ECONOMIC TRANSITION AND GROWTH

By

Peter C.B. Phillips and Donggyu Sul

June 2005

COWLES FOUNDATION DISCUSSION PAPER NO. 1514

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

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

June 14, 2005

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 and measuring them econometrically are developed in the paper. A new regression test of convergence is proposed, its asymptotic properties are derived and some simulations of its finite sample properties are reported. 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: Economic growth, Growth convergence, Heterogeneity, Neoclassical growth, Relative transition, Transition curve, Transitional divergence.

JEL Classification Numbers: 030; 040; C33.
First Completed Draft: January 2005

*Phillips gratefully acknowledges research support from a Kelly Fellowship at the Business School, University of Auckland, and the NSF under Grant No. SES 04-142254.

"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 occured 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. These arguments are supported 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 start by considering the impact of cross sectional and temporal heterogeneity on the speed of (beta) convergence in a traditional growth convergence setting such as that considered in Barro and Sala-i-Martin (1992). In such a setting, the transition dynamics of log per capita real income y_{it} in country i at time t has the following simple form

$$\log y_{it} = a_i + b_i e^{-\beta t} + xt, \tag{1}$$

where β represents the speed of convergence and may be taken to be a function of the growth rate of technological progress x, amongst other factors. The definition of 'technological progress' can, of course, be rather broad and may include social, political, cultural, scientific, engineering and economic factors. The first term, a_i, in (1) incorporates initial conditions and steady state levels. In this model, under homogeneity of β and x, 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 (leading to 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 the cross country income inequality transient, as argued by Lucas.

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 addressed in the present paper, it is hoped that the methods developed here for studying empirical economic transitions in growth performance will be relevant in addressing similar issues regarding the transition behavior of the many 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 conjuction 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 $h_{it} = \log y_{it} / \log \bar{y}_t$, where $\log \bar{y}_t$ denotes the cross section average of log per capita real income in the panel or a suitable subgroup of the panel[1]. 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 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 3 derives some stylized facts concerning long term growth patterns across countries based on average real per capita income for 18 Western OECD countries over the past 500 years. This section also considers the effects of technological adaptation and learning on the time forms of economic transition. 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. Section 5 develops an econometric formulation of this concept, which provides the time profile of transition for one economy relative to a group average. This relative transition curve is identified and can be fitted using various smoothing methods, which we discuss. The fitted transition curves can be used to reveal evidence on

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

central issues such as growth convergence and the possibility of transient divergent behavior. A regression test is developed to conduct formal econometric tests of this behavior. Empirical applications of these methods are reported in Section 6, 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 7. Further technical material, asymptotic justifications, and information on the data are given in the Appendices.

2 Heterogeneous Progress of Technology and Growth

We start from the neoclassical theory of growth convergence and attempt 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

$$\tilde{y} = f(\tilde{k}), \ \tilde{y} = Y/LHA, \ \tilde{k} = K/LHA, \ y = \tilde{y}HA = \tilde{y}A \tag{2}$$

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 (2), 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

$$A_{it} = A_{i0}e^{xt}, \tag{3}$$

where the growth rate of technology is common across countries. The latter condition 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 $t \to \infty$. In such a case, the evolution of A_{it} is inevitably more complex than (3), thereby accommodating a wider range of possible growth behavior. This motivation underlies the framework for empirical analysis that we develop later in this paper.

Let us assume that technology is a public good, is widely available and is represented at time t by a common technology variable C_t whose time profile follows

$$C_t = C_0e^{\xi t}. \tag{4}$$

For developed countries, the full extent of common technology C_t is taken to be instantly accessible. Indeed, it may be presumed that these countries created C_t and are materially

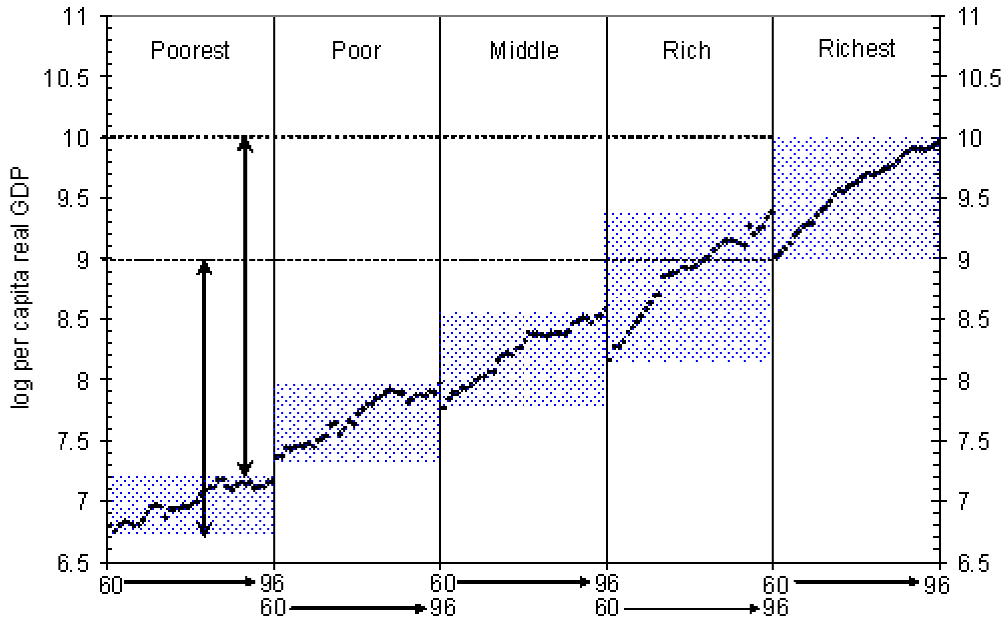


Figure 1: **A Taxonomy of World Economic Growth over 1960-1996.** The arrowed distances measure $\sigma-$ convergence and show little evidence supporting this form of the convergence hypothesis. The thickness of the shadowed areas can be used to assess $\beta-$convergence. There is apparently some evidence of $\beta-$convergence between the rich and richest country groupings. Taking all the country groupings together, the entire path has a similar form to the time path of average OECD income over the historical time frame of 500 years shown in Fig. 3. Details of the country groupings used in the figure are given in the Appendix.

involved in determining its future time path. Followers, like the developing nations, generally have to learn earlier technology first and develop an infrastructure to absorb and utilize it. As a result, it may be assumed that such countires cannot fully share in the present level of C_t. Depending on the speed of learning in these countries and the time form of their 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

$$A_{it} = C_t^{\delta_{it}} = A_{i0} e^{x_{it} t} = A\left(x_{it}, t, A_{i0}\right), \quad \text{say.} \tag{5}$$

The technological growth rate of economy i is now $x_{it} + t\dot{x}_{it}$ and is time dependent. This formulation means that technological learning differs across countries and over time even though there is a common underlying technology. These differences among economies allow for phenomena such as technological catch-up and slow-down, which are known to be important in empirical work.

Using this framework and a Cobb-Douglas technology in (2), the transitional growth path for country i is shown in Phillips and Sul (2005) to be

$$\log y_{it} = \log \tilde{y}_i^* + \left[\log \tilde{y}_{i0} - \log \tilde{y}_i^*\right] e^{-\beta_{it} t} + \log A_{it}, \tag{6}$$

which is an extension of (1), where y_i is per capita real income, $\tilde{y}_i^*$ is the corresponding steady state level, and the speed of convergence parameter β_i is functionalized as

$$\beta_{it} = \beta(\underset{-}{\alpha_i}, \underset{+}{\delta_i}, \underset{+}{v_i}, \underset{+}{x_{it}}; t). \tag{7}$$

In this formulation, α_i denotes the technology parameter in the Cobb-Douglas function, δ_i is the rate of depreciation and v_i is the population growth rate. Appropriate sign effects are indicated beneath these parameters in (7).

The term involving $e^{-\beta_{it} t}$ in (6) decays as $t \to \infty$ provided the condition

$$\beta_{it} t \to \infty \tag{8}$$

holds, in which case the path of log per capita real income is primarily dependent on the term $x_{it} t$ and may therefore be substantially affected by heterogeneity in technology progress over time and across economies. Following Bernard and Durlauf (1996), growth convergence may be defined as[2]

$$\lim_{k \to \infty} \left(\log y_{it+k} - \log y_{jt+k}\right) = 0. \tag{9}$$

[2]For the time being, all variables are taken to be non-stochastic, so a simple limiting operation is used in (9) in place of the limit of a conditional expectation as in Bernard and Durlauf (1996).

Thus, growth convergence requires that log per capita real income be the same across countries in the long run. A necessary condition for (9) in a model with heterogeneous transitional technology like (5) is

$$x_{it+k} \to x, \quad \text{for all } i \text{ as } k \to \infty. \tag{10}$$

Even when there is ultimate growth convergence of the type implied by (10), transient divergence in growth patterns may still occur – for instance, when an initially poor country adopts technology more slowly than rich countries create new technology. During this period the speed of learning in the poor economy is slower than the speed of technological creation in the rich economy and transitional divergence may occur. Subsequently, as x_{it} rises and the speed of learning picks up, the poor economy may begin to catch up with the richer economies.

3 Empirical Regularities and Economic Transition

We now proceed to develop some stylized empirical regularities concerning long term growth patterns across economies and use these to demonstrate the practical import of temporal and cross section heterogeneity in the progress of technology. We look at two separate bodies of evidence dealing with world and OECD economic growth and provide some graphical analysis of hypothetical transition effects, relating these to the actual growth paths of certain regional groupings of world economies.

First, studies by Durlauf and Quah (1999), Easterly (2001), Sun (2001) and others have recently raised doubts about the empirical evidence for growth convergence across the world economies. In place of convergence, there is evidence of emerging convergence clubs. In particular, the richest countries appear to be growing more slowly than some of the newly developed countries, such as the Asian Dragons and some other rapidly growing developing countries like China, whereas the remaining world economies appear to be growing at similar rates or slower rates than the rich countries. These ongoing differences in growth rates make the idea of convergence clubs and emerging multi-modality in the world distribution of income appealing.

Fig. 1 provides a new way of looking at some of the evidence on convergence and growth. This figure shows five groupings of cross sectional averages of log per capita real income for 88 countries from the PWT over 1960 to 1996 (a data set that has frequently been used in empirical work).

The subgroupings are based on initial income and the number of countries in each of the first four groups is 17 while the richest group number 20 countries. The time paths of the subgroup averages are shown over the five successive panels in the figure. Each panel covers the same 37 year period. While each panel restarts the time profile from 1960 onwards, the arrangement of the panels produces an escalator effect from the poorest to the richest groups

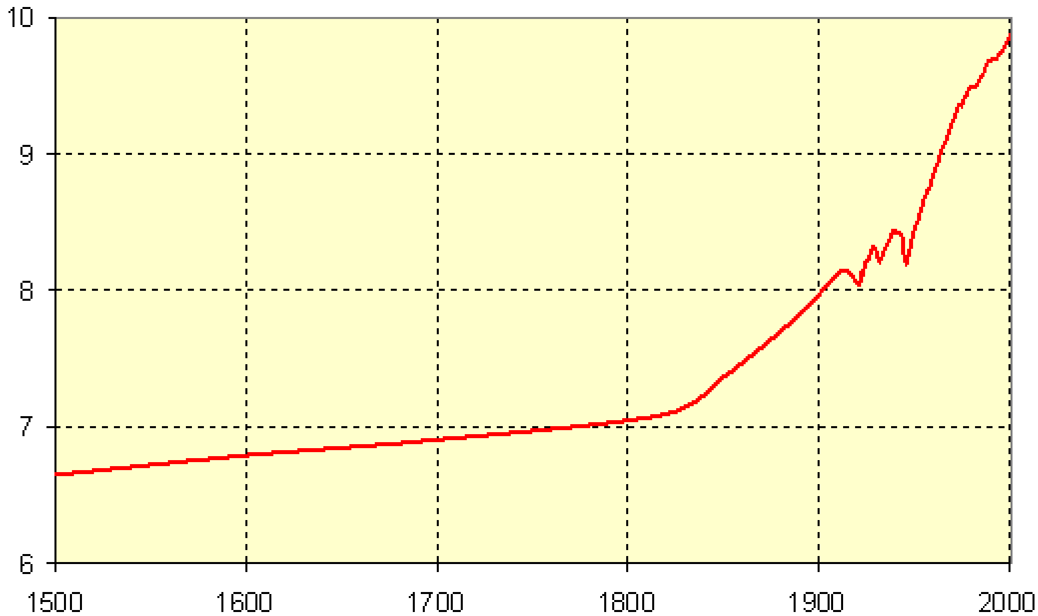


Figure 2: Cross sectional average of log per capita real income for 18 Western OECD countries over 500 years.

that is surprisingly connected in form. The escalator begins with a stair that has a fairly flat shape corresponding to the slow growth of the poorest nations and the stairs generally become steeper as the nation groups become richer and grow faster.

The content of Fig. 1 also provides summary evidence on growth convergence over 1960-1996. First, the heights of the shadowed areas in the figure measure total economic growth between 1960 and 1996 for the average in each country group and therefore provide information about β- convergence, according to which initially poor countries will grow faster than initially rich countries enabling the poor countries to catch up. Evidence on the heights of the shadowed areas in the figure does not seem to give any general support to β- convergence. While the 'rich' countries do appear to be catching up with the 'richest' nations, the rest of world economies appear not to be growing fast enough to catch up with either the 'rich' or the 'richest' countries. The so-called σ-convergence does not seem to have much empirical support either. The heights of the two arrows in the figure indicate cross sectional income disparity in 1960 and 1996, respectively. Evidently, using this criterion, cross sectional income dispersion in the world seems to be widening rather than narrowing, at least over this time frame. Thus, Fig. 1's broad visual evidence on world economic growth and its disparities appears to be consonant with the conclusions reached by Durlauf and Quah (1999) and others in the articles cited

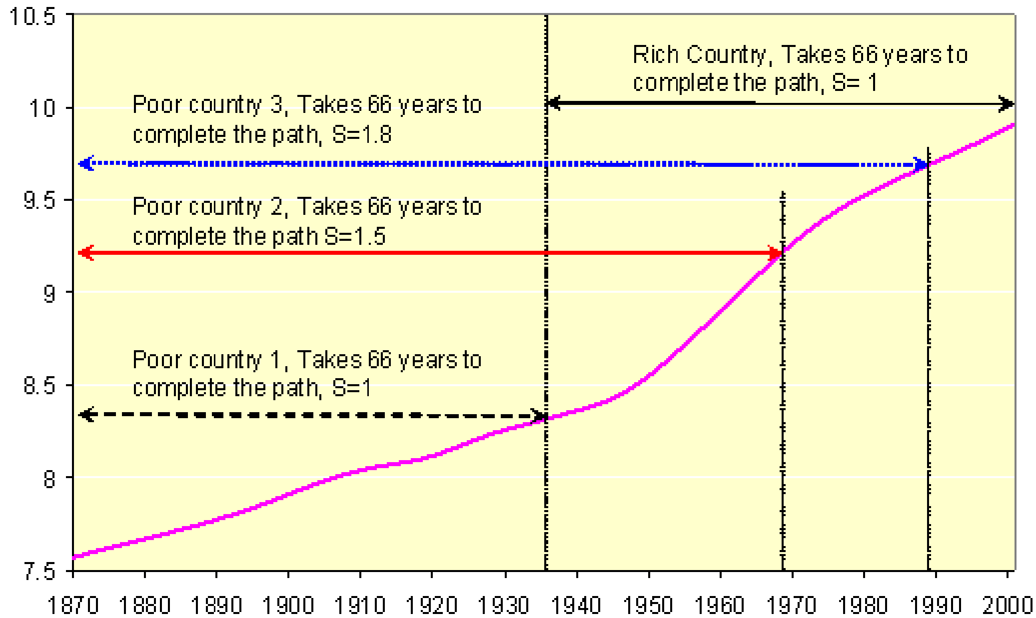


Figure 3: **Anatomy of Historical OECD Growth over 1870-2001.** The curve shows the effects of two actual (rich country and poor country 1) and two hypothetical (poor countries 1 and 2) learning rates for technological progress.

above.

Notwithstanding this evidence, we may still address the question whether cross section economic divergence is a transient phenomenon. Transitional divergence in economic growth may ocur when there is temporal heterogeneity of technology progress. That is, when x_{it} evolves over time. As is clear from Fig. 1, the heights of the shadowed areas – which measure economic growth over a 37 year time span for each group average – are positively correlated with the level of initial income in the group. Moreover, the ordering of total economic growth within the first four groups (i.e. all groups except the 'richest') over 1960-1996 matches exactly the ordering of initial income in the group. If such a pattern of growth were to persist over the next 37 years, then the lower income groups (i.e., the 'poorest', 'poor', and 'middle' groups) may grow faster than they have over the last 37 years (as they transition to a higher category), whereas the 'rich' countries may not grow faster as they transition into the 'richest' category (just as the 'richest' group have grown on average slower than the 'rich' group over 1960-1996). If this process were to continue, then eventually per capita real income across the world economies would narrow and recent evidence of cross sectional heterogeneity would then be viewed as transitional. This reasoning appears to support the observation made by Lucas

(2002) in the headnote of this article about the transience of income inequality. But much remains uncertain in this calculation, including the transitioning between groups and the time frame of the transitions.

As a second body of evidence, we take the cross section average of log per capita real incomes for 18 OECD countries from 1500 to 2001 and plot the data[3] against time in Fig. 2. Between 1500 and 1800, economic growth was very slow compared to the subsequent periods of the industrial revolution in the nineteenth century and the scientific revolution over the twentieth century. It is intriguing that this new figure which is based on 500 years of data bears more than a passing resemblance to Fig. 1 which is based on only 37 years of data but involves a much wider distribution of world economies. This resemblance suggests that there may be information in the long historical economic performance of (now advanced) OECD countries that is manifest in the world income distribution in recent decades.

We offer one possible answer to this question by considering the pattern of average OECD growth over the last two centuries more closely. We will use the observed historical pattern to suggest some hypothetical examples that shed light on possible forms of transitional economic performance. The examples are based on modifying the actual historical pattern to produce hypothetical economies with differing rates of technological progress. We may focus on (temporal and cross section) heterogeneity in x_{it} because model (6) itself suggests that for long historical series the term involving the exponential decay factor $e^{-\beta_{it}t}$ may be neglected and the growth path may be regarded as being principally determined by technological progress x_{it} through the term $\log A_{it}$.

To proceed, we consider Fig. 3, which plots the cross section average of $\log y_{it}$ across 18 OECD countries[4] over 132 years between 1870 and 2001 after eliminating business cycle components. This time profile of historical OECD growth is used to explain how transitory divergence in economic growth patterns can occur. Suppose that we observe two groups of economies (rich and poor) between time q and $T+q$ where $T = 66$. We split the cross sectional average of $\log y_{it}$ into two parts. The first part starts in $q = 1870$ (which we now use to represent the poor group) while the second part starts from $q = 1936$ (which we use to represent the rich group). We denote the speed of technological learning or creation by S. Our normalization is that if $S = 1$ then either the rich or the poor country takes $T = 66$ years to complete the given growth path which is based on the historical OECD record. The initial incomes for the two groups are approximately \$2,000 and \$4,000 dollars, respectively. The rich group is assumed to grow along the given growth path shown in Fig. 3. Now consider three hypothetical poor countries in the poor group, two of which will involve time dilation effects to capture differing rates of technological progress. Poor country 1 (which is based on actual OECD performance) takes 66 years to complete the growth path shown, so that this country's speed of learning is

[3] The Data Appendix provides details of the data and data sources.

[4] See the Data Appendix for details of the countries included.

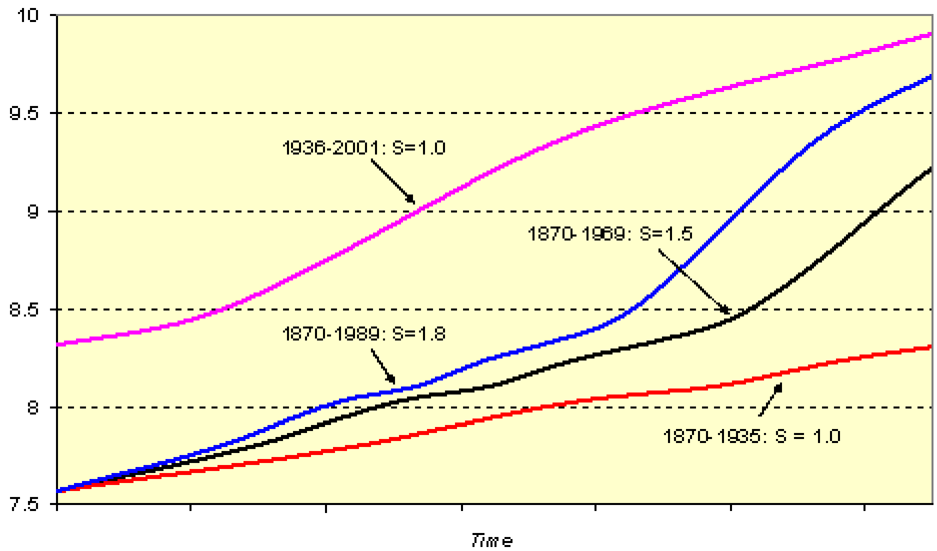


Figure 4: **Transition, Catch up and Divergence relative to the Speed of Learning.** The curves show the effects of the differing technological learning speeds (taken from Fig. 3) on growth performance, replacing time dilation by accelerated technological learning.

also just $S = 1$. Poor countries 2 and 3 are assumed to learn faster than poor country 1 and their speeds of learning are approximately 1.5 and 1.8, respectively, which are obtained by time dilation. Hence, poor country 3 grows further than both poor country 2 and poor country 1 because of its higher rate of technological progress. Both countries 2 and 3 will eventually catch up to the rich group in a finite period of time because they experience accelerated learning and growth.

Fig. 4 re-draws the hypothetical growth paths of these countries against the same time horizon with no time dilation. The effects of a learning speed $S > 1$ are now assumed to be transmitted by way of a faster rate of technological progress. Evidently, poor country 1 seems to diverge since this country's speed of learning does not exceed the speed of technological creation of the rich group. Poor country 2 also seems to diverge initially from the rich group and yet begins to catch up towards the end of the period. Poor country 3, on the other hand, appears to converge to the rich group. While poor countries 2 and 3 are hypothetical and are constructed using time dilation effects, all countries are based on the actual OECD record and poor country 1 is the actual record over chronological time with no temporal distortion.

This simple hypothetical example shows the importance of heterogeneity in the progress of technology on economic performance. Since we have removed business cycle effects and worked with average OECD income, the resulting curves are smooth, but they indicate a variety of possible forms of transitional behavior that include both convergence and apparent divergence. Actual patterns of transition, including transitionally divergent behavior, will inevitably be more complicated than that shown in Figs. 3 and 4. Our econometric objective, taken up in Section 6 below, is to learn about actual transitional behavior by estimating individual patterns of transition.

Table 1: Economic Performance and Speed of Learning based on Long Run historical OECD Growth

	y_{1950}	y_{2001}	Base Trajectory Year of $y_{G,1950}$	Base Trajectory Year of $y_{G,2001}$	Speed of Learning
18 OECD	5,150	20,110	1950	2001	1.00
Africa	949	1,796	1727	1860	2.61
Middle East	1,737	5,808	1858	1950	1.80
Asian Dragons	1,533	18,289	1849	1995	2.86
East Asian	620	1,454	before 1500	1847	>6.80
NIEs	1,072	4,952	1802	1942	2.75
India	619	1,957	before 1500	1866	>8.18
China	438	3,583	before 1500	1917	>7.18
Carribean	1,801	4,674	1861	1940	1.55
Japan	1,920	20,683	1920	2001	1.59
Latin American	3,673	6,947	1919	1957	0.75

Next we proceed to relate these hypothetical transition curves to the actual growth paths of certain regional groupings of the world economies. We start by supposing that Fig. 2 provides a model for world economic growth in the sense that the actual economic performance of the 18 Western OECD countries over the historical time frame 1500-2001 is a base trajectory for economic evolution that other countries in the world follow, albeit over different time frames, thereby allowing for differing speeds of technological learning across countries. Accordingly, we seek to estimate where other countries lie on the long run base OECD trajectory and to calculate the approximate speed of learning within those countries that is necessary to have achieved their given economic performance.

For many countries, data on annual per capita real income is available only from around 1950, so it is convenient to use the period 1950-2001 to represent recent trends in world economic growth. The OECD historical data base provides estimates of per capita income over the long historical period 1500-2001 but not for all 18 constituent OECD countries and not on an annual basis prior to 1870. Therefore, to obtain a complete base trajectory series for OECD trend growth, we infilled observations over the 1500-1870 using a combination of linear interpolation and coordinate trend fitting, as proposed in Phillips (2004)[5]. Next, we formed 9 geographical subgroups from 88 countries in the world economy and used the cross sectional average of their log per capita income to approximate their common growth component. Using this data, we matched the initial and last period incomes of the 9 subgroups in 1950 and 2001, respectively, with the base OECD trajectory and estimated the corresponding speed of learning for each subgroup (G).

Table 1 reports the results. Base OECD trajectory income in 1950 was \$5,150 and takes 52 years to reach \$20,110 in 2001. Columns 4 and 5 of the table show the base OECD trajectory year corresponding to the observed income levels $y_{G,1950}$ and $y_{G,2001}$ for each country grouping G. For example, the average initial and final period incomes for the African countries group in 1950 and 2001 were \$949 and \$1,796 (columns 2 and 3). These figures correspond to base OECD trajectory income in years 1727 and 1860, respectively. We deduce from this calculation that the speed of learning for the African group is approximately 2.61 (column 6). Thus, the African countries can be said to be undergoing growth comparable to that of the OECD base group over 133 years during the industrial revolution in the eighteenth and nineteenth centuries, but the growth experience of these countries is actually compressed into a 52 year time frame during the twentieth century. In an analogous way, initial year 1950 income and final period 2001 income for the Asian Dragons are \$1,533 and \$18,289, respectively, which places this group on the base OECD trajectory over the period from 1849 to 1995. The Asian Dragons have therefore experienced in 52 years economic growth that is comparable to that of 146 years of growth for the OECD base group up to 1995. The learning speed of the Asian Dragons is 2.86 and is therefore faster than that of the African group. Moreover, the Asian Dragon experience co-relates to a much more recent period of the base OECD trajectory. Similarly, Japan's speed of technology learning is 1.59 and this implies that it has compressed the last 81 years of base trajectory OECD growth up to 2001 into 52 years.

The table also reveals that the fastest learning countries are China, India and the East Asian group. Remarkably, China has experienced over four centuries of base trajectory OECD growth in the last 52 years taking it to year 1917 levels on the OECD trajectory. India and the East Asian group of countries have experienced more than three and a half centuries of base trajectory growth in 52 years, taking them to mid-nineteenth century OECD levels of income.

[5] From the historical OECD series described in the Data Appendix, there is century data over 1500-1820. From 1820 onwards, there is annual data for four OECD countries and from 1870 onwards there is annual data for 18 countries. The gaps in these series are interpolated, first by a linear interpolation and subsequently by a smoothing infill filter based on a fitted coordinate trend (Phillips, 2004).

4 Economic Transition Curves

We start by developing some econometric formulations of the neoclassical model that allow for heterogeneity in the speed of convergence and transition effects over time. 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 will be manifest in their growth performance and the phenomenon of transition.

From (5) and (6), the transition path of log per capita real income can be written as follows

$$\log y_{it} = \log \tilde{y}_i^* + \log A_{i0} + [\log \tilde{y}_{i0} - \log \tilde{y}^*] \, e^{-\beta_{it} t} + x_{it} t = a_{it} + x_{it} t, \tag{11}$$

where

$$a_{it} = \log \tilde{y}_i^* + \log A_{i0} + [\log \tilde{y}_{i0} - \log \tilde{y}^*] \, e^{-\beta_{it} t}. \tag{12}$$

Under (8), (12) is a decay model for a_{it} which captures the evolution $a_{it} \to \log \tilde{y}_i^* + \log A_{i0}$ as $t \to \infty$. Correspondingly for large t, $\log y_{it}$ eventually follows a long run path determined by the term $x_{it} t$ in (11).

Following the discussion in Section 2, 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 (5), 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, for instance, the industrial and scientific revolutions.

We may then write (11) in the following form

$$\log y_{it} = \left(\frac{a_{it} + x_{it} t}{\mu_t} \right) \mu_t = b_{it} \mu_t, \tag{13}$$

where b_{it} 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 growth rate of technical progress parameter x_{it} and the initial technical endowment and steady state levels through the parameter a_{it}.

For example, 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 (4), for example. Then, according to (13), $b_{it} = x_{it} + \frac{a_{it}}{t}$ and, further, under the growth convergence conditions (8) and (10) we have the convergence $b_{it} \to x$ as $t \to \infty$. When the economies have heterogeneous technology and x_{it} converges to x_i, we have

$$b_{it} = x_{it} + \frac{a_{it}}{t} \to x_i, \quad \text{as } t \to \infty, \tag{14}$$

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

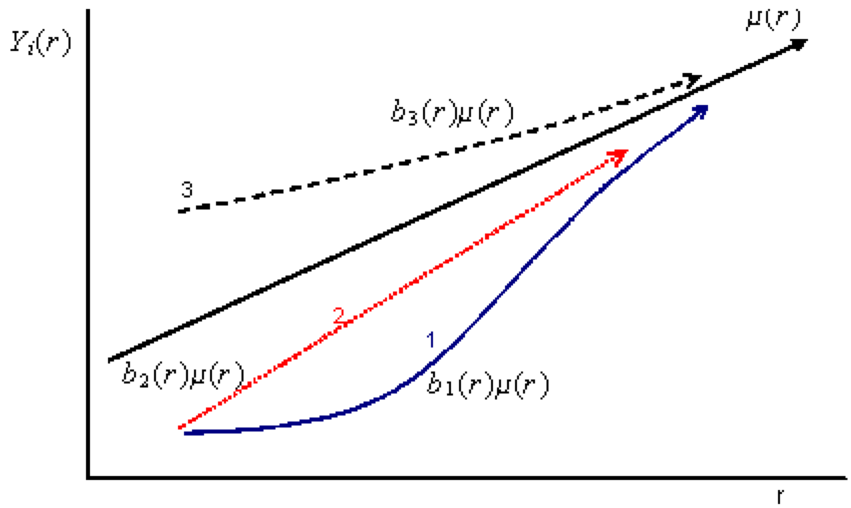


Figure 5: Possible Transition Paths for $Y_i(r) = b_i(r)\mu(r)$

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 (14) 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 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. Appendix A provides some mathematical details of how these extensions and standardizations can be accomplished.

In brief, we proceed as follows. Our purpose is to standardize $\log y_{it}$ in (13) so that the standardized quantity approaches a limit function that embodies both the common growth component and the transition path. To do so, it is convenient to assume that there is a suitable overall normalization of $\log y_{it}$ for which we may write equation (13) in the standardized form given by (15) below. Suppose the standardization factor is $d_{iT} = T^{\gamma_i} W_i(T)$, for some $\gamma_i > 0$ and some slowly varying function[6] $W_i(T)$, so that $\log y_{it}$ grows for large t according to the power law t^{γ_i} up to the effect of $W_i(t)$ and stochastic fluctuations. We may similarly suppose that the common trend component μ_t grows according to $t^{\gamma} Z(t)$ for some $\gamma > 0$ and where Z is another slowly varying factor. Then, we may write

$$\begin{aligned}
\frac{1}{d_{iT}} \log y_{it} &= \frac{1}{T^{\gamma_i} W_i(T)} \left(\frac{a_{it} + x_{it} t}{\mu_t} \right) \mu_t \\
&= \frac{a_{it}}{T^{\gamma_i} W_i(T)} + \left(\frac{x_{it} t}{T^{\gamma_i} W_i(T)} \frac{T^{\gamma} Z(T)}{\mu_t} \right) \left(\frac{\mu_t}{T^{\gamma} Z(T)} \right) \\
&= o(1) + \left(\frac{x_{it} t}{T^{\gamma_i} W_i(T)} \frac{T^{\gamma} Z(T)}{\mu_t} \right) \left(\frac{\mu_t}{T^{\gamma} Z(T)} \right) \\
&= b_{iT}\left(\frac{t}{T}\right) \mu_T\left(\frac{t}{T}\right) + o(1),
\end{aligned} \tag{15}$$

where we may define the sample functions μ_T and b_{iT} as

$$\mu_T\left(\frac{t}{T}\right) = \left(\frac{t}{T}\right)^{\gamma} \frac{Z\left(\frac{t}{T}T\right)}{Z(T)}, \quad \text{and } b_{iT}\left(\frac{t}{T}\right) = \left(\frac{t}{T}\right)^{\gamma_i - \gamma} \frac{W_i\left(\frac{t}{T}T\right) Z(T)}{W_i(T) Z\left(\frac{t}{T}T\right)}, \tag{16}$$

as shown in Appendix A.

Now suppose that $t = [Tr]$, the integer part of Tr, so that r is effectively the fraction of the sample T corresponding to observation t. Then, for such values of t, (15) leads to the following asymptotic characterization

$$\frac{1}{d_{iT}} \log y_{i[Tr]} \sim b_{iT}\left(\frac{[Tr]}{T}\right) \mu_T\left(\frac{[Tr]}{T}\right) \sim b_{iT}(r) \mu_T(r). \tag{17}$$

In (17), $\mu_T(r)$ is the sample growth curve, $b_{iT}(r)$ is the sample transition path (given T observations) for economy i at time T. It is further convenient to assume that these functions converge in some sense to certain limit functions as $T \to \infty$. For instance, the requirement that b_{iT} and μ_T satisfy

$$\mu_T(r) \to_p \mu(r), \quad b_{iT}(r) \to_p b_i(r), \quad \text{uniformly in } r \in [0,1], \tag{18}$$

where the limit functions $\mu(r)$ and $b_i(r)$ are continuous or, at least, piecewise continuous, seems fairly weak. By extending the probability space in which the functions b_{iT} and μ_T

[6] That is $W_i(xT)/W_i(T) \to 1$ as $T \to \infty$ for all $x > 0$. For example. the constant function, $\log(T)$, and $1/\log(T)$ are all slowly varying functions.

are defined, (18) also includes cases where the functions may converge to limiting stochastic processes[7]. The limit functions $\mu(r)$ and $b_i(r)$ represent the common steady state growth curve and limiting transition curve for economy i, respectively. Further discussion, examples and some general conditions under which the formulations (17) and (18) apply are given in Appendix A.

Combining (17) and (18), we have the following limiting behavior for the standardized version of log per capita income for economy i

$$\frac{1}{d_{iT}} \log y_{i[Tr]} \to_p Y_i(r) = b_i(r)\,\mu(r). \tag{19}$$

With this limiting decomposition, we may think about $\mu(r)$ as the limiting form of the common growth path and $b_i(r)$ as the limiting representation of the transition path of economy i as this economy moves towards the growth path $\mu(r)$. The representation (19) is sufficiently general to allow for cases where economies approach the common growth path in a monotonic way either from below or above $\mu(r)$, just as economies 2 and 3 in the stylized paths shown in Fig. 5, or in a much more indirect manner where there may be periods of transitional divergence, as shown in the path of economy 1 in Fig. 5.

To illustrate (19), when μ_t is a stochastic trend with positive drift, we have the simple standardization factor $d_{iT} = T$ and then

$$T^{-1}\mu_{t=[Tr]} = g\frac{[Tr]}{T} + O_p\left(T^{-1/2}\right) \to_p gr,$$

for some constant $g > 0$. If b_{it} satisfies (14), then the limit function $b_i(r) = b_i = x_i$ is a constant function of r. Combining the two factors gives the limiting path $Y_i(r) = b_i gr$ for economy i, so that the long run growth paths are linear and parallel across economies. When there is convergence across economies, we have limit transition curves $b_i(r)$ each with the property that $b_i(1) = b$, for some constant $b > 0$, but which may differ for intermediate values (i.e., $b_i(r) \neq b_j(r)$ for some and possibly all $r < 1$). In this case, each economy may transition in its own way towards a common limiting growth path given by the linear function $Y(r) = bgr$. In this way, the framework permits a family of potential transitions to a common steady state.

Fig. 5 illustrates some possible transition paths for $Y_i(r)$ of this type. Paths 2 and 3 in the figure show monotonic convergence to the common growth path $\mu(r)$, whereas path 1 involves transient divergence of $Y_1(r)$ from $\mu(r)$ with subsequent catch up and convergence. In this case, the more complex approach to the common growth path is reflected in the transition curve $b_1(r)$ for this economy.

[7] For example, if μ_t is a unit root process, then under quite general conditions we have the weak convergence $T^{-1/2}\mu_{[Tr]} = \mu_T(r) \Rightarrow B(r)$ to a limit Brownian motion B (e.g., Phillips and Solo, 1992). After a suitable change in the probability space, we may write this convergence in probability, just as in (18).

5 Relative Transition and Convergence

5.1 Some Stylized Models and Asymptotics

Taking ratios to cross-sectional averages in (15) removes the common trend μ_t and leaves the following standardized quantity

$$h_{iT}\left(\frac{t}{T}\right) = \frac{d_{iT}^{-1}\log y_{it}}{\frac{1}{N}\sum_{j=1}^{N} d_{jT}^{-1}\log y_{jt}} = \frac{b_{iT}\left(\frac{t}{T}\right)}{\frac{1}{N}\sum_{j=1}^{N} b_{jT}\left(\frac{t}{T}\right)}, \tag{20}$$

which describes the relative transition of economy i against the benchmark of a full cross sectional average. Other benchmarks are possible (including subgroup averages or even a single advanced economy like the USA) and these will be considered below and in our empirical work. Clearly, h_{iT} depends on N also but we omit the subscript for simplicity because this quantity often remains fixed in the calculations. In view of (18), we have

$$h_{iT}\left(\frac{[Tr]}{T}\right) \rightarrow_p h_i(r) = \frac{b_i(r)}{\frac{1}{N}\sum_{j=1}^{N} b_j(r)}, \quad \text{as } T \rightarrow \infty, \tag{21}$$

and the function $h_i(r)$ then represents the limiting form of the relative transition curve for economy i.

The curve $h_i(r)$ shows the time profile of transition for economy i relative to the average. At the same time, $h_i(r)$ measures economy i's relative departure from the common steady state growth path $\mu(r)$. Thus, any divergences from $\mu(r)$ are reflected in the transition $h_i(r)$. 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 curve to measure the extent of the divergent behavior and to assess whether or not the divergence is transient.

When there is common (limiting) transition behavior across economies, we have $h_i(r) = h(r)$ across i, and when there is ultimate growth convergence we have

$$h_i(1) = 1, \quad \text{for all } i. \tag{22}$$

This framework of growth convergence admits a family of relative transitions, where the curves $h_i(r)$ may differ across i for $r < 1$, while allowing for ultimate convergence when (22) holds. Removing the common (steady state) trend function $\mu(r)$ that appears in Fig. 5, Fig. 6 shows the corresponding relative transition curves, each satisfying the growth convergence condition (22).

While the criterion for the ultimate convergence of economy i to the steady state is given by (22), the manner of economic transition and convergence can be very different across economies. Fig 6 shows three different stylized paths. Economies 2 and 3 have quite different initializations

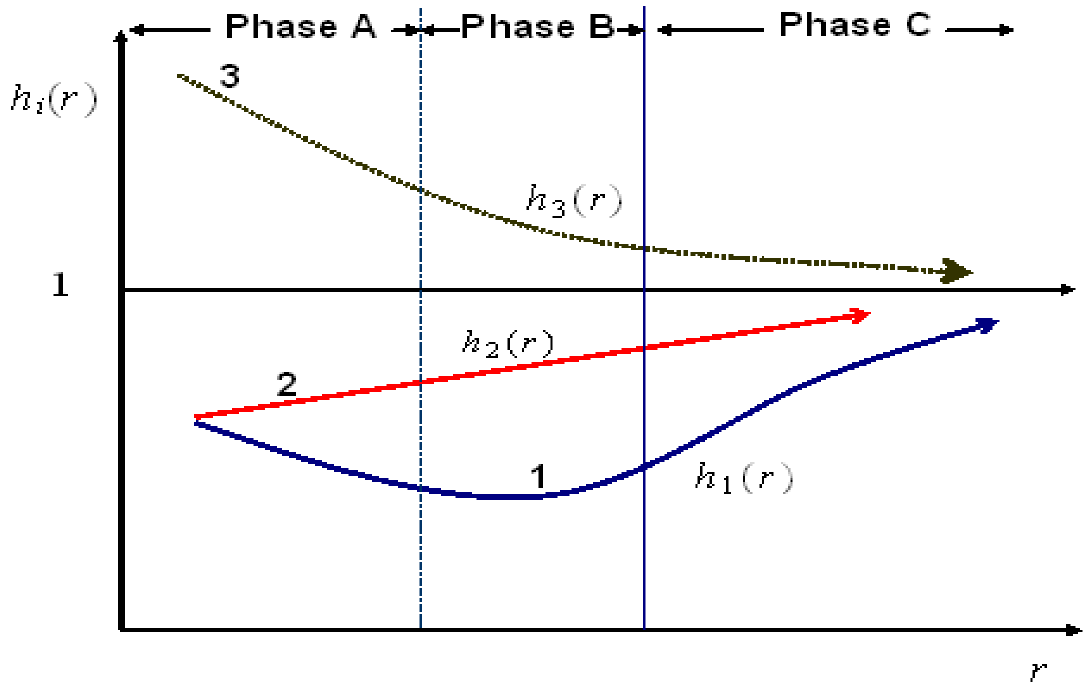


Figure 6: Relative Transition Curves $h_i(r)$ and Phases of Transition

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).

The framework in (21) is compatible with a situation where there is an infinite population. In this event, as N passes to infinity, if a law of large numbers applies to the transition functions b_i, so that $N^{-1}\sum_{j=1}^{N} b_j \to \bar{b}$, we would define $h_i(r) = b_i(r)/\bar{b}(r)$. Convergence for economy i would then apply if $b_i(1) = \bar{b}(1)$, leading to the same criterion as that given in (22) above. In this extension, some economies may converge (when $b_i(1) = \bar{b}(1)$ for some i) while others remain outliers and do not converge to the average (when $b_j(1) \neq \bar{b}(1)$ for some j). In such situations, there is per capita income inequality in the limit distribution (or full population of economies), yet still the possibility of convergence to the mean for some economies, much as in Galton's (1886, 1889) work on "regression to mediocrity" in human physical characteristics like height. So, this form of growth path convergence is analogous to Galton's empirical finding that the deviation of children's heights from the mean regresses toward zero over time (relative to their parents' heights), while the overall distribution of heights in the population does not necessarily narrow over time. Correspondingly, there is some advantage in not insisting on a requirement of the form that in the full population $\int (h-1)^2 \, dP_h = 0$, where $h = h(1)$ and P_h is the limit measure of $h(r)$, a requirement that would correspond to Galton's fallacy (Quah, 1993, and Hart, 1995).

On the other hand, in the case of economic growth, we might reasonably ask whether there is a narrowing in dispersion over time, for example by comparing the variation of h at points r and s as in a calculation of the form

$$\int \left(h(r) - \int h(r) \right)^2 dP_h < \int \left(h(s) - \int h(s) \right)^2 dP_h \text{ for } r > s.$$

Alternatively, we could ask whether for some subgroup of N_G economies G there is convergence in the sense that we have

$$h_{iT}^G \to_p 1 \quad \text{for } i \in G, \tag{23}$$

or $\sigma_{TG}^2 = N_G^{-1} \sum_{i \in G} (h_{iT}^G - 1)^2 \to_p 0$, as $T \to \infty$, where

$$h_{iT}^G = \frac{b_{iT}(1)}{N_G^{-1} \sum_{j \in G} b_{jT}(1)} \tag{24}$$

is the relative transition curve for economy i in group G.

A regression test of convergence in the transition function h_{iT}^G is developed later in this section. When the time series sample T is large, we have the prospect of estimating the relative transition curve $h_i(r)$ for each individual economy and testing the convergence criterion (22), possibly for subgroups of economies like G as indicated above or against the benchmark of a single economy like the US. For short time series with small T, transitions are generally still occuring because of temporal heterogeneity in the key parameters β_{it} and x_{it}, and it is therefore difficult and less meaningful to test for growth convergence. Even in such cases, however, the fitted transition curve may still reveal interesting empirical properties of the individual economies in transition.

Also, as Figs. 5 and 6 illustrate in a stylized way, when there is temporal and cross section heterogeneity, there exists an infinite number of possible transition paths even in cases where there is ultimate convergence. Moreover, as discussed earlier, there are good reasons for thinking that there will be income inequality across economies even in the limit distribution as $T \to \infty$. Appendix E provides some discussion of these issues in the context of models with heterogeneous stochastic trends with drift.

5.2 Fitting Transition Curves

There are several possible approaches to fitting transition curves. In its most general form the problem has two nonparametric elements, involving the unknown transition function $b_i(r)$ and the growth curve $\mu(r)$. Without further assumptions it is not possible to separately identify these two functions. However, as shown above, many of the important elements are embodied in the relative transition function $h_i(r)$. So one of the primary goals of the econometric analysis is to estimate the limiting form of the transition function, $h_i(r)$, which may then be used for various purposes, including an analysis of convergence. In practical work, the time series will often be short and the task therefore presents some difficulties.

Of course, the quantity $\log y_{it}/\frac{1}{N}\sum_{i=1}^{N}\log y_{it}$ can be calculated directly from observations of log per capita real income. But it will often be preferable to remove business cycle components from the data first as interest centres on the long run component. Extending (13) to incorporate a business cycle effect κ_{it}, we can write

$$\log y_{it} = b_{it}\mu_t + \kappa_{it}.$$

Smoothing methods offer a convenient mechanism for separating out the cycle κ_{it}, and we can employ filtering, smoothing and regression methods to achieve this. In our empirical work, we have used two methods to extract the trend component $b_{it}\mu_t$. The first is the Whittaker-Hodrick-Prescott (WHP) smoothing filter[8]. The procedure is flexible, requires only the input of a smoothing parameter, and does not require prior specification of the nature of the common trend μ_t in $\log y_{it}$. The method is also suitable when the time series are short. In addition to the WHP filter, we employed a coordinate trend filtering method (Phillips, 2004). This is a series method of trend extraction that uses regression methods on orthonormal trend components to extract an unknown trend function. Again, the method does not rely on a specific form of μ_t and is applicable whether the trend is stochastic or deterministic.

The empirical results reported below were little changed by the use of different smoothing techniques. The coordinate trend method has the advantage that it produces smooth function estimates and standard errors can be calculated for the fitted trend component. Kernel methods, rather than orthonormal series regressions, provide another general approach to smooth trend extraction and would also give standard error estimates. Kernel methods were not used in our practical work here because some of the time series we use are very short and comprise as few as 30 time series observations. Moreover, kernel method asymptotics for estimating stochastic processes are still largely unexplored and there is no general asymptotic theory to which we may appeal, although some specific results for Markov models have been obtained in work by Phillips and Park (1998) and Guerre(2004).

[8] Whittaker (1923) first suggested this penalized method of smoothing or 'graduating' data and there has been a large subsequent literature on smoothing methods of this type (e.g. see Kitagawa and Gersch, 1996). The approach has been used regularly in empirical work in time series macroeconomics since Hodrick and Prescott (1982/1997).

Using the trend estimate $\hat{f}_{it} = \widehat{b_{it}\mu_t}$ from the smoothing filter, the estimates

$$\hat{h}_{it} = \frac{\hat{f}_{it}}{\frac{1}{N}\sum_{i=1}^{N}\hat{f}_{it}} \tag{25}$$

of the transition coefficients $h_{it} = b_{it}/\left(N^{-1}\sum_{i=1}^{N} b_{it}\right)$ are obtained by taking ratios to cross-sectional averages. Assuming a common standardization[9] $d_{iT} = d_T$ for simplicity and setting $t = [Tr]$ we then have the estimate $\hat{h}_i(r) = \hat{h}_{i[Tr]}$ of the limiting transition curve $h_i(r)$ in (21). We can decompose the trend estimate $\hat{f}_{it}$ as

$$\hat{f}_{it} = f_{it} + e_{it} = \left[b_{it} + \frac{e_{it}}{\mu_t}\right]\mu_t, \tag{26}$$

where e_{it} is the error in the filter estimate of f_{it}. Since μ_t is the common trend component, the condition $\frac{e_{it}}{\mu_t} \rightarrow_p 0$ uniformly in i seems reasonable[10]. Then,

$$\hat{h}_i(r) = \frac{\left[b_{i[Tr]} + \frac{e_{i[Tr]}}{\mu_{[Tr]}}\right]}{\frac{1}{N}\sum_{i=1}^{N}\left[b_{j[Tr]} + \frac{e_{j[Tr]}}{\mu_{[Tr]}}\right]} = \frac{\frac{1}{d_T}\left(b_{i[Tr]} + \frac{e_{i[Tr]}}{\mu_{[Tr]}}\right)}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_T}\left[b_{j[Tr]} + \frac{e_{j[Tr]}}{\mu_{[Tr]}}\right]}$$

$$= \frac{b_{iT}\left(\frac{t}{T}\right)}{\frac{1}{N}\sum_{j=1}^{N} b_{jT}\left(\frac{t}{T}\right)} + o_p(1) \rightarrow_p \frac{b_i(r)}{\frac{1}{N}\sum_{j=1}^{N} b_j(r)} = h_i(r),$$

so that the relative transition curve is consistently estimated by $\hat{h}_i(r)$.

5.3 Testing Divergence

We may use the fitted transition curves $\hat{h}_i(r)$ to reveal evidence on key issues such as possible transitional divergence for some individual economies and the ultimate convergence of others. Under the growth convergence criterion (22) we have $\hat{h}_i(1) \rightarrow_p 1$ for such convergent economies. In a manner similar to (24) we can construct for a group G of N_G economies the consistent estimate

$$\hat{h}_{it}^{G} = \frac{\hat{f}_{it}}{N_G^{-1}\sum_{j\in G}\hat{f}_{jt}} \rightarrow_p h_{it}^{G} = \frac{b_{it}}{N_G^{-1}\sum_{j\in G} b_{jt}}.$$

[9] Alternatively, if the standardizations d_{iT} were known (or estimated) and were incorporated directly into the estimates $\hat{f}_{it}$ then $\hat{h}_{it} = \hat{f}_{it}/\left(N^{-1}\sum_{i=1}^{N}\hat{f}_{it}\right)$ would correspondingly build in the individual standardization factors. Accordingly, $\hat{h}_{it}$ is an estimate of $h_{it} = h_{iT}\left(\frac{t}{T}\right)$ as given in (20).

[10] Primitive conditions under which $\frac{e_{it}}{\mu_t} \rightarrow_p 0$ holds will depend on the properties of μ_t and the selection of the bandwidth/smoothing parameter/regression number in the implementation of the filter. In the case of the WHP filter, this turns on the choice of the smoothing parameter (λ) in the filter and its asymptotic behavior as the sample size increases. For instance, if μ_t is dominated by a linear drift and $\lambda \rightarrow \infty$ sufficiently quickly as $T \rightarrow \infty$, then the WHP filter will consistently estimate the trend effect. Phillips and Jin (2002) provide some asymptotic theory for the WHP filter under various assumptions about λ and the trend function.

Now, if $h_i^G = 1$ for all $i \in G$, it follows that

$$\hat{\sigma}_G^2 = \frac{1}{N_G^{-1}} \sum_{j \in G} \left(\hat{h}_j(1) - 1\right)^2 \rightarrow_p 0, \quad \text{as } T \rightarrow \infty, \tag{27}$$

which provides a mechanism for studying convergence characteristics of subgroups of economies.

Under more specific conditions that are discussed in Appendix B, we may assess the evidence in support of a transition to convergence by regression methods. In particular, defining

$$H_t^G = N_G^{-1} \sum_{j \in G} \left(\hat{h}_{jt}^G - 1\right)^2, \tag{28}$$

for $t = [Tr]$ for some $r > 0$, we have the regression equation

$$\log H_t^G = \log A - 2\alpha \log t + u_t^G, \tag{29}$$

whose error u_t^G is shown in the Appendix to be asymptotically equivalent to a zero mean, weakly dependent time series whose explicit expression in terms of model components is given in (51).

Equation (29) is a logarithmic regression corresponding to a decay model for the sample variance of the relative transition that has the explicit parametric form $H_t^G \sim A/t^{2\alpha}$ in the limit as $t \rightarrow \infty$. In this formulation, the parameter α governs the rate at which the cross section variation over the transitions decays to zero over time. Appendix B investigates some stochastic models for which the transitions are of this type and for which the "$\log t$" regression equation (29) may be derived with an error of $o_p(1)$ for large T and N_G. Conditions under which (29) may be validly fitted by least squares regression and appropriate procedures for the construction of standard errors and regression tests in this model are also discussed in this Appendix. Least squares regression on (29) is shown to produce $\sqrt{TN_G}$ consistent estimates of the parameter α, autocorrelation robust standard errors may be calculated using conventional procedures and formula (62) based on the limit theory for the regression coefficient and econometric tests of hypotheses about α may be implemented from this regression.

The form of (29) suggests a simple regression test of the hypothesis of convergence. When $\alpha > 0$, $\log H_t^G$ and hence H_t^G decreases with t and (27) is satisfied in the limit as $t \rightarrow \infty$. Conversely, when $\alpha = 0$, H_t^G does not converge to zero and so there is no convergence in the subgroup G. More specifically, we may test $H_0 : \alpha = 0$ against the alternative $H_1 : \alpha > 0$ in a least squares regression of (29) using observations $t = [Tr], [Tr] + 1, ..., T$ for some $r > 0$. The regression may employ either h_{it}^G computed directly from $\log y_{it}$ as in (28) or, as discussed above, use the filtered data $\hat{h}_{it}^G$ based on the fitted values $\widehat{\log y_{it}}$ from a coordinate trend regression or smoothing filter. The growth convergence hypothesis (27) corresponds to the alternative H_1. So, rejecting H_0 provides empirical evidence in favor of convergence among the economies in subgroup G and the value of α is a measure of the rate of convergence. Further details of the construction of this test and its asymptotic properties under the null and local alternatives are provided in Appendix B. The test is consistent and has non trivial power in local neighborhoods of width $O\left(1/\sqrt{TN_G}\right)$. The use of this test is illustated in the simulations and the empirical applications that follow.

5.4 Transitioning between Subgroups and Subsample Convergence Testing

Suppose there are N economies overall in the cross section and that these economies are divided into $\mathcal{G}$ subgroups, each of size N_G, so that $N = \mathcal{G}N_G$. Fig. 1 illustrates such a collection of $\mathcal{G} = 5$ subgroups of national economies ordered according to income and traces the evolution of these subgroups over the period 1960-1996. The historical experience shown in Fig. 1 indicates that there is some transitioning over time between these economic subgroups in which the upper members of one subgroup can catch up with a higher subgroup. To consider such transitioning we develop the following subsample version of the $\log t$ test.

Suppose that the $\mathcal{G}$ subgroups are ordered from the richest to the poorest groups according to income. If the speed of technological creation and learning is different across these subgroups, some transitional divergence can be expected between the richest and some of the other subgroups as indicated in the stylized paths of Fig. 5 and the empirical time forms shown in Fig. 1. On the other hand, when the speed of technological creation and learning is comparable across subgroups, we would not expect to see transitional divergence of this type. To find empirical evidence of these effects, we may cross-fertilize the original groups by dropping the first ℓ richest countries from the top subgroup and forming $\mathcal{G} - 1$ new subgroups based on income ordering in the same way with each group having the same number of economies, N_G, as before. By repeating this exercise, we effectively subsample the full-group and thereby obtain evidence on linkage performance over time among the original subgroups. This information sheds light on whether there is transitional divergence or catching up between groups of countries. More specifically, the new subgroups overlap the “joints” in the original subgroups and provide a natural way of focusing attention on transitioning between subgroups. In particular, performing a $\log t$ test on the new subgroups provides a direct test of whether one group of economies is catching up to the next group.

Subsampling along these lines also helps to reduce size distortion (that is, a tendency to reject the null H_0 of divergence even when there is no convergence) in the $\log t$ test that shows up in simulations of the finite sample performance of this test. The reason for this improvement is that subsampling raises the hurdle for rejecting the null hypothesis and rejection of H_0 occurs only when the evidence is convincing across the various subgroups that there is convergence. More details of the way the procedure is implemented is provided in Appendix D. As in other subsampling procedures, we have to choose the number ℓ that determines how many economies to discard in reforming the subgroups. In simulations, we have found $\ell = N_G/2$ to work well.

6 Empirical Illustrations of Transition

As suggested in the stylized examples considered earlier, 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 analyzed in the growth convergence literature in the past, but our application focuses attention on the phenomenon of economic transition as part of the larger empirical story about 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 using $\log t$ regressions.

6.1 Transitional Divergence: Graphical Illustrations

The first illustration is based on regional economic growth among the 48 contiguous U.S. states[11]. 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. 7 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. Evidently, there is heterogeneity across states, but 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 (22).

Fig. 8 shows the cross-sectional average of the fitted transition curves over 9 separate geographical regions for the contiguous U.S. States data. The shapes of these regional transition curves are similar to those for the full 48 States shown in Fig. 7, but in the new figure it is easier to distinguish the regional transition patterns. The Mid-Atlantic, New England and Pacific regions start the period above average and transition in a downward direction, whereas the South Altantic, West South Central and East South Central start the period below average and transition upwards. The time profile of these cross sectional averages shows that the transition curves have been steadily converging toward unity over the last 70 years, as indicated in criterion (22). Interestingly, the regional convergence does not seem to be completed yet and there is remaining regional inequality, although there is also evidence of some subgroup convergence of the form (23).

The second illustration involves a panel of real per capita income for 18 OECD countries taken from the OECD historical data set. Fig. 9 displays the relative transition parameters for log per capita income in these 18 OECD countries between 1929 and 2001. The countries were

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

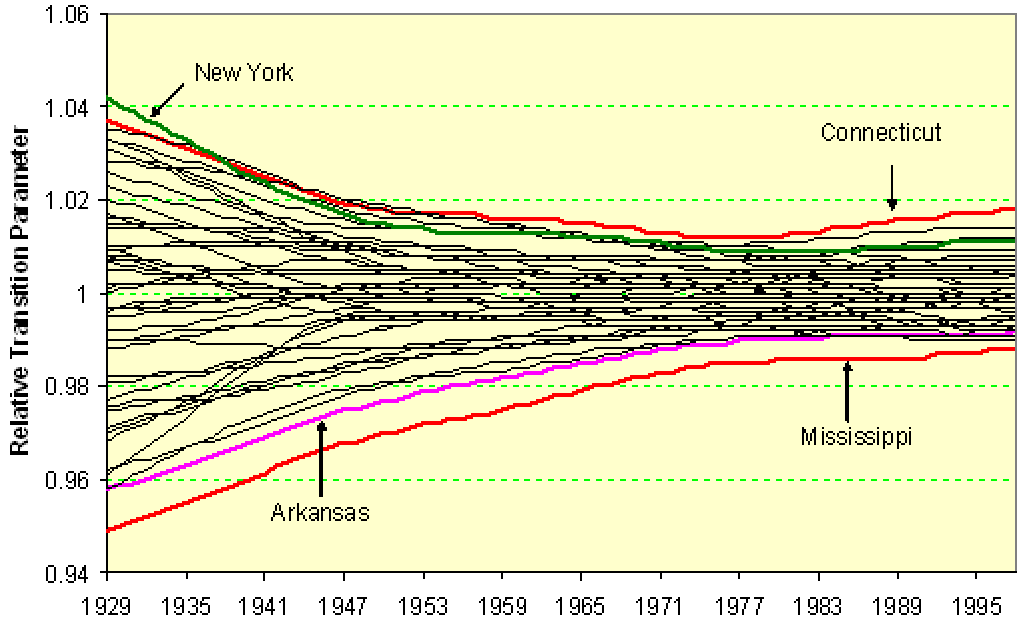


Figure 7: Transition Paths for the 48 Contiguous US States.

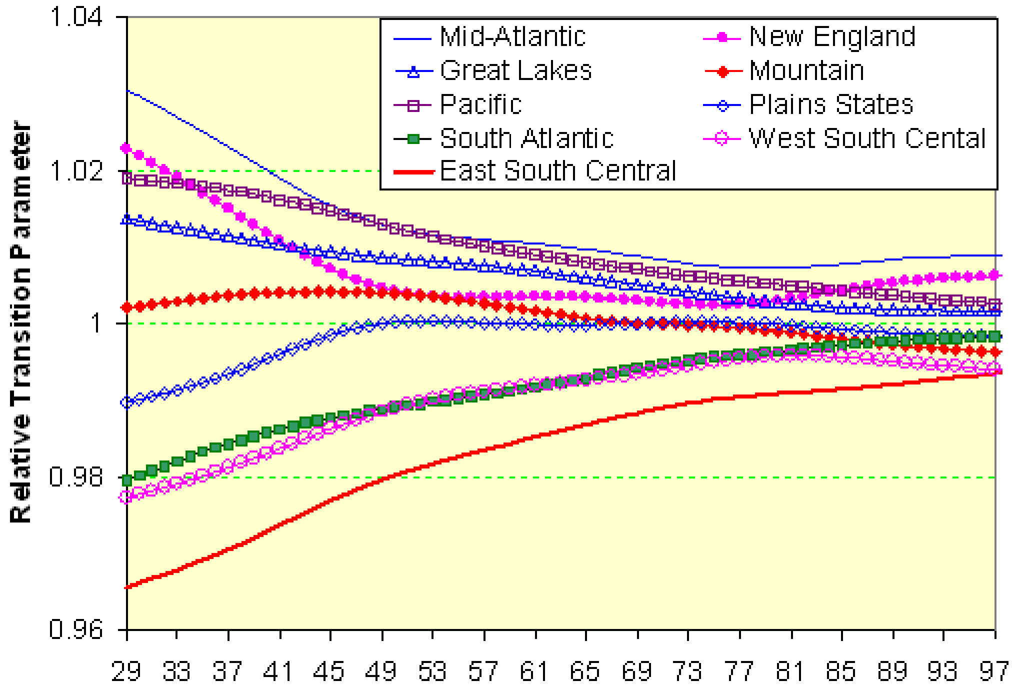


Figure 8: Transition Paths for Regional Groups of the Contiguous US States

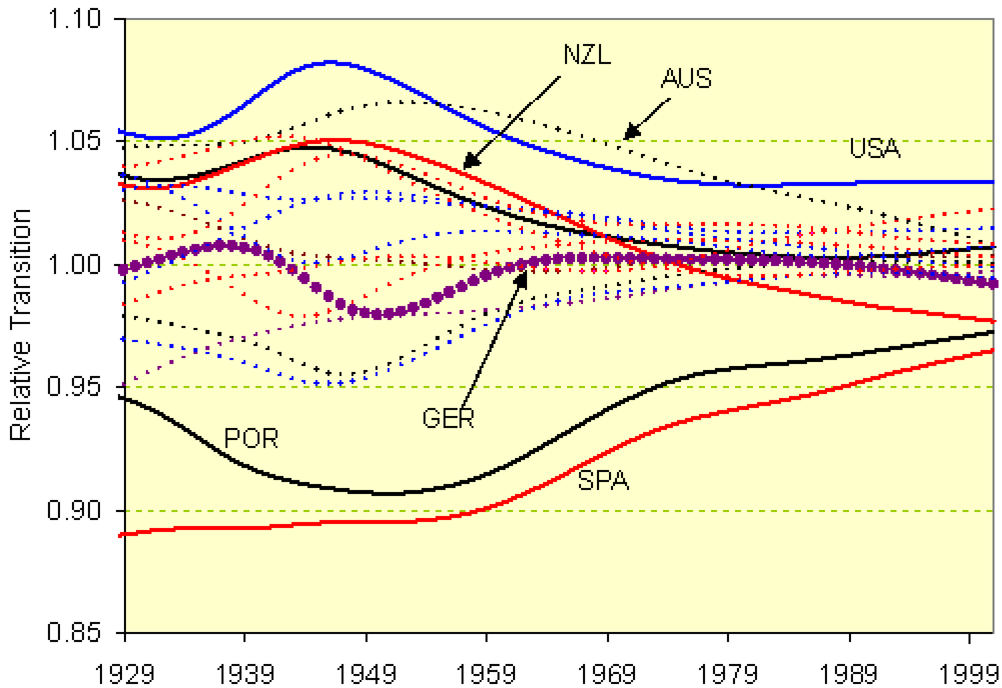


Figure 9: OECD Transition Paths: 18 Western OECD countries from 1929-2001

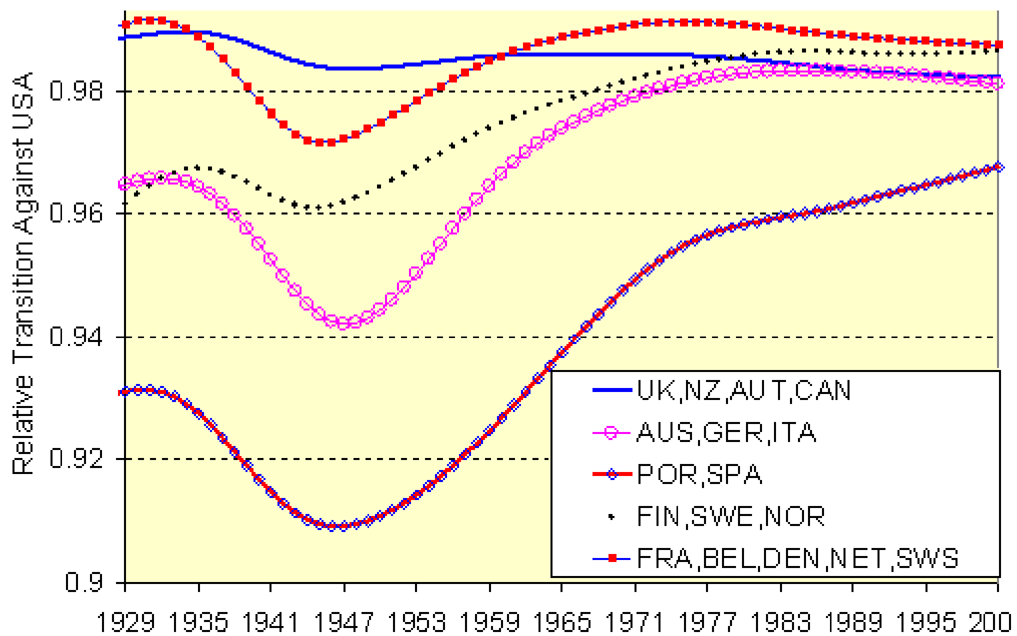


Figure 10: Evidence of Phase B & C Transitions in Historical OECD Data

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. 7, 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 to diverge before World War II. After around 1950, however, the pattern of transition seems similar to that of Fig. 7. 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.

Fig. 10 shows the relative transition parameters for certain subgroups of countries against the benchmark of the U.S. . 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, Fig. 10 provides an empirical illustration of the stylized patterns of economic performance characterized as phases B and C of Fig. 6. Extending the panel back to 1870 and through to 1930, Fig. 11 shows transition curves that are similar in form to phase A (transitional divergence) in the stylized patterns of Fig. 6, with evidence of the phase B turn-around coming towards the end of the period.

The final illustration is based on log per capita income in 88 PWT countries in the world economy over the period 1960 to 1996. The country selection is mainly based on data avail-

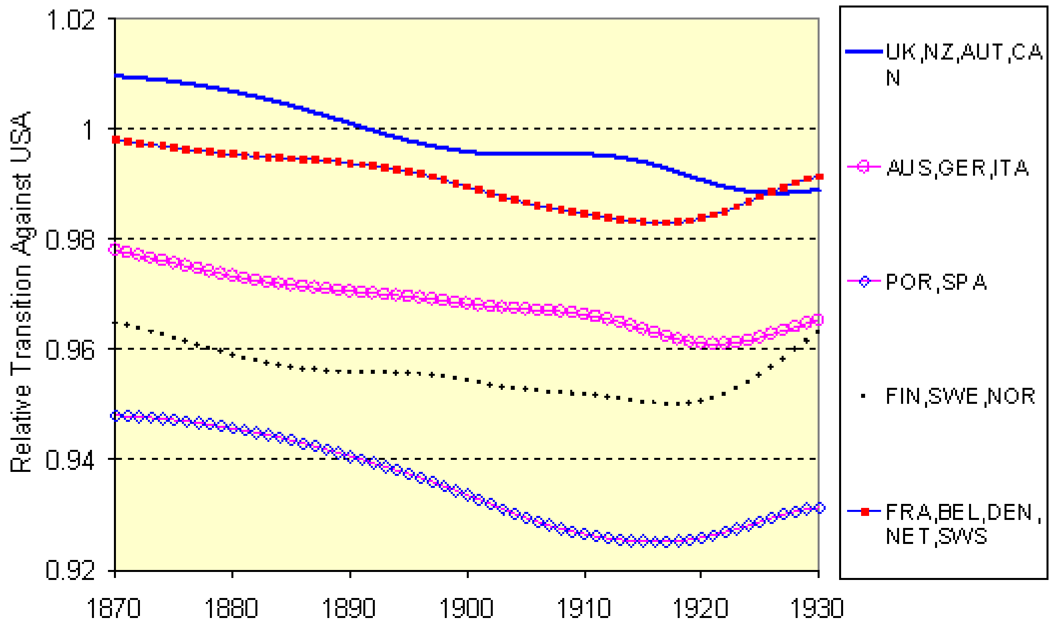


Figure 11: Evidence of Phase A and B Transitions in Historical OECD Data

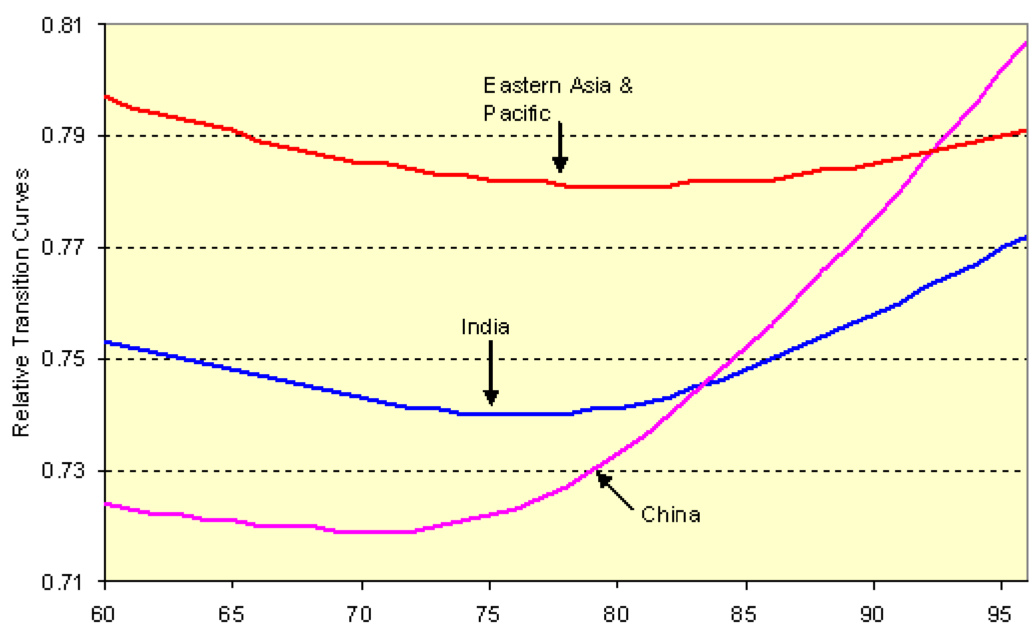


Figure 12: Examples of Phase B Transitions among the World Economies

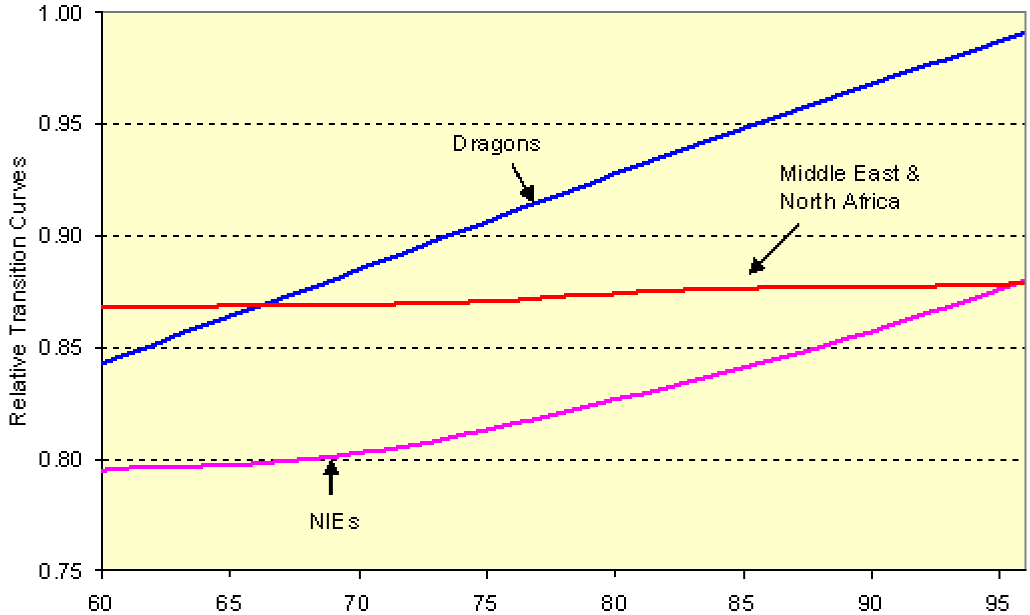


Figure 13: Examples of Phase C Transitions among the World Economies

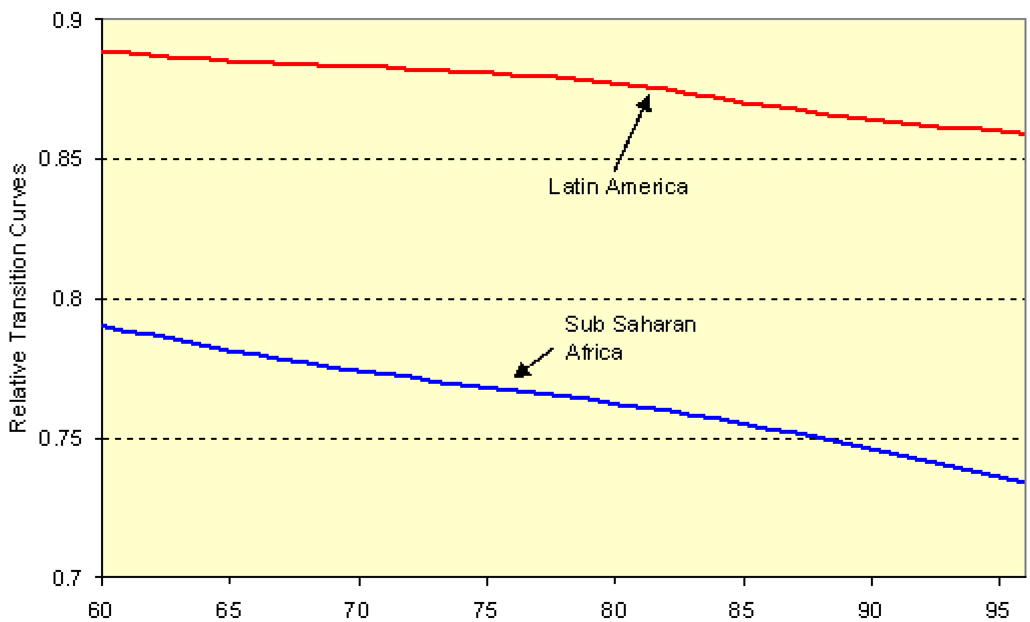


Figure 14: Examples of Phase A Transitions among the World Economies

ability. 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 18 OECD countries. 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 Africa, and the Latin American & Carribean economies (Fig. 14). Phase B transitions occur in three subgroups – India, China and the Eastern Asian countries (Fig. 12). Finally, phase C transitions are evident in two subgroups – the Asian dragons and the newly industrialized economies (NIEs) shown in Fig. 13. 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 and the Eastern Asian countries will continue to grow faster over the next decade than the OECD nations as they move into phase C transition; and, soon or later, we can expect to witness the Sub-Saharan and Latin American countries entering phase B transition where they begin to turn around economic performance and start to catch up with the 18 OECD countries. However, from the evidence to date in these figures, we cannot yet distinguish, especially for the Sub-Saharan Africa and Latin American countries, whether or when such changes will occur. We now provide a formal statistical test of this issue.

6.2 Testing Growth Convergence

We perform the $\log t$ regression test based on (29) with three panel data sets – the 48 contiguous United States, 18 OECD countries and 88 PWT countries. The results are shown in Table 4. With the full-group test, we reject the null of no convergence among the 48 contiguous United States and the 18 OECD countries. However, we cannot reject the null for the 88 PWT countries. In fact, in the latter case the fitted coefficient $\hat{\alpha}$ is positive and significant.

As shown in the simulations reported in Appendix D, the $\log t$ test may not perform well in terms of the accuracy of its nominal size when there are a large number of heterogeneous cross sectional units, as in the case of the 88 PWT nations. In particular, the test shows a tendency to overreject in such situations. We therefore perform additional tests using the subsampling approach described in section 5.4 to improve size and provide more information about transitioning to convergence. In the first step, we create 4 subgroups based on a panel ordering with respect to last period income, with subgroup A having the highest last period income, subgroup B the next highest, and so on. Table 2 shows the results of these tests. For each subgroup (A through D), the null hypothesis of no convergence is rejected at the 5% level. Importantly, this rejection of the null does not mean that all countries in the 88 PWT panel are converging over the sample period, but it does imply evidence of subgroup convergence.

Next, we proceed with the second step test, forming three subgroups from the original groupings. Subgroup E comprises the lower 11 income countries from original subgroup A and the higher 11 income countries from original subgroup B. Similar constructions apply to give new subgroups F and G. The outcome of the $\log t$ test applied to these new subgroups can be interpreted in the following way. In view of the overlapping nature of the data (relative to the original subgroups), rejection of the null hypothesis of no convergence for a new subgroup provides evidence in favor of catch-up growth performance between the original subgroups. In effect, the top income countries in the lower group are catching up with the lower income countries in the upper group. As seen in the final panel of Table 2, the $\log t$ test rejects the null decisively for each of the subgroups E, F and G. The rejection of the null for subgroup E (which comprises countries from both subgroup A and subgroup B) therefore suggests convergence amongst these countries, so that the upper income countries in subgroup B are catching up with subgroup A. The same conclusion holds for subgroups G and H. This finding supports the catch-up hypothesis across the original subgroups and thereby indicates some evidence in favor of these countries being on a transitional path towards ultimate convergence, akin to that of the long run historical trajectory of the OECD countries over the past five centuries.

Table 2: Empirical Evidence for Transitional Divergence

Regression: $\log H_t^G = \log A - 2\alpha \log t + u_t^G$

Cases	**Sample Period**	$\hat{\alpha}$	t **ratio**
$\log t$ **Test within a panel**			
48 United States	1929-1998	1.174	94.4
18 OECD	1870-2001	0.643	5.29
88 PWT	1960-1996	-0.200	-22.49
Sub group $\log t$ **Test for 88 PWT : Step 1**			
Sub A: 22 Highest	1960-1996	1.318	16.36
Sub B: 22 High	1960-1996	0.822	36.52
Sub C: 22 Poor	1960-1996	1.041	11.22
Sub D: 22 Poorest	1960-1996	0.136	2.85
Sub group $\log t$ **Test for 88 PWT : Step 2**			
Sub E: 22 countries linking A and B	1960-1996	0.882	15.01
Sub G: 22 countries linking B and C	1960-1996	1.126	14.60
Sub H: 22 countries linking C and D	1960-1996	0.316	9.31

The same regression tests (both all-group and subgroup tests) were run with the 127 country OECD historical data covering the longer time period 1950-2001. These regressions gave very similar results to those reported in Table 2 for the PWT data, so they are not repeated here.

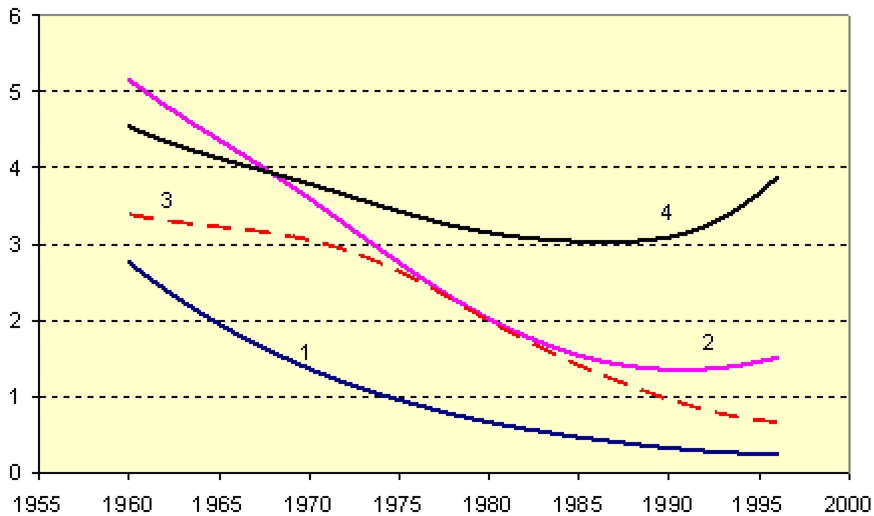


Figure 15: Four subgroups and convergence within each group. (88 PWT countries, $N_G = 22$, G=4)

6.3 Transitioning in Sample Variation

Fig. 15 shows the time path of the sample variance H_t^G for 4 subgroups of the 88 PWT countries, so that $N_G = 22$. The groups are arranged in descending order of income with group 1 having the highest income countries and group 4 those with the lowest income. Evidently, there is some strong evidence favorable to convergence within each group, although H_t^G turns upward for group 4 towards the end of the sample and appears to flatten out for group 2. Fig. 16 shows H_t^G for 3 mixed subgroups obtained by taking the intermediate country groupings after discarding the 11 highest income countries and the 11 poorest countries. For each of these new subgroups, the total number of countries included is again 22. Apparently, there is evidence of transitioning to convergence within each group. The results imply that group 4 is catching up with group 3, which in turn is catching up with group 2, and so on. In other words, overall there is some evidence of a transitioning towards long run convergence among the 88 PWT countries. This is so, even though a full-group test rejects convergence, indicating that the process of transition is more complex than what can be captured in a single test.

7 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

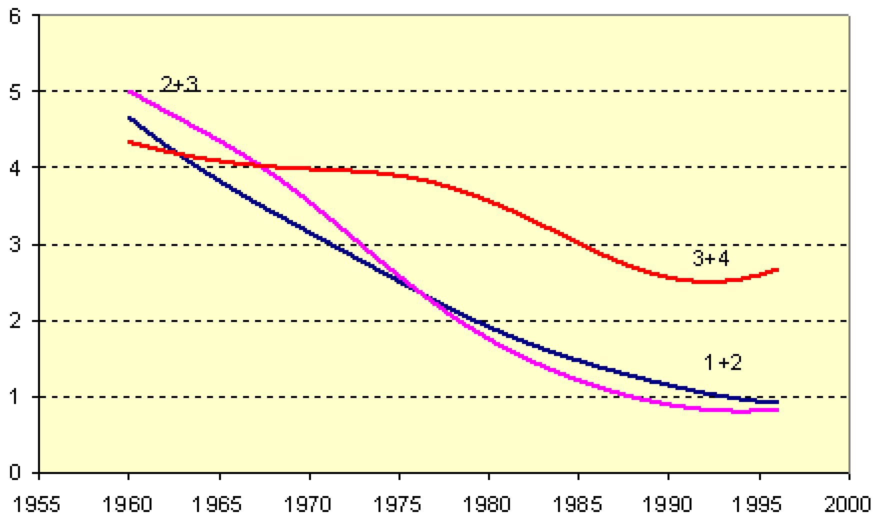


Figure 16: The second subgrouping. Group 1+2 consist of 11 low income in group 1 and 11 high income in group 2, and so on.

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 seeks to provide 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 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. 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, but also including the possibility of transitional divergence from that 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 characterise individual countries. In other work (Phillips and Sul, 2005), 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 undertanding 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 Sala-i-Martin (1995), *Economic Growth.* New York: McGraw-Hill.

[4] Becker, G.S., K.M. Murphy, and R. Tamura, (1990) "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy* 98 (5) 12–37.

[5] Bernard, A.B, and S.N. Durlauf (1995), "Convergence in International Output", *Journal of Applied Econometrics* 10(2): 97-108.

[6] Bernard, A.B, and S.N. Durlauf (1996), "Interpreting Tests of the Convergence Hypothesis", *Journal of Econometrics*, 71(1-2): 161-174.

[7] DeLong, J.B (1988), "Productivity Growth, Convergence, and Welfare: A Comment", *American Economic Review* 78(5):1138-55.

[8] 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.

[9] Durlauf, S. N., P. A. Johnson and J. Temple (2004). Growth Econometrics. University of Wisconsin, mimeographed.

[10] Durlauf, S.N. and D.T. Quah (1999), "The New Empirics of Economic Growth", Ch. 4 in Handbook of Macroeconomics 1A, (Elsevier)

[11] Easterly, W. (2001) "The Middle Class Consensus and Economic Development," Journal of Economic Growth, Vol. 6 (4), 317–336.

[12] Galton, F. (1886). "Regression toward mediocrity in hereditary stature." *Journal of the Anthropological Institute of Great Britain and Ireland,.15, 246-263.*

[13] Galton, F. (1989) *National Inheritance.* London: Macmillan.

[14] Guerre, E. (2004). "Design-Adaptive Pointwise Nonparametric Regression Estimation For Recurrent Markov Time Series," mimeo, Université Pierre et Marie

[15] Hart, P. E. (1995) "Galtonian regression across countries and the convergence of productivity." Oxford Bulletin of Economics and Statistics, 57, 287-293.

[16] Hodrick, R. and E. Prescott (1997). "Post war business cycles: An empirical investigation". *Journal of Money Credit and Banking*, 29, 1-16.

[17] Kitagawa, G. and W. Gersch (1996). *Smoothness Priors Analysis of Time Series*. New York: Springer Verlag

[18] 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

[19] Phillips, P. C. B. (2004). "Challenges of Trending Time Series Econometrics". *Mathematics and Computers in Simulation* (forthcoming)

[20] Phillips, P.C.B. and H.R. Moon (1999) : Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, 67, 1057–1111.

[21] Phillips, P. C. B. and S. Jin, (2002). "Limit Behavior of the Whittaker (HP) Filter". Yale University, mimeographed.

[22] Phillips, P. C. B. and D. Sul (2003). "The elusive empirical shadow of growth convergence", Cowles Foundation Discussion Paper #1398, Yale University.

[23] Phillips, P. C. B. and D. Sul (2005). "Alternative Interpretation of the Augmented Solow Regressions", mimeo, University of Auckland

[24] Quah, D. (1993), "Galton's Fallacy and Tests of the Convergence Hypothesis," *Scandinavian Journal of Economics* 95, 427-443.

[25] Sala-i-Martin, X. X. (1997). "I just ran two million regressions" American Economic Review, 87, 178-183.

[26] Sun, Y. (2001), "Catching up, Forging ahead, and Falling Behind: A Panel Structure Analysis of Convergence Clubs". UCSD mimeographed.

[27] Whittaker, E. T. (1923). On a new method of graduation. Proceedings of the Edinburgh Mathematical Association, 78, 81-89.

Technical Appendix

Appendix A: Standardizing the Growth Components

This Appendix provides an analysis of how the growth components in (11), i.e., the elements in the decomposition

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

may be standardized to yield the transition and growth curves. We let $t \to \infty$ and characterize the limiting behavior of the components b_{it} and μ_t.

We first proceed as if the growth components were nonstochastic. Suppose $x_{it} t = f_i(t)$ is regularly varying at infinity with power exponent γ_i (e.g. Seneta, 1976) so that

$$f_i(t) = t^{\gamma_i} W_i(t), \tag{30}$$

where $W_i(t)$ is slowly varying at infinity, viz. $W_i(\lambda t)/W_i(t) \to 1$ as $t \to \infty$ for all $\lambda > 0$. Similarly, let μ_t be regularly varying at infinity with power exponent $\gamma > 0$ so that

$$\mu_t = t^{\gamma} Z(t) \tag{31}$$

for some slowly varying function $Z(t)$. The regular variation requirement means that $f_i(t)$ and μ_t both behave asymptotically like power functions for large t. In the simplest case where the common growth component is a linear drift (i.e., $\mu_t = t$) and $x_{it} \to x$ for all i as $t \to \infty$, there is growth convergence and we have $\gamma_i = \gamma = 1$ and $W_i(t) = Z(t) = 1$. Conditions (30) and (31) allow for a much wider variety of asymptotic behavior, including the possibility that individual i economy's growth may deviate from the common path (when $\gamma_i \neq \gamma$).

Set $t = [Tr]$ for some $r > 0$ representing the fraction of the overall sample T corresponding to observation t. Then under (30)

$$T^{-\gamma_i} x_{it} t = T^{-\gamma_i} [Tr]^{\gamma_i} \frac{W_i(Tr)}{W_i(T)} W_i(T) \sim r^{\gamma_i} W_i(T), \tag{32}$$

and

$$T^{-\gamma} \mu_t = T^{-\gamma} [Tr]^{\gamma} \frac{Z(Tr)}{Z(T)} Z(T) \sim r^{\gamma} Z(T).$$

We deduce from this asymptotic behavior and (11) that

$$\begin{aligned} T^{-\gamma_i} \log y_{it} &= \frac{a_{it} + x_{it} t}{T^{\gamma_i}} = \frac{a_{it}}{T^{\gamma_i}} + \frac{x_{it} t}{T^{\gamma_i}} \sim r^{\gamma_i} W_i(T), \\ T^{-\gamma} \mu_t &\sim r^{\gamma} Z(T) = \mu(r) Z(T). \end{aligned}$$

where $\mu(r) = r^{\gamma}$.

Writing, as in (11),

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

we then have

$$\begin{aligned}
\frac{1}{T^{\gamma_i}}\left(\frac{a_{it}+x_{it}t}{\mu_t}\right)\mu_t &= \frac{a_{it}}{T^{\gamma_i}}+\frac{x_{it}t}{T^{\gamma_i}}\frac{T^{\gamma}}{\mu_t}\left(\frac{\mu_t}{T^{\gamma}}\right)\\
&= o(1)+\frac{x_{it}t}{T^{\gamma_i}}\frac{T^{\gamma}}{\mu_t}\left(\frac{\mu_t}{T^{\gamma}}\right)\\
&\sim \left\{r^{\gamma_i-\gamma}J_i(T)\right\}\left\{r^{\gamma}Z(T)\right\}\\
&= b^J_{iT}(r)\,\mu^Z_T(r),
\end{aligned}$$

where $J_i(T) = W_i(T)/Z(T)$ is slowly varying at infinity. Thus, the functions $b^J_{iT}(r)$ and $\mu^Z_T(r)$ are regularly varying and behave asympotically like the power functions $r^{\gamma_i-\gamma}$ and r^{γ}, at least up to slowly varying factors.

Next set $d_{iT} = T^{\gamma_i}J_i(T)Z(T) = T^{\gamma_i}W_i(T)$, so that the slowly varying components are factored into the standardization. Then, for $t=[Tr]$, we have

$$\begin{aligned}
\frac{1}{d_{iT}}\log y_{it} &= \frac{1}{T^{\gamma_i}J_i(T)Z(T)}\left(\frac{a_{it}+x_{it}t}{\mu_t}\right)\mu_t\\
&= \frac{a_{it}}{T^{\gamma_i}W_i(T)}+\frac{x_{it}t}{T^{\gamma_i}W_i(T)}\frac{T^{\gamma}Z(T)}{\mu_t}\left(\frac{\mu_t}{T^{\gamma}Z(T)}\right)\\
&= o(1)+\left(\frac{x_{it}t}{T^{\gamma_i}W_i(T)}\frac{T^{\gamma}Z(T)}{\mu_t}\right)\left(\frac{\mu_t}{T^{\gamma}Z(T)}\right)\\
&= o(1)+b_{iT}\left(\frac{t}{T}\right)\mu_T\left(\frac{t}{T}\right) &(33)\\
&\sim b_{iT}(r)\,\mu_T(r). &(34)
\end{aligned}$$

In (33) we define

$$\mu_T\left(\frac{t}{T}\right)=\frac{\mu_t}{T^{\gamma}Z(T)}=\frac{\mu_t}{t^{\gamma}Z(t)}\frac{t^{\gamma}Z(t)}{T^{\gamma}Z(T)}=\left(\frac{t}{T}\right)^{\gamma}\frac{Z\left(\frac{t}{T}T\right)}{Z(T)}, \tag{35}$$

and in a similar manner

$$b_{iT}\left(\frac{t}{T}\right)=\left(\frac{t}{T}\right)^{\gamma_i-\gamma}\frac{W_i\left(\frac{t}{T}T\right)Z(T)}{W_i(T)Z\left(\frac{t}{T}T\right)}. \tag{36}$$

Then, for $t=[Tr]$ we have

$$b_{iT}(r)\to b_i(r)=r^{\gamma_i-\gamma}, \tag{37}$$

and

$$\mu_T(r)\to\mu(r)=r^{\gamma}. \tag{38}$$

Relations (34) to (38) lead to a nonstochastic version of the stated result (18). For a stochastic version, we may continue to assume that the standardized representation (34) applies with an $o_p(1)$ error uniformly in $t \leq T$ and require that

$$\begin{aligned} b_{iT}(r) &\rightarrow_p b_i(r) = r^{\gamma_i - \gamma}, \\ \mu_T(r) &\rightarrow_p \mu(r) = r^{\gamma}, \end{aligned}$$

uniformly in $r \in [0,1]$, so that the limit transition function $b_i(r)$ and growth curve $\mu(r)$ are non random functions.

More generally, the limit functions $b_i(r)$ and $\mu(r)$ may themselves be stochastic processes. For example, if the common growth component μ_t in $\log y_{it}$ is a unit root stochastic trend, then by standard functional limit theory on a suitably defined probability space

$$T^{-1/2}\mu_{[Tr]} = \mu_T(r) \rightarrow_p B(r), \tag{39}$$

for some Brownian motion $B(r)$. In place of (30), suppose that $f_i(t) = x_{it}t/\mu_t$ is stochastically regularly varying at infinity in the sense that $f_i(t)$ continues to follow (30) for some power exponent γ_i but with $W_i(t)$ stochastically slowly varying at infinity, i.e., $W_i(\lambda t)/W_i(t) \rightarrow_p 1$ as $t \rightarrow \infty$ for all $\lambda > 0$. Then, in place of (32) we have

$$\frac{1}{T^{\gamma_i}}\frac{x_{it}t}{\mu_t} = \frac{[Tr]^{\gamma_i}}{T^{\gamma_i}}\frac{W_i(Tr)}{W_i(T)}W_i(T) \sim r^{\gamma_i}W_i(T).$$

Then, setting $d_{iT} = T^{\gamma_i + 1/2}W_i(T)$, $t = [Tr]$ and working in the same probability space where (39) holds, we have

$$\begin{aligned} d_{iT}^{-1}\log y_{it} &= \frac{a_{it}}{T^{\gamma_i+1/2}W_i(T)} + \left(\frac{1}{T^{\gamma_i}W_i(T)}\frac{x_{it}t}{\mu_t}\right)\left(\frac{\mu_t}{\sqrt{T}}\right) \\ &= o_p(1) + b_{iT}(r)\,\mu_T(r) \rightarrow_p b_i(r)\,B(r), \end{aligned}$$

with $b_i(r) = r^{\gamma_i}$. In this case the limiting common trend function is the stochastic process $\mu(r) = B(r)$ and the transition function is the non random function $b_i(r) = r^{\gamma_i}$.

Appendix B: Transition Curve Regression and Convergence Testing

A general theory for the calculation of asymptotic standard errors of fitted curves of the type $\hat{h}_i(r)$ that allow for deterministic and stochastic trend components of unknown form is presently not available in the literature and is beyond the scope of the present paper. Instead, we will confine ourselves to the important special case where the trend function involves a stochastic trend with linear drift and we focus attention in this Appendix on the development of a test of subgroup convergence.

We start by considering a model in which the common trend component μ_t has the form of a stochastic trend with drift. More specifically, let $\mu_t = at + \sum_{j=0}^{t} \varepsilon_j$ for some constant $a > 0$ and where ε_t is a zero mean weakly dependent sequence for which the functional law $T^{-1/2} \sum_{j=0}^{[Tr]} \varepsilon_j \Rightarrow B_\varepsilon (r)$ holds as $T \to \infty$, where B is Brownian motion with variance ω_ε^2, the long run variance of ε_t. We further suppose that (15) holds with the common standardization $d_T = T$ so that

$$T^{-1} \log y_{it} = b_{it} \mu_T \left(\frac{t}{T} \right), \quad \text{with } \mu_T \left(\frac{t}{T} \right) = a \frac{t}{T} + \frac{1}{T} \sum_{s=0}^{t} \varepsilon_s, \tag{40}$$

and that $b_{it} = b_i + b_{it}^0 \to_p b_i$ as $t \to \infty$ for some constants $b_i > 0$. In what follows, we assume that the b_{it}^0 are independent across i and may be written in the form $b_{it}^0 = \sigma_t \xi_{it}$, where $\sigma_t \to 0$ as $t \to \infty$ and ξ_{it} is a zero mean, variance unity, weakly dependent time series for each i. The limit theory that follows makes use of the mean zero sequence

$$\zeta_{jt} = \left(\xi_{jt}^2 - 1 \right) - 2\sigma_b \xi_{jt},$$

and it is convenient to assume that ζ_{jt} satisfies the functional law

$$T^{-1/2} \sum_{t=1}^{[Tr]} \zeta_{jt} \Rightarrow B_j (r) , \tag{41}$$

for some Brownian motions B_j with variances $\omega_j^2 = \text{lrvar}\left(\zeta_{jt}\right)$. Primitive conditions under which (41) holds are quite general and may be derived in terms of linear process models as in Phillips and Solo (1992). Clearly, ω_j^2 depends on fourth moments of the original sequence $\xi_{jt.}$ The variates ζ_{jt}, like ξ_{jt}, are independent across j, so that the limit Brownian motions B_j are also independent across j.

We also make use of the rate condition

$$\frac{1}{T} + \frac{1}{N_G} + \frac{T}{N_G} \to 0,$$

so that both T and N_G pass to infinity and $T = o\left(N_G\right)$.

According to this formulation, in the long run as $t \to \infty$, economy i shares in common growth μ_t up to the proportionality factor b_i (c.f. (45) below). Since $\sigma_t \to 0$ as $t \to \infty$, the component $b_{it}^0 = \sigma_t \xi_{it}$ in b_{it} decays in importance as t increases. In what follows, it is convenient to employ the following parametric formulation

$$\sigma_t = \frac{\sigma_b}{t^\alpha}, \quad \text{for some } \sigma_b, \alpha \geq 0, \tag{42}$$

which is a form of evaporating trend or decay model (when $\alpha > 0$) of the type studied in Phillips (2000). As shown below, this parametric form is helpful in the development of convergence tests and a suitable limit theory that can be used in practical work.

Proceeding with this set up, we have the explicit standardized form

$$\begin{aligned} T^{-1}\log y_{it} &= (b_i+\sigma_t\xi_{it})\left(a\frac{t}{T}+\frac{1}{T}\sum_{s=0}^{t}\varepsilon_s\right) \\ &= ab_i\frac{t}{T}+\sigma_t a\frac{t}{T}\xi_{it}+\frac{b_i}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s+\frac{\sigma_t}{\sqrt{T}}\xi_{it}\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s, \end{aligned} \tag{43}$$

so that for $t=[Tr]$ with $r>0$

$$T^{-1}\log y_{i[Tr]} = ab_i\frac{[Tr]}{T}+\sigma_{[Tr]}\xi_{i[Tr]}a\frac{[Tr]}{T}+\left(\frac{b_i}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{s=0}^{[Tr]}\varepsilon_s\right)+o_p\left(\frac{1}{\sqrt{T}}\right) \tag{44}$$

$$\rightarrow_p \; ab_i r, \tag{45}$$

giving economy i's share in the common growth limit function ar. Note that (45) becomes abr under the hypothesis of growth convergence, where $b_i=b$ for all i and $\alpha>0$.

Averaging (43) and setting $\bar{b}_G = N_G^{-1}\sum_{j\in G}b_j$, $q_{Nt} = N_G^{-1}\sum_{j\in G}\xi_{jt}$ and $\tilde{q}_{it} = \xi_{it} - N_G^{-1}\sum_{j\in G}\xi_{jt}$, we have

$$\begin{aligned} N_G^{-1}\sum_{j\in G}\left(T^{-1}\log y_{jt}\right) &= a\bar{b}_G\frac{t}{T}+\sigma_t a\frac{t}{T}\left(N_G^{-1}\sum_{j\in G}\xi_{jt}\right)+\frac{\bar{b}_G}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s \\ &\quad +\frac{\sigma_t}{\sqrt{T}}\left(N_G^{-1}\sum_{j\in G}\xi_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s\right) \\ &= a\bar{b}_G\frac{t}{T}+\sigma_t a\frac{t}{T}q_{Nt}+\frac{\bar{b}_G}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s+\frac{\sigma_t}{\sqrt{T}}q_{Nt}\left(\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s\right), \end{aligned}$$

and then

$$\begin{aligned} & T^{-1}\log y_{it}-N_G^{-1}\sum_{j\in G}\left(T^{-1}\log y_{it}\right) \\ = \; & a\left(b_i-\bar{b}_G\right)\frac{t}{T}+\frac{\left(b_i-\bar{b}_G\right)}{\sqrt{T}}\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s+\sigma_t a\frac{t}{T}\tilde{q}_{it}+\frac{\sigma_t}{\sqrt{T}}\tilde{q}_{it}\frac{1}{\sqrt{T}}\sum_{s=0}^{t}\varepsilon_s. \end{aligned}$$

It follows that for large t and N_G we have

$$\begin{aligned} h_{it}^G-1 &= \frac{T^{-1}\log y_{it}-N_G^{-1}\sum_{j\in G}\left(T^{-1}\log y_{it}\right)}{N_G^{-1}\sum_{j\in G}\left(T^{-1}\log y_{it}\right)} \\ &= \frac{b_i-\bar{b}_G+\sigma_t\left(\xi_{it}-N_G^{-1}\sum_{j\in G}\xi_{jt}\right)}{\bar{b}_G+\sigma_t\left(N_G^{-1}\sum_{j\in G}\xi_{jt}\right)} = \frac{b_i-\bar{b}_G+\sigma_t\tilde{q}_{it}}{\bar{b}_G+\sigma_t q_{Nt}} \\ &= \frac{\left(b_i-\bar{b}_G\right)}{\bar{b}_G}+\frac{\sigma_t}{\bar{b}_G}\xi_{it}+O_p\left(\left\{\left|b_i-\bar{b}_G\right|+\sigma_t\right\}\frac{1}{\sqrt{N_G}}\right) \end{aligned} \tag{46}$$

The factors $\left|b_i - \bar{b}_G\right|$ and σ_t are retained in the error term because the error magnitude in (46) is influenced by whether there is group homogeneity in the b_i, in which case $b_i - \bar{b}_G = 0$, and by the value of α in the decay model (42) for σ_t.

Squaring (46), we have

$$\left(h_{it}^G - 1\right)^2 = \frac{\left(b_i - \bar{b}_G\right)^2 + \sigma_t^2 \tilde{q}_{it}^2 + 2\sigma_t \left(b_i - \bar{b}_G\right) \tilde{q}_{it}}{\left\{\bar{b}_G + \sigma_t q_{Nt}\right\}^2} \tag{47}$$

Setting

$$\begin{aligned}
\tilde{b}_j &= b_j - \bar{b}_G,\ \sigma_{\xi G}^2 = \lim_{N_G \to \infty} N_G^{-1} \sum_{j \in G} E\xi_{jt}^2 = 1, \quad \sigma_{bG}^2 = \lim_{N_G \to \infty} N_G^{-1} \sum_{j \in G} \tilde{b}_j^2, \\
\sigma_{b\xi G} &= \lim_{N_G \to \infty} N_G^{-1} \sum_{j \in G} E\left(\xi_{jt}\right) \tilde{b}_j, \quad H_t^G = N_G^{-1} \sum_{j \in G} \left(h_{jt}^G - 1\right)^2,
\end{aligned}$$

and summing (47) leads to

$$\begin{aligned}
& H_t^G \\
= \ & N_G^{-1} \sum_{j \in G} \left\{\tilde{b}_j^2 + \sigma_t^2 \tilde{q}_{it}^2 + 2\sigma_t \tilde{b}_j \tilde{q}_{it}\right\} \div \left\{\bar{b}_G + \sigma_t q_{Nt}\right\}^2 \\
= \ & \left\{\left(\sigma_{bG}^2 + \left[N_G^{-1} \sum_{j \in G} \left(\tilde{b}_j^2 - \sigma_{bG}^2\right)\right]\right) + \sigma_t^2 \left(\sigma_{\xi G}^2 + \left[N_G^{-1} \sum_{j \in G} \left(\xi_{jt}^2 - \sigma_{\xi G}^2\right)\right] - q_{Nt}^2\right)\right. \\
& \left. + 2\sigma_t \left(\sigma_{b\xi G} + \left[N_G^{-1} \sum_{j \in G} \left(\tilde{b}_j \tilde{q}_{it} - \sigma_{b\xi G}\right)\right]\right)\right\} \div \left\{\bar{b}_G^2 + 2\sigma_t \bar{b}_G q_{Nt} + \sigma_t^2 q_{Nt}^2\right\} \\
= \ & \bar{b}_G^{-2} \left\{\left(\sigma_{bG}^2 + \left[N_G^{-1} \sum_{j \in G} \left(\tilde{b}_j^2 - \sigma_{bG}^2\right)\right]\right) + \sigma_t^2 \left(\sigma_{\xi G}^2 + \left[N_G^{-1} \sum_{j \in G} \left(\xi_{jt}^2 - \sigma_{\xi G}^2\right)\right]\right)\right. \\
& \left. + 2\sigma_t \left(\sigma_{b\xi G} + \left[N_G^{-1} \sum_{j \in G} \left(\tilde{b}_j \tilde{q}_{it} - \sigma_{b\xi G}\right)\right]\right) + O_p\left(\frac{\sigma_t^2}{N_G}\right)\right\} \times \left\{1 - \frac{2\sigma_t}{\bar{b}_G} q_{Nt} + O_p\left(\frac{\sigma_t^2}{N_G}\right)\right\} \\
= \ & \left\{\frac{\sigma_{bG}^2}{\bar{b}_G^2} + \frac{N_G^{-1} \sum_{j \in G} \left(\tilde{b}_j^2 - \sigma_{bG}^2\right)}{\bar{b}_G^2}\right\} \left\{1 - \frac{2\sigma_t}{\bar{b}_G} q_{Nt}\right\} \\
& + \frac{\sigma_t^2 \left(\sigma_{\xi G}^2 + \left[N_G^{-1} \sum_{j \in G} \left(\xi_{jt}^2 - \sigma_{\xi G}^2\right)\right]\right)}{\bar{b}_G^2} \left\{1 - \frac{2\sigma_t}{\bar{b}_G} q_{Nt}\right\} \\
& + \frac{2\sigma_t}{\bar{b}_G^2} \left(\sigma_{b\xi G} + \left[N_G^{-1} \sum_{j \in G} \tilde{b}_j \tilde{q}_{it} - \sigma_{b\xi G}\right]\right) \left\{1 - \frac{2\sigma_t}{\bar{b}_G} q_{Nt}\right\} \\
& + O_p\left(\frac{\sigma_t^2}{N_G}\right)
\end{aligned} \tag{48}$$

as $N_G \to \infty$.

When $b_j = b = \bar{b}_G$ for all $j \in G$, we have $\sigma^2_{bG} = \sigma_{b\xi G} = 0$ and (48) then becomes

$$\begin{aligned} H^G_t &= \sigma^2_b \frac{\sigma^2_{\xi G} + \left[N_G^{-1}\sum_{j\in G}\left(\xi^2_{jt} - \sigma^2_{\xi G}\right)\right]}{\bar{b}^2_G t^{2\alpha}}\left\{1 - \frac{2\sigma_t}{\bar{b}_G} q_{Nt}\right\} + O_p\left(\frac{\sigma^2_t}{N_G}\right) \\ &= \frac{\sigma^2_b \sigma^2_{\xi G}}{b^2 t^{2\alpha}} + \frac{\sigma^2_b}{b^2 t^{2\alpha}}\left\{N_G^{-1}\sum_{j\in G}\left(\xi^2_{jt} - \sigma^2_{\xi G}\right)\right\} - 2\frac{\sigma^3_b \sigma^2_{\xi G}}{b^2 t^{3\alpha}} q_{Nt} + O_p\left(\frac{\sigma^2_t}{N_G}\right) \\ &= \frac{A}{t^{2\alpha}}\left\{1 + \frac{1}{\sigma^2_{\xi G}}\left\{N_G^{-1}\sum_{j\in G}\left(\xi^2_{jt} - \sigma^2_{\xi G}\right)\right\} - 2\frac{\sigma_b}{t^\alpha} q_{Nt}\right\} + O_p\left(\frac{\sigma^2_t}{N_G}\right) \end{aligned} \tag{49}$$

where

$$A = \frac{\sigma^2_b \sigma^2_{\xi G}}{b^2} = \frac{\sigma^2_b}{b^2},$$

since $\sigma^2_{\xi G} = 1$.

Taking logarithms of (49) gives

$$\log H^G_t = \log A - 2\alpha \log t + u^G_t, \tag{50}$$

where

$$\begin{aligned} u^G_t &= \log\left\{1 + \frac{1}{\sigma^2_{\xi G}}\left\{N_G^{-1}\sum_{j\in G}\left(\xi^2_{jt} - \sigma^2_{\xi G}\right)\right\} - 2\frac{\sigma_b}{t^\alpha} q_{Nt} + O_p\left(\frac{1}{N_G}\right)\right\} \\ &= \left\{N_G^{-1}\sum_{j\in G}\left(\xi^2_{jt} - 1\right)\right\} - 2\frac{\sigma_b}{t^\alpha} q_{Nt} + O_p\left(\frac{1}{N_G}\right) \end{aligned} \tag{51}$$

giving equation (29) in the text.

As indicated in the text, we may test the null hypothesis $H_0 : \alpha = 0$ against the alternative $H_1 : \alpha > 0$ by means of a least squares regression of (50) using observations $t = [Tr], [Tr] + 1, ..., T$ for some $r > 0$. The alternative H_1 therefore corresponds to the case of subgroup G growth convergence under the maintained hypothesis $b_j = b$ for all $j \in G$. We now proceed to develop the limit theory for this regression.

Write (50) as

$$z_t = f + d \log t + u^G_t = f_T + d \log\left(\frac{t}{T}\right) + u^G_t \tag{52}$$

where $f_T = f + d \log T = \log A - 2\alpha \log T$. Define

$$\tau_t = \log \frac{t}{T} - \overline{\log \frac{t}{T}}.$$

where $\overline{\log \frac{t}{T}} = \frac{1}{T-[Tr]}\sum_{t=[Tr]}^{T} \log \frac{t}{T}$. Then regression over $t = [Tr], [Tr]+1, ..., T$ for some $r > 0$

yields

$$\hat{d} - d = \frac{\sum_{t=[Tr]}^{T} \tau_t u_t^G}{\sum_{t=[Tr]}^{T} \tau_t^2} = \frac{N_G^{-1} \sum_{j \in G} \sum_{t=[Tr]}^{T} \tau_t \left(\xi_{jt}^2 - 1 \right)}{\sum_{t=[Tr]}^{T} \tau_t^2}$$
$$-2 \frac{\sigma_b \left(N_G^{-1} \sum_{j \in G} \sum_{t=[Tr]}^{T} \tau_t \frac{\xi_{jt}}{t^\alpha} \right)}{\sum_{t=[Tr]}^{T} \tau_t^2} + O_p\left(\frac{1}{N_G} \right), \tag{53}$$

Next observe that

$$\begin{aligned}
\sum_{t=[Tr]}^{T} \tau_t^2 &= \sum_{t=[Tr]}^{T} \left(\log \frac{t}{T} - \overline{\log \frac{t}{T}} \right)^2 = T \int_r^1 \left(\log s - \frac{1}{1-r} \int_r^1 \log p dp \right)^2 ds \\
&= T \left\{ \int_r^1 \log^2 s ds - \frac{1}{1-r} \left(\int_r^1 \log p dp \right)^2 \right\} \\
&= T \left\{ (1-r) - \left(\frac{r}{1-r} \right) \log^2 r \right\},
\end{aligned} \tag{54}$$

by Euler summation and direct evaluation of the integral

$$\int_r^1 \log^2 s ds - \frac{1}{1-r} \left(\int_r^1 \log p dp \right)^2 = (1-r) - \left\{ \frac{r}{1-r} \right\} \log^2 r.$$

Scaling (53) by $\sqrt{TN_G}$ we have

$$\sqrt{TN_G} \left(\hat{d} - d \right) = \frac{N_G^{-1/2} \sum_{j \in G} \left\{ T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \left(\xi_{jt}^2 - 1 \right) \right\}}{T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2}$$
$$-2 \frac{\sigma_b \left(N_G^{-1/2} \sum_{j \in G} \left\{ T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \frac{\xi_{jt}}{t^\alpha} \right\} \right)}{T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2} + O_p\left(\sqrt{\frac{T}{N_G}} \right), \tag{55}$$

Under the null hypothesis where $\alpha = 0$, this expression becomes

$$\sqrt{TN_G} \left(\hat{d} - d \right) = \frac{N_G^{-1/2} \sum_{j \in G} \left\{ T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \zeta_{jt} \right\}}{T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2} + O_p\left(\sqrt{\frac{T}{N_G}} \right) \tag{56}$$

where $\zeta_{jt} = \left(\xi_{jt}^2 - 1 \right) - 2\sigma_b \xi_{jt}$ is a zero mean time series satisfying the functional law (41). It follows that

$$\begin{aligned}
T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \zeta_{jt} &= T^{-1/2} \sum_{t=[Tr]}^{T} \left(\log \frac{t}{T} - \overline{\log \frac{t}{T}} \right) \zeta_{jt} \\
&\Rightarrow \int_r^1 \left(\log s - \frac{1}{1-r} \int_r^1 \log p dp \right) dB_j(s) \\
&\equiv N\left(0, \omega_j^2 \int_r^1 \left(\log s - \frac{1}{1-r} \int_r^1 \log p dp \right)^2 \right) \\
&\equiv N\left(0, \omega_j^2 \left\{ (1-r) - \left(\frac{r}{1-r} \right) \log^2 r \right\} \right),
\end{aligned} \tag{57}$$

by standard weak convergence arguments. The Brownian motions $B_j(r)$ are independent across j in view of the independence of the ζ_{jt}. Note that the limit (57) is well defined for all $r \in [0,1)$.

Using independence across j and applying sequential limits in which $T \to \infty$, followed by $N_G \to \infty$, we obtain

$$\begin{aligned}
& N_G^{-1/2} \sum_{j\in G} T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \zeta_{jt} \\
\sim \quad & N_G^{-1/2} \sum_{j\in G} \left\{ \int_r^1 \left(\log s - \frac{1}{1-r} \int_r^1 \log p dp \right) dB_j(s) \right\} \\
\sim \quad & N\left(0, \left(N_G^{-1} \sum_{j\in G} \omega_j^2\right) \left\{ (1-r) - \left(\frac{r}{1-r}\right) \log^2 r \right\} \right). \qquad (58)
\end{aligned}$$

It now follows from (56) and (58) that

$$\begin{aligned}
\sqrt{TN_G}\left(\hat{d} - d\right) \quad &= \quad \frac{N_G^{-1/2} \sum_{j\in G} \left\{ T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \zeta_{jt} \right\}}{T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2} + O_p\left(\sqrt{\frac{T}{N_G}}\right) \\
&\Rightarrow \quad N\left(0, \frac{\omega_G^2}{\left\{ (1-r) - \left(\frac{r}{1-r}\right) \log^2 r \right\}}\right). \qquad (59)
\end{aligned}$$

in sequential limits as $T \to \infty$, followed by $N_G \to \infty$, where $\omega_G^2 = \lim_{N_G \to \infty} N_G^{-1} \sum_{j\in G} \omega_j^2$. Using the methods of Phillips and Moon (1999) and some additional regularity conditions, this result can presumably be extended to hold under joint limiting arguments where T and N_G may pass to infinity jointly.

Using the original parameterization of (50), we have $d = -2\alpha$ and it follows from (59) that

$$\sqrt{TN_G}\left(\hat{\alpha} - \alpha\right) \Rightarrow N\left(0, V(r)\right), \qquad (60)$$

where the limiting variance

$$V(r) = \frac{\omega_G^2}{4\left\{ (1-r) - \left(\frac{r}{1-r}\right) \log^2 r \right\}} \qquad (61)$$

is graphed as a function of r in Fig. 17 for $\omega_G^2/4 = 1$. $V(r)$ increases monotonically with r and is unbounded as $r \to 1$ because in that event the regression uses less than $O(T)$ time series observations and the rate of convergence is correspondingly slower than $\sqrt{TN_G}$.

Inference can be conducted using (60), the expression (61) and an estimate of ω_G^2, which

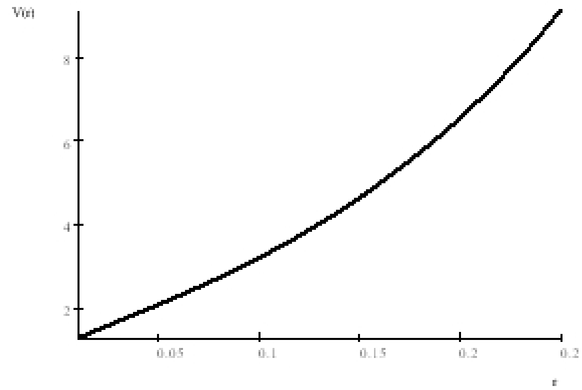


Figure 17: Limiting Variance $V(r)$ of $\sqrt{T}\left(\hat{\alpha}-\alpha\right)$ for $\omega_{\varepsilon}^{2}=1$

involves the average long run variance ω_G^2 of ζ_{jt}. In view of (51) we have

$$\begin{aligned} u_t^G &= \frac{1}{\sigma_{\xi G}^2}\left\{N_G^{-1}\sum_{j\in G}\left(\xi_{jt}^2-\sigma_{\xi G}^2\right)\right\}-2\frac{\sigma_b}{t^\alpha}\left(N_G^{-1}\sum_{j\in G}\xi_{jt}\right)+O_p\left(\frac{1}{N_G}\right) \\ &= N_G^{-1}\sum_{j\in G}\left\{\left(\frac{\xi_{jt}^2}{\sigma_{\xi G}^2}-1\right)-2\frac{\sigma_b}{t^\alpha}\xi_{jt}\right\}+O_p\left(\frac{1}{N_G}\right) \\ &= N_G^{-1}\sum_{j\in G}\left\{\left(\frac{\xi_{jt}^2}{\sigma_{\xi G}^2}-1\right)-2\sigma_b\xi_{jt}\right\}+O_p\left(\frac{1}{N_G}\right) \\ &= N_G^{-1}\sum_{j\in G}\zeta_{jt}+O_p\left(\frac{1}{N_G}\right), \end{aligned}$$

under the null hypothesis with $\alpha=0$. Since the ζ_{jt} are independent across j, direct calculation gives

$$\sum_{h=-\infty}^{\infty}E\left(u_t^G u_{t+h}^G\right)=N_G^{-2}\sum_{j\in G}\left(\sum_{h=-\infty}^{\infty}E\left(\zeta_{jt}\zeta_{jt+h}\right)\right)=N_G^{-2}\sum_{j\in G}\omega_j^2,$$

so that

$$N_G\sum_{h=-\infty}^{\infty}E\left(u_t^G u_{t+h}^G\right)=N_G^{-1}\sum_{j\in G}\omega_j^2\rightarrow\omega_G^2,$$

as $N_G\rightarrow\infty$. It follows that N_G times the long run variance of u_t^G is a consistent estimate of ω_G^2. Hence, ω_G^2 may be consistently estimated by computing a conventional long run variance estimate $\hat{\omega}_{\hat{u}_t^G}^2$ from the residuals, $\hat{u}_t^G$, of the regression (50) and scaling this estimate by N_G, giving $\hat{\omega}_G^2=N_G\hat{\omega}_{\hat{u}_t^G}^2$.

Thus, to test the null hypothesis $H_0:\alpha=0$ of no convergence we fit the regression model

(50) over $t = [Tr], ..., T$ and compute the t- statistic $t_{Gr} = \frac{\hat{\alpha}}{s_{\hat{\alpha}}}$,where

$$s_{\hat{\alpha}}^2 = \frac{1}{N_G T} \frac{\hat{\omega}_G^2}{4\left\{(1-r) - \left(\frac{r}{1-r}\right)\log^2 r\right\}} = \frac{1}{T} \frac{\hat{\omega}_{\hat{u}_t^G}^2}{4\left\{(1-r) - \left(\frac{r}{1-r}\right)\log^2 r\right\}}. \tag{62}$$

If the effective sample size, $T - [Tr]$, were used in computing the standard error, we would have

$$\tilde{s}_{\hat{\alpha}}^2 = \frac{1}{N_G (T - [Tr])} \frac{\hat{\omega}_G^2}{4\left\{(1-r) - \left(\frac{r}{1-r}\right)\log^2 r\right\}} \sim \frac{1}{T} \frac{\hat{\omega}_{\hat{u}_t^G}^2}{4\left\{(1-r)^2 - r\log^2 r\right\}}.$$

We reject H_0 in favor of $H_1 : \alpha > 0$ for a one-sided significant statistic t_{Gr}.

The power of this test can be derived using (55). From this equation, we have

$$\begin{aligned}
\sqrt{TN_G}(\hat{\alpha} - \alpha) &= -\frac{N_G^{-1/2} \sum_{j\in G} \left\{T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \left(\xi_{jt}^2 - 1\right)\right\}}{2T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2} \\
&\quad + \frac{\sigma_b \left(N_G^{-1/2} \sum_{j\in G} \left\{T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \frac{\xi_{jt}}{t^\alpha}\right\}\right)}{T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2} + O_p\left(\sqrt{\frac{T}{N_G}}\right) \\
&= -\frac{N_G^{-1/2} \sum_{j\in G} \left\{T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \left(\xi_{jt}^2 - 1\right)\right\}}{2T^{-1} \sum_{t=[Tr]}^{T} \tau_t^2} + O_p\left(\frac{1}{T^\alpha} + \sqrt{\frac{T}{N_G}}\right), \quad (63)
\end{aligned}$$

since

$$\begin{aligned}
T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \frac{\xi_{jt}}{t^\alpha} &= \frac{1}{T^\alpha} \sum_{t=[Tr]}^{T} \frac{\tau_t}{\left(\frac{t}{T}\right)^\alpha} \left(\frac{\xi_{jt}}{\sqrt{T}}\right) \\
&\sim \frac{1}{T^\alpha} \int_r^1 \frac{\log s - \frac{1}{1-r}\int_r^1 \log p dp}{s^\alpha} dB_{\xi_j}(s) \quad (64) \\
&= O_p\left(T^{-\alpha}\right),
\end{aligned}$$

where $T^{-1/2} \sum_{t=1}^{[Tr]} \xi_{jt} \Rightarrow B_{\xi_j}(r)$, for some Brownian motion B_{ξ_j} and where the integral in (64) is finite for all $r > 0$ and $\alpha > 0$. We deduce from (63) that for all $\alpha > 0$

$$\begin{aligned}
\sqrt{TN_G}(\hat{\alpha} - \alpha) &\sim -N_G^{-1/2} \sum_{j\in G} \left\{\int_r^1 \left(\log s - \frac{1}{1-r}\int_r^1 \log p dp\right) dB_{\xi_j^2}(s)\right\} \\
&\sim N\left(0, \left(N_G^{-1} \sum_{j\in G} \omega_{\xi_j^2}^2\right)\left\{(1-r) - \left(\frac{r}{1-r}\right)\log^2 r\right\}\right),
\end{aligned}$$

where $T^{-1/2} \sum_{t=1}^{[Tr]} \left(\xi_{jt}^2 - 1\right) \Rightarrow B_{\xi_j^2}(r)$, for some Brownian motion $B_{\xi_j^2}$ with variance $\omega_{\xi_j^2}^2$. It follows that for $\alpha > 0$ we have

$$t_{\hat{\alpha}} = \frac{\hat{\alpha}}{s_{\hat{\alpha}}} = \frac{\sqrt{TN_G}(\hat{\alpha} - \alpha)}{\left(TN_G s_{\hat{\alpha}}^2\right)^{1/2}} + \frac{\alpha}{s_{\hat{\alpha}}} = O_p\left(\sqrt{TN_G}\right),$$

and the test is consistent. Some further calculations reveal that the test is powerful against local alternatives of the form $\alpha = \frac{c}{\sqrt{TN_g}}$ for some constant $c > 0$. In this case, we find that

$$\begin{aligned} T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \frac{\xi_{jt}}{t^{\alpha}} &\sim T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \left(\xi_{jt} + O_p\left(\frac{\log T}{\sqrt{TN_G}} \right) \right) \\ &\sim T^{-1/2} \sum_{t=[Tr]}^{T} \tau_t \xi_{jt}, \end{aligned}$$

so that (56) continues to hold, as in the null $\alpha = 0$ case. It follows that the localizing sequence $\alpha = \frac{c}{\sqrt{TN_g}}$ provides a shift in the location of the limit distribution so that

$$\begin{aligned} t_{\hat{\alpha}} &= \frac{\hat{\alpha}}{s_{\hat{\alpha}}} = \frac{\sqrt{TN_G}\,(\hat{\alpha} - \alpha)}{\left(TN_G s_{\hat{\alpha}}^2\right)^{1/2}} + \frac{c}{\left(TN_G s_{\hat{\alpha}}^2\right)^{1/2}} \\ &\sim N\left(c \left\{ \frac{\omega_G^2}{4\left\{(1-r) - \left(\frac{r}{1-r}\right)\log^2 r\right\}} \right\}^{-1/2}, 1 \right), \end{aligned}$$

which gives the local asymptotic power function of the $\log t$ test.

Appendix C: Transitioning and Subsample Convergence Testing

We propose the following two-step approach to subsample testing of transitioning to convergence, where members of one economic subgroup catch up to members of a higher income subgroup.

Step 1 For each group $G \in \mathcal{G}$, run the $\log t$ regression and obtain the t-ratio, t_G, of the coefficient of $\log t$. Under the null hypothesis of no convergence, as $T \to \infty$, all the t_G statistics have a limiting standard normal distribution as derived in Appendix B. For finite T, the distribution of the t_G will be much more complex. Take the maximum value of the statistic across groups, viz. $t_{\mathcal{G}}^{MAX} = \max_{G \in \mathcal{G}} t_G$.Under the null hypothesis of no convergence (that is the full-group of economies do not converge), then we will accept H_0 if $t_{\mathcal{G}}^{MAX}$ exceeds the left tail critical value (e.g. $t_{\mathcal{G}}^{MAX} > -1.65$ in the case of a 5% test). In this event, some groups may have regression coefficients in $\log t$ regression

$$\ln \sigma_{Gt}^2 = a_G + b_G \ln t + \epsilon_{Gt} \quad \text{for } t = t_0, ..., T \text{ and } G \in \mathcal{G} = N/N_G. \tag{65}$$

that are significantly negative by the usual criterion but others are not significant. We therefore "accept" the null of no convergence over the full set of economies in this case. Simulations indicate that in finite samples the probability of rejecting the null of no convergence for all groups, i.e. $P\left(t_{\mathcal{G}}^{MAX} < -1.65\right)$ in the nominal 5% case, is often too large, giving upwards size distortion or distortion towards acceptance of convergence. In view of this bias towards rejection of the null, we suggest a second step to reduce the size distortion, which is based on a subsampling - search procedure to see if there is more evidence in favor of the null hypothesis.

Step 2 If $t_{\mathcal{G}}^{MAX} < -1.65$ in step 1, form new subgroups (overlapping the original subgroups) by eliminating $\ell = N_G/2$ economies from the highest income group and adding ℓ countries from the second highest income group to form a new subgroup with the same number of economies as before. Repeat this procedure with the remaining economies to form $\mathcal{G}-1$ new subgroups in total. With this new collection of economies, repeat Step 1. We reject H_0 only when the evidence is convincing that across all these subgroups there is rejection of the null. This subsampling modification sets a higher hurdle for rejection of the null or acceptance of convergence and thereby helps to control size distortion. As discussed in the body of the paper, this second step test also provides information about transitional effects between the original subgroups.

Appendix D: Some Simulations of Transition and Convergence

As is apparent from the above discussion, many different data generating mechanisms are compatible with the transition and growth curve formulations. These include models where there is growth convergence as well as models where there is growth divergence and various forms of transitional divergence combined with ultimate convergence. This section provides some brief simulation evidence based on panel data generating processes (dgp's) that include some of these possibilities. These specific models also allow us to develop an asymptotic theory for the $\log t$ regression discussed in the last section and to evaluate the statistical performance of the proposed methods in finite samples of panel data.

We consider the following two panel dgp's. The first involves a nonlinear panel specification which allows for convergence and divergence, while the second dgp is a simple panel of unit root processes with drift.

DGP 1: Nonlinear Panel Model with Common Growth

The following model is a nonlinear panel growth model with factor structure that involves a common trend component μ_t (comprising a random walk with drift) and a loading parameter (b_{it}) that is time varying and whose variance (σ_t^2) may be subject to decay over time depending on whether the convergence parameter $\alpha > 0$. Convergence occurs when $b_i = b_j$ for all $i \neq j$ and $\alpha > 0$. The model has the form

$$
\begin{aligned}
\ln y_{it} &= b_{it}\mu_t, \quad \mu_t = a + \mu_{t-1} + e_t \\
b_{it} &= b_i + b_{it}^*, \quad b_{it}^* = \rho_i b_{it-1}^* + \epsilon_{it}
\end{aligned}
$$

$$
\epsilon_{it} \equiv iid\ N\left(0, \sigma_t^2\right), \quad \sigma_t^2 = \sigma_b^2 / t^{2\alpha}
$$

$$
\rho_i \equiv U\left[0.5, 0.95\right], \quad e_t \equiv N\left(0, 1\right), \quad a = 0.04.
$$

Under the null hypothesis of no convergence, we set $b_i \equiv iid\ N\left(1, \omega_b^2\right)$ with $\omega_b \in (0.1, 0.2, 0.3)$ while under the alternative of convergence $b_i = 1$ for all i and $\alpha > 0$. With HP filtered data, we found a range of empirical values for σ_b around 0.05 along with various values of ω_b under the null hypothesis of no convergence. Four different speeds of convergence are considered by taking $\alpha \in \{1, 0.5, 0.3, 0.0\}$. As α decreases, ultimate convergence takes longer, thereby allowing for the possibility of transitionally divergent behavior for finite T. By contrast, as α increases, some evidence of transitionally convergent behavior is possible for finite T even under the null of no convergence where $b_i \neq b_j$ for $i \neq j$. Hence as α increases (decreases), this leads us to expect that there may be some upwards size distortion in the $\log t$ test for this model under the null of no convergence. When $\alpha = 0$ but $b_i = 1$, then $\ln y_{it}$ does not converge since the error variance σ_t^2 does not tend to zero.

DGP 2: Panel Random Walks with Drift

The following model is a panel of unit root processes driven by autoregressions with some panel cross section dependence structure delivered by a simple factor model with common shocks θ_t over time and idiosyncratic loadings δ_i for each economy.

$$
\begin{aligned}
\ln y_{it} &= a_i + \ln y_{it-1} + u_{it} \\
u_{it} &= \rho_i u_{it-1} + \epsilon_{it} \\
\epsilon_{it} &= \delta_i \theta_t + e_{it}
\end{aligned}
$$

The parameters for this dgp are set by taking estimates from empirical growth regressions with 88 PWT countries. From these estimates we calculate an empirically plausible range of values for the parameters and then draw the parameter settings randomly from uniform distributions set with these limits. Accordingly, we set $a_i \equiv iid\ U\left(0.001, 0.061\right)$, $\delta_i \equiv iid\ U\left(1, 4\right)$, $\theta_t \equiv iid N(0,1)$, $u_{it} \equiv iid\ N(0,1)$ and $e_{it} \equiv iid\ N\left(0, \sigma_{e_i}^2\right)$ with $\sigma_{e_i} \equiv iid\ U\left(0.00, 0.064\right)$. With HP filtered data, we found a range of empirical values for ρ_i in the interval $(0.94, 1.05)$, while with the raw data we found values of ρ_i in the much wider interval $(-0.50, 0.65)$, in both cases making no adjustments for small sample autoregressive bias. Accordingly, for the simulations we excluded negative ρ_i values and considered the following four cases: Case I: $\rho_i \in (0, 0.1)$, Case II: $\rho_i \in (0, 0.3)$, Case III: $\rho_i \in (0, 0.5)$, Case IV: $\rho_i \in (0, 0.7)$ and Case V: $\rho_i \in (0, 0.9)$.

Results

The results are reported in Appendix Tables A1 and A2, which show performance for both dgp's and both the all-group $\log t$ test and the sequential subgroup tests outlined above.

We consider the results for DGP 1 first. Here, the $\log t$ test applied to all countries in a single group shows size distortion, with actual size in the range $(0.33, 0.62)$ for a nominal 5% test. However, in the case where $\alpha = 0$ and $b_i = 1$, for which the test is specifically designed, actual size is much better and is 0.12 for the all-group test. The sequential subgroup tests are conservatively sized without showing any serious lowering of power (only in the case $\alpha = 0.3$ does power drop below unity).

Under DGP 2, the actual size of the test is lower in cases where there are more large ρ_i in the panel. In such (near unit root) cases, we can reasonably expect the divergence behavior to be more marked in the panel, so that there is less size distortion in consequence. Note that under DGP 2, as equation (75) shows, when there are large numbers of economies H_t^G may appear to decrease over time (although not to zero), and this reduction results in some size distortion in the test. Again, the subgroup tests help to reduce this size distortion.

Table A1: Rejection Rate for DGP 1: Nonlinear Panel Specification. (Nominal 5% Test), N=88, T=36. First 6 observations discarded ($r = 1/6$).

ω_b	$\alpha = 0.5$			$\alpha = 0.3$		
	Step 1	Step 2	log t	Step 1	Step 2	log t
0.0	1.000	1.000	1.000	0.924	0.903	1.000
0.1	0.331	0.276	0.623	0.333	0.281	0.592
0.2	0.115	0.040	0.438	0.112	0.055	0.400
0.3	0.048	0.004	0.367	0.052	0.009	0.332

$\alpha = 0, \omega_b = 0$		
Step1	Step2	log t
0.006	0.003	0.124

Table A2: Rejection Rate for DGP 2: Random Walks with Drift. (Nominal 5% Test), N=88, T=36. First 6 observations discarded ($r = 1/6$).

Case	**Step1**	**Step2**	$\log t$
1	0.238	0.155	0.373
2	0.240	0.152	0.370
3	0.220	0.134	0.337
4	0.188	0.098	0.282
5	0.171	0.077	0.313

Appendix E: Limit Behavior of the Transition Parameter under Divergence

The ultimate requirement for convergence in a neoclassical framework that allows for transition is that $x_{it} \to x$. Then, log per capita real income follows a linear trend (stationary process) in the steady state. During transition, however, the behavior could be stochastically nonstationary depending on the manner in which x_{it} and β_{it} evolve over time. For instance, x_{it} might follow a martingale that converged almost surely to a constant value x.

There are even more ways of characterizing divergent behavior. A simple mechanism that fits in with research by Bernard and Durlauf (1995, 1996) is to allow for log per capita real income to follow individual random walks with drift, as in

$$\log y_{it} = \xi_{it}, \ \xi_{it} = c_i + \xi_{it-1} + \varepsilon_{it}, \tag{66}$$

where c_i is an idiosyncratic drift parameter and the shocks ε_{it} are zero mean, stationary and ergodic time series with finite variance σ_ε^2 over t. The ε_{it} may well be correlated across i provided the unit root processes have full rank N and there is no cointegration between ξ_{it} and ξ_{jt} for $i \neq j$. Each $\log y_{it}$ has stochastic trend behavior and and there is growth divergence between $\log y_{it}$ and $\log y_{jt}$ for all $i \neq j$. As in (11, we can write (66) in the form[12]

$$\log y_{it} = c_i t + a_i + \sum_{s=1}^{t} \varepsilon_{is} = \left[c_i + \frac{a_i}{t} + \frac{1}{t} \sum_{s=1}^{t} \varepsilon_{is} \right] t = b_{it} \mu_t, \tag{67}$$

where a_i involves initial conditions, which we take to be $O_p(1)$ for all i, where $\mu_t = t$ is the common trend growth, and where

$$b_{it} = c_i + \frac{a_i}{t} + \frac{1}{t} \sum_{s=1}^{t} \varepsilon_{is} \to_{a.s.} c_i \quad \text{as } t \to \infty, \tag{68}$$

by ergodicity. For any economies i and j with $c_i \neq c_j$, $\log y_{it} - \log y_{jt}$ is clearly divergent as $t \to \infty$, even though there may be subperiods of time when their growth paths may be close – for instance, if the partial sum process $\sum_{s=1}^{t} (\varepsilon_{is} - \varepsilon_{js})$ crosses the trend line $(c_i - c_j)\, t + (a_i - a_j)$. In such a case, $\log y_{it}$ and $\log y_{jt}$ might seem to converge as the economies evolve towards such a crossing time but they will diverge subsequently.

Under growth divergence, $\lim_{t \to \infty} b_{it} = c_i$, with $c_i \neq c_j$ for $i \neq j$. Define

$$\bar{c}_N = \frac{1}{N} \sum_{i=1}^{N} c_i, \quad \overline{c^2}_N = \frac{1}{N} \sum_{i=1}^{N} c_i^2.$$

[12] We can also standardize $\log y_{it}$ and write (67) in the form

$$T^{-1} \log y_{it} = T^{-1} \left[c_i t + a_i + \sum_{s=1}^{t} \varepsilon_{is} \right] = \left[c_i + \frac{a_i}{t} + \frac{1}{t} \sum_{s=1}^{t} \varepsilon_{is} \right] \frac{t}{T} = b_{it} \mu \left(\frac{t}{T} \right)$$

with $\mu\left(\frac{t}{T}\right) = \frac{t}{T}$ in a format analogous to (15).

Then, with N fixed, we have $h_{itN} = b_{it}/\frac{1}{N}\sum_{j=1}^{N} b_{jt} \rightarrow_{a.s.} c_i/\bar{c}_N$ as $t \rightarrow \infty$, and, setting $H_{tN} = N^{-1}\sum_{i=1}^{N}(h_{itN}-1)^2$ to be the sample mean squared error of h_{itN} about unity, we get

$$H_{tN} \rightarrow_{a.s.} \frac{1}{N}\sum_{i=1}^{N}\left(\frac{c_i}{\bar{c}_N}-1\right)^2 = \frac{\overline{c^2}_N - \bar{c}_N^2}{\bar{c}_N^2} = d_N > 0.$$

The nonzero limit d_N represents the heterogeneity in the coefficients c_i which, in turn, reflects heterogeneity in the fundamental parameters β_{it} and x_{it}.

Next, we consider the properties of h_{itN} for fixed (large) t and asymptotically large N, corresponding to a full population of economies. We allow ε_{is} to be cross sectionally correlated but it is useful to require laws of large numbers to apply as $N \rightarrow \infty$ in order to study limit behavior. In particular, we require that the moments $\bar{c}_N, \overline{c^2}_N$ have finite limits, viz.,

$$\bar{c}_N \rightarrow \bar{c}, \quad \overline{c^2}_N \rightarrow \overline{c^2} \quad \text{as } N \rightarrow \infty, \tag{69}$$

and assume that as $N \rightarrow \infty$

$$\frac{1}{N}\sum_{i=1}^{N}\varepsilon_{is}, \frac{1}{N}\sum_{i=1}^{N}c_i\varepsilon_{is}, \frac{1}{N}\sum_{i=1}^{N}a_i\varepsilon_{is} \rightarrow_{a.s.} 0, \quad \frac{1}{N}\sum_{i=1}^{N}\varepsilon_{is}\varepsilon_{is+h} \rightarrow_{a.s.} \gamma(h) \tag{70}$$

and that as $t \rightarrow \infty$

$$\sum_{h=-t+1}^{t-1}\gamma(h) \rightarrow \sum_{h=-\infty}^{\infty}\gamma(h) = \omega^2 > 0. \tag{71}$$

Then,

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{t}\sum_{s=1}^{t}\varepsilon_{is}\right)^2 &= \frac{1}{t^2}\sum_{s,p=1}^{t}\frac{1}{N}\sum_{i=1}^{N}\varepsilon_{is}\varepsilon_{ip} \rightarrow_{a.s.} \frac{1}{t^2}\sum_{s,p=1}^{t}\gamma(p-s) \\
&= \frac{1}{t}\sum_{h=-t+1}^{t-1}\left(1-\frac{|h|}{t}\right)\gamma(h) = \frac{\omega^2}{t}\left[1+o(1)\right],
\end{aligned}$$

as $t \rightarrow \infty$. Primitive conditions under which sequential limits such as (70) - (71) hold are provided in Phillips and Moon (1999). Under (69) - (71), we deduce that as $N \rightarrow \infty$ and for large t

$$\frac{1}{N}\sum_{i=1}^{N}b_{it}^2 \rightarrow_{a.s.} \overline{c^2} + \frac{\omega^2}{t} + o\left(\frac{1}{t}\right), \tag{72}$$

$$\frac{1}{N}\sum_{i=1}^{N}h_{itN}^2 \rightarrow_{a.s.} \frac{\overline{c^2}+\frac{\omega^2}{t}}{\bar{c}^2} + o\left(\frac{1}{t}\right), \tag{73}$$

$$\frac{1}{N}\sum_{i=1}^{N}(h_{itN}-1)^2 \rightarrow_{a.s.} \frac{\overline{c^2}-\bar{c}^2+\frac{\omega^2}{t}}{\bar{c}^2} + o\left(\frac{1}{t}\right). \tag{74}$$

We may use the last property to compare the sample mean squared error, σ_{tN}^2, of h_{itN} for large N at two points in time $T_1 = [Tr_1]$ and $T_2 = [Tr_2]$ with $r_2 > r_1$. Result (74) leads to the approximate relation

$$\frac{H_{T_2N}}{H_{T_1N}} \simeq \frac{\overline{c^2} - \bar{c}^2 + \frac{\omega^2}{T_2}\bar{c}^2}{\overline{c^2} - \bar{c}^2 + \frac{\omega^2}{T_1}\bar{c}^2} < 1, \quad \text{for } T_2 > T_1, \tag{75}$$

suggesting that for large N the sample mean squared error H_{tN} may appear to decrease as t increases, but not to zero, when there is growth divergence.

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 for 18 Western OECD countries covering the period from 1500 to 2001. The third panel (C) includes 88 countries (Source: PWT 6.1). 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-Altantic: 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 Dakoda, South Dakoda

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

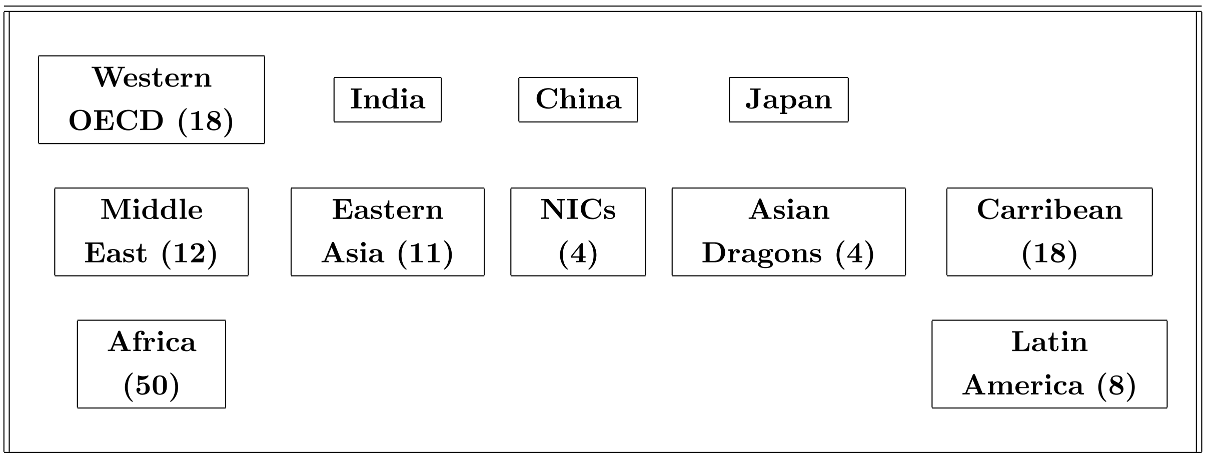
West South Central: Arkansas, Louisiana, Oklahoma, Texas

East South Central: Alabama, Kentucky, Mississippi, Tennesse

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

We form the following subgroups based on geographical location and population.

Regional Subgrouping of 127 World Economies



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

8 Latin American Countries: Argentina, Brazil, Chile, Colombia, Mexico, Peru, Uruguay, Venezuela

15 Carribean Countries: Bolivia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Haïti, Honduras, Jamaica, Nicaragua, Panama, Paraguay, Puerto Rico, Trinidad and Tobago

China, India, Japan

3 Newly Industrialized Countries (NICs): Indonesia, Thailand, Malaysia

11 Eastern Asian Countries: Bangladesh, Burma, Nepal, Pakistan, Sri Lanka, Afghanistan, Cambodia, Laos, Mongolia, Philippines, Vietnam

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

12 Middle East: Bahrain, Iran, Iraq, Israel, Jordan, Lebanon, Oman, South Arabia, Syria, Turkey, Yemen, Palestine and Gaza,

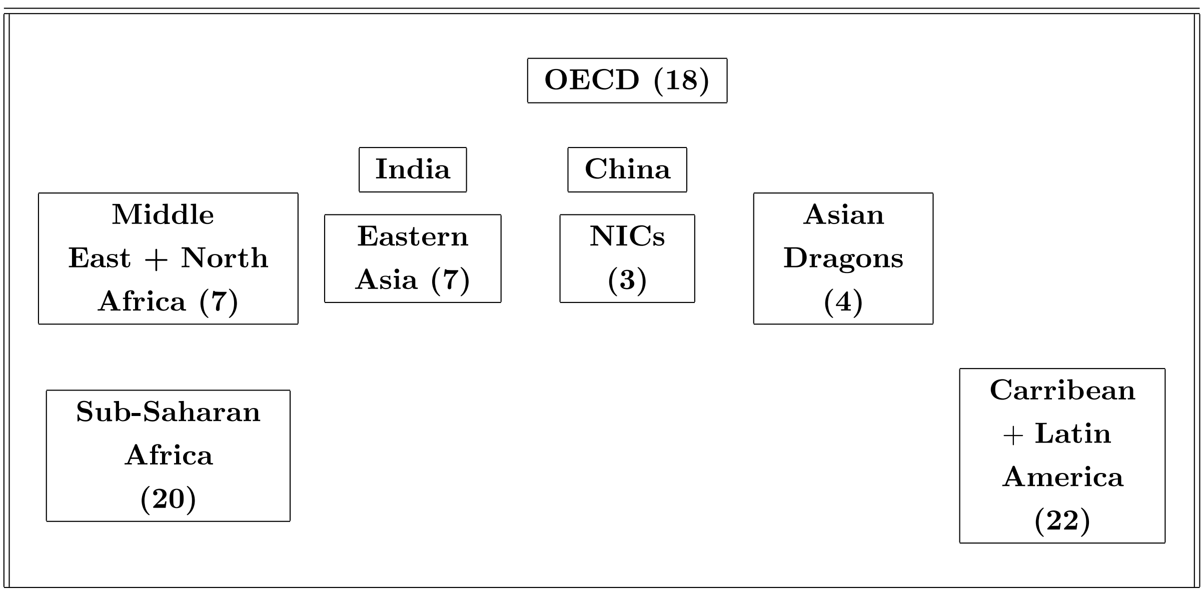
50 African Countries: Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoro Islands, Congo, Côte d'Ivoire, Djibouti, Egypt, Eritrea and Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea, Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Reunion, Rwanda, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zaire, Zambia, Zimbabwe.

127 World Economies: Includes all countries in the above 11 subgroups plus Albania, Greece and Ireland.

Panel Data Set C: 88 PWT Version 6.1 Countries

We form the following subgroups, again based on geographical location and population.

Regional Subgrouping of 88 PWT Countries



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

20 Sub-Saharan Africa Countries: Benin, Botswana, Cameroon, Central African Rep., Congo (D.R.), Gambia, Ghana, Lesotho, Malawi, Mali, Mauritius, Mozambique, Niger, Rwanda, Senegal, South Africa, Togo, Uganda, Zambia, Zimbabwe

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

India, China,

7 Eastern Asian Countries: Bangladesh, Fiji, Nepal, Pakistan, Papua New Guinea, Philippines, Sri Lanka

18 OECD Countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, 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

88 World Economies: The dataset includes all countries in the above 7 subgroups plus Romania, Greece, Iceland, Ireland and Turkey.

5 Subgroups in Figure 1 (Initial Income Ordering)

The country ordering is based on initial income in year 1960. Malawi was the poorest country, while Swizterland was the richest in 1960.

Subgroup 1 (Poorest): Malawi, Uganda, Pakistan, China, Lesotho, Nepal, Rwanda, India, Indonesia, Ghana, Gambia, Botswana, Mali, Congo, Dem. Rep., Togo, Romania, Benin

Subgroup 2 (Poor): Bangladesh, Thailand, Zimbabwe, Zambia, Sri Lanka, Taiwan, South Korea, Egypt, Syria, Cameroon, Honduras, Dominican Rep., Mozambique, Niger, Senegal, Ecuador, Guyana

Subgroup 3 (Middle Income): Philippines, Malaysia, Central African Rep., Guatemala, Paraguay, Brazil, Bolivia, Algeria, Jordan, Singapore, Panama, Papua New Guinea, Colombia, Fiji, Iran, Turkey, Jamaica

Subgroup 4 (Rich): Hong Kong, Cyprus, Nicaragua, Barbados, Costa Rica, Peru, El Salvador, Portugal, Mauritius, Chile, Mexico, Greece, Trinidad &Tobago, Japan, South Africa, Spain, Ireland

Subgroup 5 (Richest): Israel, Uruguay, Italy, Argentina, Austria, Finland, Belgium, France, Iceland, Norway, Venezuela, Netherlands, United Kingdom, Canada, Sweden, Australia, Denmark, New Zealand, USA, Switzerland

4 Subgroups in Log t tests (Last Income Ordering)

The ordering for the countries is based on the final income year 1996. Congo (D.R.) was the poorest country in 1996, while the U.S.A. was the richest in 1996.

Subgroup 1: Congo (D.R), Malawi, Mali, Uganda, Rwanda, Niger, Mozambique, Zambia, Togo, Central African Rep., Benin, Gambia, Ghana, Nepal, Lesotho, Senegal, Bangladesh, Nicaragua, Cameroon, Pakistan, India, Honduras

Subgroup 2: Guyana, Bolivia, Zimbabwe, China, Philippines, Sri Lanka, Papua New Guinea, Jordan, Dominican Rep., Guatemala, Syria, Egypt, Indonesia, Ecuador, Jamaica, El Salvador, Peru, Iran, Algeria, Romania, Costa Rica, Fiji

Subgroup 3: Paraguay, Colombia, Panama, Turkey, Botswana, Brazil, Venezuela, Thailand, South Africa, Mexico, Chile, Malaysia, Uruguay, Trinidad &Tobago, Argentina, Mauritius, Greece, Portugal, South Korea, Barbados, Spain, Taiwan

Subgroup 4: Israel, Cyprus, New Zealand, Ireland, Finland, United Kingdom, France, Italy, Sweden, Netherlands, Belgium, Iceland, Austria, Australia, Canada, Denmark, Singapore, Norway, Japan, Switzerland, Hong Kong, USA.