\{ 1 - d_{i1} \int_0^q e^{\beta s} (x_{is} - x) ds \right\}, \end{aligned}$$

we have

$$\beta_{it}^+ = \beta - t^{-1} \log \left(\left\{ 1 - d_{i1} \int_0^{t+q} e^{\beta s} (x_{is} - x) ds \right\} \left\{ 1 - d_{i1} \int_0^q e^{\beta s} (x_{is} - x) ds \right\}^{-1} \right). \quad (39)$$

In cases where technology is convergent and (9) holds, the deviations $x_{is} - x$ are bounded and so $\beta_{it}^+ = \beta + O\left(\frac{\log t}{t}\right) \rightarrow \beta$ as $t \rightarrow \infty$.

2. The Inconsistency of Augmented Solow Regression

Suppose we work, as in (38) above, from time q as the first observation. The transition path for $\log y_{it}$ is given, as in (36), by

$$\log y_{iq} = \log \tilde{y}_i^* + \log A_{iq} + [\log \tilde{y}_{i0} - \log \tilde{y}^*] \exp(-\beta_{iq} q)$$

or

$$\log y_{iq} - \log \tilde{y}_i^* - \log A_{iq} = [\log \tilde{y}_{i0} - \log \tilde{y}^*] \exp(-\beta_{iq} q).$$

Using this expression, we can now express the transition path for $\log y_{it+q}$ in terms of the initialization $\log y_{iq}$ as follows:

$$\begin{aligned} \log y_{it+q} &= \log \tilde{y}_i^* + \log A_{it+q} + [\log \tilde{y}_{i0} - \log \tilde{y}^*] \exp[-\beta_{it+q}(t+q)] \\ &= [1 - \exp(-\beta_{it}^+ t)] \log \tilde{y}_i^* + \log A_{it+q} - \exp(-\beta_{it}^+ t) \log A_{iq} + \exp(-\beta_{it}^+ t) \log y_{iq}, \end{aligned}$$

where

$$\beta_{it}^+ = \beta_{it+q} \left(\frac{t+q}{t} \right) - \beta_{iq} \frac{q}{t}.$$

Hence we have

$$\begin{aligned} \frac{\log y_{it+q} - \log y_{iq}}{t} &= \left[\frac{1 - \exp(-\beta_{it}^+ t)}{t} \right] \log \tilde{y}_i^* - \left[\frac{1 - \exp(-\beta_{it}^+ t)}{t} \right] \log y_{iq} \\ &\quad + \frac{1}{t} \log A_{it+q} - \frac{1}{t} \exp(-\beta_{it}^+ t) \log A_{iq}. \end{aligned} \quad (40)$$

Next, let $t = 1$, and define the augmented speed of convergence parameter

$$\beta_i^+ = \beta_{iq+1} + (\beta_{iq+1} - \beta_{iq}) q. \quad (41)$$

The transition equation (40) then has the form

$$\log \left(\frac{y_{iq+1}}{y_{iq}} \right) = [1 - e^{-\beta_i^+}] \log \tilde{y}_i^* - [1 - e^{-\beta_i^+}] \log y_{iq} + \log A_{iq+1} - e^{-\beta_i^+} \log A_{iq}. \quad (42)$$

Now we assume that, for given t , the β_{it} comprise independent and identically distributed draws from some distribution with finite moment generating function. In particular, suppose that over i , the β_{it} are

$$\beta_{it} \sim iid(\beta, \sigma_t^2),$$

where the variance, σ_t^2 , of β_{it} is time varying and satisfies $\sigma_t^2 \rightarrow 0$ as $t \rightarrow \infty$, so that $\beta_{it} \rightarrow_p \beta$ and $x_{it} \rightarrow_p x$ as $t \rightarrow \infty$. Correspondingly, for β_i^+ we have

$$\beta_i^+ \sim iid(\beta, \sigma^2),$$

where $\sigma^2 = (q+1)^2 \sigma_{q+1}^2 + q^2 \sigma_q^2 - 2q(q+1) \sigma_{q,q+1}$ and $\sigma_{q,q+1} = cov(\beta_{iq+1}, \beta_{iq})$. Let the moment generating function of β_i^+ be $m(r) = E(e^{\beta_i^+ r})$ and define

$$b = E(e^{-\beta_i^+}) = m(-1).$$

In the case where the distribution of β_i^+ is normal we have $b = E(e^{-\beta_i^+}) = \exp(-\beta + \frac{\sigma^2}{2})$.

We may rewrite (42) as

$$\log\left(\frac{y_{iq+1}}{y_{iq}}\right) = (1-b) \log \tilde{y}_i^* - (1-b) \log y_{iq} + \log A_{iq+1} - b \log A_{iq} + u_i \quad (43)$$

where

$$\begin{aligned} u_i &= [\exp(-\beta_i^+) - b] [\log y_{iq} - \log \tilde{y}_i^* - \log A_{iq}] \\ &= [\exp(-\beta_i^+) - b] [\log \tilde{y}_{i0} - \log \tilde{y}_i^*] \exp(-\beta_{iq} q) \end{aligned} \quad (44)$$

Suppose that an econometrician has proxy variables z_i for $\log \tilde{y}_i^*$ based on the following reduced form model for the steady state log levels:

$$\log \tilde{y}_i^* = \mathbf{z}_i \mathbf{a}_2 + v_i, \quad (45)$$

where v_i is independently and non-identically distributed as *i.n.i.d*(0, V_i). Combining (43) and (45), we obtain the following cross section augmented regression equation

$$\log\left(\frac{y_{iq+1}}{y_{iq}}\right) = a_0 + (1-b) z_i' a_2 - (1-b) \log y_{iq} + e_i, \quad (46)$$

where the error in the regression equation is given by

$$e_i = \log A_{iq+1} - b \log A_{iq} + u_i + (1-b) v_i. \quad (47)$$

Model (46) has the form of a traditional augmented Solow regression that has become popular in empirical applications. Both the proxy regressors z_i and the regressor $\log y_{iq}$ in (46) are correlated with the regression error e_i in view of (44) and (45). Note also that, in addition to this endogeneity effect, there is an omitted variable effect and a further endogeneity effect due to the presence of $\log A_{iq}$, and $\log A_{iq+1}$ in (47). In view of these sources of contamination, least squares regression on (46) suffers from both endogeneity and omitted variable bias.

3. Club Convergence Limits

Consider the case where there are two clubs of economies G_c and G_d with N_c and $N_d = N - N_c$ economies in each club and suppose $\frac{N_1}{N} \rightarrow \lambda \in (0, 1)$ as $N \rightarrow \infty$. Under club convergence, we have

$$b_{it} \rightarrow \begin{cases} b_c & i \in G_c \\ b_d & i \in G_d \end{cases},$$

for some b_c and b_d , so that

$$b_c = \lim_{N \rightarrow \infty} N_c^{-1} \sum_{i \in G_c} b_{it}, \quad b_d = \lim_{N \rightarrow \infty} N_d^{-1} \sum_{i \in G_d} b_{it},$$

and

$$h_{it} = \frac{b_{it}}{N^{-1} \sum_i b_{it}} \rightarrow \begin{cases} \frac{b_c}{\lambda b_c + (1-\lambda) b_d} & i \in G_c \\ \frac{b_d}{\lambda b_c + (1-\lambda) b_d} & i \in G_d \end{cases}.$$

Then

$$H_t = N^{-1} \sum_{i=1}^N (h_{it} - 1)^2 \rightarrow \frac{\lambda(1-\lambda) \{ \lambda b_c^2 + (1-\lambda) b_d^2 \}}{\{ \lambda b_c + (1-\lambda) b_d \}^2},$$

for all $\lambda \neq 0, 1$ and $b_c \neq b_d$. A similar weighted limit is obtained in the case of multiple clubs.

Data Appendix

Three panel data sets of log per capita real income are used in the paper. The first panel (A) relates to the 48 contiguous United States from 1929 to 1998 (Source: Bureau of Economic Analysis). The second panel (B) consists of 127 countries from 1950 to 2001. (Source: OECD The World economy: historical statistics). From the same OECD data source we also collected the long historical data set (the Maddison Data set) for 18 Western OECD countries covering the period from 1500 to 2001. The third panel (C) includes 152 countries from 1970 to 2003 and 98 countries from 1960 to 2003 (Source: PWT 6.2). The remaining three subsections of this Appendix show how subgroups for these panels were created based on regional location and geographical distance.

Panel Data Set A: Regional US State Classifications

Mid-Atlantic: Delaware, Maryland, New Jersey, New York, Pennsylvania

New England: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont

Great Lakes: Illinois, Indiana, Michigan, Ohio, Wisconsin

Mountain: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming

Pacific: California, Oregon, Washington

Plain States: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota

South Atlantic: Florida, Georgia, North Carolina, South Carolina, Virginia, West Virginia

West South Central: Arkansas, Louisiana, Oklahoma, Texas

East South Central: Alabama, Kentucky, Mississippi, Tennessee

Panel Data Set B: OECD The world economy – historical statistics.

18 Western OECD Countries: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, United Kingdom, Portugal, Spain, Australia, New Zealand, Canada, United States

Panel Data Set C: PWT Version 6.2 Countries

The following subgroups are formed based on geographical location.

Latin American and Caribbean Countries: Argentina, Barbados, Brazil, Chile, Colombia, Costa Rica, Dominican Republic, Ecuador, El Salvador, Jamaica, Mexico, Panama, Paraguay, Trinidad & Tobago, Uruguay, Venezuela, Bolivia, Guatemala, Honduras, Nicaragua, Peru

Middle East and North African Countries: Algeria, Egypt, Israel, Iran, Morocco, Jordan, Syria

Sub-Saharan Africa Countries: Benin, Burkina Faso, Burundi, Cape Verde, Chad, Comoros, Rep. Congo, Cote d'Ivoire, Equatorial Guinea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Niger, Nigeria, Rwanda, Senegal, South Africa, Tanzania, Togo, Uganda, Zambia, Zimbabwe

19 OECD Countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Iceland, Italy, Japan, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom, USA

4 Dragons: Hong Kong, South Korea, Singapore, Taiwan

3 NICs: Indonesia, Malaysia, Thailand

Table 1: Convergence Tests

Cases	Time	$\hat{\beta}_1$	$s.e.(\hat{\beta})$
48 U.S. States	1929-1998	0.46	0.04
18 Western OECD	1870-2001	1.71	0.07
	1870-1929	-0.42*	0.04
	1911-1970	-0.11	0.09
	1940-2001	1.14	0.01
152 PWT	1970-2003	-0.88*	0.01

Note: (*) significant at the 5% level.

**Table 2: Convergence Club Classification:
152 PWT Countries from 1970 to 2003**

Initial Classification β (s.e)	Tests of Club Merging β (s.e)				Final Classification β (s.e)
Club 1 [50] 0.382 (0.041)	Club 1+2 -0.051* (0.023)	Club 2+3 -0.104* (0.016)	Club 3+4 -0.192* (0.038)	Club 4+5 -0.044 (0.070)	Club 1 [50] 0.382 (0.041)
Club 2 [30] 0.240 (0.035)					Club 2 [30] 0.240 (0.035)
Club 3 [21] 0.110 (0.032)	Club 3 [21] 0.131 (0.064)				
Club 4 [24] 0.131 (0.064)	Club 4 [38] -0.044 (0.070)				
Club 5 [14] 0.190 (0.111)	Club5+Group 6 -0.881* (0.027)				
Group 6 [13] -1.116* (0.060)		Group 5 [13] -1.116* (0.060)			

Notes:

*) Reject the null hypothesis of growth convergence at the 5% level

The numbers in brackets, [], stand for the number of countries in a group.

The initial clustering suggests 5 sub convergence groups and 1 divergent group.

Tests for club mergers lead to the final classification of 4 convergence clubs and 1 divergence group.

Table 3: Economic Growth Transitions

Panel A: Transition between Clubs (152 PWT Countries from 1970 to 2003)

Club numbers	β (s.e)
Club 1 [lower 25] + Club 2 [upper 15]	0.465 (0.049)
Club 2 [lower 15] + Club 3 [upper 10]	0.554 (0.055)
Club 3 [lower 11] + Club 4 [upper 19]	-0.153* (0.019)

**Panel B: Club Transition over Time
(98 PWT Countries from 1960 to 2003)**

From 1960 to 1985: β (s.e)		From 1970 to 2003: β (s.e)	
Club 1 [25]	0.343 (0.060)	Club 1 [33]	0.317 (0.057)
Club 2 [10]	0.229 (0.090)	Club 2 [18]	0.228 (0.038)
Club3 [12]	0.255 (0.072)	Club 3 [14]	0.165 (0.034)
Club 4 [19]	0.174 (0.072)	Club 4 [26]	0.045 (0.075)
Group 5 [32]	-0.687 (0.007)*	Club 5 [7]	0.170 (0.055)

Notes:

*) Reject the null hypothesis of growth convergence at the 5% level
The numbers in brackets, [], stand for the number of countries in a group.

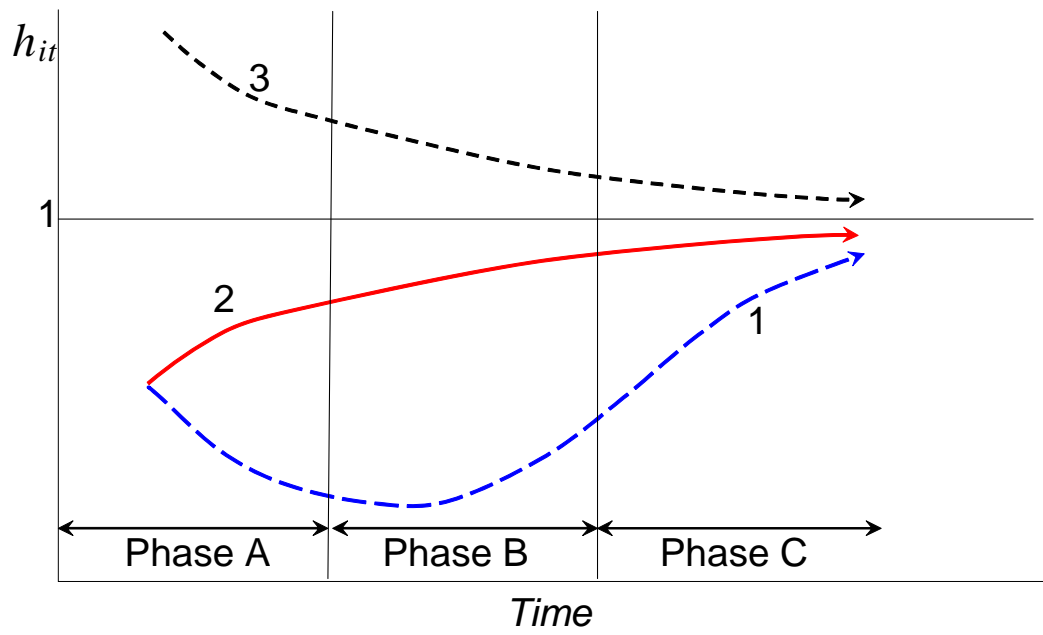


Figure 1: Relative Transition Curves h_{it} and Phases of Transition

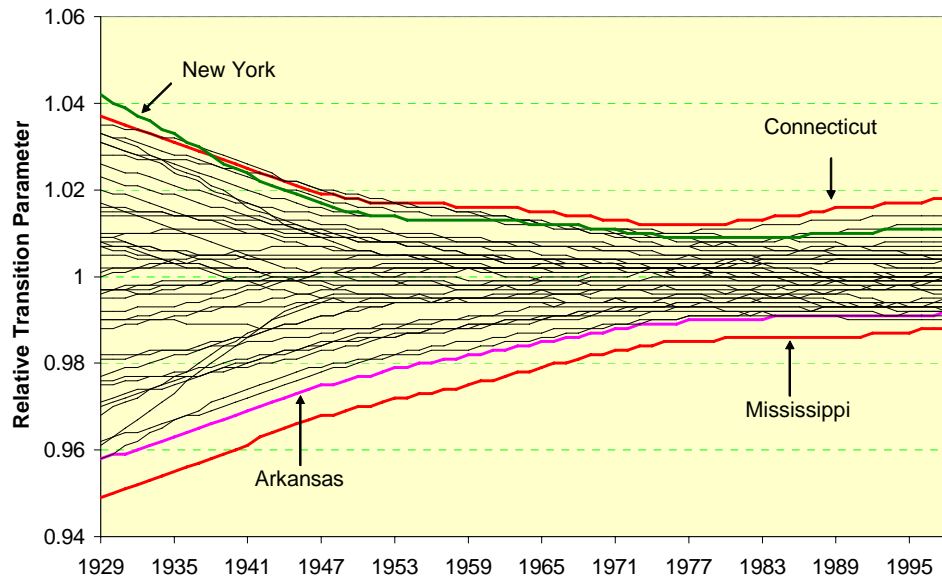
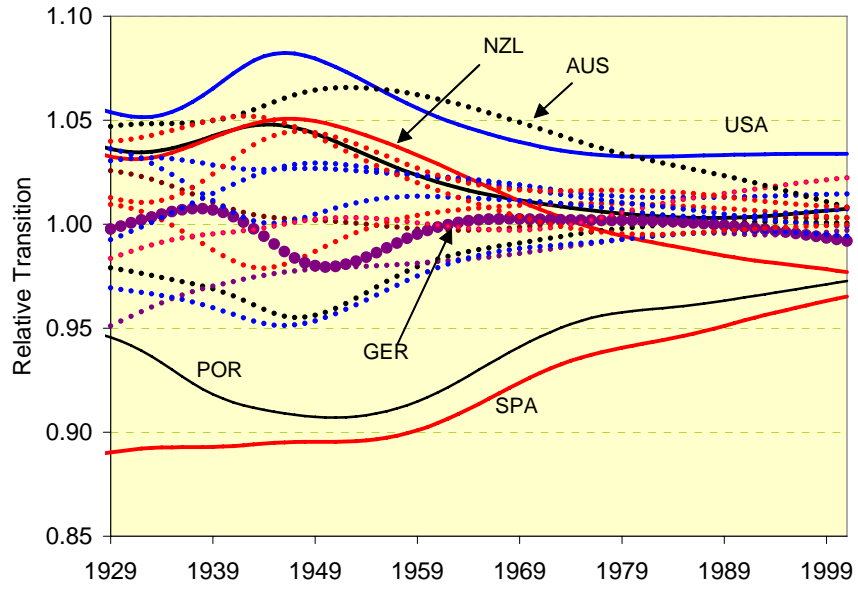
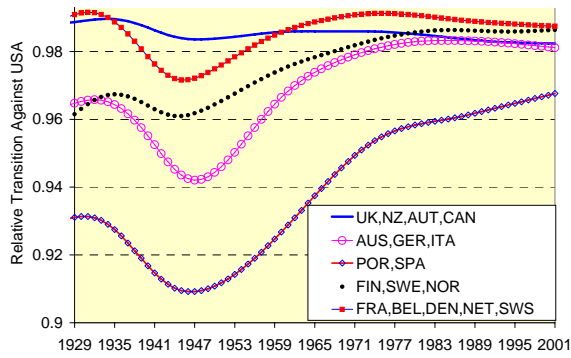


Figure 2: Transition Paths for the 48 Contiguous US States.

Panel A: 18 Western OECD countries from 1929-2001



Panel B: Evidence of Phase B & C



Panel C: Evidence of Phase A and B

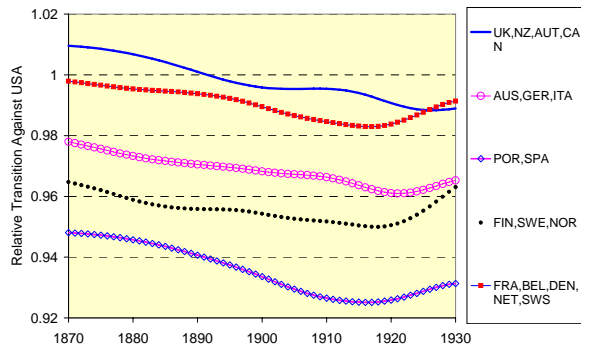
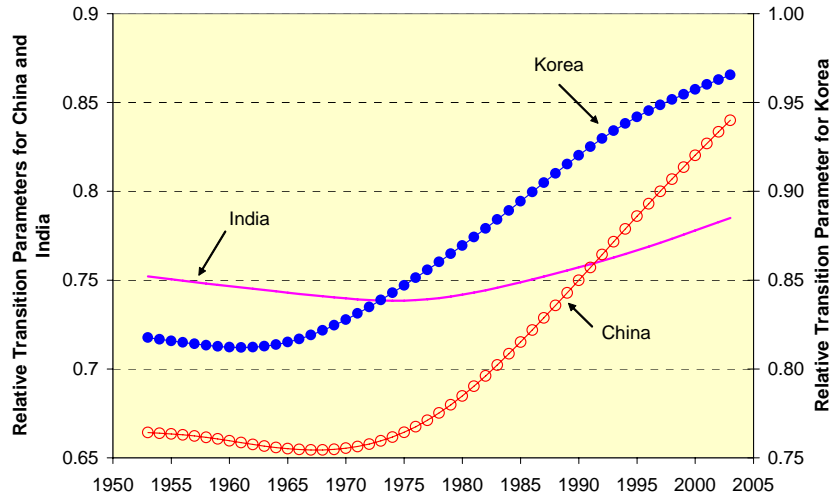
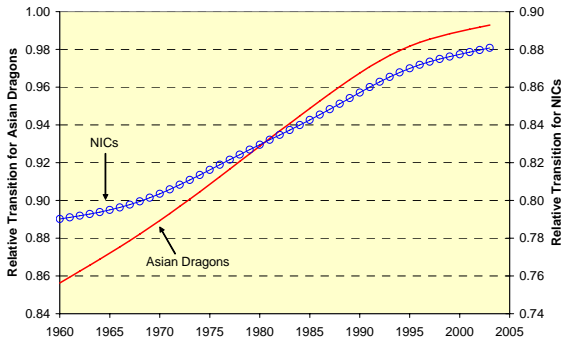


Figure 3: OECD Transition Paths

Panel 1: Examples of Phase B & C Transitions



Panel 2: Examples of Phase C Transitions



Panel 3: Examples of Phase A Transitions

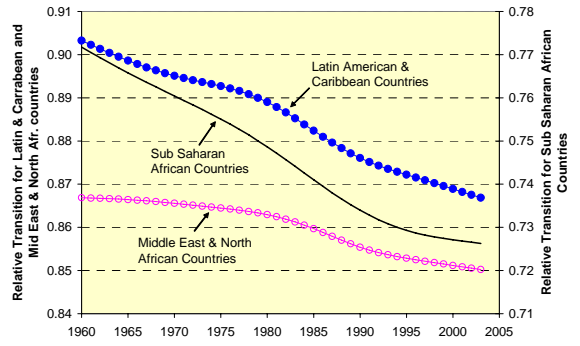


Figure 4 World Transition Paths

Figure 5: Clustering Convergence Clubs: (152 PWT Countries from 1970 to 2003)

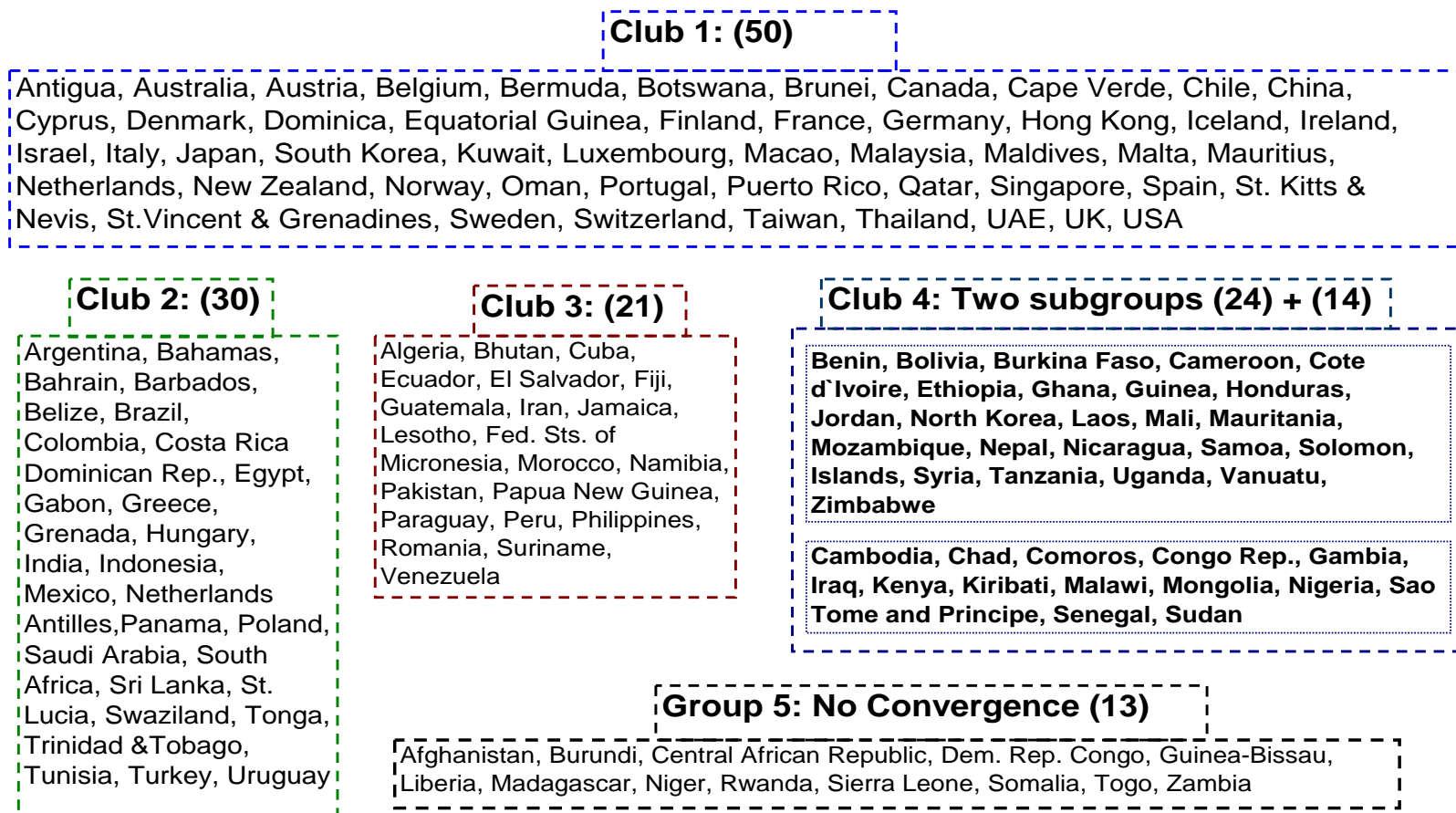


Figure 6: Transitioning to Club 1

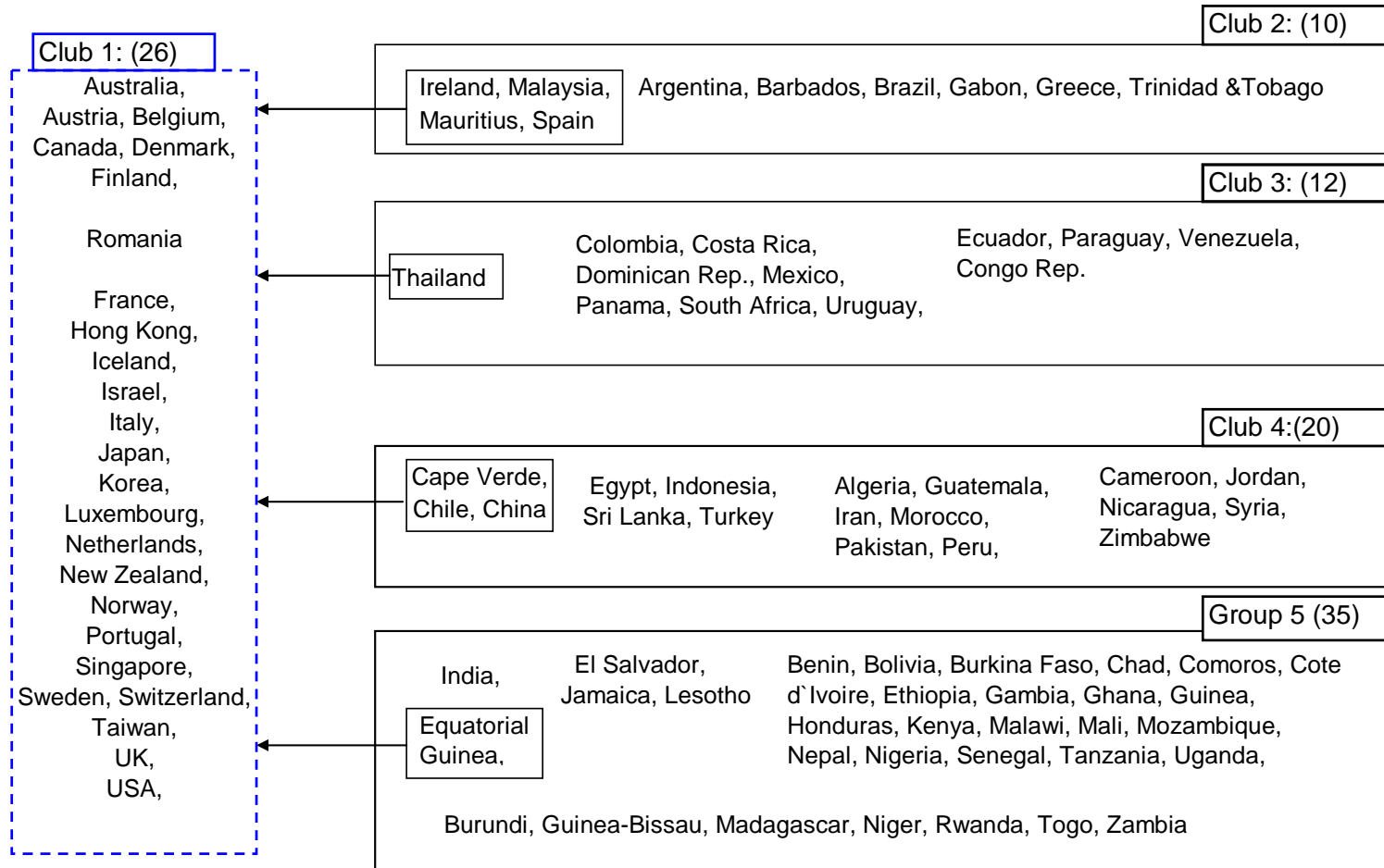


Figure 7 –Transitioning to Clubs 2, 3 and 4

