The New Empirics of Economic Growth

by

Steven N. Durlauf and Danny T. Quah*
University of Wisconsin, Madison and LSE

CENTRE FOR ECONOMIC PERFORMANCE
DISCUSSION PAPER NO. 384
January 1998

This paper is produced as part of the Centre's Programme on National Economic Performance.

---

* This paper was prepared for the *Handbook of Macroeconomics* (John Taylor and Michael Woodford, eds.). We thank the British Academy, the ESRC, the John D. and Catherine T. MacArthur Foundation, the NSF, and the Santa Fe Institute for financial support. Kim-Sau Chung, Donald Hester, Brian Krauth, Rodolfo Manuelli, and John Taylor provided valuable comments. Kim-Sau Chung provided outstanding research assistance. The word "empiric" has, in Webster's Unabridged Dictionary, two definitions: first, "that depending upon the observation of phenomena"; second, "an ignorant and unlicensed pretender; a quack; a charlatan". We have chosen, nevertheless, to use the word in the title as we think it conveys the appropriate meaning.

The New Empirics of Economic Growth

by
Steven N. Durlauf and Danny T. Quah
University of Wisconsin, Madison and LSE

January 1998

ABSTRACT

We provide an overview of recent empirical research on patterns of cross-country growth. The new empirical regularities considered differ from earlier ones, e.g., the well-known Kaldor stylized facts. The new research no longer makes production function accounting a central part of the analysis. Instead, attention shifts more directly to questions like, Why do some countries grow faster than others? It is this changed focus that, in our view, has motivated going beyond the neoclassical growth model.

Keywords: classification, convergence, cross-section regression, distribution dynamics, endogenous growth, neoclassical growth, regression tree, threshold, time series, panel data

JEL Classification: C21, C22, C23, D30, E13, O30, O41

Communications to: D. T. Quah, LSE, Houghton Street, London WC2A 2AE.

[Tel: +44-171-955-7535, Email: `dquah@lse.ac.uk`]

[(URL) `http://econ.lse.ac.uk/~dquah/`]

1. Introduction

Economists study growth across countries for at least three reasons. First, understanding the sources of varied patterns of growth is important: Persistent disparities in aggregate growth rates across countries have, over time, led to large differences in welfare. Second, the intellectual payoffs are high: The theoretical hypotheses that bear on economic growth are broad and, perhaps justifiably, ambitious in scale and scope. Third, the first wave of new empirical growth analyses, by making strong and controversial claims, have provoked yet newer ways of analyzing cross-country income dynamics. These newer techniques are, in turn, generating fresh stylized facts on growth with important implications for theory.

This paper provides one overview of the current state of macroeconomists' knowledge on cross-country growth. Since a number of excellent summaries on this subject already exist (e.g., Barro and Sala-i-Martin [11], Jones [58], Pritchett [86], D. Romer [97]), it is useful to clarify how our presentation differs. First, our emphasis is empirical: We develop different growth models focusing on their observable implications for cross-country income data. To bring out key ideas, we eschew overly-restrictive and detailed parametric assumptions on the theoretical models that we develop below. We seek only restrictions on data that follow from a general class of models. At the same time, we show that it is relatively easy to specialize from our analysis to the various empirical specifications that have become standard in the literature. This allows assessing the generality and robustness of earlier empirical findings.

Second, we provide an organizing framework for the different econometric approaches—time-series, cross-section, panel-data, and distribution dynamics—used by researchers. We survey what we take to be the important and econometrically sound findings, and we attempt to explain the different conclusions found across some of these studies. We describe the links between alternative econometric specifications used in the literature and different observable implications of growth models. By organizing the discussion around a single general framework, we seek to gauge how far the empirical literature has succeeded in discriminating

across alternative theories of growth.

The questions studied in the new empirical growth literature differ from those in earlier empirical work embodying Kaldor's stylized facts [60] or those in a production function (Solow [107]-Denison [32]) accounting exercise. The new literature emphasizes understanding cross-country patterns of income, not the stability within a single economy of factor shares or "great ratios" (the ratio of output to capital, consumption, or investment).[1] It eschews understanding growth exclusively in terms of factor inputs. It freely uses all kinds of auxiliary explanatory factors, thus no longer making the production function residual a primary part of the analysis, as was previously done.

The remainder of this paper is organized as follows. Section 2 develops some initial stylized facts: They differ from those typically given in empirical growth papers. We begin with them as they seem natural from the perspective of the theoretical framework we adopt. Sections 3 and 4 sketch some theoretical models that we use to organize the subsequent presentation of empirical results and models. Our goal is to provide a structure sufficiently rich to accommodate a range of theoretical perspectives and, at the same time, to allow comparing different empirical growth studies.

[1] Some might wish to interpret the original growth models only to explain within-country dynamics over time. Cross-country evidence, therefore, should not be taken to refute or support those theoretical models—especially with parameters and circumstances being so different across economies. There are at least two arguments against this position. First, even accepting the premise, it is long part of scientific analysis that theories be tested by going beyond their original domain and without liberally adding free parameters in the process. Looking rigorously at cross-country evidence to assess growth models is simply part of that research tradition. Second, and more specifically on the topic, economists from at least Kaldor [60] on have marshalled cross-country stylized facts as compelling starting points for discussions about economic growth. Indeed, Lucas [74] and Romer [98, 101] use exactly income comparisons across countries to motivate their endogenous growth analyses.

Section 5 presents empirical models and critically evaluates the empirical findings and methodologies in the literature. Section 6 provides conclusions. Sections 7 and 8 are the Technical and Data Appendices covering material omitted from the main text for expositional convenience.

2. Preliminaries and stylized facts

Theoretical growth models typically analyze the behavior of a single representative national economy. However, turning to the observed historical experiences of national economies in the twentieth century, what is most striking instead is how no single national economy is usefully viewed as representative. Rather, understanding cross-country growth behavior requires thinking about the properties of the cross-country distribution of growth characteristics. What properties are most salient?

A first set of stylized facts relates to the world population distribution. Most of the world's economies are *small*. Over the period 1960–4, the largest 5% of the world's economies contained 59.0% of the world's population; the largest 10% contained 70.9%.[2] A quarter-century later, over the period 1985–9, the largest 5% of economies held 58.3% of the population; the largest 10%, 70.2%. In both periods, the lower 50% of the world's economies ranked by population held in total less than 12.5% of the world's population.

A second set of facts relates to the stability of these cross-country population distributions. For the last 35 years, the percentiles associated with the distribution of population across countries have been remarkably stable. This is not to say that those countries now highly populated have always been highly populated, rather that the distribution of cross-section differences has changed little. Indeed, churning within a stable cross-section distribution will figure prominently in discussions below.

---

[2] Hereafter, "the world's economies" refers to the 122 countries with essentially complete income and population data for 1960–1989 in the Summers-Heston [110] V.6 database. These countries are identified in the Data Appendix below.

Economists have typically been most interested in growth models as a way to understand the behavior of per capita income or per worker output (labor productivity). What are the stylized facts here? From 1960 through 1989, world income per capita grew at an annual average rate of 2.25%. However, the time paths of per capita incomes in individual economies varied widely around that of the world average. Averaged over 1960–4, the poorest 10% of the world's national economies (in per capita incomes, taken at the beginning of the interval) each had per capita incomes less than 0.22 times the world average; those economies contained 26.0% of the world's population. Poor economies therefore appear to be also large ones, although it is actually China alone accounting for most of that population figure. By contrast, the richest 10% of national economies each had per capita incomes exceeding 2.7 times the world average, while all together containing 12.5% of the world's population. By 1985–9 the 10th percentile per capita income level had declined to only 0.15 times the world average—those economies in that poorest 10% then held only 3.3% of the world's population as China became relatively richer and became no longer a member of this group. At the same time the 90th percentile per capita income level *increased* to 3.08 times the world average; the share of the world population in those 10% richest economies fell to 9.3%.

In contrast to the stability of population size distributions, the cross-country distributions of per capita incomes seem quite volatile. The extremes appear to be diverging away from each other—with the poor becoming poorer, and the rich richer. However, that is not the entire picture. In 1960–4, the income distance between the 15th and 25th percentiles was 0.13 times world per capita income; by 1985–9, this distance had fallen to 0.06. Over this same time period, the income distance between the 85th and 95th percentiles fell from 0.98 times world per capita income to 0.59. Thus, while the overall spread of incomes across countries increased over this 25 year period, that rise was far from uniform. Within clusters, one sees instead a fall in the spread between (relatively) rich and (relatively) poor.

Fig. 1 plots a stylized picture of the empirical regularities just described. The figure shows the distribution of income across national economies at two different

points in time. It caricatures the increase in overall spread together with the reduction in intra-distribution inequalities by an emergence of distinct peaks in the distribution. Fig. 1 also shows, to scale, the historical experiences of some relative growth successes and failures. Singapore and South Korea experienced high growth relative to the world average, Venezuela the opposite.

The above constitutes an initial set of stylized facts around which we organize our discussion of economic growth in this paper. We focus on the dynamics of per capita incomes as providing the background against which to assess alternative empirical analyses on growth. In this we depart from, say, Kaldor's [60] stylized facts—the stability of factor shares, the variability of factor input quantities, the stability of time-averaged growth rates in income and in physical capital investment, and so on. Recent empirical analyses of growth and convergence study how alternative conditioning economic variables or different economic hypotheses imply differing behavior for time paths of per capita incomes. We think it useful, therefore, to focus on exactly those dynamics.

3. Theoretical models

This section develops a growth model on which we will base our analysis of the empirical literature. The model is designed to ease comparison across different studies, and to clarify the lessons from empirical work for theoretical reasoning.

Consider a closed economy, with total output denoted Y. Let the quantity of labor input be N, and assume that the stock of human capital H is embodied in the labor force so that the effective labor input is $\widetilde{N} = NH$. There are different kinds of physical capital; write them as the vector $K = (K_1, K_2, \dots)$. Finally, let A be the (scalar) state of technology.

We use two different production technologies in the discussion:

$$Y = \widetilde{F}(K, \widetilde{N}, A)$$

where either

$$\widetilde{F}(K, \widetilde{N}, A) = F(K, \widetilde{N}A) \tag{1a}$$

or

$$\widetilde{F}(K, \widetilde{N}, A) = AF(K, \widetilde{N}). \tag{1b}$$

The distinction between these is whether technical change is labor-augmenting (1a) or Hicks-neutral (1b). We will generally employ (1a), but will draw on (1b) to provide certain links to the literature.

Initially, we assume that F is twice differentiable, homogeneous of degree 1, increasing, and jointly concave in all its arguments and strictly concave in each. Different combinations of these assumptions will be relaxed when we consider endogenous growth models. In addition, we require some Inada-type conditions on F such that

$$\begin{aligned} &\forall l \text{ and } \forall A,\ \widetilde{N},\ K_1^\dagger,\ K_2^\dagger,\ \ldots,\ K_{l-1}^\dagger,\ K_{l+1}^\dagger,\ \ldots \text{ greater than } 0: \\ &\qquad \lim_{K_l \to 0} \widetilde{F}(K_1^\dagger,\ \ldots,\ K_{l-1}^\dagger,\ K_l,\ K_{l+1}^\dagger,\ \ldots,\ \widetilde{N},\ A) \geq 0 \end{aligned} \tag{2}$$

and

$$\forall l: \qquad \partial \tilde{F}/\partial K_l \to \infty \text{ as } K_l \to 0. \tag{3}$$

The homogeneity of degree 1 and concavity assumptions rule out increasing-returns endogenous growth. However, as we will see below, they can nevertheless generate observations usually taken as evidence for endogenous growth models with technological nonconvexities.

Define quantities in per effective labor unit terms as $\tilde{y} \stackrel{\text{def}}{=} Y/\widetilde{N}A$ and vector $\tilde{k} \stackrel{\text{def}}{=} (\widetilde{N}A)^{-1}K$. These are unobservable, however, and so we write their measured counterparts as:

$$\begin{aligned} y &\stackrel{\text{def}}{=} HA \times \tilde{y} = Y/N, \\ k &\stackrel{\text{def}}{=} (\, k_1, \quad k_2, \quad \cdots) = HA \times \tilde{k} = N^{-1}K. \end{aligned}$$

The definitions imply $y = F(k, HA)$ under (1a) and $y = AF(k, H)$ under (1b). In turn, under assumption (1a) total output can be rewritten:

$$Y = \widetilde{N}A \times F((\widetilde{N}A)^{-1}K, 1) \implies \tilde{y} = f(\tilde{k}),$$

where

$$f(\cdot) \stackrel{\text{def}}{=} F(\cdot, 1).$$

This gives growth rate in per worker output y as

$$\dot{y}/y = (\dot{H}/H + \dot{A}/A) + f(\tilde{k})^{-1} \left[\nabla f(\tilde{k})\right]' \frac{d\tilde{k}}{dt} ,$$

with ∇f denoting the gradient of f:

$$\nabla f = \begin{pmatrix} \partial f/\partial \tilde{k}_1 \\ \partial f/\partial \tilde{k}_2 \\ \vdots \end{pmatrix}.$$

But

$$\left[\nabla f(\tilde{k})\right]' \frac{d\tilde{k}}{dt} = (\, \tilde{k}_1 \partial f(\tilde{k})/\partial \tilde{k}_1, \quad \tilde{k}_2 \partial f(\tilde{k})/\partial \tilde{k}_2, \quad \cdots\,) \begin{pmatrix} \dot{k}_1/k_1 - \dot{H}/H - \dot{A}/A \\ \dot{k}_2/k_2 - \dot{H}/H - \dot{A}/A \\ \vdots \end{pmatrix},$$

so that defining

$$s_l(\tilde{k}) = \left[\frac{\tilde{k}_l \times \partial f(\tilde{k})/\partial \tilde{k}_l}{f(\tilde{k})}\right]$$

(necessarily, $s_l(\tilde{k}) \in [0, 1]$ and $\sum_l s_l(\tilde{k}) \leq 1$) we have the growth equation

$$\begin{aligned} &\dot{y}/y = (\dot{H}/H + \dot{A}/A) + \sum_l s_l(\tilde{k}) \times \left\{\dot{k}_l/k_l - \dot{H}/H - \dot{A}/A\right\}, \text{ or} \\ &\dot{\tilde{y}}/\tilde{y} = \sum_l s_l(\tilde{k}) \times \dot{\tilde{k}}_l/\tilde{k}_l. \end{aligned} \tag{4a}$$

(Equation (4a) refers to both the expressions above, as they are logically identical.)

Applying similar reasoning to specification (1b) we obtain the growth equation

$$\dot{y}/y = (\dot{H}/H + \dot{A}/A) + \sum_l s_l(\tilde{k}A) \times \left\{ \dot{k_l}/k_l - \dot{H}/H \right\}, \tag{4b}$$

where functions s_l are defined as before, only here they are evaluated at $\tilde{k}A$ rather than $\tilde{k}$. But no matter where they are evaluated, each s_l is nonnegative, and their sum is bounded from above by 1. When F is Cobb-Douglas, each s_l is constant. More generally, nonnegativity and boundedness of s_l's follow from the assumptions that F is increasing, homogeneous, and concave. The terms in braces on the right-hand side of equations (4a) and (4b) can therefore have only similarly bounded impact on growth rates $\dot{y}/y$ (i.e., the impact of $\dot{k_l}/k_l$ on $\dot{y}/y$ is never more than one-for-one).

To study the dynamics of this system under different economic assumptions, we first provide some definitions. We say *balanced growth* is a collection of time paths in observable per capita quantities (y, k) with

$$\dot{y}/y = \dot{k_l}/k_l = \text{ a constant} \qquad \forall\ l. \tag{5}$$

A *balanced-growth equilibrium* is a collection of time paths in (y, k) satisfying balanced growth (5) and consistent with the decisions of all economic agents in a specific model. Finally, *equilibrium tending towards balanced growth* is a collection of time paths in (y, k) consistent with a specific economic model and satisfying

$$\lim_{t\to\infty} \dot{y}(t)/y(t) \text{ exists, and } \lim_{t\to\infty} \left(\dot{y}(t)/y(t) - \dot{k_l}(t)/k_l(t) \right) = 0 \quad \forall\ l. \tag{6}$$

Conditions (5) and (6) are appropriate to use when working with observable quantities y and k. Translating them to the technology-adjusted $\tilde{y}$ and $\tilde{k}$ is trivial and often convenient when discussing theoretical models. We will do so freely below. Also, (5) and (6) are, again, appropriate when the model is deterministic. For stochastic models, they can be modified to be, for instance, statements on expectations. We consider some of those below in Section 5.

In the best-known case—the neoclassical growth model with exogenous technical progress—technology is assumed to be

$$A(t) = A(0)e^{\xi t},$$

so that ξ is the exogenously-given constant rate of technical progress. Balanced-growth equilibrium then occurs with (y, k) growing at rate ξ, and therefore implying $(\tilde{y}, \tilde{k})$ constant and finite. That equilibrium, under the standard assumptions we have made here, is approached from almost all initial values of k. In other situations, such as endogenous growth, there is no guarantee that a balanced-growth equilibrium exists. We will then be interested in whether there are equilibria that tend towards balanced growth, and if so, what characteristics those show.

Distinguishing balanced-growth equilibrium and balanced growth is useful to understand how adding economic structure to equations (4a) and (4b) can produce new insights. For instance, suppose technical change is labor augmenting so that growth follows (4a). Suppose further F is Cobb-Douglas, so that

$$F(K, \tilde{N}A) = \left(\prod_l K_l^{\alpha_l}\right)(\tilde{N}A)^{1-\sum_l \alpha_l} \quad \text{with } \alpha_l > 0 \text{ and } \sum_l \alpha_l \in (0, 1)$$

giving

$$f(\tilde{k}) = \prod_l \tilde{k_l}^{\alpha_l}.$$

Equation (4a) then becomes

$$\dot{\tilde{y}}/\tilde{y} = \sum_l \alpha_l \times \dot{\tilde{k}}_l/\tilde{k}_l,$$

so that under balanced growth (5)

$$\dot{\tilde{y}}/\tilde{y} = \left(\sum_l \alpha_l\right) \times \dot{\tilde{k}}_1/\tilde{k}_1.$$

Since the multiplier $(\sum_l \alpha_l)$ is strictly less than 1, equality between $\dot{\tilde{y}}/\tilde{y}$ and $\dot{\tilde{k}}_1/\tilde{k}_1$ can occur only at

$$\dot{\tilde{k}}_l/\tilde{k}_l = \dot{\tilde{k}}_1/\tilde{k}_1 = \dot{\tilde{y}}/\tilde{y} = 0,$$

independent of any other economic structure beyond the technology specification. (We will see this below when we study the Solow-Swan model [106, 107, 111], its general equilibrium Cass-Koopmans version [23, 65], and the modification due to Mankiw, Romer, and Weil [78].)

The reasoning just given extends naturally to production technologies beyond Cobb-Douglas when the counterpart to $\sum_l \alpha_l$ (or, more generally, $\sum_l s_l(\tilde{k})$) is not constant but always remains strictly less than 1. The reasoning fails, instructively, in the following counter example. Suppose $\tilde{k}$ is scalar but F is CES with

$$F(K, \tilde{N}A) = \left[\gamma_K K^\alpha + \gamma_N (\tilde{N}A)^\alpha\right]^{1/\alpha}, \quad 0 < \alpha < 1 \text{ and } \gamma_K,\ \gamma_N > 0,$$

so that

$$f(\tilde{k}) = \left[\gamma_K \tilde{k}^\alpha + \gamma_N\right]^{1/\alpha}.$$

Then,

$$\sum_l s_l(\tilde{k}) = \frac{\gamma_K}{\left[\gamma_K + \gamma_N \tilde{k}^{-\alpha}\right]} \nearrow 1 \qquad \text{as } \tilde{k} \to \infty.$$

Here, it is possible to have $\dot{\tilde{k}}_l/\tilde{k}_l$ and $\dot{\tilde{y}}/\tilde{y}$ always positive and tending towards positive balanced growth in a way that varies with economic parameters. This behavior occurs also in endogenous growth models that exploit externalities and increasing returns (Romer [98]) or in models with the production technology "asymptotically linear" (Jones and Manuelli [59], Rebelo [96]).

Our definition of balanced-growth equilibrium compares the growth rates $\dot{y}/y$ and $\dot{k}_l/k_l$. This is not sensible for technology (1b) where we see that $\dot{A}/A$ appears with $\dot{y}/y - \dot{H}/H$ but not with $\dot{k}_l/k_l - \dot{H}/H$. The definition of balanced growth is,

then, not generally useful for such technologies, although special cases exist when it is—for instance where A is suitably endogenized.

If factor input markets are competitive and F fully describes the contribution of factor inputs to production, then s_l is the factor share of total output paid to the owners of the l-th physical capital good. However, the discussion thus far has made no assumptions about market structure, the behavior of economic agents, the processes of capital accumulation and technological progress, and so on. Production functions (1a) and (1b) imply, respectively, (4a) and (4b) regardless of whether savings rates are endogenous (as in the Cass-Koopmans approach) or exogenous (as in the Solow-Swan formulation). The implications hold independent of whether technology A evolves exogenously, or endogenously through physical capital accumulation or R&D investment. Thus, growth theories whose substantive differences lie in alternative F specifications can be compared by studying the different restrictions they imply for dynamics (4a) and (4b).

This reasoning provides a useful insight for empirically distinguishing endogenous and neoclassical growth models. In so far as many models differ substantively only through alternative specifications of the production technology, formulating them within a general equilibrium framework might have only limited payoff empirically. To be clear, doing so is important for issues such as existence or optimality, and sometimes can place further qualitative restrictions on the behavior of particular aggregates. However, it provides no fundamentally new empirical perspective. Indeed, studies such as Barro and Sala-i-Martin [9, 10], while using general equilibrium formulations to justify their empirical analyses, typically consider regression models observationally equivalent to the Solow-Swan model with exogenous savings rates.

Many approaches to studying growth empirics can be viewed as tracing out implications of either (4a) or (4b). For example, under (4a) a researcher investigating the determinants of long-run economic growth might consider situations where the last summand—the term involving the different capital stocks—vanishes, and seek only to understand the economic forces driving $\dot{H}/H$ and $\dot{A}/A$. Alterna-

tively, a researcher interested in the dynamics *surrounding* the time path implied by $\dot{H}/H + \dot{A}/A$ might seek to model only $\sum_l s_l(\tilde{k}) \times \left\{\dot{k}_l/k_l - \dot{H}/H - \dot{A}/A\right\}$ or $\sum_l s_l(\tilde{k}A) \times \left\{\dot{k}_l/k_l - \dot{H}/H\right\}$, taking as given (conditioning on) $\dot{H}/H$ and $\dot{A}/A$. This is exactly what is done in studies of conditional β convergence (defined in Section 5 below): see, e.g., Barro and Sala-i-Martin [10] or Mankiw, Romer, and Weil [78].

Finally, this formulation highlights how certain terminologies have been used inconsistently in the literature. For example, while Lucas [74] uses a definition of human capital that is H in our formulation, Mankiw, Romer, and Weil [78] use a definition of human capital that is one of the components in vector K. Of course, both definitions are consistent with higher human capital improving labor productivity, but they do so in conceptually distinct ways.

While interesting exceptions exist, a wide range of growth models can be cast as special cases of our framework. We use it then as an organizing structure for the analysis of empirical work that follows.

4. From theory to empirical analysis

In this section, we consider a number of growth models in the literature, and study how they restrict observations on growth dynamics.

The neoclassical model: One capital good, exogenous technical progress

The first specific structure we consider is the neoclassical growth model, as developed in Barro and Sala-i-Martin [10], Cass [23], Koopmans [65], Solow [106, 107], and Swan [111].

As argued in Section 3, the key empirical implications of the neoclassical model depend solely on the assumed production funcion. However, some quantitative features of the dynamics do depend on preferences. To clarify those, we study a general equilibrium formulation here.

The neoclassical model assumes the production function (1a) supplemented with the following:

$$\dot{H}/H = 0, \quad \text{normalizing } H(0) = 1, \tag{7a}$$

$$\dot{A}/A = \xi \geq 0, \quad \text{given } A(0) > 0, \tag{7b}$$

$$\dot{N}/N = \nu \geq 0, \quad \text{given } N(0) > 0, \tag{7c}$$

$$K \text{ scalar}, \quad \text{given } K(0) > 0. \tag{7d}$$

These assumptions say that only physical capital is accumulated, and population growth and technical change are exogenous. In addition, assume that

$$\forall \widetilde{N}A > 0: \qquad \lim_{K \to \infty} F(K, \widetilde{N}A)/K = 0. \tag{8}$$

Let physical capital depreciate exponentially at rate $\delta > 0$. Physical capital accumulation will be assumed to follow one of two possibilities. First, as in Solow [106] and Swan [111], suppose savings is a constant fraction $\tau \in (0, 1)$ of income. Then,

$$\dot{\tilde{k}}/\tilde{k} = \tau f(\tilde{k})/\tilde{k} - (\delta + \nu + \xi). \tag{9a}$$

As the second possibility, suppose as in Cass [23] and Koopmans [65], that economy-wide savings is determined by the optimization problem:

$$\max_{\{c(t),\ K(t)\}_{t \geq 0}} N(0) \int_0^\infty U(c(t)) e^{-(\rho - \nu)t}\, dt, \qquad \rho > \nu + \xi \geq 0$$

subject to

$$\dot{K}(t) = Y(t) - c(t)N(t) - \delta K(t), \tag{10}$$

$$U(c) = \frac{c^{1-\theta} - 1}{1 - \theta}, \qquad \theta > 0,$$

and (1a), (7a–d).

The maximand in (10) is the number of people multiplied by what each enjoys in present discounted value of utility from consumption c. The $\dot{K}$ constraint says

that capital accumulates from the output left over after total consumption and depreciation. Coefficient θ parametrizes the intertemporal elasticity of substitution in consumption, while ρ is the discount rate. We emphasize that we have restricted ρ to be not just non-negative but to exceed the sum of the rates of population growth and technical change,

$$\rho > \nu + \xi. \tag{11}$$

Problem (10) determines consumption and thus savings and investment to maximize social welfare. Define $\tilde{c}$ to be per capita consumption normalized by technology, i.e., $\tilde{c} = c/A$. The Technical Appendix shows that the necessary first order conditions to (10) are:

$$\begin{aligned} \dot{\tilde{k}}/\tilde{k} &= \frac{f(\tilde{k}) - \tilde{c}}{\tilde{k}} - (\delta + \nu + \xi), \\ \dot{\tilde{c}}/\tilde{c} &= \left(\nabla f(\tilde{k}) - [\rho + \delta + \theta\xi]\right) \theta^{-1} \\ \lim_{t \to \infty} \tilde{k}(t) e^{-(\rho - \nu - \xi)t} &= 0 \end{aligned} \tag{9b}$$

A balanced-growth equilibrium is a positive time-invariant technology-normalized capital stock $\tilde{k}$ (together with implied $\tilde{y} = f(\tilde{k})$) such that under (9a)

$$\dot{\tilde{y}} = \dot{\tilde{k}} = 0$$

and under (9b)

$$\dot{\tilde{c}} = 0$$

where

$$\tilde{c} = f(\tilde{k}) - (\delta + \nu + \xi)\tilde{k} \in (0, f(\tilde{k})).$$

(Our balanced-growth equilibrium definition implies that we can specialize to time-invariant $\tilde{k}$.)

Balanced-growth predictions are identical under either accumulation assumptions (9a) and (9b). To see this, note that at balanced-growth equilibrium under (9b) we can find τ in $(0, 1)$ such that

$$\tilde{c} = f(\tilde{k}) - (\delta + \nu + \xi)\tilde{k} = (1 - \tau)f(\tilde{k})$$

as both $\tilde{k}$ and $\tilde{c}$ are constant through time; equation (9b) thus reduces to (9a).

Two questions arise from this formulation. First, does a balanced-growth equilibrium always exist? And, second, even if both formulations have the same empirical implications in long-run steady state, do transitions to steady state differ?

Fig. 2 shows that a unique balanced-growth equilibrium exists and that $\tilde{k}$ satisfying (9a) is dynamically stable everywhere in the region $\tilde{k} > 0$ (the Technical Appendix also proves this). Since $\tilde{y} = f(\tilde{k})$, we immediately have that output per effective worker too has a unique, globally stable steady state.

The dynamics of this model can be understood further by taking a Taylor series expansion in $\log \tilde{k}$ about steady-state $\tilde{k}^*$,

$$\dot{\tilde{k}}/\tilde{k} \stackrel{a}{=} \tau \left(\nabla f(\tilde{k}) - f(\tilde{k})\tilde{k}^{-1} \right) \Big|_{\tilde{k}=\tilde{k}^*} \times (\log \tilde{k} - \log \tilde{k}^*).$$

For F Cobb-Douglas,

$$\begin{aligned} & F(K, NA) = K^{\alpha}(NA)^{1-\alpha}, \qquad \alpha \in (0, 1) \\ & \implies f(\tilde{k}) = \tilde{k}^{\alpha}, \end{aligned} \tag{12}$$

this first-order series expansion becomes

$$\begin{aligned} \frac{d}{dt} \log \tilde{k} &\stackrel{a}{=} -(1 - \alpha)(\delta + \nu + \xi) \times (\log \tilde{k} - \log \tilde{k}^*) \\ &= \lambda \times (\log \tilde{k} - \log \tilde{k}^*) \end{aligned}$$

where we have defined

$$\lambda \stackrel{\text{def}}{=} -(1 - \alpha)(\delta + \nu + \xi) < 0. \tag{13}$$

Solving this differential equation gives

$$\begin{aligned} & \log\tilde{k}(t) - \log\tilde{k}^* = (\log\tilde{k}(0) - \log\tilde{k}^*)e^{\lambda t} \\ \Longrightarrow\ & \log\tilde{y}(t) - \log\tilde{y}^* = (\log\tilde{y}(0) - \log\tilde{y}^*)e^{\lambda t} \to 0 \text{ as } t \to \infty, \end{aligned} \tag{14a}$$

i.e., $\log\tilde{k}$ and $\log\tilde{y}$ converge to their respective steady state values $\log\tilde{k}^*$ and $\log\tilde{y}^* \stackrel{\text{def}}{=} \log f(\tilde{k}^*)$ exponentially at rate $|\lambda|$. As α increases to 1 this rate of convergence approaches 0: thus, the larger is the Cobb-Douglas coefficient on physical capital, the slower does $\log\tilde{y}$ converge to its steady state value.

Under the Cobb-Douglas assumption (12), the accumulation equation (9a) and Fig. 2 imply the steady state level

$$\begin{aligned} \tilde{y}^* = (\tilde{k}^*)^\alpha &= \left[(\tilde{k}^*)^{-(1-\alpha)}\right]^{-\alpha/(1-\alpha)} \\ &= \left[(\delta+\nu+\xi)^{-1}\tau\right]^{\alpha/(1-\alpha)} \end{aligned} \tag{15}$$

Equation (15) gives steady state income levels as depending positively on the saving rate and negatively on the labor force growth rate.

Before discussing in detail the empirical implications of (14a), we turn to how the Solow-Swan and the general equilibrium Cass-Koopmans versions of this model differ in their observable predictions. First, rewrite the first two equations in (9b) as:

$$\frac{d}{dt}\begin{pmatrix}\log\tilde{k}\\ \log\tilde{c}\end{pmatrix} = \begin{pmatrix} \frac{f(\tilde{k})-\tilde{c}}{\tilde{k}} - (\delta+\nu+\xi) \\ \left(\nabla f(\tilde{k}) - [\rho+\delta+\theta\xi]\right)\theta^{-1}\end{pmatrix}. \tag{16}$$

Define the zero of $(\dot{\tilde{k}}/\tilde{k}, \dot{\tilde{c}}/\tilde{c})$ by $(\tilde{k}^*, \tilde{c}^*)$. (The Technical Appendix establishes that this is well-defined.) Then the first-order Taylor series expansion of $(\log\tilde{k},\ \log\tilde{c})'$ about $(\log\tilde{k}^*,\ \log\tilde{c}^*)$ is:

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix}\log\tilde{k}\\ \log\tilde{c}\end{pmatrix} &\stackrel{a}{=} \left.\begin{pmatrix} \nabla f(\tilde{k}) - (f(\tilde{k})-\tilde{c})\tilde{k}^{-1} & -\tilde{c}\tilde{k}^{-1} \\ \nabla^2 f(\tilde{k})\tilde{k}\theta^{-1} & 0\end{pmatrix}\right|_{(\tilde{k}^*,\ \tilde{c}^*)} \times \begin{pmatrix}\log\tilde{k}-\log\tilde{k}^*\\ \log\tilde{c}-\log\tilde{c}^*\end{pmatrix} \\ &\stackrel{\text{def}}{=} M \times \begin{pmatrix}\log\tilde{k}-\log\tilde{k}^*\\ \log\tilde{c}-\log\tilde{c}^*\end{pmatrix}. \end{aligned} \tag{17}$$

Coefficient matrix M in (17) has determinant $\nabla^2 f(\tilde{k})\tilde{c}\theta^{-1} < 0$ so its eigenvalues are real and of opposite sign. Moreover, its trace is

$$\begin{aligned}\nabla f(\tilde{k}^*) - (f(\tilde{k}^*) - \tilde{c}^*)/\tilde{k}^* &= (\rho + \delta + \theta\xi) - (\delta + \nu + \xi) \\ &= \rho - (\nu + \xi) + \theta\xi > 0.\end{aligned}$$

Denote the eigenvalues of M by $\lambda_1 > 0 > \lambda_2$. Fig. 3 uses these determinant and trace properties to establish how λ_1 and λ_2 vary with the parameters of the model. For the Cobb-Douglas technology $f(\tilde{k}) = \tilde{k}^\alpha$, eigenvalue λ_2 increases towards 0 as α rises towards 1.

Eigenvalue λ_2 determines dynamics local to the steady state as:

$$\begin{aligned}\log\tilde{k}(t) - \log\tilde{k}^* &= (\log\tilde{k}(0) - \log\tilde{k}^*)e^{\lambda_2 t} \\ \log\tilde{c}(t) - \log\tilde{c}^* &= (\log\tilde{c}(0) - \log\tilde{c}^*)e^{\lambda_2 t}\end{aligned} \tag{18}$$

with $[\log\tilde{k}(0) - \log\tilde{k}^*]$ and $[\log\tilde{c}(0) - \log\tilde{c}^*]$ satisfying a specific proportionality condition described in the Technical Appendix. Then for technology (12), with $\tilde{y}^* = (\tilde{k}^*)^\alpha$, the first equation in (18) gives

$$\log\tilde{y}(t) - \log\tilde{y}^* = (\log\tilde{y}(0) - \log\tilde{y}^*)e^{\lambda_2 t} \to 0 \text{ as } t \to \infty. \tag{14b}$$

Comparing equations (14a) and (14b) we see that assumptions (9a) and (9b) deliver identical observable implications—not just in steady-state balanced growth, but also locally around steady state. The convergence rates λ and λ_2 have different interpretations as they depend on different economic parameters. However, they vary in the same way when the technology parameter α changes.

How are these common observable implications useful for understanding patterns of cross-country growth? Parallel to the theoretical development above, we interpret the bulk of the empirical literature as concerned with two sets of implications: first, steady-state balanced-growth predictions and, second, (convergence) predictions local to steady state.

Without loss, write the convergence coefficient as λ in both (14a) and (14b). From observed per capita income $y = \tilde{y}HA = \tilde{y}A$ we have:

$$\begin{aligned}\log y(t) &= \log \tilde{y}(t) + \log A(t) \\ &= \log \tilde{y}^* + [\log \tilde{y}(0) - \log \tilde{y}^*]e^{\lambda t} + \log A(0) + \xi t.\end{aligned}$$

Moreover, since $\tilde{y}^* = f(\tilde{k}^*)$ and $f(\tilde{k}^*)/\tilde{k}^* = (\delta + \nu + \xi)\tau^{-1}$, there is some function g such that $\tilde{y}^* = g((\delta + \nu + \xi)^{-1}\tau)$. We can therefore write the implied sample path in observable per capita income as

$$\begin{aligned}\log y(t) = \log(g((\delta + \nu + \xi)^{-1}\tau)) + \log A(0) + \xi t \\ + [\log y(0) - (\log(g((\delta + \nu + \xi)^{-1}\tau)) + \log A(0))]e^{\lambda t},\end{aligned} \tag{19}$$

and its time derivative

$$\frac{d}{dt}\log y(t) = \xi + \lambda \times [\log y(0) - (\log(g((\delta + \nu + \xi)^{-1}\tau)) + \log A(0))]e^{\lambda t} \tag{19$'$}$$

From (19) $\log y$ can be viewed as having two components: a convergence component (the term involving $e^{\lambda t}$) and a levels component (the rest of the right-hand side).

Fig. 4 graphs (19) for two possible values of $\log(g((\delta + \nu + \xi)^{-1}\tau)) + \log A(0)$. The figure shows two different possible steady state paths—corresponding to two possible values for the sum $\log \tilde{y}^* + \log A(0) = \log(g((\delta + \nu + \xi)^{-1}\tau)) + \log A(0)$.

Relative to typical claims in the literature, Fig. 4 conveys a negative message. As long as $\log \tilde{y}^* + \log A(0)$ remains unobserved or unrestricted, *any* pattern of cross-country growth and convergence is consistent with the model. As drawn in Fig. 4, the a value applies to economies at $y_1(0)$ and $y_2(0)$ while the b value to $y_3(0)$ and $y_4(0)$. Economies 1 and 2 converge towards each other, as do economies 3 and 4. At the same time, however, economies 2 and 3, although each obeying the neoclassical growth model, are seen to approach one another, criss-cross, and then diverge.

We can now organize those empirical studies that use the neoclassical growth model for their theoretical underpinnings. Cross-section regression analyses, such

as Barro and Sala-i-Martin [10], Baumol [12], DeLong [29], Mankiw, Romer, and Weil [78], and Sachs and Warner [102] estimate variants of (19). Mankiw, Romer, and Weil [78], in particular, consider two versions of (19): First, when the term in $e^{\lambda t}$ is already at its limiting value, then the first component of the expression is taken to "explain" the steady-state cross section distribution of income.[3] Second, when the term in $e^{\lambda t}$ is taken to be central—and the rest of the right-hand side of (19) is given (or are taken to be nuisance parameters)—the equation is viewed to "explain" convergence in income. This second interpretation motivates the convergence analyses of the other papers mentioned above.[4]

In our reading of the empirical literature, there is some confusion over the goals of the analysis. On the one hand, a researcher might study (19) to estimate the coefficients of interest in it. But the only parameters related to the economic reasoning in (19) are those in the function g, i.e., parameters of the production function. Thus, standard econometric techniques applied to this equation might be useful for recovering such parameters. A reseacher might go further and seek, in an ad hoc way, to parameterize $A(0)$ and ξ as functions of other economic variables. While this might be useful for regression fitting, its results are difficult to interpret in terms of the original economic analysis. After all, $A(0)$ and ξ played no integral role in the theoretical reasoning and it is unclear that a structural model incorporating these other variables would produce a regression of the type typically

[3] The Mankiw-Romer-Weil formulation, of course, includes human capital accumulation. That feature is ignored for expositional convenience here as it does not affect our basic point. We return to it below.

[4] An earlier literature (e.g., Grier and Tullock [52]) studied similar regression equations with growth on the left-hand side and explanatory variables on the right. We distinguish this from the work described in the text only because that earlier research did not show any preoccupation with convergence. It instead investigated, using exploratory empirical techniques, only the determinants of growth—an important question, certainly, but distinct from the simultaneous interest in convergence that characterizes the newer literature.

estimated.

A second goal of an empirical analysis of (19) is to address questions of cross-country patterns of growth. We think, however, that all such analyses, even at their most successful, are silent on those questions. From Fig. 4, as long as $A(0)$ is unrestricted or omitted from the analysis, no study of (19) can reveal how cross-country incomes evolve.

One interpretation of the preceding is that the basic model's key implications are both too strong and too weak. If $A(0)$ were required to be identical across economies, then the growth and convergence predictions in Fig. 2 would be likely inconsistent with the inequality dynamics in cross-country incomes we described in Section 2. If, on the other hand, a researcher goes to the opposite extreme and allows $A(0)$ to differ arbitrarily across economies, then the theoretical model says little about cross-country patterns of growth. The free parameters $A(0)$ carry the entire burden of explanation. Finally, should a researcher take a middle path, and restrict $A(0)$ to depend on specific economic variables in an ad hoc manner, then that researcher might well end up fitting the data satisfactorily. However, the results of such a procedure can be difficult to interpret within the Solow-Swan (or Cass-Koopmans) growth model.[5]

Empirical studies such as Bernard and Durlauf [16, 17], Durlauf and Johnson [39], and Quah [94] seek to circumvent some of the criticisms we have just described. One strand of this work estimates models that explicitly nest the traditional neoclassical setup. Another strand seeks to identify those features of the long-run behavior of cross-country incomes that are invariant with respect to finely-detailed structural assumptions.

Before turning to more detailed empirics, however, we describe models that depart from the basic set of assumptions in the neoclassical growth model. This

[5] Mankiw, Romer, and Weil [78] is a key exception. Those authors focus on that part of the steady-state path that depends on savings and population growth rates, not on $A(0)$, and suggest that their human-capital modification of the Solow-Swan model does fit the data. We discuss that model below.

is easy to do given the structure we have set up. Again, our goal is not to repeat discussion already found elsewhere, but to survey in a unified way the empirical implications of the different classes of models.

The neoclassical model: Multiple capital goods

A well-known model due to Mankiw, Romer, and Weil [78] (hereafter MRW) adds human capital to the Solow-Swan model, and develops empirics that potentially better explain the cross-country income data than models that account only for physical capital accumulation following Solow's original work. The MRW model fits in our framework as follows.

Again, take production technology (1a), and assume (7a–c). In place of (7d), let K have two components, the first called *physical* capital K_p and the second *human* capital K_h:

$$K = (K_p, K_h)'. \tag{7d$'$}$$

(Distinguish K_h from that concept of human capital that is H—the latter multiplies the labor input N to produce effective labor input $\tilde{N}$, while the former is an entry in the vector of capital stocks, and thus is better viewed as analogous to physical capital K_p.) Extend the accumulation assumption (9a) to

$$\begin{aligned} \dot{K}_p &= \tau_p Y - \delta_p K_p, \qquad \tau_p,\ \delta_p > 0 \\ \dot{K}_h &= \tau_h Y - \delta_h K_h, \qquad \tau_h,\ \delta_h > 0 \\ \tau_p + \tau_h &< 1. \end{aligned} \tag{9a$'$}$$

Then technology-intensive effective capital stocks $\tilde{k} = (\tilde{k}_p, \tilde{k}_h)'$ with $\tilde{k}_p = K_p/\tilde{N}A$ and $\tilde{k}_h = K_h/\tilde{N}A$ satisfy

$$\begin{aligned} \dot{\tilde{k}}_p/\tilde{k}_p &= \tau_p \tilde{y}/\tilde{k}_p - (\delta_p + \nu + \xi) \\ \dot{\tilde{k}}_h/\tilde{k}_h &= \tau_h \tilde{y}/\tilde{k}_h - (\delta_h + \nu + \xi). \end{aligned}$$

A balanced-growth equilibrium is a positive time-invariant triple $(\tilde{y}, \tilde{k}_p, \tilde{k}_h)^*$ such that

$$\begin{aligned}\tilde{y} &= f(\tilde{k}_p, \tilde{k}_h)\\ \tau_p \tilde{y}/\tilde{k}_p &= \delta_p + \nu + \xi\\ \tau_h \tilde{y}/\tilde{k}_h &= \delta_h + \nu + \xi.\end{aligned}$$

When F is Cobb-Douglas so that

$$f(\tilde{k}_p, \tilde{k}_h) = (\tilde{k}_p)^{\alpha_p}(\tilde{k}_h)^{\alpha_h}, \qquad \alpha_p,\ \alpha_h > 0 \text{ and } \alpha_p + \alpha_h < 1, \tag{20}$$

straightforward calculation establishes that a balanced-growth equilibrium has:

$$\begin{aligned}\begin{pmatrix}\log \tilde{k}_p^*\\ \log \tilde{k}_h^*\end{pmatrix} &= (1-\alpha_p-\alpha_h)^{-1}\begin{pmatrix}-(1-\alpha_h) & -\alpha_h\\ -\alpha_p & -(1-\alpha_p)\end{pmatrix}\\ &\qquad \times \begin{pmatrix}\log\left((\delta_p+\nu+\xi)\tau_p^{-1}\right)\\ \log\left((\delta_h+\nu+\xi)\tau_h^{-1}\right)\end{pmatrix}\\ &= (1-\alpha_p-\alpha_h)^{-1}\\ &\qquad \times \begin{pmatrix}(1-\alpha_h)\log\left((\delta_p+\nu+\xi)^{-1}\tau_p\right) + \alpha_h \log\left((\delta_h+\nu+\xi)^{-1}\tau_h\right)\\ \alpha_p \log\left((\delta_p+\nu+\xi)^{-1}\tau_p\right) + (1-\alpha_p)\log\left((\delta_h+\nu+\xi)^{-1}\tau_h\right)\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}\log \tilde{y}^* = (1-\alpha_p-\alpha_h)^{-1}\big[&\alpha_p \log\left((\delta_p+\nu+\xi)^{-1}\tau_p\right)\\ &+ \alpha_h \log\left((\delta_h+\nu+\xi)^{-1}\tau_h\right)\big].\end{aligned} \tag{15$'$}$$

Equation (15′) is the MRW counterpart to the Solow-Swan levels prediction (15). It specializes to the latter when α_h is set to 0; otherwise, it comprises a geometric average of contributions from physical and human capital.

It is easy to show in state space $(\tilde{k}_p, \tilde{k}_h)$ that this system is globally stable and converges to balanced-growth equilibrium. In general, then, all dynamics—including those of $\tilde{y}$—depend on the bivariate state vector $(\tilde{k}_p, \tilde{k}_h)$. This would

suggest that, in a growth regression, studying the (one-dimensional) coefficient on initial income alone, with or without auxiliary ad hoc conditioning, gives a misleading picture of dynamics local to steady state. However, with additional restrictions on model parameters, conditioning on the level of $\tilde{y}(t)$ can render the local convergence behavior of $\tilde{y}$ independent of the state $(\tilde{k}_p(t), \tilde{k}_h(t))$.

Mankiw, Romer, and Weil [78] achieve this by setting equal the depreciation rates on human and physical capital, i.e., $\delta_p = \delta_h$. From (20), and taking the first-order Taylor series expansion in $\log \tilde{y}$, $\log \tilde{k}_p$, and $\log \tilde{k}_h$, we have:

$$\begin{aligned}
\dot{\tilde{y}}/\tilde{y} &= \alpha_p \dot{\tilde{k}}_p/\tilde{k}_p + \alpha_h \dot{\tilde{k}}_h/\tilde{k}_h \\
&= \alpha_p \left[\tau_p \tilde{y}/\tilde{k}_p - (\delta_p + \nu + \xi)\right] \\
&\quad + \alpha_h \left[\tau_h \tilde{y}/\tilde{k}_h - (\delta_h + \nu + \xi)\right] \\
&\stackrel{a}{=} \alpha_p \left[(\delta_p + \nu + \xi)\left((\log \tilde{y} - \log \tilde{y}^*) - (\log \tilde{k}_p - \log \tilde{k}_p^*)\right)\right] \\
&\quad + \alpha_h \left[(\delta_h + \nu + \xi)\left((\log \tilde{y} - \log \tilde{y}^*) - (\log \tilde{k}_h - \log \tilde{k}_h^*)\right)\right]
\end{aligned}$$

so that $\delta_p = \delta_h = \delta$ then gives

$$\dot{\tilde{y}}/\tilde{y} = -(1 - \alpha_p - \alpha_h)(\delta + \nu + \xi) \times (\log \tilde{y} - \log \tilde{y}^*). \tag{21}$$

Under this MRW specification the sample path (19) changes so that the levels and convergence components include terms in τ_h and α_h. The observable implications remain unchanged: Observed per capita income evolves in balanced-growth equilibrium as $A(t)$; away from steady state, observed per capita income converges towards that balanced-growth path. The dynamics are still as given in Fig. 4.

The MRW model has been used as the basis for numerous empirical studies. To aid our subsequent discussion of those studies, we develop a more explicit representation for the model's predictions. From (21) now let

$$\lambda \stackrel{\text{def}}{=} -(1 - \alpha_p - \alpha_h)(\delta + \nu + \xi) < 0, \tag{22}$$

so that

$$\begin{aligned} \log \tilde{y}(t) - \log \tilde{y}^* &= [\log \tilde{y}(0) - \log \tilde{y}^*]\, e^{\lambda t} \\ \Longrightarrow\ \log \tilde{y}(t+T) - \log \tilde{y}^* &= [\log \tilde{y}(t) - \log \tilde{y}^*]\, e^{\lambda T}. \end{aligned}$$

Transforming to get observable $\log y(t)$, this becomes:

$$\begin{aligned} \log y(t+T) - [\log A(0) + (t+T)\xi] &= (1 - e^{\lambda T}) \log \tilde{y}^* \\ &\quad + [\log y(t) - \log A(0) - t\xi]\, e^{\lambda T} \\ \Longrightarrow\ \log y(t+T) - \log y(t) &= (1 - e^{\lambda T}) \log \tilde{y}^* + (e^{\lambda T} - 1) \log y(t) \\ &\quad + (1 - e^{\lambda T}) \log A(0) + (t + T - e^{\lambda T} t)\xi \end{aligned}$$

Substituting in (15′) for steady state $\log \tilde{y}^*$ gives

$$\begin{aligned} \log y(t+T) - \log y(t) &= (1 - e^{\lambda T}) \log A(0) + (t + T - e^{\lambda T} t)\xi \\ &\quad + (e^{\lambda T} - 1) \log y(t) \\ &\quad + (1 - e^{\lambda T}) \frac{\alpha_p}{1 - \alpha_p - \alpha_h} \log \tau_p \\ &\quad + (1 - e^{\lambda T}) \frac{\alpha_h}{1 - \alpha_p - \alpha_h} \log \tau_h \\ &\quad - (1 - e^{\lambda T}) \frac{\alpha_p + \alpha_h}{1 - \alpha_p - \alpha_h} \log(\delta + \nu + \xi). \end{aligned} \tag{23}$$

In words, growth depends on some (exogenously given) constants, the initial level $\log y(t)$, savings rates, technological parameters, and the population growth rate. Since $\lambda < 0$, the coefficient on the initial level $\log y(t)$ should be negative.

Comparing MRW's convergence rate (22) with Solow-Swan's (13), the only difference is the addition of α_h in the former. Thus, keeping fixed α_p (physical capital's coefficient), δ, ν, and ξ, MRW's addition of human capital to the neoclassical model implies λ closer to zero, or a slower rate of convergence, than in the Solow-Swan model.

In both the MRW and traditional neoclassical models the levels of balanced-growth income time paths can vary with the parameters of preferences and technology (τ, ρ, θ, and α). However, the rate of change in those balanced-growth time paths in incomes is always just the exogenously given $\xi = \dot{A}/A$. This is useful to remember when working with representations such as (23)—although the dependent variable in the regression equation is a growth rate, these models do *not* explain growth rates over long time horizons. It is this that makes it useful to label these models of *exogenous growth*.

Endogenous growth: Asymptotically linear technology

We now consider a range of models that generate long-run growth from other than exogenous technical change. When possible, we will show how such models can be derived by straightforward perturbations of the parameterizations we have used to describe the neoclassical model.[6]

Assume, as in the standard one-capital neoclassical model, (1a) and (7a–d), but instead of (8), suppose that

$$\forall \tilde{N}A > 0: \qquad \lim_{K\to\infty} F(K, \tilde{N}A)/K > 0. \tag{24}$$

For instance, the CES production function

$$F(K, \tilde{N}A) = \left[\gamma_K K^\alpha + \gamma_N (\tilde{N}A)^\alpha\right]^{1/\alpha}$$

is homogeneous of degree 1, concave, and satisfies (2), (3), and (24) with

$$\forall \tilde{N}A > 0: \qquad \lim_{K\to\infty} F(K, \tilde{N}A)/K = \gamma_K^{1/\alpha} > 0.$$

[6] Such a strategy is inspired by Solow [106, Example 3]; see also Jones and Manuelli [59].

Call a production function satisfying (24) *asymptotically linear*. The motivation for this terminology comes from $f(\tilde{k})$ varying linearly with $\tilde{k}$ as the latter gets large.[7]

By l'Hôpital's rule, (24) gives

$$\lim_{\tilde{k}\to\infty} \nabla f(\tilde{k}) = \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} > 0 \implies \lim_{\tilde{k}\to\infty} s(\tilde{k}) = 1,$$

so that, following the reasoning in Section 3, balanced-growth equilibria with positive $\dot{\tilde{y}}/\tilde{y}$ are now possible.

Let capital accumulation follow (9a) as before. Whereas previously Fig. 2 established existence of a unique balanced-growth equilibrium with finite $(\tilde{y}^*, \tilde{k}^*)$ and $\dot{\tilde{k}}/\tilde{k} = 0$, Fig. 5 now shows a range of possibilities. Taking technology parameters as fixed, define the threshold savings rate

$$\underline{\tau} = \frac{\delta + \nu + \xi}{\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1}}.$$

The numerator is the rate at which technology-adjusted physical capital per worker naturally "dissipates", given the rates of discount, population growth, and exogenous technology development. The denominator is physical capital's limiting average product, which equals the limiting marginal product. This expression thus displays a tension between two opposing forces: The more productive physical

[7] Of course, even if the limiting $f(\tilde{k})\tilde{k}^{-1}$ were zero rather than positive, we would still have asymptotic linearity (albeit trivially), but we hereafter ignore this possibility when using the phrase. A useful alternative is to say that (24) implies $f(\tilde{k})$ is $O(\tilde{k})$ (or big-oh $\tilde{k}$), following standard terminology in statistics and elsewhere. Duffy and Papageorgiou [35] find that a CES specification for the aggregate production function fits cross-country data better than a Cobb-Douglas, and moreover that the elasticity of substitution between capital and labor exceeds one. This evidence implies the possibility for endogenous growth of the kind described in Jones and Manuelli [59] and this subsection.

capital is in the limit, the lower is the threshold savings rate, whereas the faster capital naturally dissipates, the higher is the threshold. If $\underline{\tau}$ is at least 1, then all feasible savings rates $\tau \in (0,1)$ imply the same behavior as the Solow-Swan outcome: Growth in y occurs in the long run at rate ξ. However, if $\underline{\tau}$ is less than 1, more intricate long-run dynamics can manifest. When an economy has τ less than $\underline{\tau}$, again, the result is the Solow-Swan outcome.

But when economies have sufficiently high savings rates, i.e., $\tau \in (\underline{\tau}, 1)$, then $\dot{\tilde{k}}/\tilde{k}$ always exceeds a time-invariant positive quantity, and has limiting behavior given by

$$\lim_{t\to\infty} \dot{\tilde{k}}(t)/\tilde{k}(t) = \left(\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1}\right)\tau - (\delta+\nu+\xi) > 0.$$

Moreover, such (y,k) paths tend towards balanced-growth equilibrium since

$$\dot{\tilde{k}}(t)/\tilde{k}(t) - \dot{\tilde{y}}(t)/\tilde{y}(t) = \left[1 - \frac{\nabla f(\tilde{k}(t))}{f(\tilde{k}(t))\tilde{k}(t)^{-1}}\right]\dot{\tilde{k}}(t)/\tilde{k}(t) \to 0 \quad \text{as } t\to\infty.$$

As long-run growth rates are then

$$\dot{y}/y = \xi + \left[(\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1})\tau - (\delta+\nu+\xi)\right] > \xi,$$

they increase in τ, meaning that economies saving a higher fraction of their income grow faster in the long run. It is this growth effect that makes the current specification an "endogenous growth" model. Compare this with the standard neoclassical growth model where savings rates affect only the levels of balanced-growth sample paths, not growth rates.

This relation between savings and long-run income growth applies only to those economies with savings rates exceeding the threshold value $\underline{\tau}$. All economies with savings rates below this value cannot influence long-run income growth rates by changing their savings behavior (unless they move savings rates above that threshold). What observable implications follow from this? If savings rates were

uniformly distributed across countries, there should be one cluster of economies around the same low per capita income growth rate and a different group with scattered income growth rates increasing in savings rates; see, for instance, Fig. 6.

As in the standard neoclassical model, this asymptotically linear technology model can be given a general equilibrium interpretation. Recall assumption (9b), and assume the preference parameter θ satisfies:

$$\frac{\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho+\delta)}{\xi} > \theta > \frac{\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho+\delta)}{\rho-\nu} > 0. \tag{25}$$

From (10) the parameter θ is the inverse of the intertemporal elasticity of substitution. Thus, (25) states that that elasticity can be neither too high nor too low—it must respect bounds varying with technology parameters.

From $\rho > \nu$, (25) implies that $\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} > \rho+\delta$. For the interval of feasible values for θ to exist, it suffices that $\xi < \rho - \nu$, which in turn follows from (11). Finally, these relations imply

$$\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} > \delta + \nu + \xi,$$

which had been used earlier to guarantee $\underline{\tau} < 1$. Thus, (25) is related to but strengthens the assumption underlying Fig. 5.

In the Technical Appendix, we show that (25) implies there exists a balanced-growth equilibrium with a positive growth rate given by

$$\lim_{t\to\infty} \dot{\tilde{y}}(t)/\tilde{y}(t) = \left(\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - [\rho+\delta+\theta\xi]\right)\theta^{-1} > 0,$$

and that for every initial $k(0)$ there exists an equilibrium tending towards balanced growth. If, however, θ is too large, then the unique balanced-growth equilibrium has $\lim_{t\to\infty} \dot{\tilde{y}}(t)/\tilde{y}(t) = 0$. The equilibria have exactly the character described above in the discussion surrounding Fig. 5, only with θ^{-1} replacing τ.

The models in Rebelo [96] and Romer [98] differ from those above in several important ways. Rebelo [96] uses a linear AK specification in place of the usual convex production technologies. (Linearity, of course, implies asymptotic linearity.) Equilibrium in that model tends towards balanced growth.

Romer [98] distinguishes the productive effects of individual-specific physical capital from economy-wide externalities induced by private accumulation. Romer's model uses the production technology (1b) with the arguments to F identified as the actions of private agents, and lets A depend on K, but with K defined as the social or aggregate outcome. Private agents ignore the effects of their actions on A; there is an externality in private agents' decisions to accumulate physical capital.

In Romer's model, as far as private agents are concerned, A still evolves exogenously. In equilibrium, of course, A depends on the purposeful actions of economic agents, and thus is properly viewed as endogenous. Private agents' optimizing decisions on consumption and savings remain identical to those in the standard neoclassical model. At the same time, the equilibrium aggregate outcome can display ongoing, endogenously-determined growth differing from the standard model. Moreover, the model also allows evaluating the efficiency properties of particular decentralized economic equilibria. Some versions of Romer's model imply equilibria tending towards balanced growth; others display ongoing growth but with no tendency towards balanced growth.[8]

Essential economic features therefore differ. However, the model of Rebelo [96] and certain versions of the general model in Romer [98] resulting in ongoing endogenous growth have, in essence, the same mathematical structure as that described earlier in this section. Their observable implications, therefore, are also the same.

One apparently natural conclusion from these models is that the researcher should now calculate regressions across economies of income growth rates on savings rates, tax rates, and so on—variables that in the analyses of Jones and

[8] A suitably parameterized model following Example 1 in Romer [98, p. 1028] yields equilibria tending towards balanced growth.

Manuelli [59], Rebelo [96], and Romer [98] potentially affect long-run growth rates. Such regressions would resemble the MRW regression (23) except that there is now no reason for the initial condition $\log y(t)$ to appear with a negative coefficient. This line of reasoning suggests that what distinguishes exogenous and endogenous growth models is whether the initial condition $\log y(t)$ enters negatively in an equation explaining growth rates. Note, though, that this endogenous growth analysis does not imply that the initial condition $\log y(t)$ should *never* appear in an estimated regression. By contrast, that initial condition is absent only in the balanced-growth limit, i.e., with $\tilde{k}$ infinite. But in any balanced-growth limit, even the exogenous-growth neoclassical model has the initial condition vanish from the right of (19), (19′), or (23).

Nonconvexities and poverty traps

An alternative class of models has focused on specific nonconvexities in the aggregate production function.[9] This research has analyzed the implications of such nonconvexities for the relation between initial conditions and the steady state behavior of aggregate output. Models with nonconvexities, unlike the neoclassical model, lead to long-run dependence in the time-series properties of aggregate output. Specifically, nonconvex models can display poverty traps, where economies with low initial incomes or capital stocks converge to one steady-state level of per capita output, while economies with high initial incomes or capital stocks converge to a different steady-state level.

Examples of such models include those by Durlauf [37], Galor and Zeira [51], and Murphy, Shleifer, and Vishny [80]. The model due to Azariadis and Drazen [4] is particularly convenient for illustrating the empirical differences between this framework and the neoclassical approach. The Azariadis-Drazen model works off

[9] Increasing returns to scale, of the kind studied in Romer [98], is also a nonconvexity, of course. What we mean instead are those nonconvexities associated specifically with certain threshold effects we will describe below.

thresholds in the accumulation of human and physical capital. These thresholds stem from spillovers between individual investments arising when aggregate capital is sufficiently high. In effect, economies with insufficient aggregate capital have different production functions from those with sufficiently high aggregate capital.

We present the basic ideas of Azariadis and Drazen [4] in our framework as follows. Modify the MRW production technology (20) to:

$$\begin{aligned} \tilde{y}(t) &= \tilde{k}_p(t)^{\alpha_p(t)} \tilde{k}_h(t)^{\alpha_h(t)} \\ \alpha_p(t) &= \begin{cases} \overline{\alpha}_p & \text{if } \tilde{k}_p(t) > \kappa_p(t) \\ \underline{\alpha}_p & \text{otherwise;} \end{cases} \\ \alpha_h(t) &= \begin{cases} \overline{\alpha}_h & \text{if } \tilde{k}_h(t) > \kappa_h(t) \\ \underline{\alpha}_h & \text{otherwise;} \end{cases} \end{aligned} \tag{26}$$

where the explicit (t) indicates variables changing through time and the coefficients $\alpha_p(t)$, $\alpha_h(t)$ vary with the underlying state $(\tilde{k}_p, \tilde{k}_h)$. The quantities $\kappa_p(t)$ and $\kappa_h(t)$ denote thresholds for physical and human capital respectively. They are written to depend on t to allow the aggregate production function possibly evolving through time.

The nonconvexities associated with these threshold effects can generate multiple steady-state equilibria, depending on the dynamics of $\kappa_p(t)$ and $\kappa_h(t)$. For instance, when $\tilde{k}_p(t)$ is low, the $\underline{\alpha}_p$ branch in (26) is activated, which in turn can imply the same steady-state equilibrium (low) value of $\tilde{k}_p(t)$. However, when $\tilde{k}_p(t)$ is high instead, then the $\overline{\alpha}_p$ branch is activated, so the high value of $\tilde{k}_p(t)$ can now be a steady-state equilibrium as well.

This description clarifies an important general point. When the aggregate production function contains threshold effects, there will not exist a linear cross-section growth relationship of the kind conventionally studied. Even if over a fixed time period no economies moved across capital thresholds, an economy with production technology (26) will follow one of four distinct Solow-Swan laws of motion, depending on the configuration of values in $\tilde{k}_p(t)$, $\tilde{k}_h(t)$, $\kappa_p(t)$, and $\kappa_h(t)$.

Thus, across economies, four classifications exist, with the Solow-Swan dynamics differing across each classification.

Under these assumptions, the law of motion for economy j changes from (23) to have α_p, α_h, and thus λ depend on time and state:

$$\begin{aligned}\log y_j(t+T) - \log y_j(t) \\ &= T\xi + (1 - e^{\lambda_j T}) \left[\log A_j(0) + t\xi\right] \\ &\quad + \frac{(1 - e^{\lambda_j T})}{1 - \lambda_p - \lambda_h} \left[\alpha_{p,j}\tau_{p,j} + \alpha_{h,j}\tau_{h,j} - (\alpha_{p,j} + \alpha_{h,j}) \log(\delta + \nu_j + \xi)\right] \\ &\quad - (1 - e^{\lambda_j T}) \log y_j(t).\end{aligned} \tag{27}$$

Durlauf and Johnson [39] study (27) and find evidence for multiple regimes in cross-country dynamics. They conclude that initial conditions matter, and that the MRW extension of the neoclassical model does not successfully explain the patterns of growth across countries. We discuss their findings in greater detail below in Section 5, under *Clustering and classification*.

Dynamics similar to those in the Durlauf-Johnson equation (27) also obtain in the model of Galor and Zeira [51]. Quah [92] applies Galor and Zeira's ideas to study empirically cross sections of economies (rather than cross sections of families as in the original model). Fig. 7—a two-regime counterpart to equation (27)—is used to motivate analysis of the distribution dynamics in cross-country incomes.

This formulation gives an interpretation different from that in Azariadis and Drazen [4] and Durlauf and Johnson [39]. Here, only one law of motion exists across economies—that given in Fig. 7. However, that law of motion displays a *polarization* effect, namely, economies evolve towards one of two distinct steady states (see, e.g., Esteban and Ray [42]). Regardless of the interpretation, however, the observable implications are the same. Already-rich economies converge to a high steady-state level; already-poor ones, to a low steady-state level.

Endogenous growth: R&D and endogenous technical progress

Yet a different class of endogenous growth models turns to features of the production technology (1) thus far unconsidered.

We have already described Romer's model [98] with accumulation externalities, where in (1b) A is taken to depend on the social outcome in capital investment. While A—the ultimate cause of growth—evolves endogenously, it is not the consequence of a deliberate action by any economic agent. One class of endogenous growth models makes A directly the result of such choices. Our immediate concern is: How do the empirical implications then differ?

Certain key details differ, but the models of Aghion and Howitt [1], Grossman and Helpman [53], Jones [56], and Romer [99] all associate the evolution of A with a measurable input such as research and development expenditure, the number of scientists and engineers, and so on. By contrast, models such as those in Lucas [74, 75] focus on improvement in H—human capital embodied in the labor force—as the source for endogenous growth. When the production technology is (1a) the resulting dynamics in measured per capita income will be indistinguishable across A and H improvements.

The empirical approach suggested by this reasoning focuses on variables that proxy the effects and economic costs of research activity. Jones [57] notes that the US, for one, has seen neither permanent changes in growth rates nor trend path levels of per capita GDP since 1880. Yet, resources devoted to R&D, by almost any measure, have increased dramatically in the last half century alone. Thus, in Jones's analysis, R&D-based growth models (or, indeed, all growth models with "scale effects") are at odds with empirical evidence.

This conclusion has to be tempered somewhat in light of results from two distinct lines of research. Recall that the empirical evidence in Jones [57] takes two forms: his Figure 1, indicating stability of an (ex-ante estimated) deterministic time trend; and his Table 1, showing the time-series stability properties of US GDP per capita growth rates. This should be compared with that line of research beginning from the unit-root analyses of Nelson and Plosser [82], extending through the

breaking-trend research of Perron [84] (and numerous others since), arguing that, over different timespans, the time-series properties of different income measures do show permanent changes. We do not suggest here that the evidence is decisive one way or the other, merely that circumspection is called for in these univariate time-series analyses. The second line of research, e.g., Ben-David [13], documents permanent growth and trend path changes—across time samples comparable to that in Jones's work—for a wide range of countries other than the US. The subtlety of statistical tests on these growth series, and the wide range of variation observable in the data had, indeed, formed part of the empirical motivation in the early endogenous growth discussion in Romer [98].

Coe and Helpman [27] investigate the dependence of a country's A levels on domestic and foreign R&D capital. They relate their estimates of such cross-country spillovers to openness of an economy to trade. Their findings are two-fold: first, beneficial cross-country R&D spillovers are stronger, the more open is an economy. Across the G7, in particular, up to one quarter of the total benefits of R&D investment can accrue to one's trade partners. Second, the estimated effects on A of R&D—both foreign and domestic—are large. Coe and Helpman chose to conduct their analysis entirely in terms of productivity and income levels.

The Coe-Helpman and Jones analyses, although substantively interesting, raise issues that differ from our focus in this chapter. We therefore do not discuss them further below.

Growth with cross-country interactions

Lucas [75] presents a growth model with empirical implications that differ markedly from those we have considered above. The model shows how taking into account patterns of cross-country interaction—in this case, human capital spillovers—alters conclusions on patterns of growth, even when one considers fixed and quite standard production technologies.[10]

[10] To emphasize, it is spillovers *across* economies that will be of interest here, not spillovers within an economy, such as one might find in models with externalities.

In the notation of equation (1) take A and N to be constant and equal to 1, but let there now be *work effort* $w \in [0,1]$ so that:

$$Y = F(K, wH) \implies y = F(k, wH),$$

with F satisfying assumptions (2), (3), and (8) as in the Solow-Swan model. The harder the labor force works, the higher is w, and thus the more output can be produced for a given quantity of human capital H.

Assume there is no depreciation and adopt the Solow-Swan savings assumption so that:

$$\dot{k} = \tau y. \tag{28}$$

Begin by letting

$$\dot{H}/H = G(w), \qquad G(w) > 0 \text{ for } w > 0, \tag{29}$$

so that how fast human capital accumulates depends on work effort w. If the economy shows learning by doing, then $G' > 0$; on the other hand, schooling effects or resting effects (where having rested, labour is subsequently more efficient) give $G' < 0$.

A balanced-growth equilibrium is a configuration of time paths (y, k, H, w) satisfying (28) and (29) such that

$$\dot{y}/y = \dot{k}/k = \dot{H}/H \quad \text{and} \quad w = \bar{w} \text{ constant}.$$

Since w varies in a bounded interval, it is natural to take it constant in balanced growth. Further, assuming identical preferences across economies, all countries then select a common constant effort level $\bar{w}$. A theory of differing cross-country growth rates can be constructed from allowing $\bar{w}$ to vary, but that is not considered here.

From (28), we have in balanced growth

$$\begin{aligned} \dot{k}/k &= \tau y/k = \tau F(k, \bar{w}H)/k \\ &= \tau F(k/\bar{w}H, 1)(k/\bar{w}H)^{-1} = \tau f(k/\bar{w}H)(k/\bar{w}H)^{-1} \end{aligned}$$

(using homogeneity of degree 1 in F). Moreover, subtracting $\dot{H}/H = G(\bar{w})$ from both sides yields

$$\dot{k}/k - \dot{H}/H = \tau f(k/\bar{w}H)(k/\bar{w}H)^{-1} - G(\bar{w}).$$

The right hand side of this generates the same graph as Fig. 2 substituting $G(\bar{w})$ for $\delta+\nu+\xi$ and $k/\bar{w}H$ for $\tilde{k}$. Thus, we see that balanced-growth equilibrium exists, is unique, and is globally stable. Indeed, once again, Fig. 4 describes equilibrium time paths in y, and all the previous remarks apply. The substantive difference between the two models is that in the interactions model

$$H(t) = H(0)e^{G(\bar{w})t}$$

replaces the neoclassical technical progress term $A(t)$. Because k/H is constant across economies in balanced growth, economies evolve with per capita incomes following parallel paths. These levels of per capita income are determined by the initial level of human capital $H(0)$. As before, for a given economy, per capita income converges to its balanced-growth path. However, the balanced-growth paths of different economies will not be the same, unless those economies are identical in all respects, including initial conditions.

Next, suppose there are cross-country spillovers in accumulating human capital. Write world average human capital as $\mathbf{H}$, and suppose each economy is small relative to the rest of the world. Change (29) to

$$\text{for economy } j: \qquad \dot{H}_j = G(w)H_j^{1-\pi}\mathbf{H}^{\pi}, \quad \pi \in [0,1]. \tag{29$'$}$$

The parameter π measures the strength of cross-country spillovers in human capital. The larger is this parameter, the more does economy j's human capital evolve in step with the world average. Conversely, when π is zero, (29$'$) reduces to (29) where no cross-country spillover occurs.

Write from (29$'$)

$$\dot{H}_j/H_j - G(w) = \left[(\mathbf{H}/H_j)^{\pi} - 1\right] G(w).$$

This says that when H_j exceeds the world average $\mathbf{H}$, then growth in human capital in economy j slows below $G(w)$. On the other hand when H_j is low relative to $\mathbf{H}$, then growth speeds up and $\dot{H}_j/H_j$ exceeds $G(w)$. Applying this to balanced growth with $w = \bar{w}$—and recalling that each economy j is small relative to the world average—we see that the ratio $\mathbf{H}/H_j$ is globally stable around a unique steady-state value of unity, so that eventually $H_j = \mathbf{H}$ for all j.

But then all equilibrium observed time paths in Fig. 4 must coincide, so that the distribution of incomes across economies eventually converges to a point mass, as in Fig. 8.

What are the principal empirical conclusions to take away from this discussion? Whether or not convergence happens—in the sense that all economies converge to a common level of per capita output (illustrated in Fig. 8)—is a matter here of accounting for the interactions across countries, not only of assumptions on the form of the production function. Whether the cross section distribution piles up at a single value, as in Fig. 8, depends on the nature of those interactions. It is easy to see that if we allowed natural groupings of economies to form, so that economies within a group interact more with each other than with those outside, then the "average" $\mathbf{H}$ that they converge to will, in general, vary across groups. Depending on other assumptions one can construct models where convergence takes the form of convergence-club dynamics, as in Fig. 9 (e.g., Quah [94]).[11] The empirical intuition emerging from these models matches well that from the stylized facts discussed in Section 2.

5. Empirical techniques

This section describes a variety of empirical approaches that have been used in growth analysis.

[11] Models displaying persistent inequality between families due to neighborhood spillover effects (e.g., Bénabou [14] and Durlauf [38]) are also driven by endogenous formation of interaction networks.

Cross section regression: β-convergence

The most common approach to growth and convergence applies cross-section regression analysis to variants of (19) and (19′).[12] Taking δ, ν, ξ, and τ to be time-averaged measures for each country, the term $g((\delta + \nu + \xi)^{-1}\tau)$ is determined up to unknown parameters in an assumed production function. When the researcher tacks on a least-squares residual on the right of (19) or (19′) then cross-section least-squares regression with hypothesized steady-state levels or time-averaged growth rates in income potentially recovers the unknown parameters in these equations.

Barro and Sala-i-Martin [10] focus on the initial condition $\lambda \times e^{\lambda t} \times [\log \tilde{y}(0) - \log \tilde{y}^*(0)]$ in (19′), and ask if the coefficient λ is negative. If so, then the data are said to satisfy *β-convergence* (β in their paper is $-\lambda$ in this paper).

In Barro and Sala-i-Martin [9] the leading term in (19′), the common technology growth rate ξ, is constrained to be identical across regional economies in the cross section. If the same assumption is made in our model, a negative λ implies *unconditional β-convergence*. Following Barro and Sala-i-Martin [10], when this leading term depends on auxiliary economic variables—measures of democracy, political stability, industry and agriculture shares in countries, rates of investment—a negative λ implies *conditional β-convergence*.[13]

In most empirical studies, the choices of additional control variables are ad hoc across datasets and political units. As one example, the data appendix in Levine and Renelt [72] lists over 50 possibilities. Among the range of controls that have appeared in the literature are the growth of domestic credit, its standard

[12] Well-known examples include Barro and Sala-i-Martin [9, 10], Baumol [12], and Mankiw, Romer, and Weil [78], but the list is legion.

[13] Some researchers use the phrase *absolute β-convergence* to mean unconditional β-convergence. We prefer just to contrast conditional and unconditional. Thus, we also do not distinguish situations where the conditioning uses variables appearing in the original Solow-Swan model from where the conditioning uses yet a broader range of variables.

deviation, inflation and its standard deviation, an index of civil liberties, numbers of revolutions and coups per year, rates of primary and secondary enrollment, and measures of exchange-rate distortion and outward orientation.[14] Following the publication of Levine and Renelt's paper, yet other control variables have been introduced. We discuss further below the issues raised by these additional regressors.

Barro and Sala-i-Martin [10] and Sala-i-Martin [103] assert that with the right conditioning variables, a rate of convergence of 2% per year is uniformly obtained across a broad range of samples. They draw two implications: First, in a Cobb-Douglas production function for aggregate output, physical capital's coefficient is over 0.9, appreciably larger than the 0.4 implied by factor shares in national income accounts. Second, convergence occurs: the poor do catch up with the rich.

Mankiw, Romer, and Weil [78] provide an essentially equivalent β-convergence analysis when they add human capital investment as an additional control. Their analysis differs from the vast majority of such studies in that their modification of the basic growth regression is justified by an explicit economic model; namely, they estimate the exact law of motion generated by the Solow model with Cobb-Douglas technology.

The first column of results in Table 1 presents a baseline MRW estimate. From the estimated coefficient on $\log y_j(1960)$, the implied convergence rate $|\lambda|$ is 0.014, similar to Barro and Sala-i-Martin's 2%. However, the estimate of α_p is only 0.43, in line with physical capital's factor share in national income accounts.

Recalling the earlier comparison between (13) and (22), we note that the key contribution in Mankiw, Romer, and Weil [78] is to alter Barro and Sala-i-Martin's first conclusion. In MRW a low estimated rate of convergence does not imply a large coefficient α_p for physical capital. Indeed, as seen in Tables IV, V, and VI of their paper, Mankiw, Romer, and Weil find convergence rates similar to Barro and Sala-i-Martin's estimates. The difference between the two papers is the structural

[14] Of course, none of these is explicitly modelled in either neoclassical or endogenous growth analyses.

interpretation of that 2% rate of convergence.[15]

Researchers have identified a number of econometric problems with conditional β-convergence analysis. Den Haan [31] and Kocherlakota and Yi [64] argue that how one augments the growth model with stochastic disturbances profoundly affects the inference to be drawn from the data.[16] Their point resembles the classical econometric result where serially correlated disturbances in distributed lag equations lead to regression estimators that are inconsistent for the parameters of interest.

A more fundamental interpretive difficulty for β convergence analysis arises from recalling Fig. 4, where cross-country growth patterns can exhibit highly nonlinear dynamics. Suppose that the a and b values there index multiple steady-state equilibria in the sense of, say, Azariadis and Drazen [4]. The figure then graphically illustrates the point in Bernard and Durlauf [17] and Durlauf and Johnson [39] that models having multiple steady states can display convergence of the kind studied in Barro and Sala-i-Martin [10], Mankiw, Romer, and Weil [78], and others. Thus, for discriminating between models having widely different policy implications, standard cross-country tests of convergence need not provide great insight. While, under the neoclassical model, the conventional cross-country growth equation is

[15] Cohen [28] takes this "deconstruction" exercise a step further, and in a different direction. He argues that, typically-constructed stocks of human and physical capital show unconditional β convergence, even if per capita income does not. He concludes that it is the dynamics of the Solow residual across countries that account for this, and suggests a vintage human-capital model to explain it.

[16] This result on the importance of the stochastic specification is related to but different from that in Kelly [61] and Leung and Quah [70]. These authors show that an appropriate stochastic specification can distort, not just statistical inference, but the underlying relation between physical capital's coefficient in the production function and the convergence or divergence properties of observed per capita income. In some of the examples they construct, even technologies displaying increasing returns to scale can give convergence of the cross-section distribution to a degenerate point mass.

(approximately) linear, under many endogenous growth models, it is profoundly nonlinear. As shown in Bernard and Durlauf [17], using a linear specification to test one model versus another is then of limited use. Put differently, relative to the class of endogenous growth models, no uniformly most powerful test exists under the null hypothesis of the neoclassical model. To emphasize the point, recall from Section 4 that while the Romer [98] model produces observations not satisfying (conditional) β-convergence, data generated by the Azariadis-Drazen [4] model might—even though in both kinds of endogenous growth models, global convergence fails.

The linear/nonlinear distinction we have just drawn is not mere nitpicking. The lack of attention to the implications of nonlinear alternatives to the neoclassical growth model in assessing empirical results is one basis for our rejecting the commonly held position summarized in Barro [7, p. x]: "It is surely an irony that one of the lasting contributions of endogenous growth theory is that it stimulated empirical work that demonstrated the explanatory power of the neoclassical growth model." If the explanatory power of a model means, as we think it should, demonstrating that greater understanding of some phenomenon derives from that model as opposed to its alternatives, rather than merely compatibility with some empirical observations, then evidence of β convergence simply does not provide the sort of corroboration of the neoclassical model claimed by Barro and many others.[17]

Barro and Sala-i-Martin [9] recognize that part of the importance of the convergence-rate estimate lies in its ability to shed light on whether and how rapidly poorer economies are catching up with the richer ones. They attempt to analyze this question through use of their concept of σ-convergence. They define *σ-convergence* to occur when the cross-section standard deviations of per capita incomes diminish over time. This type of convergence differs from β-convergence; that they are not the same illustrates some of the conceptual difficulties associated with statistical convergence measures in general and cross-country growth

[17] See Galor [50] for further discussion.

regressions in particular.

But σ-convergence too is problematic. To understand those difficulties, it is convenient to begin with a further look at β-convergence. For simple stochastic models constructed around (19), quite elaborately varied behavior for the cross-section distribution is consistent with even well-behaved (unconditional) β-convergence. Fig. 10a–c, similar to those in Quah [93], show three possibilities. It is easy to generate all three from a single fixed model satisfying the same transition dynamics as given in (19), varying only $y(0)$ and the variance of the regression residual term (itself ad hoc and not suggested by any explicit economic structure). Thus, the same β-convergence statistics are found in all three cases, even though implications on the poor catching up with the rich differ across them.

We can make this argument explicit by drawing on reasoning given in Quah [89]. Remove from each observed y its upward-sloping steady state growth path in Fig. 10a–c, so that all y's have mean zero. Suppose, moreover, that in the long run these transformed y's satisfy two conditions:

(i) Holding the cross-sectional economy j fixed, the time-series process y_j is stationary with finite second moments. This holds for all j.

(ii) Holding the time point t fixed, the collection of random variables $\{y_j(t) : \text{integer } j\}$ is independently and identically distributed. This holds for all t.

These restrictions are innocuous, given the points we wish to make here: essentially the same conclusions hold under quite general conditions.

For an arbitrary pair of time points t_1 and t_2 with $t_1 < t_2$, the population cross-section regression of $\log y(t_2)$ on a constant and $\log y(t_1)$ is, by definition, the projection

$$P\left[\log y(t_2) \mid 1,\ \log y(t_1)\right] = E_C \log y(t_2) + b\left(\log y(t_1) - E_C \log y(t_1)\right),$$

where

$$b = \mathrm{Var}_C^{-1}\left(\log y(t_1)\right) \cdot \mathrm{Cov}_C\left(\log y(t_2), \log y(t_1)\right),$$

and the C subscript denotes *cross-section*. Rearranging the projection so that growth rates appear on the left gives

$$P\left[\log y(t_2) - \log y(t_1) \mid 1,\ \log y(t_1)\right] \\ = \left[E_C \log y(t_2) - bE_C \log y(t_1)\right] - (1-b)\log y(t_1). \quad (30)$$

The sign of the coefficient on $\log y(t_1)$ in this regression depends on whether b exceeds 1. The projection coefficient b, in turn, depends on how large the covariance between growth and initial income is relative to the variance of initial income. Suppose that we are in the situation described by Fig. 10c, where long-run stationary steady state has been reached and $\log y(t)$ has its cross sectional variances invariant in time. Since $t_2 > t_1$, equation (30) is a regression of growth rates on initial conditions. The Cauchy-Schwarz inequality

$$|\mathrm{Cov}_C(\log y(t_2), \log y(t_1))| \leq \mathrm{Var}_C^{1/2}(\log y(t_2))\mathrm{Var}_C^{1/2}(\log y(t_1))$$

(with the inequality strict except in degenerate cases) then implies that $-(1-b)$ in (30) is negative. In words, the conditional average—for that is what is represented by a cross-section regression—shows its growth rate negatively related to its initial level. That might, at first, suggest that we should see converging cross-section dynamics like those in Fig. 10b, where the poor eventually attain the same income levels as the rich. However, recall that this negative relation between growth rates and initial levels has been constructed precisely when the cross-section dynamics are instead those in Fig. 10c, where the gap between poorest and richest is always constant.

More elaborate examples are easily constructed. For one, we need not consider situations only at long-run steady state. Since—outside of degenerate cases—the Cauchy-Schwarz inequality is strict, it is easy to find examples where $-(1-b)$ is negative even when $\mathrm{Var}_C(\log y(t_2))$ is bigger than $\mathrm{Var}_C(\log y(t_1))$, i.e., the cross-section dispersion is increasing even as the regression representation is suggesting dynamics like Fig. 10b. Moreover, if one perturbs the regressor so that it is not

$\log y(t_1)$ but instead some other $\log y(t_0)$ then the same argument shows that the regression coefficient on the "initial" level can be positive regardless of whether the cross-section distribution is expanding, diminishing, or unchanged in dispersion.

Different interpretations can be given to the effects we have just described—one early manifestation of these is known in the statistics literature as Galton's fallacy or Galton's paradox (see, e.g., Friedman [48], Maddala [76, 3.12], Stigler [108, Ch. 8], or Quah [89]).[18] We prefer to regard the situation constructed above as one where knowledge of what happens to the conditional average (the regression representation) is uninformative for what happens to the entire cross section. In this interpretation, further β-convergence regression analysis of the growth equation (23)—be it with cross section data, panel data, or any other structure; be it conditional or unconditional—cannot reveal whether the poor will catch up with the rich. These considerations suggest instead directly analyzing the dynamics of the cross-section distribution. Doing so goes beyond studying just σ convergence, as the latter studies only one aspect of the distribution at each point in time. Moreover, σ convergence is silent on whether clusters form within the cross section (as in the emerging twin peaks of Fig. 1) and on whether transitions occur within the distribution: both Fig. 10c and Fig. 10d show the same σ-convergence dynamics, yet economic behavior across them must differ dramatically.

[18] This connection had been impressed on Quah by G. S. Maddala and Marc Nerlove separately in private communication.

Augmented cross section regression

More recent empirical growth studies have tried to go beyond the original cross-section regressions and, instead, emphasize identifying those factors that explain international differences. Relative to the neoclassical growth model of Section 4, these exercises can be interpreted as parameterizing A.

Table 2 surveys those regressors that, in the literature, have been used in cross-country regressions.[19] In addition to the four variables suggested by the augmented Solow-Swan model (initial income and the rates of human capital investment, physical capital investment, and population growth), the table includes 36 different categories of variables and 87 specific examples. Recall that the sample to which nearly all these additional control variables have been applied has only about 100 observations (the size of the subsample typically used from the Heston-Summers dataset).

While these augmented cross-section regression studies have suggested some insightful extensions of the neoclassical growth model, we find problematic the lessons drawn from some of the empirical findings.

First, many studies fail to make clear whether the regressions they consider can be interpreted within some economic model. It is certainly always possible to let A be a linear function of arbitrary control variables. But exploiting that hypothesized linear function need not be a useful way of studying the control in question. For example, the threshold externality in the Azariadis-Drazen [4] model can be viewed as a latent variable indexing the aggregate production function. Such an interpretation is plausible for factors ranging from international market access to political regime—the ability of a society to innovate and to exploit readily available opportunities is influenced by political culture, with well documented historical examples going as far back as Athens and Sparta. However, we conclude from the model that these factors induce nonlinearities in the growth relation.

[19] Temple [112] provides an excellent literature overview discussing some of these studies in greater detail.

Linear regressions are, not surprisingly, unable to get at the features of interest.

Moreover, it is unclear what exercise a researcher conducts by adding a particular control variable, even when the variable is motivated by a particular economic theory. The basic Solow-Swan model admits an immense range of extensions through factors such as inequality, political regime, or trade openness. These are often highly correlated with one another, and are neither mutually exclusive nor prioritized as possible explanations of growth. Hence, it is difficult to assign much import to the statistical significance of an arbitrarily chosen subset of possible controls. We therefore find unpersuasive claims that these regressions are able to identify economic structure.

The problem of open-ended alternative models also extends to various attempts in the literature to find instruments for the various baseline and augmented Solow-Swan regressors, which are of course typically endogenous themselves. Frankel and Romer [45] use geographic variables to instrument their measure of trade openness. However, that these variables are exogenous with respect to trade openness does not make them legitimate instruments. For example, from the perspective of European and Asian history it is wholly plausible that land mass correlates with military expenditures and military strength, which themselves correlate with tax rates and political regime—two alternative augmentations of the Solow model which have been proposed. Because growth explanations are so broad, it is especially easy to construct plausible reasons why "exogenous" instruments are less useful than they might first appear. The failure of the growth model to naturally generate useful instruments contrasts with rational expectations models whose structure produces such instruments automatically from the orthogonality of forecast errors and available information.

This reasoning has led to a re-examination of the empirical conclusions from this line of work. The issue has been addressed in two ways. First, Levine and Renelt [72] have challenged many of the findings in cross-country growth regressions. They emphasized that findings of statistical significance may be fragile due to dependence on additional controls whose presence or absence is not strongly

motivated by any theory. By applying Leamer's [68] extreme bounds analysis (thereby identifying the range of coefficient estimates for a given regressor generated by alternative choices of additional regressors) they found that only the physical capital investment rate and, to a weaker degree, initial income are robustly related to cross-country growth rate differentials.

Levine and Renelt [72] have identified a serious problem with the empirical growth literature. However, their procedure for dealing with the problem is itself problematic. The difficulty may be most easily seen in the following example. Suppose that one is interested in the coefficient b_0 relating variables X and Y, where the true data generating process is given by

$$Y_j = X_j b_0 + \epsilon_j,$$

with X deterministic and ϵ normally distributed $\mathbb{N}(0, \sigma_\epsilon^2)$. Suppose the researcher considers a set of controls $\{Z_l : \text{ integer } l\}$, each Z_l being separately entered in the regression:

$$Y_j = X_j b + Z_{l,j} c_l + \epsilon_j. \tag{31}$$

Assume that the Z_l's are nonstochastic and that, in sample, have zero cross-product with X. Denote the sample second moments of X and Z_l by $\|X\|^2$ and $\|Z_l\|^2$ respectively. Then OLS on (31) produces $\hat{b}$ estimates that are draws from the normal distribution $\mathbb{N}(b_0\,, (\|X\|^2 + \|Z_l\|^2)^{-1}\sigma_\epsilon^2)$. Since the Z_l's are deterministic, as researchers increase the number of separate Z_l's used in different regression analyses, so does the probability increase that some draw on $\hat{b}$ will have sign opposite to that on the true b.[20]

The problem is that the $\hat{b}$ distribution has support that can become unbounded due to sampling variation induced by the arbitrarily chosen regressors. Without a theory on how to control this problem, it is difficult to draw strong conclusions

[20] The basic argument clearly still applies when the Z_l's are stochastic, even though then the $\hat{b}$ distributions are not typically normal.

about the fragility of regression coefficients. Hence, while we find the Levine and Renelt analysis suggestive, the import of the challenge is unclear.

Sala-i-Martin [104] has attempted to deal with this limitation by calling "robust" only those variables found statistically significant in 95% of a group of regressions in a wide range of possible combinations of controls. This work finds that many more variables appear to be robust. These variables fall into 9 categories: 1) region (dummy variables for Sub-Saharan Africa and Latin America), 2) political structure (measures of rule of law, civil liberties and political instability), 3) religion, 4) market distortions (measured with reference to official and black market exchange rates), 5) equipment investment, 6) natural resource production, 7) trade openness, 8) degree of capitalism, and 9) former Spanish colonies.

However, it is again unclear how to interpret such results. Suppose that one were to take a given regression relationship and begin to include alternative sets of right hand side variables which were in each case orthogonal to the original regressors. The presence or absence of these regressors would have by assumption no effect on estimated coefficient size or estimated standard errors. Hence, one could always generate an arbitrarily large number of regressions with the same significant coefficient but with no implications as to whether the coefficient estimate is or is not robust. It is impossible to know whether Sala-i-Martin's exercise actually reveals something about robustness, or merely something about the covariance structure of the controls which he studies. Further, the exercise assumes that robustness is interesting outside of the context of which variables are under study. The fact that the presence of one variable in a growth regression renders another insignificant is not vitiated by the fact that others do not do so, when the first is of economic interest, and the others are not.

The problem with both these approaches to robustness of control variables in growth regressions is that they attempt to use mechanical statistical criteria in identifying factors whose interest and plausibility is motivated by economic (or social science) theory. The dimensions along which one wants estimates to be robust are determined by the goals of the researcher, which cannot be reduced to

algorithms of the kind that have been employed.

Panel data analysis

To permit unobservable country-specific heterogeneity in growth regressions, Benhabib and Siegel [15], Canova and Marcet [21], Caselli, Esquivel, and Lefort [22], Evans [43], Islam [55], Lee, Pesaran, and Smith [69], and Nerlove [83] have used panel data methods to study cross-country income data. Following traditional motivation in panel-data econometrics (e.g., Chamberlain [24]), many such studies seek to eliminate, in the notation of Section 4, unobservable country-level heterogeneity in $A(0)$. Those heterogeneities, denoted *individual effects* in the language of panel-data econometrics, constitute nuisance parameters that within the conventional framework the researcher attempts to remove.[21]

Panel data studies proceed from the neoclassical (MRW) model (23) as follows. Assume that depreciation δ and technology growth ξ are constant across economies. Fix horizon T, append a residual on the right, and redefine coefficients to give, across economies j, the regression equation:

$$\log y_j(t+T) - \log y_j(t) = b_0 + b_1 \log y_j(t) + b_2 \log \tau_{p,j} + b_3 \log \tau_{h,j} + b_4 \log(\delta + \nu_j + \xi) + \epsilon_{j,t} \tag{32}$$

[21] Canova and Marcet [21] and Evans [43] are exceptions to this. Canova and Marcet analyze a Bayesian-motivated parameterization of the individual effects, and conclude that those effects do, indeed, differ across economies. Evans, using a different statistical technique, concludes the same. Evans follows Levin and Lin [71] and Quah [90] in taking an underlying probability model where both time and cross-section dimensions in the panel dataset are large. This contrasts with standard panel-data studies where the time dimension is taken to be relatively small. The large N, large T framework then allows inference as if the individual effects are consistently estimated, and permits testing for whether they differ across countries.

with

$$
\begin{aligned}
b_0 &\stackrel{\text{def}}{=} (1 - e^{\lambda T}) \log A(0) + (t + T - e^{\lambda T} t)\xi, \\
b_1 &\stackrel{\text{def}}{=} e^{\lambda T} - 1, \\
b_2 &\stackrel{\text{def}}{=} (1 - e^{\lambda T}) \frac{\alpha_p}{1 - \alpha_p - \alpha_h}, \\
b_3 &\stackrel{\text{def}}{=} (1 - e^{\lambda T}) \frac{\alpha_h}{1 - \alpha_p - \alpha_h}, \\
b_4 &\stackrel{\text{def}}{=} -(1 - e^{\lambda T}) \frac{\alpha_p + \alpha_h}{1 - \alpha_p - \alpha_h}.
\end{aligned}
$$

Let $T = 1$ and assume that b_0 is a random variable with unobservable additive components varying in j and t:

$$
\begin{aligned}
\log y_j(t+1) - \log y_j(t) = \mu_j + \kappa_t + b_1 \log y_j(t) \\
+ b_2 \log \tau_{p\,j} + b_3 \log \tau_{h\,j} + b_4 \log(\delta + \nu_j + \xi) + \epsilon_{j,t}.
\end{aligned} \tag{33}
$$

This formulation differs from the original MRW specification in two ways. First, the law of motion for output is taken in one-period adjustments. This is inessential, however, and the researcher is free to recast equation (32) with T set to whatever is deemed appropriate. Second, the (originally) constant b_0 is decomposed into economy-specific and time-specific effects

$$
b_0 = \mu_j + \kappa_t. \tag{34}
$$

Panel data methods, applied to the model above, have produced a wide range of empirical results. While Barro and Sala-i-Martin [9, 10] defend a 2% annual rate of convergence from cross-section regressions, estimates from panel data analyses have been more varied. Lee, Pesaran, and Smith [69] conclude annual convergence rates are approximately 30% when one allows heterogeneity in all the parameters. Islam [55] permits heterogeneity only in the intercept terms, and finds annual convergence rates between 3.8% and 9.1%, depending on the subsample under study.

Caselli, Esquivel, and Lefort [22] suggest a convergence rate of 10%, after conditioning out individual heterogeneities and instrumenting for dynamic endogeneity. Nerlove [83], by contrast, finds estimates of convergence rates that are even lower than those generated by cross-section regression. He explains this difference as being due to finite sample biases in the estimators employed in the other studies using the neoclassical growth model. The disparate results across panel data studies can sometimes, but not always, be attributed to the different datasets that different researchers have employed.

The use of a panel data structure has advantages and disadvantages. One significant advance comes from clarifying the difficulties in interpreting the standard cross-section regression. In particular, the dynamic panel (33) typically displays correlation between lagged dependent variables and the unobserved residual. The resulting regression bias depends on the number of observations in time and only disappears when that number becomes infinite. Moreover, the bias does not disappear with time averaging. Thus, if the dynamic panel were the underlying structure, standard cross-section regressions will not consistently uncover the true structural parameters.

But beyond simply pointing out difficulties with the cross-section OLS formulation, the panel data structure has been argued, on its own merits, to be more appropriate for analyzing growth dynamics. For instance, Islam [55] shows how time- and country-specific effects can arise when per capita output is the dependent variable instead of output per effective worker (Islam argues this substitution to be appropriate). Alternatively, one might view the error structure as a consequence of omitted variables in the growth equation, whereupon the separate time and country effects in (34) have alternative natural interpretations. These instances of the greater flexibility (and, thus, reduced possibilities for misspecification) allowed by panel-data analyses—unavailable to cross-section regression studies—account for their broader econometric use more generally, not just in studies of economic growth.

However, the putatively greater appeal of panel data studies should not go

unchallenged. To see the potential disadvantages, consider again the decomposition in (34). For researchers used to the conventions in panel-data econometric analysis, this generalization from a constant unique b_0 is natural. But for others, it might appear to be a proliferation of free parameters not directly motivated by economic theory.

Freeing b_0 so that it can vary across countries and over time can only help a theoretical model fit the data better. Restricting b_0 to be identical across countries and over time—when, in reality, b_0 should differ—can result in a model that is misspecified, thereby lowering confidence that the researcher has correctly identified and estimated the parameters of interest. This advantage of a panel data approach applies generally, and is not specific to growth and convergence. But for convergence studies, the flexibility from decomposing b_0 into economy-specific and time-specific components can instead be problematic, giving rise to misleading conclusions.

We describe two scenarios where we think this might naturally occur. First, note that equation (32) implies that $A(0)$ (and thus b_0 through μ_j) forms part of the long-run path towards which the given economy converges (see again Fig. 10). Ignore Galton's Fallacy to sharpen the point here. If the researcher insists that $A(0)$ be identical across economies, then that researcher concludes convergence to an underlying steady-state path precisely when catching up between poor and rich takes place. Thus, the implication from a convergence finding is transparent: it translates directly into a statement about catching up (again, abstracting away from Galton's Fallacy). By contrast, when the researcher allows $A(0)$ to differ across economies, finding convergence to an underlying steady-state path says nothing about whether catching up occurs between poor and rich: Fig. 10 shows different possibilities. This is not just the distinction between conditional and unconditional convergence. In panel-data analysis, it is considered a virtue that the individual heterogeneities $A(0)$ are unobservable, and not explicitly modelled as functions of observable right-hand side explanatory variables. By leaving free those individual heterogeneities, the researcher gives up hope of examining whether

poor economies are catching up with rich ones. The use of panel-data methods therefore compounds the difficulties in interpreting convergence regression findings in terms of catchup from poor to rich.

For the second scenario, recall that the problem the panel-data regression equation (33) traditionally confronts is the possibility that the μ_j's, the individual-specific effects, are correlated with some of the right-hand side variables. If not for this, OLS on equation (33) would allow both consistent estimation and (with appropriately corrected standard errors) consistent inference.[22] One class of solutions to the inconsistency problem derives from transforming equation (33) to annihilate the μ_j's. For instance, in the so-called "fixed-effects" or *within* estimator, one takes deviations from time-averaged sample means in equation (33), and then applies OLS to the transformed equation to provide consistent estimates for the regression coefficients. But note that in applying such an individual-effects annihilating transformation, the researcher winds up analyzing a left-hand side variable purged of its long-run (time-averaged) variation across countries. Such a method, therefore, leaves unexplained exactly the long-run cross-country growth variation originally motivating this empirical research. The resulting estimates are, instead, pertinent only for higher-frequency variation in the left-hand side variable: this might be of greater interest for business cycles research than it is for understanding patterns of long-run economic growth across countries.[23]

Our point is general: It applies not just to the fixed-effects estimator, but

[22] OLS might not be efficient, of course, and GLS might be preferred where one takes into account the covariance structure of the μ_j's.

[23] This statement clearly differs from saying that fixed-effects estimators are inconsistent in dynamic models without strict exogeneity of the regressors (e.g., Chamberlain [24]). The absence of strict exogeneity characterizes equation (33), and thus is an additional problem with fixed-effects estimators. This shortcoming has motivated studies such as Caselli, Esquivel, and Lefort [22] that use techniques appropriate for such correlation possibilities. However, those techniques do nothing for the short-run/long-run issue we raise.

also to the first-difference estimator, and indeed to any panel-data technique that conditions out the individual effects as "nuisance parameters". In dealing with the correlation between individual effects and right-hand side variables—a properly-justified problem in microeconometric studies (again see, e.g., Chamberlain [24])—the solution offered by panel-data techniques ends up profoundly limiting our ability to explain patterns of cross-country growth and convergence.[24]

Interestingly, that conditioning out country-specific effects leaves only high-frequency income movements to be explained creates not only the problem just described, but also its dual. Over what time horizon is a growth model supposed to apply? Many economists (or Solow and Swan themselves in the original papers for that matter) regard growth analyses as relevant over long time spans. Averaging over the longest time horizon possible—as in cross-section regression work—comes with the belief that such averaging eliminates business cycle effects that likely dominate per capita income fluctuations at higher frequencies. By contrast, Islam [55, p. 1137] has argued that since equation (23) is "based on an approximation around the steady state ... it is, therefore, valid over shorter periods of time." However, we think this irrelevant. Different time scales for analyzing the model are mutually appropriate only if the degree of misspecification in the model is independent of time scale. In growth work, one can plausibly argue that misspecification is greater at higher frequencies. Taking Islam's argument seriously, one might attempt using the neoclassical growth model to explain even weekly or daily income fluctuations in addition to decadal movements.

[24] Quah [93, p. 1367] has also argued this.

Time series: Unit roots and cointegration

An alternative approach to long-run output dynamics and convergence based on time-series ideas has been developed in Bernard and Durlauf [16, 17], Durlauf [36], and Quah [87]. Convergence here is identified not as a property of the relation between initial income and growth over a fixed sample period, but instead of the relationship between long-run forecasts of per capita output, taking as given initial conditions.

Bernard and Durlauf [17] define *time-series forecast convergence* as the equality of long-term forecasts taken at a given fixed date. Thus, given $\mathfrak{F}_t$ the information at date t, economies j and j' show time-series forecast convergence at t when:

$$\lim_{T\to\infty} E(y_j(t+T) - y_{j'}(t+T) \mid \mathfrak{F}_t) = 0,$$

i.e., the long-term forecasts of per capita output are equal given information available at t. It is easy to show that time-series forecast convergence implies β-convergence when growth rates are measured between t and $t+T$ for some fixed finite horizon T. The critical distinction between time-series forecast convergence and β-convergence is that an expected reduction in contemporary differences (β-convergence) is not the same as the expectation of their eventual disappearance.

This dynamic definition has the added feature that it distinguishes between convergence between pairs of economies and convergence for all economies simultaneously. Of course, if convergence holds between all pairs then convergence holds for all. Some of the theoretical models we have described—in particular, those with multiple steady states—show that convergence need not be an all or nothing proposition. Subgroups of economies might converge, even when not all economies do.

To operationalize this notion of convergence, a researcher examines whether the difference between per capita incomes in selected pairs of economies can be characterized as a zero-mean stationary stochastic process. Hence, forecast convergence can be tested using standard unit root and cointegration procedures. Under

the definition, deterministic (nonzero) time trends in the cross-pair differences is as much a rejection of convergence as is the presence of a unit root.

In the literature applying these ideas, two main strands can be distinguished. The first, typified by Bernard and Durlauf [16, 17], restricts analysis to particular subgroups of economies, for instance the OECD. This allows the researcher to use long time series data, such as those constructed by Maddison [77]. Multivariate unit root and cointegration tests reject the null hypothesis that there is a single unit root process driving output across the OECD economies—thus, across all the economies in the OECD grouping, time-series forecast convergence can be rejected. At the same time, however, individual country pairs—for instance, Belgium and the Netherlands—do display such convergence.

In a second strand, Quah [87] studies the presence of common stochastic trends in a large cross section of aggregate economies. He does this by subtracting US per capita output from the per capita output of every economy under study, and then examines if unit roots remain in the resulting series. Because the number of time-series observations is the same order of magnitude as the number of countries, random-field asymptotics are used to compute significance levels. Quah's results confirm those of Bernard and Durlauf described above. He rejects the null hypothesis of no unit roots in the per capita output difference series; in other words, he finds evidence against convergence (in the sense given by the forecasting definition).

Time series approaches to convergence are subject to an important caveat. The statistical analysis under which convergence is tested maintains that the data under consideration can be described by a time-invariant data generating process. However, if economies are in transition towards steady state, their associated per capita output series will not satisfy this property. Indeed, as argued by Bernard and Durlauf [17], the time series approach to convergence, by requiring that output differences be zero-mean and stationary, requires a condition inconsistent with that implied in cross-section regressions, namely that the difference between a rich and poor economy have a nonzero mean. Time-series and cross-section approaches

to convergence rely on different interpretations of the data under consideration. Hence they can provide conflicting evidence; in practice, the two approaches commonly do.

Clustering and classification

Following Azariadis and Drazen's theoretical insights [4], Durlauf and Johnson [39] study equation (27), and find evidence for multiple regimes in cross-country growth dynamics. They do this in the dataset originally used by MRW [78] by identifying sample splits so that within any given subsample all economies obey a common linear cross-section regression equation. Durlauf and Johnson allow economies with different 1960 per capita incomes and literacy rates (LR) to be endowed with different aggregate production functions. Using a regression-tree procedure[25] to identify threshold levels endogenously, Durlauf and Johnson find the MRW dataset display four distinct regimes determined by initial conditions:

1. $y_j(1960) < \$800$;
2. $\$800 \leq y_j(1960) \leq \$4850$ and $LR_j(1960) < 46\%$;
3. $\$800 \leq y_j(1960) \leq \$4850$ and $46\% \leq LR_j(1960)$; and
4. $\$4850 < y_j(1960)$.

Thus, the regression-tree procedure partitions the cross section into low, intermediate, and high-output economies, and then further divides the intermediate group across low and high literacy rates.

These groupings or regimes are, in turn, associated with markedly different aggregate production functions. The low-income regime shows $\alpha_p = 0.31$ and $\alpha_h = -0.03$ (although not statistically significant). For intermediate incomes the low-literacy grouping has $\alpha_p = 0.19$ (not statistically significant) and $\alpha_h = 0.42$, while the high-literacy grouping has $\alpha_p = 0.79$ and $\alpha_h = -0.07$ (not statistically significant). Finally, high-income economies display $\alpha_p = 0.31$ and $\alpha_h = 0.46$.

[25] Breiman, Friedman, Olshen, and Stone [20] describe the regression-tree procedure and its properties.

Table 1 gives these regression results for two of the regimes, and compares them with the original MRW regression. We see that, depending on initial conditions, different economies face aggregate production opportunities that differ considerably. This casts doubt on the empirical validity of the neoclassical model (with or without its MRW extension): Initial conditions matter for potential long-run incomes. Cross-country differences are not explained entirely by differences in the rates of physical and human capital accumulation and population growth.

Durlauf and Johnson's analysis [39] has a classification interpretation. Interest lies in which economies belong to which subgroups. This line of reasoning has been usefully extended. Franses and Hobijn [47] attempt to identify groups of similarly-behaving economies using measures of social indicators in addition to per capita GNP. Unlike Durlauf and Johnson [39] they use clustering algorithms to partition the data set. Nevertheless, their results are qualitatively similar. Franses and Hobijn [47] also find that high and low income economies do not converge to one another, but that they do converge (to different limits). Interestingly, Franses and Hobijn additionally find that productivity convergence does not lead to convergence in social indicators like infant mortality. This work suggests that a richer notion of convergence, one accounting explicitly for the multivariate nature of aggregate socioeconomic characteristics, warrants further study.

Distribution dynamics

Bianchi [18], Desdoigts [33], Jones [58], Lamo [67], Quah [88, 89, 92, 94] have studied the predictions of the theoretical growth models in terms of the behavior of the entire cross-section distribution. While this work is often quite technical, it can be viewed as just a way to make precise the ideas previously described informally in Section 2.

Turning back to Fig. 1, label the cross-section distribution F_t at time period t, and call the associated (probability) measure ϕ_t. Fig. 1 can then be interpreted as describing the evolution of a sequence of measures $\{\phi_t : \ t \geq 0\}$. In empirical work on distribution dynamics, the researcher seeks a law of motion for the stochastic

process $\{\phi_t : \ t \geq 0\}$. With such a scheme in hand, one can ask about the long-run behavior of ϕ_t: If ϕ_t displays tendencies towards a point mass, then one can conclude that there is convergence towards equality. If, on the other hand, ϕ_t shows tendencies towards limits that have yet other properties—normality or twin peakedness or a continual spreading apart—then those too would be revealed from the law of motion. Moreover, having such a model would allow one to study the likelihood and potential causes of poorer economies becoming richer even than those already currently rich, and similarly the likelihood and potential causes of those already rich regressing to become relatively poor. Finally, a researcher with access to such a law of motion can look further to ask what brings about particular patterns of cross-country growth.

The simplest scheme for modelling the dynamics of $\{\phi_t : t \geq 0\}$ is analogous to the first-order autoregression from standard time-series analysis:

$$\phi_t = T^*(\phi_{t-1}, u_t) = T^*_{u_t}(\phi_{t-1}), \ t \geq 1, \tag{35}$$

where T^* is an operator that maps the Cartesian product of measures and generalized disturbances u to probability measures, and $T^*_{u_t}$ absorbs the disturbance into the definition of the operator. (See the Technical Appendix for the meaning of * in the two operators T^* and $T^*_{u_t}$.) This is no more than a stochastic difference equation taking values that are entire measures. Equivalently, it is an equation describing the evolution of the distribution of incomes across economies.

A first pass at equation (35) discretizes the income space, whereupon the measures ϕ_t can be represented by probability vectors. For instance, Quah [88] considers dividing income observations into five cells: the first comprising per capita incomes no greater than 1/4 the world average (at each date); the second, incomes greater than 1/4 but no more than 1/2; the third, incomes greater than 1/2 but no more than the average; the fourth, greater than the average but no more than double; and finally, in the fifth cell, all other incomes. In terms of Fig. 1, at any given date t, a five-element probability vector ϕ_t completely describes the situation.

Moreover, since we observe which economies transit to different cells in this discretization (and the cells from which they came), we can construct a matrix M_t whose rows and columns are indexed by the elements of the discretization, and where each row of M_t is the fraction of economies beginning from that row element and ending up in the different column elements. By construction M_t has the properties of a *transition probability matrix*: its entries are non-negative and its row sums are all 1. If we assume that the underlying transition mechanism is time-invariant, then one can average the M_t's to obtain a single transition probability matrix M describing the dynamics of the (discretized) distribution.

Table 3 shows such an M, as estimated in Quah [88]. Because the transitions are only over a one-year horizon, it is unsurprising that the diagonal entries are close to 1, and most of the other entries are zero. What interests us, however, is not any single one of these numbers but what the entire law of motion implies. The row labelled *Ergodic* is informative here. To understand what it says, note that by construction:

$$\phi_{t+1} = M'\phi_t,$$

so that

$$\forall s \geq 1: \qquad \phi_{t+s} = (M^s)'\phi_t. \tag{36}$$

Since M is a transition probability matrix, its largest eigenvalue is 1, and the left eigenvector corresponding to that eigenvalue can be chosen to have all entries non-negative summing to 1. Generically, that largest eigenvalue is unique, so that M^s converges to a rank-one transition probability matrix. But then all its rows must be equal, and moreover equal to that probability vector satisfying:

$$\phi_\infty = M'\phi_\infty.$$

The vector ϕ_∞ is the *Ergodic* row vector; it corresponds to the limit of (36) as $s \to \infty$. In words, ϕ_∞ is the long-run limit of the distribution of incomes across

economies.[26]

Table 3 shows that limiting distribution to be twin-peaked. Although in the observed sample, economies are almost uniformly distributed across cells—if anything, there is a peak in the middle-income classes—as time evolves, the distribution is predicted to thin out in the middle and cluster at rich and poor extremes. This polarization behavior is simply a formalization of the tendencies suggested in Fig. 1.

Such analysis leads to further questions. How robust are these findings? The discretization to construct the transition probability matrix is crude and ad hoc. Moving from a continuous income state space—Fig. 1—to a discrete one comprising cells—Table 3—aliases much of the fine details on the dynamics. Does changing the discretization alter the conclusions?

To address these issues, we get rid of the discretization. In the Technical Appendix we describe the mathematical reasoning needed to do this. The end result is a *stochastic kernel*—the appropriate generalization of a transition probability matrix—which can be used in place of matrix M in the analysis. Quah [92, 94] estimates such kernels. Fig. 11 shows the kernel for the transition dynamics across 105 countries over 1961 through 1988, where the transition horizon has been taken to be 15 years. The twin-peaked nature of the distribution dynamics is apparent now, without the aliasing effects due to discretization.

Bianchi [18] and Jones [58] eschew dealing with the stochastic kernel by considering the cross-section distribution F_t for each t in isolation. This ignores information on transition dynamics, but is still useful for getting information on the shape dynamics in F. Each F_t is estimated nonparametrically. Bianchi [18] goes further and applies to each F_t a bootstrap test for multimodality (twin-peakedness,

[26] Potential inconsistency across M matrices estimated over single- and multiple-period transitions is a well-known problem from the labor and sociology literature (e.g., Singer and Spilerman [105]). Quah [88] shows that, in the Heston-Summers cross-country application, the long-run properties of interest are, approximately, invariant to the transition period used in estimation.

after all, is just bimodality). Bianchi finds that in the early part of the sample (the early 1960s) the data show unimodality. However, by the end of the sample (the late 1980s) the data reject unimodality in favor of bimodality. Since Bianchi imposes less structure in his analysis—nowhere does he consider intradistribution dynamics, or in the language of the Technical Appendix, the structure of T^*—one guesses that his findings are more robust to possible misspecification. Here again, however, twin-peakedness manifests.

We have taken care, in building up the theoretical discussion from the previous sections, to emphasize that those models give, among other things, ways to interpret these distribution dynamics. An observed pattern in the distribution dynamics of cross-country growth and convergence can be viewed as a reduced form—and one can ask if it matches the theoretical predictions of particular classes of models. We view in exactly this way the connection between the empirics just discussed and the distribution dynamics of models such as Lucas's [75] described in Section 4 above.

The work just described, while formalizing certain facts about the patterns of cross-country growth, does not yet provide an explanation for those patterns. Putting this differently, we need to ask what it is that explains these reduced forms in distribution dynamics. In light of our discussion above on the restrictions implied by cross-country interactions, we conjecture that this "explaining distribution dynamics" needs to go beyond representative-economy analysis. Quah [94] has addressed exactly this issue: in the spirit of our discussion above on theoretical models with cross-country interaction, Quah asks for the patterns of those interactions that can explain these reduced-form stochastic kernels. He finds that the twin-peaks dynamics can be explained by spatial spillovers and patterns of cross-country trade—who trades with whom, not just how open or closed an economy is.

6. Conclusion

We have provided an overview of recent empirical work on patterns of cross-country growth. We think the profession has learned a great deal about how to match those empirical patterns to theoretical models. But as researchers have learned more, the criteria for a successful confluence of theory and empirical reality have also continued to sharpen.

In Section 2 we described some of the new stylized facts on growth—they differ from Kaldor's original set. It is this difference, together with the shift in priorities, that accounts for wishing to go beyond the original neoclassical growth model. Neither the newer empirical nor theoretical research has focused on preserving the stability of the "great ratios" or of particular factor prices. Instead, attention has shifted to a more basic set of questions: Why do some countries grow faster than others? What makes some countries prosper while others languish?

Sections 3 and 4 described a number of well-known theoretical growth models and presented their empirical implications. Although a considerable fraction of the empirical work extant has studied growth and convergence equations—whether in cross-section or panel data—we have tried to highlight first, that those equations might be problematic and second, that in any case they need not be the most striking and useful implications of the theory. Distribution-dynamics models make this particularly clear. Appropriate empirical analysis for all the different possibilities we have outlined above is an area that remains under study.

Section 5 described a spectrum of empirical methods and findings related to studying patterns of cross-country growth. The range is extensive and, in our view, continues to grow as researchers understand more about both the facts surrounding growth across countries and the novel difficulties in carrying out empirical analyses in this research area.

At the same time, we feel that the new empirical growth literature remains in its infancy. While the literature has shown that the Solow model has substantial statistical power in explaining cross-country growth variation, sufficiently many problems exist with this work that the causal significance of the model is still far

from clear. Further, the new stylized facts of growth, as embodied in nonlinearities and distributional dynamics have yet to be integrated into full structural econometric analysis. While we find the new empirics of economic growth to be exciting, we also see that much remains to be done.

7. Technical Appendix

This appendix collects together proofs and additional discussion omitted from the main presentation. It is intended to make this paper self-contained, but without straying from the empirical focus in the principal sections.

Single capital good, exogenous technical progress

The classical Cass-Koopmans [23, 65] analysis produces dynamics (9b) from the optimization program (10). To see this, notice that given (1a) and (7a–c),

$$\dot{K}(t) = Y(t) - c(t)N(t) - \delta K(t)$$

can be rewritten as

$$\dot{k} = y - c - (\delta + \nu)k,$$

The original problem (10) can then be analyzed as

$$\begin{aligned} &\max_{\{c(t),k(t)\}_{t\geq 0}} \int_0^\infty U(c(t))e^{-(\rho-\nu)t}\,dt \\ &\text{subject to } \dot{k} = F(k,A) - c - (\delta+\nu)k. \end{aligned}$$

The first-order conditions for this are:

$$\begin{aligned} \dot{c}\,U'' &= (\rho + \delta - \partial F(k,A)/\partial k)\,U', \\ \dot{k} &= F(k,A) - c - (\delta+\nu)k, \\ \lim_{t\to\infty} k(t)e^{-(\rho-\nu)t} &= 0. \end{aligned}$$

Rewrite these in growth rates and then in technology-normalized form; use the parameterized preferences U from (10); and recall that F homogeneous degree 1 means its first partials are all homogeneous degree 0. This yields (9b).

Turn now to convergence. To understand Fig. 2 note that if we define $g(\tilde{k}) \stackrel{\text{def}}{=} f(\tilde{k})\tilde{k}^{-1}$, then on $\tilde{k} > 0$ function g is continuous and strictly decreasing:

$$\begin{aligned}\nabla g(\tilde{k}) &= \nabla f(\tilde{k})\tilde{k}^{-1} - f(\tilde{k})\tilde{k}^{-2} \\ &= \left[\tilde{k}\nabla f(\tilde{k}) - f(\tilde{k})\right]\tilde{k}^{-2} < 0\end{aligned}$$

by concavity and $\lim_{\tilde{k}\to 0} f(\tilde{k}) \geq 0$ from (2). Moreover, $\lim_{\tilde{k}\to 0} g(\tilde{k}) \to \infty$ (directly if $\lim_{\tilde{k}\to 0} f(\tilde{k}) > 0$; by l'Hôpital's Rule and (3) otherwise) and $\lim_{\tilde{k}\to\infty} g(\tilde{k}) = 0$ from (8). These endpoints straddle $(\delta + \nu + \xi)\tau^{-1}$, and therefore the intersection $\tilde{k}^*$ exists, and $\tilde{k}$ satisfying $\tau^{-1}\dot{\tilde{k}}/\tilde{k} = g(\tilde{k}) - (\delta + \nu + \xi)\tau^{-1}$ is dynamically stable everywhere on $\tilde{k} > 0$.

To see that $(\tilde{k}^*, \tilde{c}^*)$, the zero of (16), is well-defined, let $\tilde{k}^*$ solve

$$\nabla f(\tilde{k}) = \rho + \delta + \theta\xi$$

($\tilde{k}^* > 0$) and notice that then

$$\tilde{c}^* \stackrel{\text{def}}{=} \left[f(\tilde{k}^*)/\tilde{k}^* - (\delta + \nu + \xi)\right]\tilde{k}^* > 0$$

since

$$f(\tilde{k}^*)/\tilde{k}^* \geq \nabla f(\tilde{k}^*) = \rho + \delta + \theta\xi > \delta + \nu + \xi$$

from the assumption $\rho > \nu + \xi$.

To see how (18) follows from (17), notice that since M's eigenvalues are distinct and different from zero, we can write its eigenvalue-eigenvector decomposition:

$$M = V_M \Lambda_M V_M^{-1},$$

with V_M full rank and having columns equal to M's right eigenvectors, and

$$\Lambda_M = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Then the unique stable solution of (17) is

$$\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} = \begin{pmatrix} \log \tilde{k}(0) - \log \tilde{k}^* \\ \log \tilde{c}(0) - \log \tilde{c}^* \end{pmatrix} e^{\lambda_2 t},$$

with

$$V_M^{-1} \times \begin{pmatrix} \log \tilde{k}(0) - \log \tilde{k}^* \\ \log \tilde{c}(0) - \log \tilde{c}^* \end{pmatrix} \text{ having 0 as its first entry.}$$

(This proportionality property can always be satisfied since $\tilde{c}(0)$ is free to be determined while $\tilde{k}(0)$ is given as an initial condition.) This timepath constitutes a solution to the differential equation (17) for it implies

$$\begin{aligned} \frac{d}{dt}\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} &= \lambda_2 \times \begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} \\ \Longrightarrow\ V_M^{-1}\frac{d}{dt}\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} &= \lambda_2 \times V_M^{-1}\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} V_M^{-1}\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} \\ \Longrightarrow\ \frac{d}{dt}\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} &= M\begin{pmatrix} \log \tilde{k}(t) - \log \tilde{k}^* \\ \log \tilde{c}(t) - \log \tilde{c}^* \end{pmatrix} \end{aligned}$$

This solution is clearly stable. Since any other solution contains an exponential in λ_1, this solution is also the unique stable one.

Endogenous growth: Asymptotically linear technology

We need to verify that (24) and (25) imply the existence of a balanced-growth equilibrium with positive $\lim_{t\to\infty} \tilde{c}/\tilde{k}$ and $\lim_{t\to\infty} \dot{\tilde{k}}(t)/\tilde{k}(t)$. Along the optimal path $\tilde{c}$ is a function of $\tilde{k}$. Consider conditions (9b) as $\tilde{k} \to \infty$ (since we are interested in equilibria with $\dot{\tilde{k}}(t)/\tilde{k}(t)$ bounded from below by a positive quantity). Then

$$\begin{aligned}
\lim_{\tilde{k}\to\infty} \dot{\tilde{k}}/\tilde{k} &= \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\delta + \nu + \xi) - \lim_{\tilde{k}\to\infty} \tilde{c}/\tilde{k}, \\
\lim_{\tilde{k}\to\infty} \dot{\tilde{c}}/\tilde{c} &= \left(\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - [\rho + \delta + \theta\xi] \right) \theta^{-1},
\end{aligned}$$

using $\lim_{\tilde{k}\to\infty} \nabla f(\tilde{k}) = \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1}$. For these to be equal,

$$\begin{aligned}
\lim_{\tilde{k}\to\infty} \tilde{c}/\tilde{k} &= \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\delta + \nu) - \left[\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho + \delta) \right] \theta^{-1} \\
&> \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\delta + \nu) - (\rho - \nu) \\
&= \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho + \delta) > 0.
\end{aligned}$$

The long-run growth rate is

$$\left(\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - [\rho + \delta + \theta\xi] \right) \theta^{-1},$$

which is positive from (25)

$$\begin{aligned}
\theta &< \frac{\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho + \delta)}{\xi} \\
\Longrightarrow\ 0 &< \lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho + \delta + \theta\xi).
\end{aligned}$$

Finally, along such balanced-growth paths we have $\lim_{t\to\infty} \tilde{k}(t)e^{-(\rho-\nu-\xi)t} = 0$ since

$$\theta > \frac{\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho+\delta)}{\rho-\nu}$$

$$\Longrightarrow \rho-\nu > \left[\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho+\delta)\right]\theta^{-1}$$

$$\Longrightarrow \rho-\nu-\xi > \left[\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} - (\rho+\delta+\theta\xi)\right]\theta^{-1}.$$

If θ is too large (exceeding the upper bound in (25)) then this model collapses to the traditional neoclassical model where balanced-growth equilibrium has finite $(\tilde{y}^*, \tilde{k}^*)$, and neither preference nor technology parameters (apart from ξ) influences the long-run growth rate.

Distribution dynamics

Rigorous expositions of the mathematics underlying a formulation like (35) can be found in Chung [26], Doob [34], Futia [49], and Stokey and Lucas (with Prescott) [109].[27] Since we are concerned here with real-valued incomes, the underlying state space is the pair $(\mathbb{R}, \mathfrak{R})$, i.e., the real line $\mathbb{R}$ together with the collection $\mathfrak{R}$ of its Borel sets. Let $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ denote the Banach space of bounded finitely-additive set functions on the measurable space $(\mathbb{R}, \mathfrak{R})$ endowed with total variation norm:

$$\forall\varphi \text{ in } \mathbf{B}(\mathbb{R}, \mathfrak{R}): \qquad |\varphi| = \sup \sum_j |\varphi(A_j)|,$$

[27] Economic applications of these tools have also appeared in stochastic growth models (e.g., the examples in Stokey and Lucas (with Prescott) [109, Ch. 16]), income distribution dynamics (e.g., Loury [73]), and elsewhere. Using these ideas for studying distribution dynamics rather than analyzing a time-series stochastic process, say, exploits a duality in the mathematics. This is made explicit in Quah [91], a study dealing not with cross-country growth but business cycles instead.

where the supremum in this definition is taken over all $\{A_j : j = 1, 2, \ldots, n\}$ finite measurable partitions of $\mathbb{R}$.

Empirical distributions on $\mathbb{R}$ can be identified with probability measures on $(\mathbb{R}, \mathfrak{R})$; those are, in turn, just countably-additive elements in $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ assigning value 1 to the entire space $\mathbb{R}$. Let $\mathfrak{B}$ denote the Borel σ-algebra generated by the open subsets (relative to total variation norm topology) of $\mathbf{B}(\mathbb{R}, \mathfrak{R})$. Then $(\mathbf{B}, \mathfrak{B})$ is another measurable space.

Note that $\mathbf{B}$ includes more than just probability measures: an arbitrary element φ in $\mathbf{B}$ could be negative; $\varphi(\mathbb{R})$ need not be 1; and φ need not be countably-additive. On the other hand, a collection of probability measures is never a linear space: that collection does not include a zero element; if ϕ_1 and ϕ_2 are probability measures, then $\phi_1 - \phi_2$ and $\phi_1 + \phi_2$ are not; neither is $x\phi_1$ a probability measure for $x \in \mathbb{R}$ except at $x = 1$. By contrast, the set of bounded finitely-additive set functions certainly is a linear space and, as described above, is easily given a norm and then made Banach.

Why embed probability measures in a Banach space as we have done here? A first reason is so that distances can be defined between probability measures; it then makes sense to talk about two measures—and their associated distributions—getting closer to one another. A small step from there is to define open sets of probability measures, and thereby induce (Borel) σ-algebras on probability measures. Such σ-algebras then allow modelling random elements drawn from collections of probability measures, and thus from collections of distributions. The data of interest when modelling the dynamics of distributions are precisely random elements taking values that are probability measures.

In this scheme, then each ϕ_t associated with the observed cross-sectional income distribution F_t is a measure in $(\mathbf{B}, \mathfrak{B})$. If $(\Omega, \mathfrak{F}, \Pr)$ is the underlying probability space, then ϕ_t is the value of an $\mathfrak{F}/\mathfrak{B}$-measurable map $\Phi_t : (\Omega, \mathfrak{F}) \to (\mathbf{B}, \mathfrak{B})$. The sequence $\{\Phi_t : t \geq 0\}$ is then a $\mathbf{B}$-valued stochastic process.

To understand the structure of operators like $T^*_{u_t}$, it helps to use the following:

STOCHASTIC KERNEL DEFINITION: *Let φ and ψ be elements of* $\mathbf{B}$

that are probability measures on $(\mathbb{R}, \mathfrak{R})$. *A* **stochastic kernel** *relating* φ *and* ψ *is a mapping* $M_{(\varphi,\psi)} : (\mathbb{R}, \mathfrak{R}) \to [0,1]$ *satisfying:*

(i) $\forall y$ *in* $\mathbb{R}$, *the restriction* $M_{(\varphi,\psi)}(y, \cdot)$ *is a probability measure;*
(ii) $\forall A$ *in* $\mathfrak{R}$, *the restriction* $M_{(\varphi,\psi)}(\cdot, A)$ *is* $\mathfrak{R}$*-measurable;*
(iii) $\forall A$ *in* $\mathfrak{R}$, *we have* $\varphi(A) = \int M_{(\varphi,\psi)}(y, A)\, d\psi(y)$.

To see why this is useful, first consider (iii). At an initial point in time, for given y, there is some fraction $d\psi(y)$ of economies with incomes close to y. Count up all economies in that group who turn out to have their incomes subsequently fall in a given $\mathfrak{R}$-measurable subset $A \subseteq \mathbb{R}$. When normalized to be a fraction of the total number of economies, this count is precisely $M(y, A)$ (where the (φ, ψ) subscript can now be deleted without loss of clarity). Fix A, weight the count $M(y, A)$ by $d\psi(y)$, and sum over all possible y's, i.e., evaluate the integral $\int M(y, A)\, d\psi(y)$. This gives the fraction of economies that end up in state A regardless of their initial income levels. If this equals $\varphi(A)$ for all measurable subsets A, then φ must be the measure associated with the subsequent income distribution. In other words, the stochastic kernel M is a complete description of transitions from state y to any other portion of the underlying state space $\mathbb{R}$.

Conditions (i) and (ii) simply guarantee that the interpretation of (iii) is valid. By (ii), the right hand side of (iii) is well-defined as a Lebesgue integral. By (i), the right hand side of (iii) is a weighted average of probability measures $M(y, \cdot)$, and thus is itself a probability measure.

How does this relate to the structure of $T^*_{u_t}$? Let $\mathbf{b}(\mathbb{R}, \mathfrak{R})$ be the Banach space under sup norm of bounded measurable functions on $(\mathbb{R}, \mathfrak{R})$. Fix a stochastic kernel M and define the operator T mapping $\mathbf{b}(\mathbb{R}, \mathfrak{R})$ to itself by

$$\forall f \text{ in } \mathbf{b}(\mathbb{R}, \mathfrak{R}),\ \forall y \text{ in } \mathbb{R}: \qquad (Tf)(y) = \int f(x) M(y, dx).$$

Since $M(y, \cdot)$ is a probability measure, the image Tf can be interpreted as a forward conditional expectation. For example, if all economies in the cross section

begin with incomes y, and we take f to be the identity map, then $(Tf)(y) = \int xM(y, dx)$ is next period's average income in the cross section, conditional on all economies having income y in the current period.

Clearly, T is a bounded linear operator. Denote the adjoint of T by T^*. By Riesz Representation Theorem, the dual space of $\mathbf{b}(\mathbb{R}, \mathfrak{R})$ is just $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ (our original collection of bounded finitely additive set functions on $\mathfrak{R}$); thus T^* is a bounded linear operator mapping $\mathbf{B}(\mathbb{R}, \mathfrak{R})$ to itself. It turns out that T^* is also exactly the mapping in (iii) of the Stochastic Kernel Definition, i.e.,

$$\forall \psi \text{ probability measures in } \mathbf{B},\ \forall A \text{ in } \mathfrak{R}: \qquad (T^*\psi)(A) = \int M(y, A)\, d\psi(y).$$

(This is immediate from writing the left-hand side as

$$\begin{aligned}
(T^*\psi)(A) &= \int \mathbf{1}_A\, d(T^*\psi)(y) = \int (T\mathbf{1}_A)(y)\, d\psi(y) && \text{(adjoint)} \\
&= \int \left[\int \mathbf{1}_A(x) M(y, dx) \right] d\psi(y) && \text{(definition of } T\text{)} \\
&= \int M(y, A)\, d\psi(y), && \text{(calculation)}
\end{aligned}$$

with $\mathbf{1}_A$ the indicator function for A.)

8. Data Appendix

The data used in Section 2 are from version V.6 of Summers and Heston [110]. Income is taken to be real GDP per capita in constant dollars using Chain Index (at 1985 international prices) (series `RGDPCH`). Economies not having data in 1960 and 1989 were excluded. The remaining sample comprised 122 economies (integers immediately before country names are the indexes in the Summers-Heston database):

1	(1)	Algeria	2	(2)	Angola
3	(3)	Benin	4	(4)	Botswana
5	(5)	Burkina Faso	6	(6)	Burundi
7	(7)	Cameroon	8	(8)	Cape Verde Islands
9	(9)	Central African Republic	10	(10)	Chad
11	(11)	Comoros	12	(12)	Congo
13	(14)	Egypt	14	(16)	Gabon
15	(17)	Gambia	16	(18)	Ghana
17	(19)	Guinea	18	(20)	Guinea Bissau
19	(21)	Ivory Coast	20	(22)	Kenya
21	(23)	Lesotho	22	(25)	Madagascar
23	(26)	Malawi	24	(27)	Mali
25	(28)	Mauritania	26	(29)	Mauritius
27	(30)	Morocco	28	(31)	Mozambique
29	(32)	Namibia	30	(33)	Niger
31	(34)	Nigeria	32	(35)	Reunion
33	(36)	Rwanda	34	(37)	Senegal
35	(38)	Seychelles	36	(39)	Sierra Leone
37	(40)	Somalia	38	(41)	South Africa
39	(43)	Swaziland	40	(44)	Tanzania
41	(45)	Togo	42	(46)	Tunisia
43	(47)	Uganda	44	(48)	Zaire
45	(49)	Zambia	46	(50)	Zimbabwe
47	(52)	Barbados	48	(54)	Canada

49	(55)	Costa Rica	50	(57)	Dominican Republic
51	(58)	El Salvador	52	(60)	Guatemala
53	(61)	Haiti	54	(62)	Honduras
55	(63)	Jamaica	56	(64)	Mexico
57	(65)	Nicaragua	58	(66)	Panama
59	(67)	Puerto Rico	60	(71)	Trinidad and Tobago
61	(72)	USA	62	(73)	Argentina
63	(74)	Bolivia	64	(75)	Brazil
65	(76)	Chile	66	(77)	Colombia
67	(78)	Ecuador	68	(79)	Guyana
69	(80)	Paraguay	70	(81)	Peru
71	(82)	Suriname	72	(83)	Uruguay
73	(84)	Venezuela	74	(86)	Bangladesh
75	(88)	China	76	(89)	Hong Kong
77	(90)	India	78	(91)	Indonesia
79	(92)	Iran	80	(94)	Israel
81	(95)	Japan	82	(96)	Jordan
83	(97)	Korean Republic	84	(100)	Malaysia
85	(102)	Myanmar	86	(105)	Pakistan
87	(106)	Philippines	88	(108)	Saudi Arabia
89	(109)	Singapore	90	(110)	Sri Lanka
91	(111)	Syria	92	(112)	Taiwan
93	(113)	Thailand	94	(116)	Austria
95	(117)	Belgium	96	(119)	Cyprus
97	(120)	Czechoslovakia	98	(121)	Denmark
99	(122)	Finland	100	(123)	France
101	(125)	Germany, West	102	(126)	Greece
103	(128)	Iceland	104	(129)	Ireland
105	(130)	Italy	106	(131)	Luxembourg
107	(132)	Malta	108	(133)	Netherlands
109	(134)	Norway	110	(136)	Portugal
111	(137)	Romania	112	(138)	Spain
113	(139)	Sweden	114	(140)	Switzerland

115	(141)	Turkey	116	(142)	UK
117	(143)	USSR	118	(144)	Yugoslavia
119	(145)	Australia	120	(146)	Fiji
121	(147)	New Zealand	122	(148)	Papua New Guinea

The clustering-classification results described in Section 5 derive from the following subsample split (taken from Durlauf and Johnson [39, Table IV]):

1. $y_j(1960) < \$800$: Burkina Faso, Burundi, Ethiopia, Malawi, Mali, Mauritania, Niger, Rwanda, Sierra Leone, Tanzania, Togo, Uganda;
2. $\$800 \leq y_j(1960) \leq \$4850$ and $LR_j(1960) < 46\%$: Algeria, Angola, Benin, Cameroon, Central African Republic, Chad, Congo (People's Republic), Egypt, Ghana, Ivory Coast, Kenya, Liberia, Morocco, Mozambique, Nigeria, Senegal, Somalia, Sudan, Tunisia, Zambia, Zimbabwe, Bangladesh, India, Jordan, Nepal, Pakistan, Syria, Turkey, Guatemala, Haiti, Honduras, Bolivia, Indonesia, Papua New Guinea;
3. $\$800 \leq y_j(1960) \leq \$4850$ and $46\% \leq LR_j(1960)$: Madagascar, South Africa, Hong Kong, Israel, Japan, Korea, Malaysia, Philippines, Singapore, Sri Lanka, Thailand, Greece, Ireland, Portugal, Spain, Costa Rica, Dominican Republic, El Salvador, Jamaica, Mexico, Nicaragua, Panama, Brazil, Colombia, Ecuador, Paraguay, Peru;
4. $\$4850 < y_j(1960)$: Austria, Belgium, Denmark, Finland, France, Germany (Federal Republic), Italy, Netherlands, Norway, Sweden, Switzerland, UK, Canada, Trinidad and Tobago, USA, Argentina, Chile, Uruguay, Venezuela, Australia, New Zealand.

References

[1] Aghion, Phillipe and Howitt, Peter. (1992), "A Model of Growth Through Creative Destruction", *Econometrica*, 60(2):323–351, March.

[2] Alesina, Alberto, Ozler, Sule, Roubini, Nouriel, and Swagel, Phillip. (1996), "Political Instability and Economic Growth", *Journal of Economic Growth*, 1(2):189–211, June.

[3] Alesina, Alberto and Rodrik, Dani. (1994), "Distributive Politics and Economic Growth", *Quarterly Journal of Economics*, 109(2):465–490, May.

[4] Azariadis, Costas and Drazen, Allan. (1990), "Threshold Externalities in Economic Development", *Quarterly Journal of Economics*, 105(2):501–526, May.

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

[6] Barro, Robert J. (1996), "Democracy and Growth", *Journal of Economic Growth*, 1(1):1–27, March.

[7] Barro, Robert J. (1997), *Determinants of Economic Growth*, MIT Press, Cambridge.

[8] Barro, Robert J. and Lee, Jong-Wha. (1994), "Sources of Economic Growth", *Carnegie-Rochester Conference Series on Public Policy*, 40:1–57.

[9] Barro, Robert J. and Sala-i-Martin, Xavier. (1991), "Convergence Across States and Regions", *Brookings Papers on Economic Activity*, 1:107–182, April.

[10] Barro, Robert J. and Sala-i-Martin, Xavier. (1992), "Convergence", *Journal of Political Economy*, 100(2):223–251, April.

[11] Barro, Robert J. and Sala-i-Martin, Xavier. (1995), *Economic Growth*, McGraw-Hill.

[12] Baumol, William J. (1986), "Productivity Growth, Convergence, and Welfare", *American Economic Review*, 76(5):1072–85, December.

[13] Ben-David, Dan. (1996), "Trade and Convergence Among Countries", *Journal of International Economics*, 40(3/4):279–298.

[14] Bénabou, Roland. (1993), "Workings of a City: Location, Education, and Production", *Quarterly Journal of Economics*, 108(3):619–52, August.

[15] Benhabib, Jess and Spiegel, Mark M. (1997), "Cross-Country Growth Regressions", Working Paper 97–20, CV Starr Center, New York University.

[16] Bernard, Andrew B. and Durlauf, Steven N. (1995), "Convergence in International Output", *Journal of Applied Econometrics*, 10(2):97–108, April.

[17] Bernard, Andrew B. and Durlauf, Steven N. (1996), "Interpreting Tests of the Convergence Hypothesis", *Journal of Econometrics*, 71(1–2):161–174.

[18] Bianchi, Marco. (1997), "Testing for Convergence: Evidence from Non-Parametric Multimodality Tests", *Journal of Applied Econometrics*, 12(4):393–409, July.

[19] Blomstrom, Magnus, Lipsey, Robert E., and Zejan, Mario. (1996), "Is Fixed Investment the Key to Economic Growth?", *Quarterly Journal of Economics*, 111(1):269–276, February.

[20] Breiman, Leo, Friedman, Jerome H., Olshen, Richard A., and Stone, Charles J. (1984), *Classification and Regression Trees*, Chapman and Hall, New York.

[21] Canova, Fabio and Marcet, Albert. (1995), "The Poor Stay Poor: Non-Convergence across Countries and Regions", Discussion Paper 1265, CEPR, November.

[22] Caselli, Francesco, Esquivel, Gerardo, and Lefort, Fernando. (1996), "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics", *Journal of Economic Growth*, 1(3):363–389, September.

[23] Cass, David. (1965), "Optimal Growth in an Aggregate Model of Capital Accumulation", *Review of Economic Studies*, 32:233–240, July.

[24] Chamberlain, Gary. (1984), "Panel Data", in Griliches, Zvi and Intriligator, Michael D., (eds.), *Handbook of Econometrics vol. II*, chapter 22, pages 1247–1318. Elsevier North-Holland, Amsterdam.

[25] Cho, Dongchul. (1996), "An Alternative Interpretation of Conditional Convergence Results", *Journal of Money, Credit, and Banking*, 28(4):669–681.

[26] Chung, Kai Lai. (1960), *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, Berlin.

[27] Coe, David T. and Helpman, Elhanan. (1995), "International R&D Spillovers", *European Economic Review*, 39(5):859–887, May.

[28] Cohen, Daniel. (1996), "Tests of the 'Convergence Hypothesis': Some Further Results", *Journal of Economic Growth*, 1(3):351–362, September.

[29] DeLong, J. Bradford. (1988), "Productivity Growth, Convergence, and Welfare: A Comment", *American Economic Review*, 78(5):1138–55, December.

[30] DeLong, J. Bradford and Summers, Lawrence H. (1993), "How Strongly Do Developing Economies Benefit from Equipment Investment", *Journal of Monetary Economics*, 32(3):395–415.

[31] den Haan, Wouter J. (1995), "Convergence in Stochastic Growth Models: The Importance of Understanding Why Income Levels Differ", *Journal of Monetary Economics*, 35(1):65–82, February.

[32] Denison, Edward F. (1974), *Accounting for United States Growth, 1929–1969*, The Brookings Institution, Washington DC.

[33] Desdoigts, Alain. *Changes in the World Income Distribution: A Non-Parametric Approach to Challenge the Neoclassical Convergence Argument.* PhD thesis, European University Institute, Florence, June 1994.

[34] Doob, Joseph L. (1953), *Stochastic Processes*, John Wiley, New York NY.

[35] Duffy, J. and Papageorgiou, C. (1997), "The Specification of the Aggregate Production Function: A Cross-Country Empirical Investigation", Working paper, University of Pittsburgh.

[36] Durlauf, Steven N. (1989), "Output Persistence, Economic Structure, and the Choice of Stabilization Policy", *Brookings Papers on Economic Activity*, 2:69–116.

[37] Durlauf, Steven N. (1993), "Nonergodic Economic Growth", *Review of Economic Studies*, 60(2):349–366, April.

[38] Durlauf, Steven N. (1996), "A Theory of Persistent Income Inequality", *Journal of Economic Growth*, 1(1):75–93, March.

[39] Durlauf, Steven N. and Johnson, Paul A. (1995), "Multiple Regimes and Cross-Country Growth Behavior", *Journal of Applied Econometrics*, 10(4):365–384, October.

[40] Easterly, William. (1993), "How Much Do Distortions Affect Growth?", *Journal of Monetary Economics*, 32(2):187–212.

[41] Easterly, William, Kremer, Michael, Pritchett, Lant, and Summers, Lawrence H. (1993), "Good Policy or Good Luck? Country Growth Performance and Temporary Shocks", *Journal of Monetary Economics*, 32(3):459–483.

[42] Esteban, Joan-María and Ray, Debraj. (1994), "On the Measurement of Polarization", *Econometrica*, 62(4):819–851, July.

[43] Evans, Paul. (1998), "Using Panel Data to Evaluate Growth Theories", *International Economic Review*, May. Forthcoming.

[44] Forbes, Kristin. (1997), "Back to the Basics: The Positive Effect of Inequality on Growth", Working paper, MIT, Cambridge.

[45] Frankel, Jeffrey A. and Romer, David. (1996), "Trade and Growth: An Empirical Investigation", Working Paper 5476, NBER, Cambridge, June.

[46] Frankel, Jeffrey A., Romer, David, and Cyrus, Teresa. (1996), "Trade and Growth in East Asian Countries: Cause and Effect?", Working Paper 5732, NBER, Cambridge, June.

[47] Franses, Philip Hans and Hobijn, Bart. (1995), "Convergence of Living Standards: An International Analysis", Technical Report 9534/A, Econometric Institute, Erasmus University Rotterdam, September.

[48] Friedman, Milton. (1992), "Do Old Fallacies Ever Die?", *Journal of Economic Literature*, 30(4):2129–2132, December.

[49] Futia, Carl. (1982), "Invariant Distributions and the Limiting Behavior of Markovian Economic Models", *Econometrica*, 50(1):377–408, January.

[50] Galor, Oded. (1996), "Convergence? Inferences from Theoretical Models", *Economic Journal*, 106(437):1056–1069, July.

[51] Galor, Oded and Zeira, Joseph. (1993), "Income Distribution and Macroeconomics", *Review of Economic Studies*, 60(1):35–52, January.

[52] Grier, Kevin B. and Tullock, Gordon. (1989), "An Empirical Analysis of Cross-National Economic Growth, 1951–80", *Journal of Monetary Economics*, 24(2):259–276, November.

[53] Grossman, Gene M. and Helpman, Elhanan. (1991), *Innovation and Growth in the Global Economy*, MIT Press.

[54] Harrison, Ann. (1995), "Openness and Growth: A Time-Series, Cross-Country Analysis for Developing Countries", Working Paper 5221, NBER, Cambridge.

[55] Islam, Nazrul. (1995), "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*, 110(443):1127–1170, November.

[56] Jones, Charles I. (1995), "R&D-based Models of Economic Growth", *Journal of Political Economy*, 103(3):759–784, August.

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

[58] Jones, Charles I. (1997), "On the Evolution of the World Income Distribution", *Journal of Economic Perspectives*, 11(3):19–36, Summer.

[59] Jones, Larry E. and Manuelli, Rodolfo. (1990), "A Convex Model of Equilibrium Growth: Theory and Implications", *Journal of Political Economy*, 98(5, part 1):1008–1038, October.

[60] Kaldor, Nicholas. (1963), "Capital Accumulation and Economic Growth", in Lutz, Friedrich A. and Hague, Douglas C., (eds.), *Proceedings of a Conference Held by the International Economics Association*. Macmillan, London.

[61] Kelly, Morgan. (1992), "On Endogenous Growth with Productivity Shocks", *Journal of Monetary Economics*, 30(1):47–56, October.

[62] King, Robert G. and Levine, Ross. (1993), "Finance and Growth: Schumpeter Might Be Right", *Quarterly Journal of Economics*, 108(3):717–737.

[63] Knowles, Stephen and Owen, P. Dorian. (1995), "Health Capital and Cross-Country Variation in Income per Capita in the Mankiw-Romer-Weil Model", *Economics Letters*, 48(1):99–106.

[64] Kocherlakota, Narayana R. and Yi, Kei-Mu. (1995), "Can Convergence Regressions Distinguish Between Exogenous and Endogenous Growth Models?", *Economics Letters*, 49:211–215.

[65] Koopmans, Tjalling C. (1965), "On the Concept of Optimal Economic Growth", in *The Econometric Approach to Development Planning*. North Holland, Amsterdam.

[66] Kormendi, Roger and Mequire, P. (1985), "Macroeconomic Determinants of Growth: Cross-Country Evidence", *Journal of Monetary Economics*, 16(2):141–163.

[67] Lamo, Ana R. *Cross-section Distribution Dynamics*. PhD thesis, LSE, February 1996.

[68] Leamer, Edward E. (1978), *Specification Searches: Ad Hoc Inference from Non-Experimental Data*, John Wiley, New York.

[69] Lee, Kevin, Pesaran, M. Hashem, and Smith, Ron P. (1997), "Growth and Convergence in a Multi-Country Empirical Stochastic Solow Model", *Journal of Applied Econometrics*, 12(4):357–392, July.

[70] Leung, Charles and Quah, Danny. (1996), "Convergence, Endogenous Growth, and Productivity Disturbances", *Journal of Monetary Economics*, 38(3):535–547, December.

[71] Levin, Andrew and Lin, Chien-fu. (1992), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", Working paper, Economics Department, UCSD, San Diego.

[72] Levine, Ross and Renelt, David. (1992), "A Sensitivity Analysis of Cross-Country Growth Regressions", *American Economic Review*, 82(4):942–963, September.

[73] Loury, Glenn C. (1981), "Intergenerational Transfers and the Distribution of Earnings", *Econometrica*, 49(4):843–867, July.

[74] Lucas Jr., Robert E. (1988), "On The Mechanics of Economic Development", *Journal of Monetary Economics*, 22(3):3–42, June.

[75] Lucas Jr., Robert E. (1993), "Making a Miracle", *Econometrica*, 61(2):251–271, March.

[76] Maddala, G. S. (1988), *Introduction to Econometrics*, Macmillan, New York.

[77] Maddison, Angus. (1989), *The World Economy in the 20th Century*, Development Centre of the OECD, Paris.

[78] Mankiw, N. Gregory, Romer, David, and Weil, David N. (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics*, 107(2):407–437, May.

[79] Mauro, Paolo. (1995), "Corruption and Growth", *Quarterly Journal of Economics*, 110(3):681–713, August.

[80] Murphy, Kevin M., Shleifer, Andrei, and Vishny, Robert W. (1989), "Industrialization and the Big Push", *Journal of Political Economy*, 97(4):1003–1026, October.

[81] Murphy, Kevin M., Shleifer, Andrei, and Vishny, Robert W. (1991), "The Allocation of Talent: Implications for Growth", *Quarterly Journal of Economics*, 106(2):503–530, May.

[82] Nelson, Charles R. and Plosser, Charles I. (1982), "Trends and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, 10:129–162, September.

[83] Nerlove, Marc. (1996), "Growth Rate Convergence, Fact or Artifact?", Working paper, University of Maryland, June.

[84] Perron, Pierre. (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis", *Econometrica*, 57(6):1361–1401, November.

[85] Persson, Torsten and Tabellini, Guido. (1994), "Is Inequality Harmful for Growth?", *American Economic Review*, 84(3):600–621, June.

[86] Pritchett, Lant. (1997), "Divergence, Big Time", *Journal of Economic Perspectives*, 11(3):3–17, Summer.

[87] Quah, Danny. (1992), "International Patterns of Growth: I. Persistence in Cross-Country Disparities", Working paper, LSE, London, October.

[88] Quah, Danny. (1993), "Empirical Cross-Section Dynamics in Economic Growth", *European Economic Review*, 37(2/3):426–434, April.

[89] Quah, Danny. (1993), "Galton's Fallacy and Tests of the Convergence Hypothesis", *The Scandinavian Journal of Economics*, 95(4):427–443, December.

[90] Quah, Danny. (1994), "Exploiting Cross Section Variation for Unit Root Inference in Dynamic Data", *Economics Letters*, 44(1):9–19, January.

[91] Quah, Danny. (1996), "Aggregate and Regional Disaggregate Fluctuations", *Empirical Economics*, 21(1):137–159, March.

[92] Quah, Danny. (1996), "Convergence Empirics Across Economies with (Some) Capital Mobility", *Journal of Economic Growth*, 1(1):95–124, March.

[93] Quah, Danny. (1996), "Empirics for Economic Growth and Convergence", *European Economic Review*, 40(6):1353–1375, June.

[94] Quah, Danny. (1997), "Empirics for Growth and Distribution: Polarization, Stratification, and Convergence Clubs", *Journal of Economic Growth*, 2(1):27–59, March.

[95] Ramey, Garey and Ramey, Valerie A. (1995), "Cross-Country Evidence on the Link Between Volatility and Growth", *American Economic Review*, 85(5):1138–1151, December.

[96] Rebelo, Sergio. (1991), "Long-Run Policy Analysis and Long-Run Growth", *Journal of Political Economy*, 99(3):500–521, June.

[97] Romer, David. (1996), *Advanced Macroeconomics*, McGraw-Hill, New York NY.

[98] Romer, Paul M. (1986), "Increasing Returns and Long-run Growth", *Journal of Political Economy*, 94(5):1002–1037, October.

[99] Romer, Paul M. (1990), "Endogenous Technological Change", *Journal of Political Economy*, 98(5, part 2):S71–S102, October.

[100] Romer, Paul M. (1993), "Idea Gaps and Object Gaps in Economic Development", *Journal of Monetary Economics*, 32(3):543–574, December.

[101] Romer, Paul M. (1994), "The Origins of Endogenous Growth", *Journal of Economic Perspectives*, 8(1):3–22, Winter.

[102] Sachs, Jeffrey D. and Warner, Andrew. (1995), "Economic Reform and the Process of Global Integration", *Brookings Papers on Economic Activity*, (1):1–95, August.

[103] Sala-i-Martin, Xavier X. (1996), "Regional Cohesion: Evidence and Theories of Regional Growth and Convergence", *European Economic Review*, 40(6):1325–1352, June.

[104] Sala-i-Martin, Xavier X. (1997), "I Just Ran Two Million Regressions", *American Economic Association Papers and Proceedings*, 87(2):178–183, May.

[105] Singer, Burton and Spilerman, Seymour. (1976), "Some Methodological Issues in the Analysis of Longtitudinal Surveys", *Annals of Economic and Social Measurement*, 5:447–474.

[106] Solow, Robert M. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70(1):65–94, February.

[107] Solow, Robert M. (1957), "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics*, 39:312–320, August.

[108] Stigler, Stephen M. (1986), *The History of Statistics: The Measurement of Uncertainty before 1900*, Belknap Press, Harvard University, Cambridge.

[109] Stokey, Nancy L. and (with Edward C. Prescott), Robert E. Lucas Jr. (1989), *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge MA.

[110] Summers, Robert and Heston, Alan. (1991), "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988", *Quarterly Journal of Economics*, 106(2):327–368, May.

[111] Swan, Trevor W. (1956), "Economic Growth and Capital Accumulation", *Economic Record*, 32:334–361, November.

[112] Temple, Jonathan. *Essays on Economic Growth*. PhD thesis, Oxford, 1996.

Table 1: Cross section regressions
Initial output and literacy-based sample breaks
Dependent variable: $\log y_j(1985) - \log y_j(1960)$

	MRW	$y_j(1960) < 1950$ and $LR_j(1960) < 54\%$	$1950 < y_j(1960)$ and $54\% \leq LR_j(1960)$
Observations	98	42	42
		Unconstrained regressions	
$\log y_j(1960)$	$-0.29^{\dagger}$	$-0.44^{\dagger}$	$-0.43^{\dagger}$
	(0.06)	(0.16)	(0.08)
$\log(\delta + \nu_j + \xi)$	-0.50	-0.38	-0.54
	(0.29)	(0.47)	(0.28)
$\log \tau_{p,j}$	$0.52^{\dagger}$	$0.31^{\dagger}$	$0.69^{\dagger}$
	(0.09)	(0.11)	(0.17)
$\log \tau_{h,j}$	$0.23^{\dagger}$	$0.21^{\dagger}$	0.11
	(0.06)	(0.09)	(0.16)
$\bar{R}^2$	0.46	0.27	0.48
		Constrained regressions	
α_p	$0.43^{\dagger}$	$0.28^{\dagger}$	$0.51^{\dagger}$
α_h	$0.24^{\dagger}$	$0.22^{\dagger}$	0.11
$\bar{R}^2$	0.42	0.28	0.50

Table reports a selection of results from Durlauf and Johnson [39, Table 2], with the notation changed to match this Chapter's. Symbol † denotes significance at 5% asymptotic level. *Constrained regressions* indicate estimation imposing $\lambda = -(1 - \alpha_p - \alpha_h)(\delta + \nu_j + \xi)$, with λ further restricted from the coefficient on the initial condition $\log y_j(1960)$. The original MRW paper never reported results using such a restriction, and thus the *MRW* column is from Durlauf and Johnson [39].

Table 2: Growth regression compilation (end of table for Notes)

Explanatory variables		Reference	Finding
Change in labor force participation rate		Blomstrom, Lipsey, and Zejan [19]	$+^*$
Corruption		Mauro [79]	$-^*$
Capitalism (level)		Sala-i-Martin [104]	$+^*$
Democracy	Some	Barro [6, 7]	$+^*$
	More	Barro [6, 7]	$-^*$
	Overall	Alesina, Ozler, Roubini, and Swagel [2]	?
Domestic credit	Growth rate	Levine and Renelt [72]	$+^f$
	Volatility of growth rate	Levine and Renelt [72]	$+^f$
Education	College level	Barro and Lee [8]	$-$
	Female	Barro and Lee [8] Barro [6] Barro [7] Caselli, Esquivel, and Lefort [22] Forbes [44]	$-^*$ $-^*$ $-$ $+^*$ $-^*$
	Female growth	Barro and Lee [8]	$-^*$
	Male	Barro and Lee [8] Caselli, Esquivel, and Lefort [22] Forbes [44]	$+^*$ $-^*$ $+^*$
	Male growth	Barro and Lee [8]	$+^*$
	Overall	Barro [5] Knowles and Owen [63] Levine and Renelt [72] Mankiw, Romer, and Weil [78]	$+^*$ $+$ $+^f$ $+^*$
	Primary	Barro [7]	$-$

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Exchange rates (real)	Black market premium	Barro [6]	$-^{*}$
		Barro and Lee [8]	$-^{*}$
		Easterly [40]	$-$
		Harrison [54]	$-^{*}$
		Levine and Renelt [72]	$-^{f}$
		Sala-i-Martin [104]	$-^{*}$
	Distortions	Easterly [40]	$-$
		Harrison [54]	$-$
		Sala-i-Martin [104]	$-^{*}$
	Terms of trade improvement	Barro [6, 7]	$+^{*}$
		Barro and Lee [8]	$+$
		Caselli, Esquivel, and Lefort [22]	$+^{*}$
		Easterly, Kremer, Pritchett, and Summers [41]	$+^{*}$
External debt (dummy)		Easterly, Kremer, Pritchett, and Summers [41]	$-$
Fertility		Barro [5, 6, 7]	$-^{*}$
		Barro and Lee [8]	$-^{*}$
Financial repression		Easterly [40]	$-^{*}$
Financial sophistication		King and Levine [62]	$+^{*}$
Fraction college students	Engineering	Murphy, Shleifer,	$+^{*}$
	Law	and Vishny [81]	$+^{*}$
Government	Consumption	Barro [5, 6, 7]	$-^{*}$
		Barro and Lee [8]	$-^{*}$
		Caselli, Esquivel, and Lefort [22]	$+^{*}$
	Growth in Consumption	Kormendi and Meguire [66]	$+$
	Deficits	Levine and Renelt [72]	$-^{f}$
	Investment	Barro [5]	$+$

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Growth rates	G7	Alesina, Ozler, Roubini, and Swagel [2]	$+^{*}$
	G7 lagged	Alesina, Ozler, Roubini, and Swagel [2]	$+$
		Easterly, Kremer, Pritchett, and Summers [41]	$+$
Health (various proxies)		Barro [7]	$+^{*}$
		Barro and Lee [8]	$+^{*}$
		Caselli, Esquivel, and Lefort [22]	$-^{*}$
		Knowles and Owen [63]	$+^{*}$
Inequality	Democracies	Persson and Tabellini [85]	$-^{*}$
	Non-democracies	Persson and Tabellini [85]	$+^{*}$
	Overall	Alesina and Rodrik [3]	$-^{*}$
		Forbes [44]	$+^{*}$
Inflation	Change	Kormendi and Meguire [66]	$-$
	Level (above 15%)	Barro [7]	$-^{*}$
	Level	Levine and Renelt [72]	$-^{f}$
	Variability	Barro [7]	$+$
		Levine and Renelt [72]	$-^{f}$

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Initial income		Barro [5, 7]	$-^{*}$
		Barro and Lee [8]	$-^{*}$
		Barro and Sala-i-Martin [10]	$-^{*}$
		Ben-David [13]	$-^{*}$
		Caselli, Esquivel, and Lefort [22]	$-^{*}$
		Cho [25]	$+^{*}$
		Kormendi and Meguire [66]	$-^{*}$
		Levine and Renelt [72]	$-^{r}$
		Mankiw, Romer, and Weil [78]	$-^{*}$
		Romer [100]	$-^{*}$
Initial income (interacted with male schooling)		Barro [7]	$-^{*}$
Investment ratio		Barro [5]	$+^{*}$
		Barro [6, 7]	$+$
		Barro and Lee [8]	$+^{*}$
		Caselli, Esquivel, and Lefort [22]	$+^{*}$
		Levine and Renelt [72]	$+^{r}$
		Mankiw, Romer, and Weil [78]	$+^{*}$
Investment	Equipment, Fixed Capital	Blomstrom, Lipsey, and Zejan [19]	$-$
		De Long and Summers [30]	$+^{*}$
		Sala-i-Martin [104]	$+^{*}$
	Non-equipment	Sala-i-Martin [104]	$+^{*}$
Latitude (absolute)		Sala-i-Martin [104]	$+^{*}$
Mining (fraction of GDP)		Sala-i-Martin [104]	$+^{*}$
Money growth		Kormendi and Meguire [66]	$+$

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Politics	Civil liberties	Barro and Lee [8]	$-^{*}$
		Kormendi and Meguire [66]	$+$
		Levine and Renelt [72]	$?^{f}$
		Sala-i-Martin [104]	$+^{*}$
	Instability	Alesina, Ozler, Roubini, and Swagel [2]	$-^{*}$
		Barro [5]	$-^{*}$
		Barro and Lee [8]	$-^{*}$
		Caselli, Esquivel, and Lefort [22]	$-^{*}$
		Levine and Renelt [72]	$-^{f}$
		Sala-i-Martin [104]	$-^{*}$
	Political rights	Barro and Lee [8]	$+^{*}$
		Sala-i-Martin [104]	$+^{*}$
Population growth		Barro and Lee [8]	$+$
		Kormendi and Meguire [66]	$-^{*}$
		Levine and Renelt [72]	$-^{f}$
		Mankiw, Romer, and Weil [78]	$-^{*}$
	≤ 15 years	Barro and Lee [8]	$-^{*}$
	≥ 65 years	Barro and Lee [8]	?
Price distortion	Consumption	Easterly [40]	$+$
		Harrison [54]	$-^{*}$
	Investment	Barro [5]	$-^{*}$
		Easterly [40]	$-^{*}$
Price levels	Consumption	Easterly [40]	$+$
	Investment	Easterly [40]	$-^{*}$

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Regions	Latitude (absolute)	Sala-i-Martin [104]	$+^*$
	East Asia	Barro [7]	$+$
		Barro and Lee [8]	$+$
	Former Spanish colony	Sala-i-Martin [104]	$-^*$
	Latin America	Barro [5]	$-^*$
		Barro [7]	$-$
		Barro and Lee [8]	$-^*$
		Sala-i-Martin [104]	$-^*$
	Sub-Saharan Africa	Barro [5]	$-^*$
		Barro [7]	$-$
		Barro and Lee [8]	$-^*$
		Sala-i-Martin [104]	$-^*$
Religion	Buddhist	Sala-i-Martin [104]	$+^*$
	Catholic	Sala-i-Martin [104]	$-^*$
	Confucian	Sala-i-Martin [104]	$+^*$
	Muslim	Sala-i-Martin [104]	$+^*$
	Protestant	Sala-i-Martin [104]	$-^*$
Rule of law		Barro [6, 7]	$+^*$
		Sala-i-Martin [104]	$+^*$
Scale effects	Total area	Sala-i-Martin [104]	?
	Total labor force	Sala-i-Martin [104]	?

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Trade	Export/import/total trade as fraction of GDP	Frankel and Romer [45]	$+^{*}$
		Frankel, Romer, and Cyrus [46]	$+^{*}$
		Harrison [54]	$-$
		Levine and Renelt [72]	$+^{f}$
	Primary products in total exports (fraction)	Sala-i-Martin [104]	$-^{*}$
	Export-GDP ratio (change)	Kormendi and Meguire [66]	$+^{*}$
	FDI relative to GDP	Blomstrom, Lipsey, and Zejan [19]	$-$
	Machinery and equipment imports	Romer [100]	$+^{*}$
Trade policy	Import penetration	Levine and Renelt [72]	$?^{f}$
	Leamer index	Levine and Renelt [72]	$-^{f}$
	Openness (change)	Harrison [54]	$+^{*}$
	Openness (level)	Harrison [54]	$+^{*}$
		Levine and Renelt [72]	$?^{f}$
	Outward orientation	Levine and Renelt [72]	$?^{f}$
	Tariffs	Barro and Lee [8]	$-$
	Years open, 1950–1990	Sala-i-Martin [104]	$+^{*}$

Table 2: Growth regression compilation (Contd.)

Explanatory variables		Reference	Finding
Variability	Growth innovations	Kormendi and Meguire [66]	$-^{*}$
		Ramey and Ramey [95]	$-^{*}$
	Money	Kormendi and Meguire [66]	$-^{*}$
War	Casualties per capita	Easterly, Kremer, Pritchett, and Summers [41]	$+$
	Duration	Barro and Lee [8]	$+$
	Occurrence	Barro and Lee [8]	$-$
		Sala-i-Martin [104]	$-^{*}$

Table compiles results from a selection of the cross-country growth regressions that have been published. Under **Finding**, a * denotes a claim of significance (authors' significance levels differ across studies, and are not always explicitly reported); a ? denotes where the author did not report the result; and ^{f} and ^{r} indicate fragility and robustness in the sense used by Levine and Renelt [72]. In this table we can give no more than a flavor of the findings extant. Detailed variable definitions can be found in the individual references.

Table 3: Cross-country income dynamics
(118 economies, relative to world per capita income, 1962–1984)
Grid: (0, 1/4, 1/2, 1, 2, ∞)

(Number)	Upper Endpoint: 1/4	1/2	1	2	∞
(456)	0.97	0.03			
(643)	0.05	0.92	0.04		
(639)		0.04	0.92	0.04	
(468)			0.04	0.94	0.02
(508)				0.01	0.99
Ergodic	0.24	0.18	0.16	0.16	0.30

Table is a portion of Table 1 from Quah [88]. The table shows transition dynamics over a single-year horizon. The cells are arrayed in increasing order, so that the lower right-hand portion of the table shows transitions from the rich to the rich. The numbers in parentheses in the leftmost column are the number of economy/year pairs beginning in a particular cell. Cells showing 0 to two decimal places are left blank; rows might not add to 1 because of rounding. The ergodic row gives the long-run distribution from transitions according to the law of motion given in the matrix.

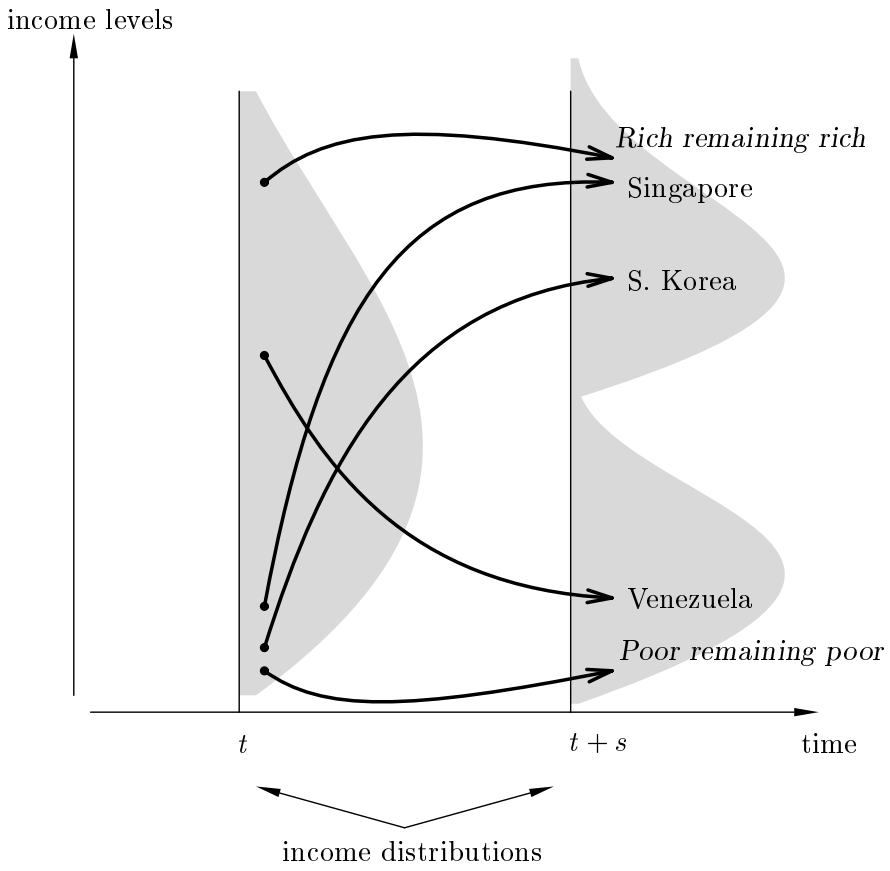


Fig. 1: Evolving cross-country income distributions Post-1960 experiences projected over 40 years for named countries are drawn to scale, relative to actual historical cross-country distributions.

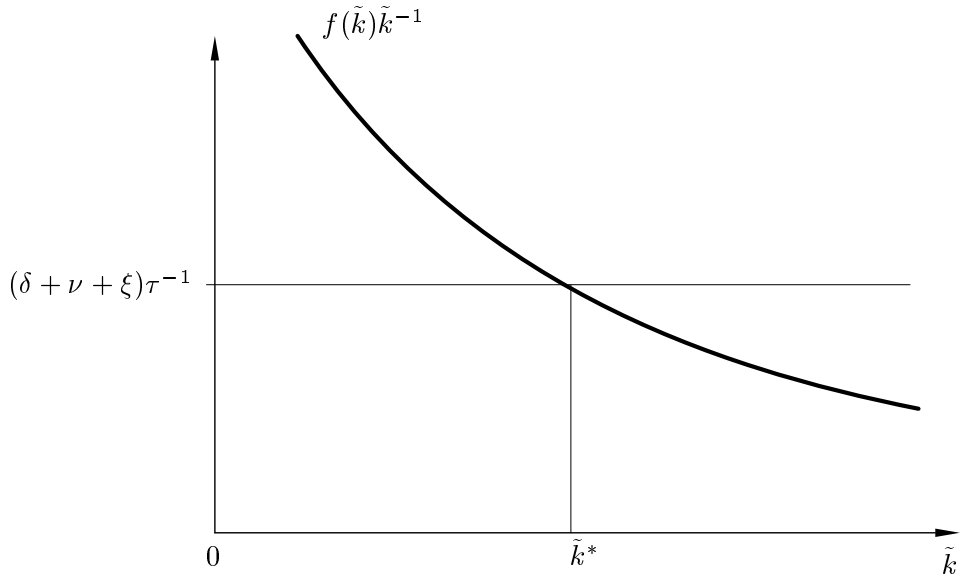


Fig. 2: Solow-Swan growth and convergence Function $f(\tilde{k})\tilde{k}^{-1}$ is continuous, and tends to infinity and zero as $\tilde{k}$ tends to zero and infinity, respectively. Moreover, it is guaranteed to be monotone strictly decreasing. The vertical distance between $f(\tilde{k})\tilde{k}^{-1}$ and $(\delta+\nu+\xi)\tau^{-1}$ is $\tau^{-1}\dot{\tilde{k}}/\tilde{k}$. Convergence to steady state $\tilde{k}^*$ therefore occurs for all initial values $\tilde{k}$.

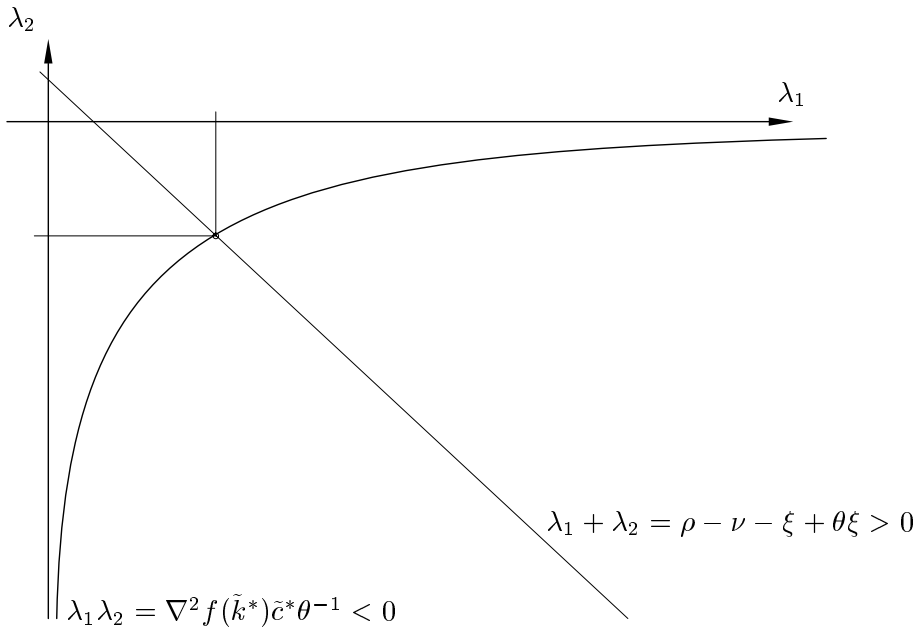


Fig. 3: Eigenvalues in the Cass-Koopmans model Since $\nabla^2 f(\tilde{k}^*)\tilde{c}^*\theta^{-1} = \left(\nabla^2 f(\tilde{k}^*)\tilde{k}^*\right)\left[f(\tilde{k}^*)/\tilde{k}^* - (\delta+\nu+\xi)\right]\theta^{-1}$, if $f(\tilde{k}) = \tilde{k}^\alpha$, with $\alpha \in (0,1)$ then as α increases towards unity the negative eigenvalue λ_2 rises towards zero.

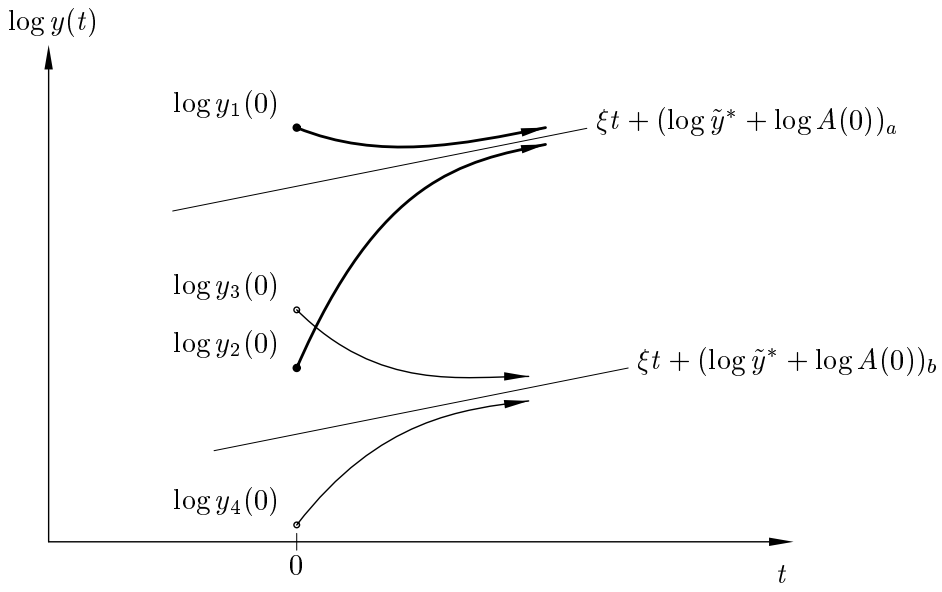


Fig. 4: Growth and convergence in the neoclassical model Figure shows two different possible steady state paths—corresponding to two possible values for the sum $\log \tilde{y}^* + \log A(0) = \log(g((\delta + \nu + \xi)^{-1}\tau)) + \log A(0)$. As long as this sum remains unobserved or unrestricted, *any* pattern of cross-country growth and convergence is consistent with the model. As drawn, the a value applies to economies at $y_1(0)$ and $y_2(0)$ while the b value to $y_3(0)$ and $y_4(0)$. Economies 1 and 2 converge towards each other, and similarly economies 3 and 4. At the same time, however, economies 2 and 3, although each obeying the neoclassical growth model, are seen to approach one another, criss-cross, and then to diverge.

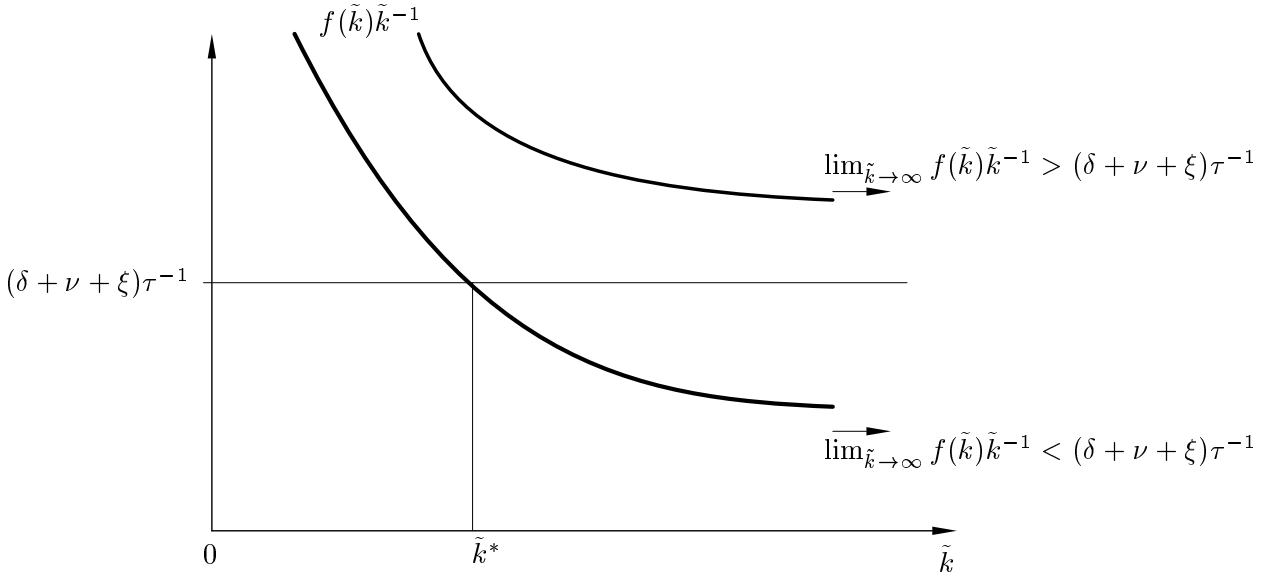


Fig. 5: Asymptotically linear $(O(\tilde{k}))$ growth and convergence The continuous function $f(\tilde{k})\tilde{k}^{-1}$ tends to infinity as $\tilde{k}$ tends to zero and to positive $\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1}$ as $\tilde{k}$ tends to infinity. Moreover, it is guaranteed to be monotone strictly decreasing for finite $\tilde{k}$. The vertical distance between $f(\tilde{k})\tilde{k}^{-1}$ and $(\delta+\nu+\xi)\tau^{-1}$ is $\tau^{-1}\dot{\tilde{k}}/\tilde{k}$. If $\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} < (\delta+\nu+\xi)\tau^{-1}$ then convergence occurs as in the Solow-Swan model with some constant finite $\tilde{k}^*$ describing balanced-growth equilibrium. However, if $\lim_{\tilde{k}\to\infty} f(\tilde{k})\tilde{k}^{-1} > (\delta+\nu+\xi)\tau^{-1}$ then $\dot{\tilde{k}}/\tilde{k}$ is always positive, and balanced growth obtains only as $\tilde{k} \nearrow \infty$. Every initial $\tilde{k}(0)$ is part of an equilibrium tending towards balanced growth.

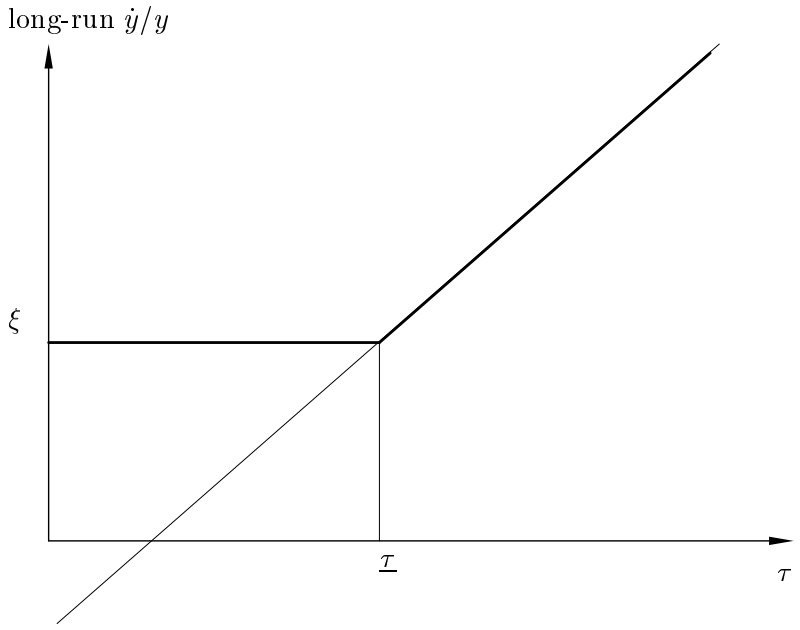


Fig. 6: Threshold effect of savings on long-run income growth rates in $O(\bar{k})$ model For economies with savings rates τ less than the threshold value $\underline{\tau}$, the long-run income growth rate is ξ independent of τ. If $\tau > \underline{\tau}$, however, then savings rates positively affect long-run growth.

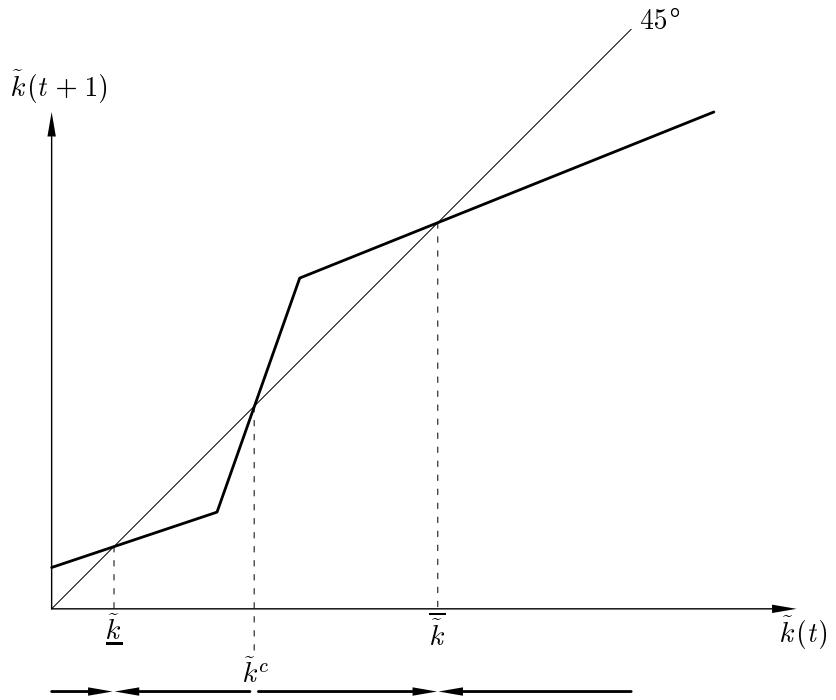


Fig. 7: Multiple locally stable steady states Either of the two possible limit points $\overline{\tilde{k}}$ or $\underline{\tilde{k}}$ obtains, depending on $\tilde{k}(0) \gtrless \tilde{k}^c$. The dark kinked line describes $\tilde{k}(t+1)$ as a function of $\tilde{k}(t)$ in the Galor-Zeira model, as applied in Quah [92] to study economies confronting imperfect capital markets. If a cross section of economies had randomly distributed initial conditions $\tilde{k}(0)$, then over time the cross-section distribution of $\tilde{k}$'s (and thus of $\tilde{y}$'s) will tend towards a clustering around $\underline{\tilde{k}}$ and $\overline{\tilde{k}}$.

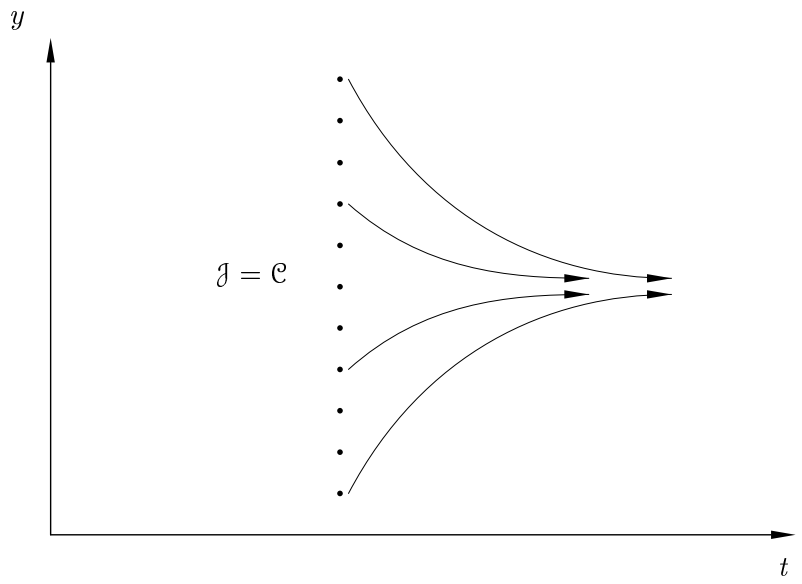


Fig. 8: Common average H Because the evolution of human capital across economies depends on the world's average—the symbol $\mathcal{J}$ denotes the entire cross section, $\mathcal{C}$ a clustering or club—convergence occurs to a degenerate point mass.

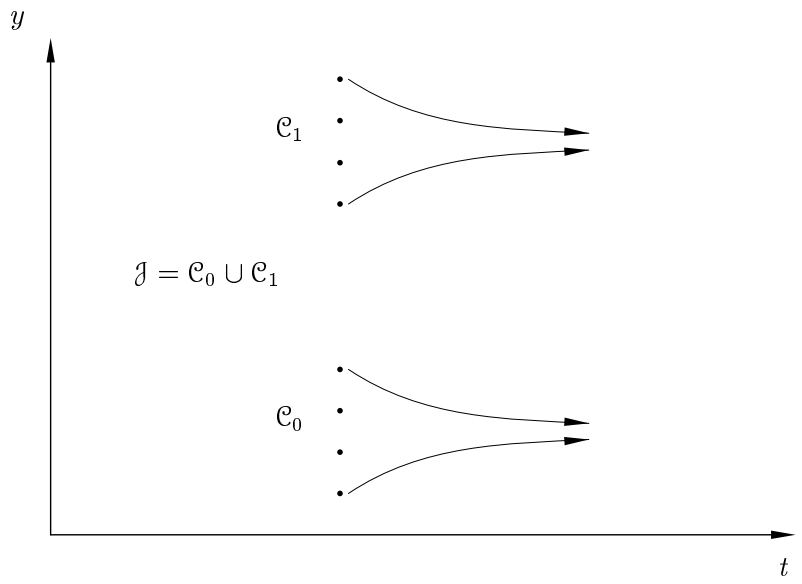


Fig. 9: Distinct average H across clubs Each economy now has a natural clustering—either $\mathcal{C}_0$ or $\mathcal{C}_1$ again with $\mathcal{J}$ the entire cross section—so that the relevant average **H** differs across economies. As drawn here convergence occurs to a two-point or twin-peaked distribution.

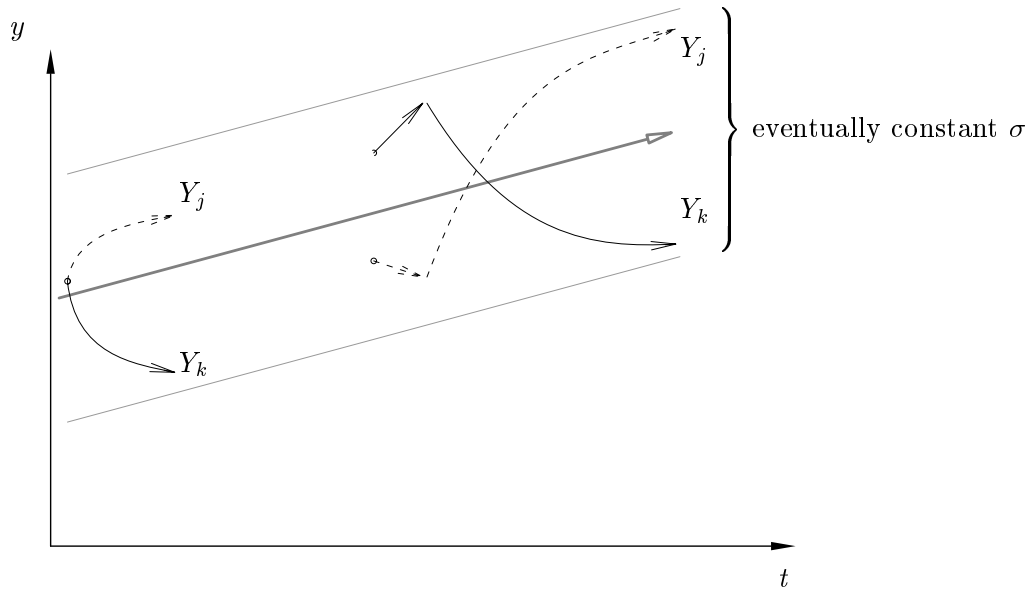


Fig. 10a: σ divergence towards σ-constant stationary state Figure shows a cross section of economies that begin close together relative to their steady state distribution and then spread out over time to converge in distribution to a well-defined steady state. Such dynamics are easy to generate, even with iid economies, each satisfying a covariance stationary linear autoregressive process.

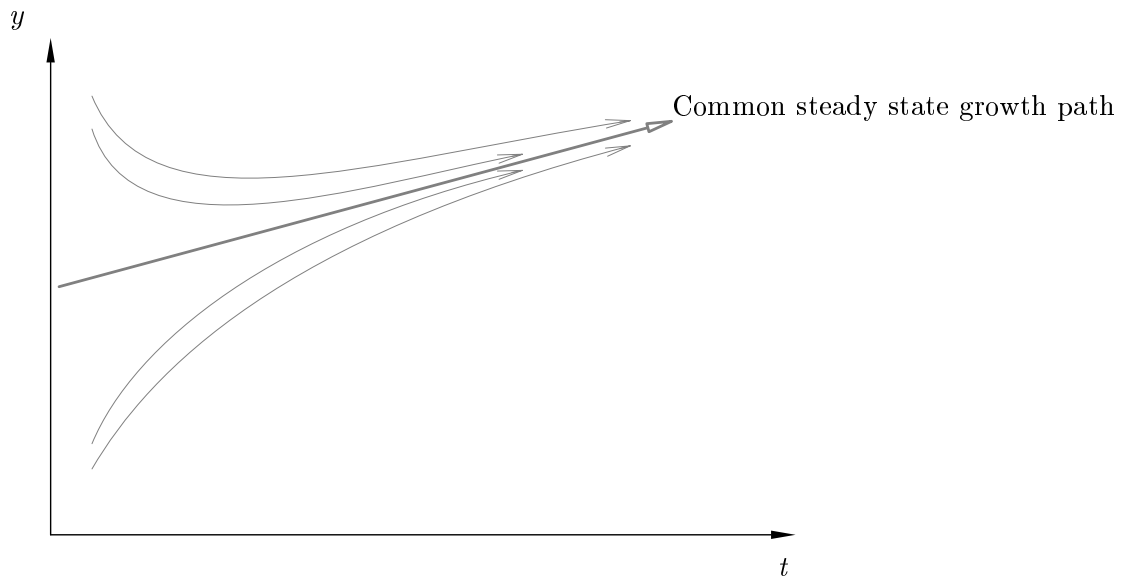


Fig. 10b: Coincident β and σ convergence Figure shows a cross section of economies where β and σ convergence coincide. All economies converge smoothly towards the common steady state growth path. Similarly, the dispersion of the cross-section distribution declines to zero.

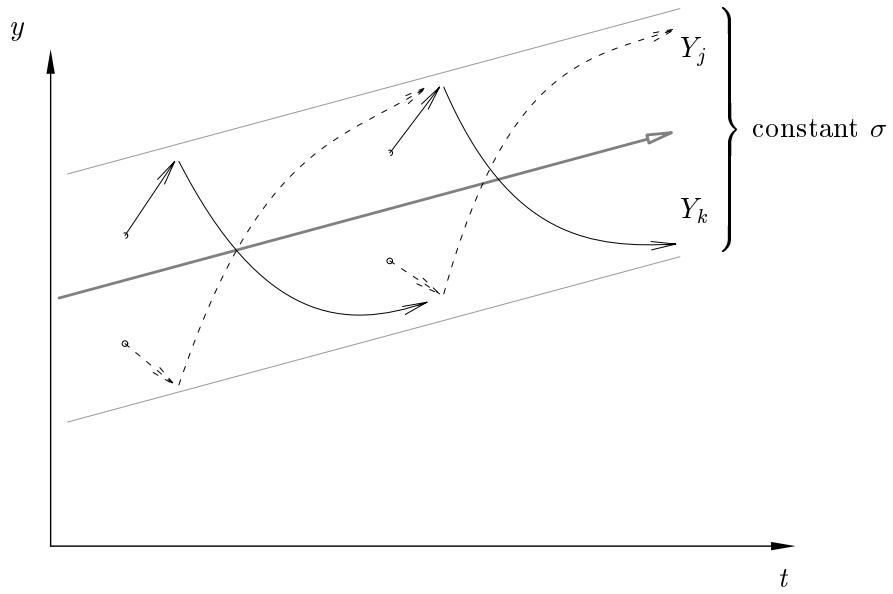


Fig. 10c: σ-convergent limit with ongoing intra-distribution churning Figure shows a cross section of economies at the steady-state distribution limit, but displaying ongoing intra-distribution dynamics. This situation might be viewed as the distributional endpoint of the earlier Fig. 10a.

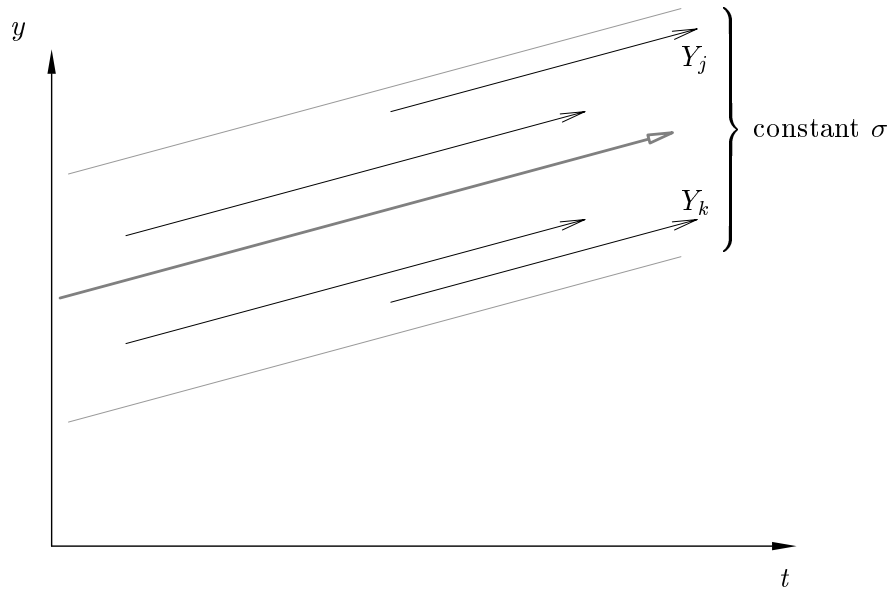


Fig. 10d: σ-convergent limit without intra-distribution churning Figure shows a cross section of economies at the steady-state distribution limit, but unlike in Fig. 10c there are no ongoing intra-distribution dynamics. All economies simply move in parallel.

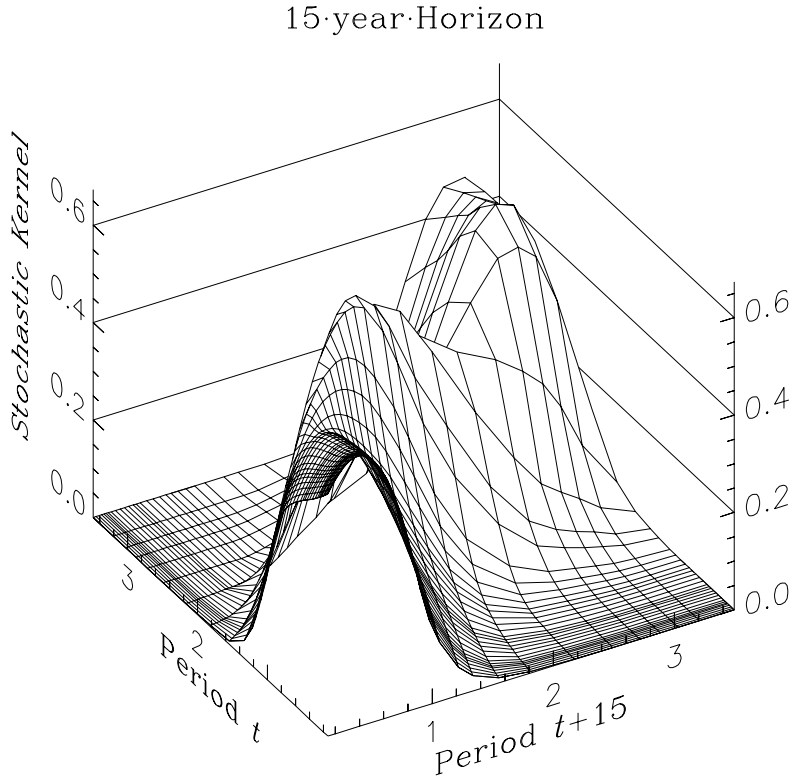
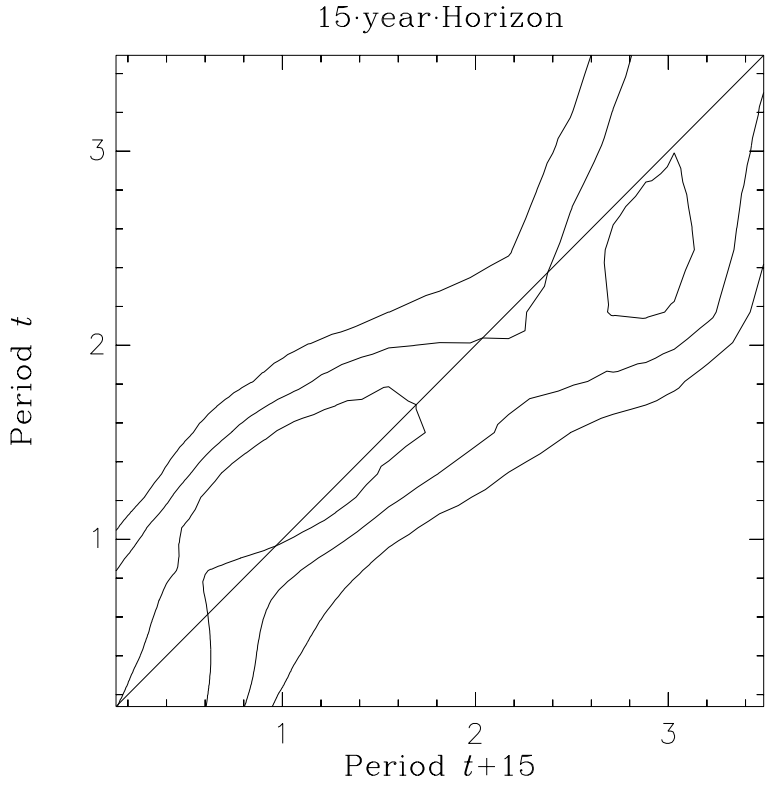


Fig. 11a: Relative Income Dynamics across 105 Countries, 1961–1988
For clarity, this stochastic kernel is one taken over a fifteen-year transition horizon. The kernel can be viewed as a continuum version of a transition probability matrix. Thus, high values along the diagonal indicate a tendency to remain. A line projected from a fixed value on the *Period* t axis traces out a probability density over the kernel, describing relative likelihoods of transiting to particular income values in *Period* $t + 15$. The emerging twin-peaks feature is evident here, now without the aliasing possibilities in discrete transition probability matrices.



Contour plot at levels 0.2, 0.35, 0.5

Fig. 11b: Relative Income Dynamics across 105 Countries, 1961–1988. Contour Plot This figure is just the view from above of Fig. 11a, where contours have been drawn at the indicated levels and then projected onto the base of the graph.