Convergence behaviour in exogenous growth models

Jochonia S. Mathunjwa
and
Jonathan R. W. Temple

Discussion Paper No. 06/590

August 2006

Department of Economics
University of Bristol
8 Woodland Road
Bristol BS8 1TN

Convergence behaviour in exogenous growth models*

Jochonia S Mathunjwa and Jonathan R W Temple†
Department of Economics, University of Bristol
8 Woodland Road, Bristol BS8 1TN

August 3, 2006

Abstract

This paper analyzes several aspects of convergence behaviour in the Solow growth model. In empirical work, a popular approach is to log-linearize around the steady-state. We investigate the conditions under which this approximation performs well, and discuss convergence behaviour when an economy is some distance from the steady-state. A formal analysis shows that convergence speeds will be heterogeneous across countries and over time. In particular, the Solow model implies that convergence to a growth path from above is slower than convergence from below. We find some support for this prediction in the data.

1 Introduction

The hypothesis of conditional convergence is central to the recent empirical growth literature. The hypothesis is that any given country can be viewed as converging to a balanced growth path, and the country's distance from this balanced growth path will influence its growth rate. Countries a long way below their steady-state path will show relatively fast growth, while countries a long way above their steady-state position will grow relatively slowly, and perhaps even see reductions in GDP per worker. In general, if we control for the determinants of the level of the steady-state path, countries that are relatively poor will grow more quickly.

From the perspective of traditional time series analysis, this is really not much more than a partial adjustment model. Models of economic growth add to this by making quantitative predictions about the speed at which economies will converge towards the long-run equilibrium. Testing these predictions may be informative about key structural parameters. The rate at which countries converge to a steady-state path also tells us whether transitional dynamics or steady-state behaviour play the dominant role in observed patterns of growth rates. For economies that take a long time to converge to their steady-state, transitional dynamics are important,

*We are grateful to Edmund Cannon for helpful suggestions. The usual disclaimer applies.
†Corresponding author. Email: jon.temple@bristol.ac.uk

while economies that converge rapidly will often be close to their steady-state positions, and differences in growth might then be attributed to steady-states that are changing over time.

Even if the rate of convergence is not the focus of interest, the partial adjustment model has implications for cross-section and panel data studies. Since the key studies of Baumol (1986), Barro (1991), Barro and Sala-i-Martin (1992), and Mankiw, Romer, and Weil (1992), empirical models of growth have routinely controlled for the initial level of GDP per capita or GDP per worker. This idea has influenced ongoing empirical debates that predate the conditional convergence literature by several decades, such as studies of the growth effects of foreign aid.

In this article, we focus on the conditional convergence effect in more detail than is usual. We study this effect for the Solow (1956) model of growth. As is now well known, the Solow model can be used to derive a conditional convergence relationship between growth and initial income. The argument relies on a first-order Taylor series approximation of the governing equations around the steady-state. This also allows the derivation of speeds of convergence and half-life times in the neighbourhood of the steady-state. The advantage of this approach is that it yields linear, analytical solutions even if the governing equations are highly non-linear and difficult to solve. The weakness of this approach is that its validity is strictly limited to within the close vicinity of the steady-state.

Although the Taylor series approximation is a standard theoretical justification for conditional convergence effects, its accuracy has rarely been studied in a systematic way, with Reiss (2000) as the leading exception. We will extend this line of research in a number of ways. We will consider several alternative definitions of the speed of convergence. Near the steady-state, these will give identical answers, but otherwise the precise definition matters. Further from the steady-state, the speed of convergence will not be constant, and the standard Taylor series approximation may then be a poor guide to growth behaviour. We will study the accuracy of this approximation for the Cobb-Douglas case. We also show that, although in principle the use of higher-order approximations should improve accuracy, in practice they do not always preserve the qualitative properties of the governing dynamic equation.

Our second main contribution is to show that these issues are more than simply a technical nicety. The standard Taylor series approximation leads to an equation for growth in which the coefficient on initial income is likely to be broadly the same across countries. This helps to justify the use of initial income in growth regressions, and these regressions usually assume that the effect of initial income is homogeneous across countries. If countries are not in the neighbourhood of the steady-state, this assumption is no longer true. We show that the coefficient on initial income can then be written as an integral of a time-varying convergence rate. Moreover, this integral depends on the distance between an economy's initial position and the steady-state.

The practical implication of this is that the coefficient on initial income will vary across countries. We investigate an especially stark example of this. The analytical results that we develop suggest that the coefficient on initial income will be higher (in absolute terms) for countries converging to a steady-state from below, compared to countries converging from above. Work by Cho and Graham (1996) has suggested that convergence from above is not a rare phenomenon. We use a standard method for dividing countries into two such groups, and find that the hypothesis of heterogeneous convergence effects has some support in the data. This is consistent with standard growth models, once we relax the assumption that economies are close to their respective steady-states.

Since much of the analysis in the paper is relatively technical, we now provide a detailed overview. In section 2, we present a detailed review of the Solow model, including the key assumptions on which it is based, its fundamental governing equations, and its main predictions. We also discuss the particular case of a Cobb-Douglas production function.

In analysing conditional convergence, one question that arises is whether to linearise or log-linearise the governing equation. In the neighbourhood of the steady-state, identical speeds of convergence are found for output and capital, regardless of whether the governing equations have been linearised or log-linearised (see for example Romer (2001) and Barro & Sala-i-Martin (2004)). But it is desirable to obtain speeds of convergence and half-life times away from the steady-state. After all, it is these departures that motivate the focus on conditional convergence in the first place. Once away from the steady-state, the precise definition of the rate of convergence becomes important, and the analysis becomes more complicated. For example, working on the Cobb-Douglas case, Reiss (2000) has shown that output and capital converge at different rates when away from the steady-state, when using certain measures of the convergence rate.

Our analysis in section 3 begins by noting that commonly used definitions of speed of convergence and half-life times can be classified into two types: those based on ordinary variables (OVs) and those based on log variables (LVs).[1] Linearising about the steady-state leads to OV-based definitions of speed of convergence while log-linearising leads to LV-based definitions. Both OV-based and LV-based measures yield one speed of convergence (and hence half-life time) for output and capital in the neighbourhood of the steady-state. Outside the neighbourhood of the steady-state, these speeds may differ, and OV-based and LV-based measures are shown to give different conclusions.

Since we focus on behaviour away from the steady-state, it is natural to ask whether significant gains are obtained by adding more terms to the widely used first-order Taylor expansions near the steady-state. Thus, in an appendix, we derive quadratic and cubic log expansions and compare their performance to that of the linear expansion. The cubic (and any higher order expansion) is relatively difficult to solve, and the solution obtained will be too complicated to offer useful insights into the transitional dynamics. The quadratic expansion yields a solution which is generally more accurate than the linear solution, but it can give results that are not consistent with the exact solution; for certain initial positions, the quadratic solution predicts that the economy will evolve away from the steady-state. The analysis therefore shows that the linear log expansion, although the least accurate away from the steady-state, always gives qualitative results that are consistent with

[1] Note that what is referred to as the OV-frame here is also called the linear scale, and the LV-frame is also called the log (or ratio) scale.

the exact solution, and is simplest to derive and solve. The linear log expansion is therefore the most useful for many purposes.

In section 4, we study convergence behaviour far from steady-state. We find that speeds of convergence and half-life definitions derived in the OV-based frame generally lead to different predictions from those derived in LV-based frames. OV-based speed-of-convergence definitions yield different speeds for output and capital, while LV-based measures give identical speeds. We derive OV-based and LV-based formulas for the half-life times of capital and output, and demonstrate that the properties of the half-life times are consistent with those of the associated OV-based and LV-based speeds of convergence, respectively.

By focusing on the case of a Cobb-Douglas production function, we confirm the finding of Reiss (2000) that OV-based definitions always yield unequal speeds of convergence (and half-life times) for output and capital. For example, if the exponent of capital is between 0 and 1, then as the economy converges to its balanced growth path from below, the speed of convergence for output is shown to increase while that for capital decreases, towards the (common) rate that obtains in the steady-state. In terms of LVs, the evolution equations for output and capital are essentially the same, and hence the LV-based definitions give identical speeds of convergence for output and capital. The LV-based speed of convergence is shown to decrease as an economy approaches its balanced growth path from below, and increase (towards the speed observed when close to the steady-state) when approaching from above.

In section 5, we derive the empirical implications of allowing economies to be far from their steady-state. In particular, as noted above, we show that the coefficient on initial income can be expressed as an integral over time of a time-varying rate of convergence. Since the convergence rate varies with an economy's distance from its steady-state, so does this integral. This implies that the coefficient on initial income will be heterogeneous across economies when some are far from their steady-state positions, contrary to the usual assumption in the empirical literature.

In section 6, we investigate the more specific hypothesis that LV-based speeds of convergence should be higher for economies converging to their steady-states from below than for economies converging from above. By modifying standard growth regressions, we show how to test this hypothesis, and find there is some support for the hypothesis in the data.

We remark that the material contained in section 2 is well-known, presented here because it forms the basis of subsequent analyses. The material in section 3 is also standard, but the approach employed here to derive the linear expansions and the speed-of-convergence definitions is relatively systematic. The rest of the material, consisting of sections 4, 5 and 6, and the appendix (the comparison of linear, quadratic and cubic log expansions), represent the original contributions of this paper.

2 The Solow model

The Solow model of economic growth forms the basis of the work presented in this article. In this section we derive the model and present its main results, first for a general production function and then for the case of a Cobb-Douglas production function.

We consider a closed economy in which output $Y(t)$ is generated according to the production function

$$Y = F(K, AL), \tag{1}$$

where $K(t)$, capital, $L(t)$, labour, and $A(t)$, the level of technology, are all functions of time. The function F is assumed to be at least twice differentiable, satisfies the Inada conditions, has positive, diminishing returns to each of its arguments, and constant returns to scale. The rates of saving, population growth, and technological progress are taken as exogenous, and L and A are assumed to grow at constant rates according to

$$\begin{aligned} \dot{L} &= nL, &(2)\\ \dot{A} &= gA, &(3) \end{aligned}$$

where $\dot{L} = \mathrm{d}L/\mathrm{d}t$. Capital, however, is taken to accumulate endogenously according to

$$\dot{K} = sY - \delta K, \tag{4}$$

where s is the rate of saving and δ is the rate of capital depreciation. The input variables of the model are capital and labour, which are assumed to be paid their marginal products. It is convenient to introduce the variables $y = Y/(AL)$ and $k = K/(AL)$, where AL is a measure of the effective units of labour. The equations (1) and (4) become

$$\begin{aligned} y &= f(k), &(5)\\ \dot{k} &= sf(k) - (n + g + \delta)k, &(6) \end{aligned}$$

respectively. The equations (5) and (6) are the fundamental equations of the Solow model and they describes the evolution of $y(t)$ and $k(t)$ in time.[2] Equation (6) says the rate of change of k is given by actual investment per unit of effective labour, $sf(k)$, less $(n + g + \delta)k$, the amount of investment required to keep k at its existing level.[3] The amount of investment required to keep k from falling due to depreciation is δk. This amount of investment is not enough to keep k constant though, because effective labour is also growing $\big($at the rate $(n + g)\big)$. Thus the total investment required to keep k at its existing level is $(n + g + \delta)k$.

When $sf(k)$ exceeds $(n + g + \delta)k$, the capital per effective unit of labour grows, and when $sf(k)$ is lower than $(n + g + \delta)k$, the capital per effective unit of labour falls. The point at which

$$sf\big(\hat{k}\big) = (n + g + \delta)\hat{k}, \tag{7}$$

represents the steady-state level at which $\dot{k} = 0$.[4] Note that since $\dot{y} = \dot{k}f'(k)$ and $f'(k)$ is finite and vanishes nowhere in $0 < k < \infty$, at the steady-state $\dot{y} = 0$ as

[2] Note that $f(k) \equiv F\big(k, 1\big)$ and, since F satisfies the Inada conditions, $\lim_{k\to 0} f'(k) = \infty$ and $\lim_{k\to\infty} f'(k) = 0$.

[3] Note that in this context, it has been assumed that saving equals investment.

[4] We use the hat to denote steady-state variables.

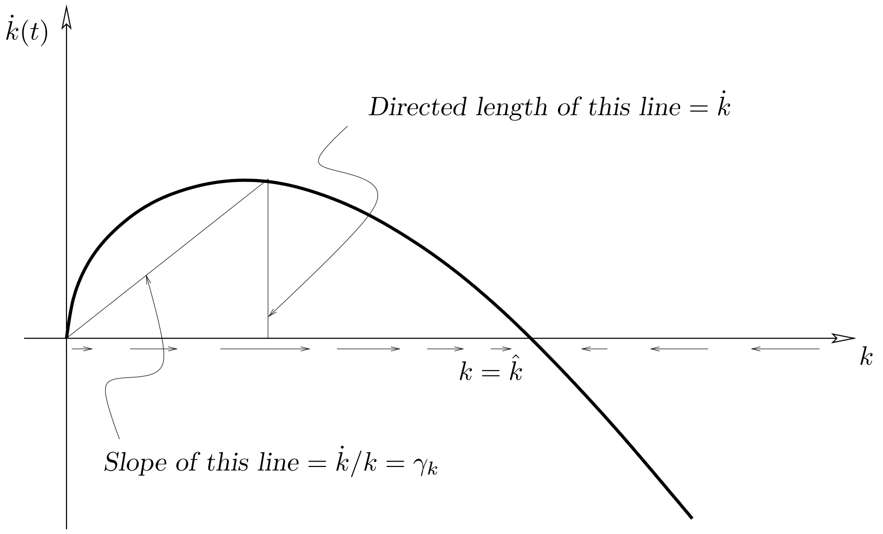

Figure 1: Phase diagram showing how $\dot{k}$ varies with k for a typical neoclassical production function. The horizontal arrows indicate the direction in which the economy evolves.

well. The situation is illustrated in figure 1 which shows a phase diagram where $\dot{k}$ is plotted against k. It shows that if the initial level of k is less than $\hat{k}$, then $k(t)$ grows towards $\hat{k}$. If it is initially greater than $\hat{k}$, then it decays towards $\hat{k}$. Thus, regardless of its initial position, the economy always evolves towards the steady-state $k = \hat{k}$, demonstrating that $k = \hat{k}$ is a stable equilibrium. In the steady-state, $K = AL\hat{k}$ and, since $\hat{k}$ is constant, the growth rate of K (and Y) is given by $(n+g)$. Also, with $K/L = A\hat{k}$ and $Y/L = A\hat{y}$, the growth rate of both output per capita Y/L and capital per capita K/L is thus g – demonstrating that, on the balanced growth path, the growth rate of per capita variables is determined only by the rate of technological progress.

From (7), the steady-state level of the economy is determined by the exogenous parameters s, n, g and δ. Economies for which these parameters are equal (that is, economies that are structurally similar) will have the same steady-state level. We analyse the effects of changes in these parameters on the steady-state level $\hat{k}$. Differentiating (7) with respect to s gives

$$f(\hat{k}) + s\frac{\partial \hat{k}}{\partial s} f'(\hat{k}) = (n + g + \delta)\frac{\partial \hat{k}}{\partial s}.$$

Then, rearranging using (7) and noting that $\hat{k}f'(\hat{k})/f(\hat{k})$ is the elasticity of output with respect to capital at steady-state, equal to capital's share of output, and denoted $\alpha(\hat{k})$ here, gives

$$\frac{\partial \log \hat{k}}{\partial \log s} = \frac{1}{1 - \alpha(\hat{k})}. \tag{8}$$

Since $0 < \alpha(\hat{k}) < 1$, this expression is always positive. In fact, as α increases from zero to one, the absolute value of (8) increases from 1 to ∞. Thus there is a positive functional dependence of $\hat{k}$ on s, indicating that an increase in the saving rate, all other parameters kept fixed, leads to a higher steady-state level.

It can also be shown that

$$\frac{\partial \log \hat{k}}{\partial \log n} = \frac{-n}{\Big(1 - \alpha\big(\hat{k}\big)\Big)(n + g + \delta)} < 0. \tag{9}$$

In a similar manner, the elasticities of $\hat{k}$ with respect to g and δ are also negative. Therefore, an increase in any of the parameters n, g or δ leads to a lower steady-state level. In the case of an increase in the rate of depreciation, for example, the steady-state level falls because more savings have to go into the replacement of worn-out capital.

Next, to gain more insight into the behaviour of the economy as it evolves towards the steady-state, we analyse the growth rate of k, denoted γ_k. From (6) and through the use of (7), this can be expressed as

$$\begin{aligned} \gamma_k(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln k(t)\Big] &= \frac{sf(k)}{k} - (n + g + \delta) \\ &= (n + g + \delta)\left[\frac{\mathcal{Y}}{\mathcal{K}} - 1\right], \end{aligned} \tag{10}$$

where $\mathcal{Y} = y/\hat{y}$ and $\mathcal{K} = k/\hat{k}$. Notice that $\gamma_k(t)$ is related to $\gamma_y(t)$ through $\alpha(k)$, as follows:

$$\gamma_y(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln y(t)\Big] = \frac{kf'(k)}{f(k)}\left[\frac{sf(k)}{k} - (n + g + \delta)\right] = \gamma_k(t)\alpha(k). \tag{11}$$

Since $0 < \alpha(k) < 1$, the growth rate $\gamma_y(t)$ will be a fraction of $\gamma_k(t)$, but the two will always bear the same sign. These expressions indicate that in the steady-state where $\mathcal{Y} = \mathcal{K} = 1$, the growth rates of $k(t)$ and $y(t)$ vanish, as expected. Away from the steady-state, using L'Hopital's rule and the Inada conditions, yields $\lim\limits_{k\to 0}(\mathcal{Y}/\mathcal{K}) = \infty$ and $\lim\limits_{k\to\infty}(\mathcal{Y}/\mathcal{K}) = 0$. Thus for economies with $k < \hat{k}$, the growth rate is positive and increases with distance from $k = \hat{k}$. In the limit $\lim\limits_{k\to 0}\gamma_k = \infty$. Above $k = \hat{k}$, growth rates are negative, ranging from $\gamma_k = 0$ when $k = \hat{k}$ to $\gamma_k = -(n + g + \delta)$ in the limit $k \to \infty$. A graphical illustration of γ_k is shown in figures 1 and 2.

The form of γ_k means that, for two structurally similar economies which differ only in their initial endowments of k, both starting with $\mathcal{K} < 1$, the poor economy will have a higher growth rate than the rich one. The poor economy will have a higher growth rate so long as it remains poorer than the other economy. The growth rates equalise only when the poor economy catches up with the initially rich economy. For $\mathcal{K} > 1$, richer economies are shown to decay faster than (and hence converge to) poor economies nearer to the steady-state. Thus, the Solow model predicts that structurally similar economies converge in the long run. The hypothesis that poor economies generally grow faster than (and hence converge to) rich ones is called absolute convergence. Empirical evidence of growth experiences of a broad selection of countries (e.g. Barro and Sala-i-Martin 2004) shows no correlation between initial output levels and subsequent rates of growth. In this case, then, the hypothesis of absolute convergence appears to be rejected by the data. The explanation for this is the presence of heterogeneities across these countries whereas the notion of absolute

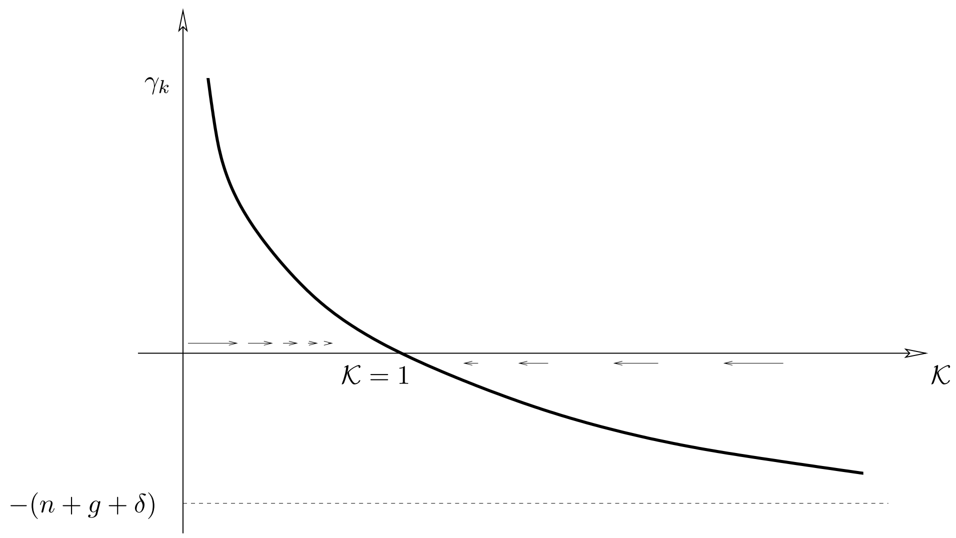

Figure 2: The growth rate γ_k plotted as a function of $\mathcal{K}$. The arrows show the direction and relative magnitude of the growth rate.

convergence is based on the assumption that all parameters, except the initial levels of capital, are identical. We note that a potential weakness of these empirical studies is the lack of a wide range of reliable data over long periods. Consequently, to avoid the problem of "selection bias", the studies use data over time periods that may not be sufficiently long.

Empirical studies of growth rates of more homogeneous groups of economies (e.g. Barro & Sala-i-Martin 2004, Sala-i-Martin 1996) show significant agreement with the hypothesis of absolute convergence. Data on the OECD economies (from 1960 to 1990), states of the United States (from 1880 to 1992), and the Japanese prefectures (from 1930 to 1990) all show that poor regions (states) generally grow faster per capita than rich ones. The convergence exhibited by these data occurs without conditioning on any other characteristics of the economies besides the initial level of per capita product or output, and hence it is absolute. This is consistent with the model because these economies may have essentially similar characteristics like technologies, tastes, and political institutions. Because of the relatively homogeneous conditions, the economies will have similar steady-state levels, and hence converge in the long-run.

For the case where heterogeneities across economies are significant, the differences in the parameters of the economies imply that they will have different steady-states. A graphical representation of the form shown in figure 2 would have multiple growth curves along which each of the economies traverse. In this situation, the notion of conditional convergence asserts that an economy's rate of growth is proportional to the distance from its own steady-state level. Thus, the growth rate of a poor economy may be lower than that of a rich economy if the poor economy is proportionately closer to its steady-state than the rich one.

Cobb-Douglas Production Function

The most widely used production function, which satisfies the properties of a neoclassical production function and may provide a reasonable description of real economies, is Cobb-Douglas. It is instructive to study the Solow model characterised by a Cobb-Douglas production function because, in this case, an exact analytical solution can be found, which greatly enhances comparative analysis. The formula (1) is thus replaced by

$$Y = K^{\alpha}(AL)^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{12}$$

The system of equations that govern the evolution of capital and output per effective unit of labour become

$$\begin{aligned} y &= k^{\alpha}, && (13)\\ \dot{k} &= sk^{\alpha} - (n+g+\delta)k, && (14)\\ \dot{y} &= \alpha s y^{2-\frac{1}{\alpha}} - \alpha(n+g+\delta)y. && (15) \end{aligned}$$

Exact analytical solutions of the evolution equations (14) and (15) can be found, and they are given by

$$k(t) = \left[\frac{s}{(n+g+\delta)}\left(1-e^{-\lambda t}\right) + k_0^{1-\alpha}e^{-\lambda t}\right]^{\frac{1}{1-\alpha}}, \tag{16}$$

$$y(t) = \left[\frac{s}{(n+g+\delta)}\left(1-e^{-\lambda t}\right) + k_0^{1-\alpha}e^{-\lambda t}\right]^{\frac{\alpha}{1-\alpha}}, \tag{17}$$

respectively, where $\lambda = (1-\alpha)(n+g+\delta)$ and $k_0 = k(0)$ is the initial level of capital.[5] These expressions show that the time-dependent components of $k(t)$ and $y(t)$ decay exponentially with time, and hence the economy approaches a steady-state level in the long-run, given by

$$\hat{k} = \left[\frac{s}{(n+g+\delta)}\right]^{\frac{1}{1-\alpha}}, \tag{18}$$

$$\hat{y} = \left[\frac{s}{(n+g+\delta)}\right]^{\frac{\alpha}{1-\alpha}}. \tag{19}$$

The terms that decay include those that depend on the initial level of capital k_0. Hence the steady-state level of the economy is independent of its initial position, but is determined by the exogenous parameters s, n, g and δ.

As the economy approaches steady-state, the growth rate of capital is given by

$$\begin{aligned} \gamma_k = \frac{\dot{k}}{k} &= (n+g+\delta)\left(\hat{k}^{1-\alpha} - k_0^{1-\alpha}\right)\left[k(t)\right]^{\alpha-1}e^{-\lambda t}\\ &= (n+g+\delta)\left(1-\mathcal{K}_0^{1-\alpha}\right)\mathcal{K}^{\alpha-1}e^{-\lambda t}. && (20) \end{aligned}$$

Since the terms $(n+g+\delta)$, $k(t)$ and $e^{\lambda t}$ are all positive, the sign of γ_k is determined by whether k_0 is less or greater than $\hat{k}$. If $k_0 < \hat{k}$, then γ_k is positive implying growth

[5] A full solution of (14), previously obtained by, for example, Williams and Crouch (1972), is presented in the appendix.

while for $k_0 > \hat{k}$, γ_k is negative and k decays in time. The absolute magnitude of the initial growth rate, given by

$$|\gamma_k(0)| = (n + g + \delta)\Big|\mathcal{K}_0^{\alpha-1} - 1\Big|,$$

is proportional to the distance of $\mathcal{K}_0$ from 1, taking small values when $\mathcal{K}_0$ is in the neighbourhood of 1 and large values when $\mathcal{K}_0$ is far from 1. With the exponent $(\alpha - 1) < 0$, when $\mathcal{K}_0 \to 0$ we have $\gamma_k(0) \to \infty$, and as $\mathcal{K}_0 \to \infty$, the initial growth rate $\gamma_k(0) \to -(n+g+\delta)$. The magnitude of γ_k at subsequent times then decreases to zero because the term $e^{-\lambda t}$, equal to one at $t = 0$, goes to zero as $t \to \infty$. Thus k approaches $\hat{k}$ (that is, the economy approaches the steady-state) asymptotically from below if $k_0 < \hat{k}$, and from above if $k_0 > \hat{k}$. Hence the level $k = \hat{k}$ is a stable equilibrium.

These results are of course in accord with the well-known findings of the previous section based on a general neoclassical production function.

3 Convergence behaviour near steady-state

As the analysis of the previous section has shown, a central result of the Solow model of economic growth is conditional convergence – that is, the farther an economy is below its steady-state level, the higher will be its growth rate (other parameters equal). A key question that arises as a consequence of this prediction is how long it takes for an out-of-equilibrium economy to adjust to steady-state. The answer to this question is pivotal because it determines whether transitional dynamics or steady-state behaviour is important in the study of an economy's time evolution. Transitional dynamics are important if an economy takes a long time to adjust to steady-state, and steady-state behaviour is important if the economy adjusts rapidly. For a variable $\mathcal{X}(t)$ which evolves from an initial value $\mathcal{X}(0) = \mathcal{X}_0$ towards a steady-state level $\hat{\mathcal{X}}$, commonly used measures of its speed of convergence can be classified into two types. First, those derived with respect to $\mathcal{X}$ and, second, those derived with respect to $\ln \mathcal{X}$. Here we refer to definitions of the first type as ordinary-variable (OV) based and those of the second type as log-variable (LV) based. OV-based definitions of speed of convergence include

$$\Lambda_{1\mathcal{X}}(t) = -\frac{\dot{\mathcal{X}}(t)}{\mathcal{X}(t) - \hat{\mathcal{X}}}, \quad \text{and} \tag{21}$$

$$\Lambda_{2\mathcal{X}}(t) = -\frac{\mathrm{d}\dot{\mathcal{X}}}{\mathrm{d}\mathcal{X}}, \tag{22}$$

where $\dot{\mathcal{X}} = \mathrm{d}\mathcal{X}/\mathrm{d}t$. Commonly used LV-based definitions of speed of convergence are

$$\Lambda_{3\mathcal{X}}(t) = -\frac{\mathrm{d}\big[\ln \mathcal{X}(t)\big]/\mathrm{d}t}{\ln \mathcal{X}(t) - \ln \hat{\mathcal{X}}}, \quad \text{and} \tag{23}$$

$$\Lambda_{4\mathcal{X}}(t) = -\frac{\mathrm{d}}{\mathrm{d}\ln \mathcal{X}}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\ln \mathcal{X}(t)\Big) = -\frac{\mathrm{d}(\dot{\mathcal{X}}/\mathcal{X})}{\mathrm{d}\ln \mathcal{X}}. \tag{24}$$

Another measure closely related to the speed of convergence is the half-life of $\mathcal{X}(t)$. Again there are two slightly different definitions, one OV-based and the other LV-based, given by

$$\begin{aligned} \tfrac{1}{2}(\mathcal{X}_0 - \hat{\mathcal{X}}) &= \mathcal{X}(T_{\mathcal{X}}) - \hat{\mathcal{X}}, \quad \text{and} \qquad (25) \\ \tfrac{1}{2}(\ln \mathcal{X}_0 - \ln \hat{\mathcal{X}}) &= \ln \mathcal{X}(\mathcal{T}_{\mathcal{X}}) - \ln \hat{\mathcal{X}}, \qquad (26) \end{aligned}$$

respectively, where $T_{\mathcal{X}}$ and $\mathcal{T}_{\mathcal{X}}$ are used to denote half-life times obtained using the two approaches.

The standard approach to studying properties of transitional behaviour is to derive OV- or LV-based linear expansions in the neighbourhood of the steady-state. With the governing equations generally difficult to solve because of the non-linear form of the production function, this approach usually yields equations that are straightforward to solve. The solutions obtained this way are in exact analytical form and hence useful to analyse the structural properties of transitional dynamics. Moreover, their linear form makes them ideal for use in linear regression empirical tests.

The formulas (21)−(26) are usually applied to the linearised equations to calculate speeds of convergence and half-life times in the neighbourhood of the steady-state. In this context, all the definitions yield the same speed of convergence (and hence half-life times) for both $k(t)$ and $y(t)$. For example, Romer (2001) has used definition (22) while Barro and Sala-i-Martin (2004) have employed definition (24), and have found the same speed of convergence for $k(t)$ and $y(t)$ near the steady-state. On the other hand, Reiss (2000) has used definition (21) and found that, outside the vicinity of the steady-state, $k(t)$ and $y(t)$ generally exhibit different convergence behaviours, while Okada (2006) has used definitions (23) and (24).[6] Since the idea of the speed of convergence is most useful for the study of out-of-equilibrium economies, the question of whether the definitions provide consistent results outside the neighbourhood of the steady-state is important. We consider convergence behaviour outside the vicinity of the steady-state in the next section.

In this section, we analyse convergence behaviour near the steady-state. We conduct a comparative study of convergence properties as predicted by the speed of convergence and half-life time definitions (21)−(26). In section 3.1, working with a general production function, we demonstrate that the different speed-of-convergence definitions arise as a result of linearising with respect to different variables (OVs or LVs) and using different approaches to measuring convergence speed. All definitions are shown to yield the same speed of convergence for both $k(t)$ and $y(t)$ near the steady-state. In an appendix, we consider the question of how useful and reliable the widely used linear expansion is. Using a Cobb-Douglas production function, we derive quadratic and cubic log expansions and compare their performances against the exact solution with that of the linear expansion. It is shown that, despite its

[6] Reiss uses definition (21) throughout his analysis. Note that, as our study demonstrates below, definition (24) is also a legitimate measure of the speed of convergence. It differs from the definition (21) because, (i), it is LV-based while (21) is an OV-based measure and (ii), it measures speed through a different mechanism. Moreover, since $\mathcal{X}$ and $\dot{\mathcal{X}}$ are functions of time, Λ_3 is also a function of time, just like Λ_1.

being the least accurate in quantitative terms, the linear log expansion is the most useful, for reasons explained in the appendix.

3.1 Linear expansions

In this section, we show how linearising in terms of OVs leads to Λ_1- and Λ_2-type definitions while linearising in terms of LVs leads to Λ_3- and Λ_4-type definitions. All definitions are shown to yield the same speed of convergence for both $y(t)$ and $k(t)$ in the neighbourhood of the steady-state.

3.1.1 Measuring speed of convergence

Both in terms of OVs and LVs, there are two different approaches to measuring the speed with which the variable $\mathcal{X}(t)$ $\big($or $\ln\mathcal{X}(t)\big)$ is converging to its steady-state level $\hat{\mathcal{X}}$. The first is to measure the proportional rate at which the gap $\big|\mathcal{X}(t)-\hat{\mathcal{X}}\big|$ $\big($or $\big|\ln\mathcal{X}(t)-\ln\hat{\mathcal{X}}\big|\big)$ is decreasing. In effect, this approach defines the speed of convergence as the growth rate of $\big|\mathcal{X}(t)-\hat{\mathcal{X}}\big|$ $\big($or $\big|\ln\mathcal{X}(t)-\ln\hat{\mathcal{X}}\big|\big)$. It will become apparent that the definitions Λ_1 and Λ_3 are based on this approach.

The second approach involves measuring the proportional rate at which the slope of the $\mathcal{X}(t)$ $\big($or $\ln\mathcal{X}(t)\big)$ curve changes in time as $\mathcal{X}(t)$ $\big($or $\ln\mathcal{X}(t)\big)$ approaches $\hat{\mathcal{X}}$ (or $\ln\hat{\mathcal{X}}$). This approach defines the speed of convergence as the growth rate of $\dot{\mathcal{X}}(t)$ $\big($or $\mathrm{d}\big[\ln\mathcal{X}(t)\big]/\mathrm{d}t\big)$. While speeds of convergence derived using the two approaches are generally unequal (unless $\dot{\mathcal{X}}$ is a linear function of $\mathcal{X}$ in the case of OV-based definitions), they simplify to the same expression in the vicinity of $\mathcal{X}=\hat{\mathcal{X}}$. Moreover, both OV-based and LV-based definitions yield identical expressions in the vicinity of the steady-state.

3.2 Speed of convergence measures

For an economy characterised by a general production function, the Solow fundamental equations that describe the time evolution of variables per unit of effective labour take the form

$$
\begin{aligned}
y &= f(k), && (27)\\
\dot{k} &= \mathcal{F}(k) = sf(k)-(n+g+\delta)k, && (28)\\
\dot{y} &= \dot{k}f'(k) = f'(k)\Big[sf(k)-(n+g+\delta)k\Big], && (29)
\end{aligned}
$$

where $f'(k)=\mathrm{d}f/\mathrm{d}k$. The condition satisfied in the steady-state is given by

$$
sf\big(\hat{k}\big) = (n+g+\delta)\hat{k}. \tag{30}
$$

Valuable insight into the behaviour of the economy near the steady-state is usually gained by approximating the governing equations using OV- or LV-based Taylor expansions. Linear expansions are the most commonly used.

To analyse behaviour predicted by the system (27)−(29) near the steady-state, we first derive an OV-based linear expansion of equation (28), to get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(k-\hat{k}\big) &= \Big[sf'(\hat{k})-(n+g+\delta)\Big](k-\hat{k}) \\
&= -\left(1-\frac{\hat{k}f'(\hat{k})}{f(\hat{k})}\right)(n+g+\delta)(k-\hat{k}) \\
&= -\Big(1-\alpha\big(\hat{k}\big)\Big)(n+g+\delta)(k-\hat{k}), \qquad (31)
\end{aligned}$$

where we have used (30) and the fact that $kf'(k)/f(k)=\alpha(k)$ is capital's share of output. Since $\hat{k}$ is time-independent, α is constant along the balanced growth path. Dividing both sides of (31) by $(k-\hat{k})$ reveals that the coefficient $-(1-\alpha)(n+g+\delta)$ is the growth rate of $(k-\hat{k})$, that is, the proportional rate at which the gap $(k-\hat{k})$ is decreasing in time. The negated growth rate of $(k-\hat{k})$ is defined as the speed of convergence of the variable $k(t)$ onto $\hat{k}$, denoted

$$\Lambda_{1k}=(1-\alpha)(n+g+\delta)=-\frac{\mathrm{d}\big(k-\hat{k}\big)/\mathrm{d}t}{k-\hat{k}}. \qquad (32)$$

By noting that $k-\hat{k}=\Delta k$ and letting $\Delta k\to 0$ in the neighbourhood of the steady-state, yields

$$\lim_{\Delta k\to 0}\frac{\mathrm{d}(\Delta k)/\mathrm{d}t}{\Delta k}=\frac{\mathrm{d}\dot{k}}{\mathrm{d}k}.$$

This is the proportional rate at which the slope of $k(t)$ changes in time, and hence provides an alternative approach to measuring the speed of convergence

$$\Lambda_{2k}=-\frac{\mathrm{d}\dot{k}}{\mathrm{d}k}. \qquad (33)$$

Therefore $\Lambda_2=\Lambda_1$ in the vicinity of the steady-state level. Away from the steady-state, $\Lambda_1(t_p)$ measures the (negative) slope of the straight line joining $\big[k(t_p),\dot{k}(t_p)\big]$ and $(\hat{k},0)$ in the phase plane, while $\Lambda_2(t_p)$ measures the instantaneous slope of $\dot{k}$ at time $t=t_p$, that is, the growth rate of $\dot{k}$. A graphical illustration is shown in figure 3.

To linearise the output equation (29), first note that $\dfrac{\partial}{\partial y}=\dfrac{\partial}{\partial f(k)}=\dfrac{1}{f'}\dfrac{\partial}{\partial k}$. Thus we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(y-\hat{y}\big)=\left\{sf'(\hat{k})-(n+g+\delta)+\frac{f''(\hat{k})}{f'(\hat{k})}\mathcal{F}(\hat{k})\right\}\big(y-\hat{y}\big),$$

and since $\mathcal{F}(\hat{k})=sf(\hat{k})-(n+g+\delta)\hat{k}=0$, this becomes

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\big(y-\hat{y}\big) &= \Big[sf'(\hat{k})-(n+g+\delta)\Big](y-\hat{y}) \\
&= -\Big(1-\alpha\big(\hat{k}\big)\Big)(n+g+\delta)(y-\hat{y}). \qquad (34)
\end{aligned}$$

Comparing this equation with (31) indicates that OV-based linear Taylor expansions yield identical evolution equations for output and capital in the neighbourhood of the steady-state.

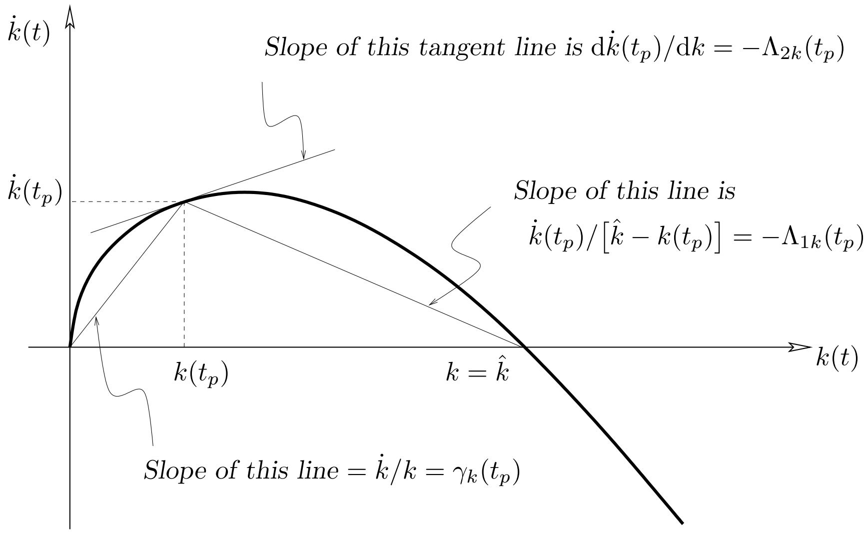

Figure 3: Graphical illustration of the speed of convergence definitions Λ_1 and Λ_2 for a neoclassical production function.

Next, we consider LV-based linear expansions. As mentioned earlier, LVs are widely used in empirical analyses because they can lead to linear equations. Moreover, coefficients of explanatory LVs in linear equations have an economic interpretation − they give the elasticity of the dependent OV with respect to the particular explanatory OV. It is thus very common to derive LV-based linear expansions to study behaviour near steady-state. We demonstrate how these expansions lead to the Λ_3- and Λ_4-definitions. Dividing equation (28) by k gives a differential equation for $\ln k(t)$

$$\frac{\dot{k}}{k} = \frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln k(t)\Big] = \mathcal{H}(k) = s\frac{f(k)}{k} - (n+g+\delta). \tag{35}$$

Then, in linearising this equation with respect to $\ln k(t)$, it is convenient to use the chain-rule formula $\dfrac{\partial}{\partial \ln k} = k\dfrac{\partial}{\partial k}$ to differentiate the terms on the right hand side, and we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln\hat{k} + \big(\ln k - \ln\hat{k}\big)\Big] &= s\left(f'(\hat{k}) - \frac{f(\hat{k})}{\hat{k}}\right)\big(\ln k - \ln\hat{k}\big) \\
\frac{\mathrm{d}}{\mathrm{d}t}\big(\ln k - \ln\hat{k}\big) &= -\left(1 - \frac{\hat{k}f'(\hat{k})}{f(\hat{k})}\right)\big(n+g+\delta\big)\big(\ln k - \ln\hat{k}\big) \\
&= -\Big(1 - \alpha\big(\hat{k}\big)\Big)(n+g+\delta)(\ln k - \ln\hat{k}). \qquad (36)
\end{aligned}$$

From this equation, the (negated) proportional rate at which the gap $\big(\ln k - \ln\hat{k}\big) = \ln\mathcal{K}$ decreases is given by

$$\Lambda_{3k} = (1-\alpha)(n+g+\delta) = -\frac{\mathrm{d}\big(\ln k - \ln\hat{k}\big)/\mathrm{d}t}{\big(\ln k - \ln\hat{k}\big)}. \tag{37}$$

This expression is the log-based counterpart of Λ_1. It can be written in the form

$$\Lambda_{3k} = -\frac{\mathrm{d}\big[\ln(k/\hat{k})\big]/\mathrm{d}t}{\ln(k/\hat{k})} = -\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln\big(\ln\mathcal{K}\big)\Big],$$

which shows that the Λ_{3k} measure is the negated growth rate of $\ln\mathcal{K}$. The last term in (37) can be written as

$$-\frac{\mathrm{d}\big[\Delta\ln k\big]/\mathrm{d}t}{\big[\Delta\ln k\big]}$$

and, in the limit $\Delta\ln k\to 0$, this yields

$$\Lambda_{4k} = -\frac{\partial\big[\mathrm{d}(\ln k)/\mathrm{d}t\big]}{\partial\ln k} = -\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln(\dot{k}/k)\Big] = -\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln\gamma_k(t)\Big]. \tag{38}$$

This shows that the Λ_{4k} definition measures the negated 'growth rate of the growth rate'. Hence $\Lambda_4 = \Lambda_3$ in the vicinity of the steady-state $k=\hat{k}$. Far from steady-state, $\Lambda_{3k}(t_p)$ gives the slope of the straight line joining the points $\Big[\ln k(t_p), \frac{\mathrm{d}}{\mathrm{d}t}\big[\ln k(t_p)\big]\Big]$ and $\big[\ln\hat{k},0\big]$ while $\Lambda_{4k}(t_p)$ measures the instantaneous slope of the $\frac{\mathrm{d}}{\mathrm{d}t}\big[\ln k(t)\big]$ curve at time $t=t_p$. The situation is illustrated in figure 4 which shows a phase diagram for the equation (35). Notice that, while for the associated OV-based equation (28), the function $\mathcal{F}(k)$ is concave down because $\mathcal{F}''(k) = sf''(k) < 0$, for the LV-based equation (35), the function $\mathcal{H}(k)$ is monotone decreasing because

$$\frac{\partial\mathcal{H}}{\partial\ln k} = \frac{sf(k)}{k}\big(\alpha(k)-1\big) < 0.$$

Use of L'Hopital's rule shows that $\lim\limits_{k\to 0}\mathcal{H}(k) = \infty$ while $\lim\limits_{k\to\infty}\mathcal{H}(k) = -(n+g+\delta)$, implying that $\mathcal{H}(k)$ is concave up. The properties of the functions $\mathcal{F}$ and $\mathcal{H}$ are thus qualitatively different, and this means that we can expect inconsistencies between OV-based and LV-based results, especially far from steady-state. Analysis of convergence behaviour far from steady-state will be presented in section 4.

Log-linearising the output equation (29) requires the use of the formula $\dfrac{\partial}{\partial\ln y} = \dfrac{\partial}{\partial\ln f(k)} = \dfrac{f}{f'}\dfrac{\partial}{\partial k}$, and yields

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln\hat{y} + \big(\ln y - \ln\hat{y}\big)\Big] &= \left\{\left(\frac{f''(\hat{k})}{f'(\hat{k})} - \frac{f'(\hat{k})}{f(\hat{k})}\right)\mathcal{F}(\hat{k}) + \mathcal{F}'(\hat{k})\right\}\big(\ln y - \ln\hat{y}\big)\\
&= \Big[sf'(\hat{k}) - (n+g+\delta)\Big]\big(\ln k - \ln\hat{k}\big)\\
\frac{\mathrm{d}}{\mathrm{d}t}\big(\ln y - \ln\hat{y}\big) &= -\Big(1-\alpha(\hat{k})\Big)(n+g+\delta)(\ln y - \ln\hat{y}),
\end{aligned} \tag{39}$$

where the relation $\hat{y} = f(\hat{k})$ has been used. Thus the LV-based linear expansions also give identical evolution equations for capital and output.

While evolution equations derived in terms of LVs generally have different properties to those derived in terms of OVs, it can be shown that the LV- and OV-based linear expansions are equivalent in the neighbourhood of the steady-state. Using the property $\ln x \simeq (x-1)$ around $x=1$, we can write $\ln k - \ln\hat{k} \simeq k/\hat{k} - 1$ near $k=\hat{k}$, and the equation (36) can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[k/\hat{k} - 1\big] \simeq -(1-\alpha)(n+g+\delta)(k/\hat{k}-1)$$

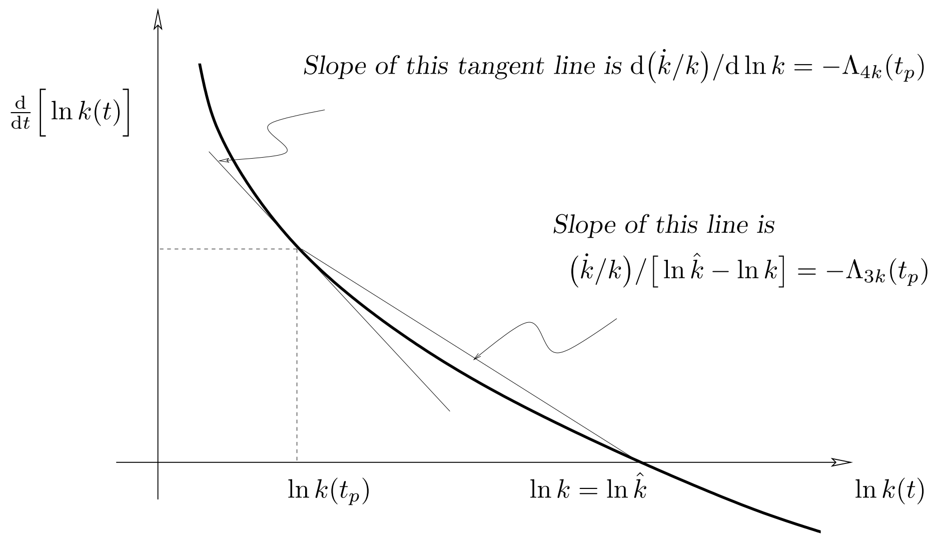

Figure 4: Graphical illustration of the speed of convergence definitions Λ_3 and Λ_4 for a neoclassical production function.

which is the same as the OV-based equation (31), for example.

Thus all four definitions of the speed of convergence yield the same value in the vicinity of steady-state, for capital and output. The expression $(1-\alpha)(n+g+\delta)$ implies that the speed of convergence is negatively related to α and positively related to n, g, and δ. The speed of convergence in this regime is independent of both the saving rate and the actual distance of the economy from steady-state.

4 Convergence behaviour far from steady-state

The conventional approach to studying transitional behaviour is first to linearise or log-linearise about the steady-state, as demonstrated in the previous section. However, measures derived using this approach are not applicable far from steady-state, and yet convergence information is most useful in the study of economies far from steady-state. Studies that analyse convergence measures far from steady-state are limited. They include the work of Reiss (2000) who has employed the OV-based Λ_1-definition and found different convergence behaviours for capital and output. In this section, we extend this analysis and study the predictions of all the OV-based and LV-based measures introduced in section 3. Using a Solow model characterised by a Cobb-Douglas production function, we calculate expressions for the speed of convergence using the Λ_1-, Λ_2-, Λ_3- and Λ_4-definitions. We demonstrate that the OV-based measures give unequal speeds of convergence for capital and output far from steady-state. The LV-based measures give equal speeds of convergence for capital and output even far from steady-state. We compare the OV-based and LV-based definitions of half-life times of convergence. The OV-based and LV-based definitions represent the times at which the variable $\mathcal{X}$ is equal to the arithmetic and geometric mean of $\mathcal{X}_0$ and $\hat{\mathcal{X}}$, respectively.

4.1 Speed of convergence

It is convenient to begin by stating the evolution equations of $k(t)$ and $y(t)$ applicable in this case, given by

$$\begin{aligned} \dot{k} &= \mathcal{B}(k) = sk^{\alpha} - (n+g+\delta)k, &(40)\\ \dot{y} &= \mathcal{D}(y) = \alpha s y^{2-\frac{1}{\alpha}} - \alpha(n+g+\delta)y. &(41) \end{aligned}$$

As already noted, in this case the exact solutions are known, and hence the exact values of convergence parameters can be found even outside the vicinity of the steady-state. We calculate expressions of speed of convergence for capital and output based on the Λ_1-, Λ_2-, Λ_3- and Λ_4-definitions.

The Λ_1-definition gives the following speeds of convergence

$$\begin{aligned} \Lambda_{1k} = \frac{\dot{k}}{\hat{k}-k} &= \frac{\gamma_k(t)}{\left(\frac{1}{\mathcal{K}}-1\right)} \\ &= (n+g+\delta)\frac{(\mathcal{K}^{\alpha-1}-1)}{\left(\frac{1}{\mathcal{K}}-1\right)} &(42)\\ \Lambda_{1y} = \frac{\dot{y}}{\hat{y}-y} &= \frac{\alpha\gamma_k(t)}{\left(\frac{1}{\mathcal{Y}}-1\right)} \\ &= \alpha(n+g+\delta)\frac{(\mathcal{K}^{\alpha-1}-1)}{\left(\frac{1}{\mathcal{K}^{\alpha}}-1\right)} &(43) \end{aligned}$$

where $\mathcal{K} = k/\hat{k}$, $\mathcal{Y} = y/\hat{y}$ and $\mathcal{Y} = \mathcal{K}^{\alpha}$. We remark that the final line of (43) is the same as the expression presented by Reiss (2000). The Λ_2-definition gives the following speeds of convergence

$$\begin{aligned} \Lambda_{2k} = -\frac{\mathrm{d}\dot{k}}{\mathrm{d}k} &= (1-\alpha)(n+g+\delta) - \alpha\gamma_k(t) \\ &= (n+g+\delta)(1-\alpha\mathcal{K}^{\alpha-1}) &(44)\\ \Lambda_{2y} = -\frac{\mathrm{d}\dot{y}}{\mathrm{d}y} &= (1-\alpha)(n+g+\delta) + (1-2\alpha)\gamma_k(t) \\ &= (n+g+\delta)\big[\alpha + (1-2\alpha)\mathcal{K}^{\alpha-1}\big]. &(45) \end{aligned}$$

Notice that in terms of LVs the evolution equation for $\ln y(t)$ is $\alpha(t)$ times the evolution equation of $\ln k(t)$ in general. In the Cobb-Douglas case, α is constant at all times, and hence the evolution equations of $\ln y(t)$ and $\ln k(t)$ are essentially the same, so that $\Lambda_{3y} = \Lambda_{3k}$ and $\Lambda_{4y} = \Lambda_{4k}$. Thus we have

$$\begin{aligned} \Lambda_{3y} = \Lambda_{3k} = -\frac{\big(\dot{k}/k\big)}{\ln k(t) - \ln\hat{k}} &= -\frac{\gamma_k(t)}{\ln\mathcal{K}} \\ &= -(n+g+\delta)\frac{(\mathcal{K}^{\alpha-1}-1)}{\ln\mathcal{K}} &(46)\\ \Lambda_{4y} = \Lambda_{4k} = -\frac{\mathrm{d}\big(\dot{k}/k\big)}{\mathrm{d}\ln k} = -\frac{\mathrm{d}\gamma_k(t)}{\mathrm{d}\ln k} &= (1-\alpha)(n+g+\delta) + (1-\alpha)\gamma_k(t) \\ &= (1-\alpha)(n+g+\delta)\mathcal{K}^{\alpha-1}, &(47) \end{aligned}$$

The last line of (47) is the same as the expression presented in Barro and Sala-i-Martin (2004).[7]

Since $\gamma_k(t) \to 0$ and $\mathcal{K} \to 1$ as the economy approaches steady-state, the formulas (42)−(47) show that all the definitions of speed of convergence give $\lambda = (1-\alpha)(n+g+\delta)$ in the steady-state − in agreement with the findings of section 3.[8] Outside the neighbourhood of the steady-state, in addition to the parameters α, n, g and δ, the speeds of convergence are shown to depend also on the economy's growth rate. For example, $\Lambda_{4k}(t) = \lambda + (1-\alpha)\gamma_k(t)$ is larger than λ when $\gamma_k > 0$, that is, if the economy is below the steady-state. Since, in this case, γ_k increases with distance from $\mathcal{K} = 1$, the farther an economy is below the steady-state, the higher the Λ_{4k} measure. Above the steady-state, $\gamma_k < 0$ and approaches $-(n+g+\delta)$ as $\mathcal{K} \to \infty$. Thus, Λ_{4k} will be less than λ in this case. It will decrease with distance from $\mathcal{K} = 1$, approaching 0 in the limit $\mathcal{K} \to \infty$.

Figures 5 and 6 show how speeds of convergence Λ_1, Λ_2, Λ_3 and Λ_4 vary with $\mathcal{K}$ and α, respectively.

OV-based definitions: For all values of α, the Λ_{1k} and Λ_{2k} speeds of convergence are less than λ whenever $\mathcal{K} < 1$ and higher than λ if $\mathcal{K} > 1$. Thus, economies that evolve from below their steady-state begin with small values of Λ_{1k} and Λ_{2k}, which then increase towards λ as the economy converges to steady-state.

The Λ_{1y} and Λ_{2y} measures are less than λ if $\mathcal{K} < 1$ and higher than λ if $\mathcal{K} > 1$, provided $\frac{1}{2} < \alpha < 1$. When $\alpha = \frac{1}{2}$, $\dot{y}$ is a linear function of y and hence the speed of convergence of y is constant at all times, given by $\Lambda_{1y} = \Lambda_{2y} = \alpha(n+g+\delta) = \lambda$. For $0 < \alpha < \frac{1}{2}$, the measures Λ_{1y} and Λ_{2y} are larger than λ when $\mathcal{K} < 1$ and less than λ when $\mathcal{K} > 1$.

The properties of the OV-based measures are driven by the concavity of the functions $\mathcal{B}(k)$ and $\mathcal{D}(y)$ in equations (40) and (41), respectively. While $\mathcal{B}(k)$ is concave down for all $0 < \alpha < 1$, the function $\mathcal{D}(y)$ is concave up, a straight line, or concave down whenever $0 < \alpha < \frac{1}{2}$, $\alpha = \frac{1}{2}$ or $\frac{1}{2} < \alpha < 1$, respectively. Consequently, whenever α is outside the interval $\frac{1}{2} < \alpha < 1$, capital and output exhibit different convergence behaviours in the OV-frame. As the economy evolves to the steady-state from above, for example, Λ_{1k} and Λ_{2k} decrease towards λ while Λ_{1y} and Λ_{2y} increase towards λ.

LV-based definitions: For all values of α, the LV-based speeds of convergence are higher than λ whenever $\mathcal{K} < 1$ and slower than λ when $\mathcal{K} > 1$. For economies that start with $\mathcal{K} > 1$, and hence are converging from above, the LV-based speeds of convergence are slower than, and gradually increase towards λ as the economies tend to their steady-states. In this case, the functions that determine the evolution of $\ln y(t)$ and $\ln k(t)$ are concave up for all values of α.

[7] Barro & Sala-i-Martin (2004), page 78.

[8] While showing this result is straightforward for the other cases, $\Lambda_{1k,y}$ and Λ_{3k} require the use of L'Hopital's rule. $\lim_{\mathcal{K}\to 1} \Lambda_{1k} = (n+g+\delta) \lim_{\mathcal{K}\to 1} \frac{(\alpha-1)\mathcal{K}^{\alpha-2}}{-\mathcal{K}^{-2}} = \lambda$, $\lim_{\mathcal{K}\to 1} \Lambda_{1y} = \alpha(n+g+\delta) \lim_{\mathcal{K}\to 1} \frac{(\alpha-1)\mathcal{K}^{\alpha-2}}{-\alpha\mathcal{K}^{-\alpha-1}} = \lambda$, $\lim_{\mathcal{K}\to 1} \Lambda_{3k} = -(n+g+\delta) \lim_{\mathcal{K}\to 1} \frac{(\alpha-1)\mathcal{K}^{\alpha-2}}{\mathcal{K}^{-1}} = \lambda$.

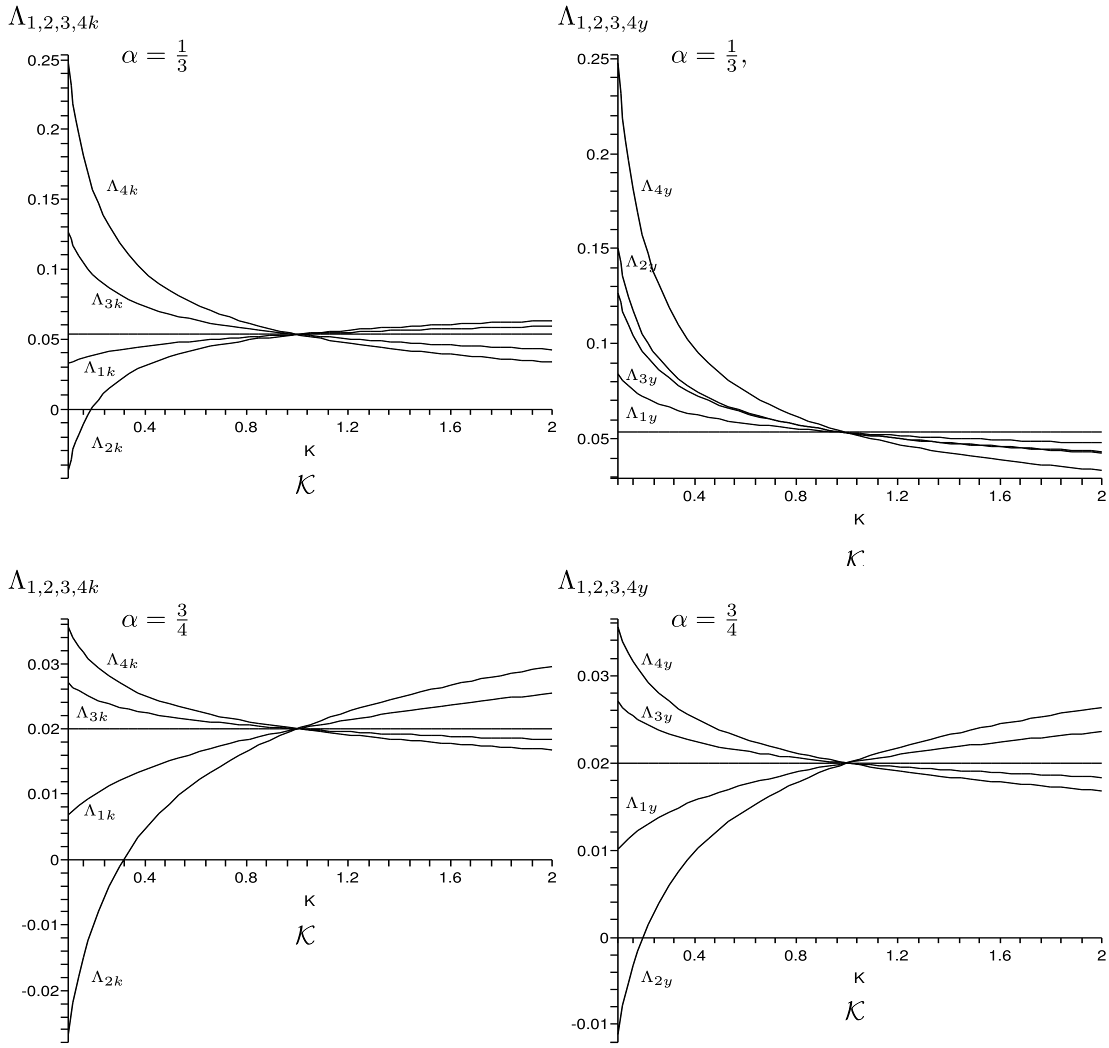

Figure 5: Speeds of convergence $\Lambda_1,\ \Lambda_2,\ \Lambda_3$ and Λ_4 plotted against $\mathcal{K}$ for $\alpha = \frac{1}{3}$ and $\alpha = \frac{3}{4}$. The curve of the speed of convergence near the steady-state, λ, is also shown (horizontal) for comparison. In all cases $n = 0.01, g = 0.02$ and $\delta = 0.05$.

The graphs in figures 5 and 6 show that speeds of convergence generally possess a negative relationship with α, that is, the higher the values of α the lower the convergence speeds. The absolute deviations (errors) between the exact speeds of convergence and λ are shown to be generally smaller for the measures based on the proportional rate of decrease of the gap $(k - \hat{k})$ $\big($ or $(\ln k - \ln \hat{k})\big)$ than for those based on the proportional rate of change of the slope of k $\big($ or $\ln k\big)$. Thus, as the economy evolves to the steady-state, the quantities $|\Lambda_1 - \lambda|$ and $|\Lambda_3 - \lambda|$ are always smaller than $|\Lambda_2 - \lambda|$ and $|\Lambda_4 - \lambda|$.

Figure 7 shows how the proportional approximation error of using λ varies with α. It shows that the proportional deviations from λ are also smaller for Λ_1 and Λ_3 than for Λ_2 and Λ_4. Proportional deviations for capital are largest (in absolute terms) as $\alpha \to 0$ and smallest as $\alpha \to 1$, in the LV-frame. In the OV-frame, the proportional deviations are largest as $\alpha \to 1$ and smallest as $\alpha \to 0$. For output, in the OV-frame, the proportional deviations are shown to be smallest around $\alpha = \frac{1}{2}$

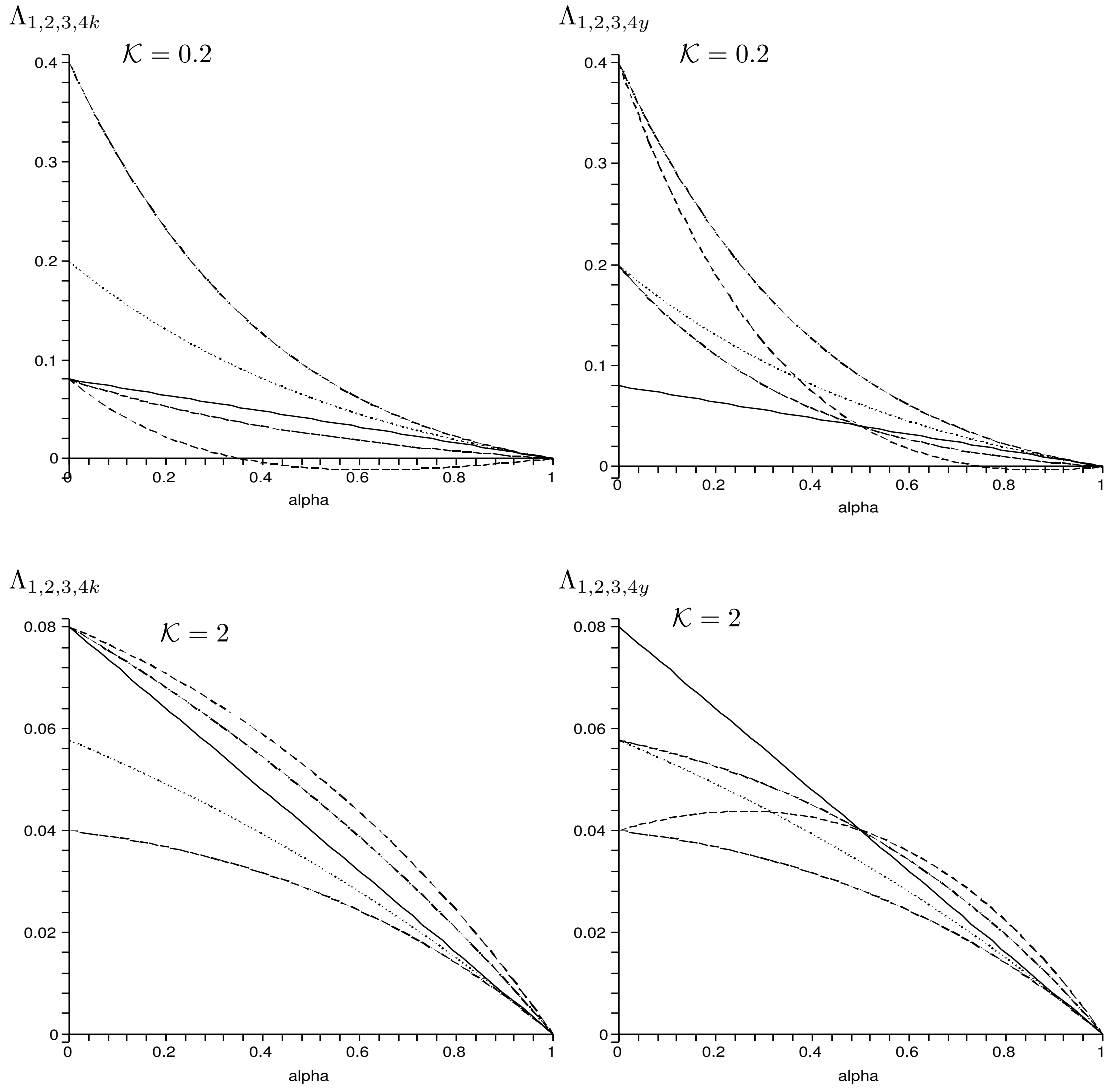

Figure 6: Speeds of convergence – – - Λ_1, - - - Λ_2, $\cdots\Lambda_3$ and –·–·– Λ_4 plotted against α for $\mathcal{K} = 0.2$ and $\mathcal{K} = 2$. In each case, the curve of the speed of convergence near the steady-state, λ, is also shown (solid) for comparison. In all cases $n = 0.01, g = 0.02$ and $\delta = 0.05$.

and maximal in the limits $\alpha \to 0$ and $\alpha \to 1$.

A weakness of the Λ_2 definition is that, away from the neighbourhood of the steady-state, it can give negative values of speed. Convergence speeds Λ_{2k} are negative whenever $(1 - \alpha\mathcal{K}^{\alpha-1}) < 0$, and Λ_{2y} values are negative whenever $\left[\alpha - (2\alpha - 1)\mathcal{K}^{\alpha-1}\right] < 0$. The reason for this is that, geometrically, Λ_{2k} is the (negative) instantaneous slope of the $\dot{k}$ curve on the phase plane $(k, \dot{k})$, and takes on negative values whenever the curve is up-sloping (see illustration in figure 3). Negative speeds of convergence are misleading because they give the (wrong) impression that the economy is evolving away from its steady-state level. Consequently, for the rest of the analysis here, we focus on the Λ_1, Λ_3, and Λ_4 definitions.

The foregoing analysis has demonstrated that, whereas all speed of convergence definitions yield the same value for both $k(t)$ and $y(t)$ near the steady-state, different definitions generally give different speeds of convergence far from steady-state. The

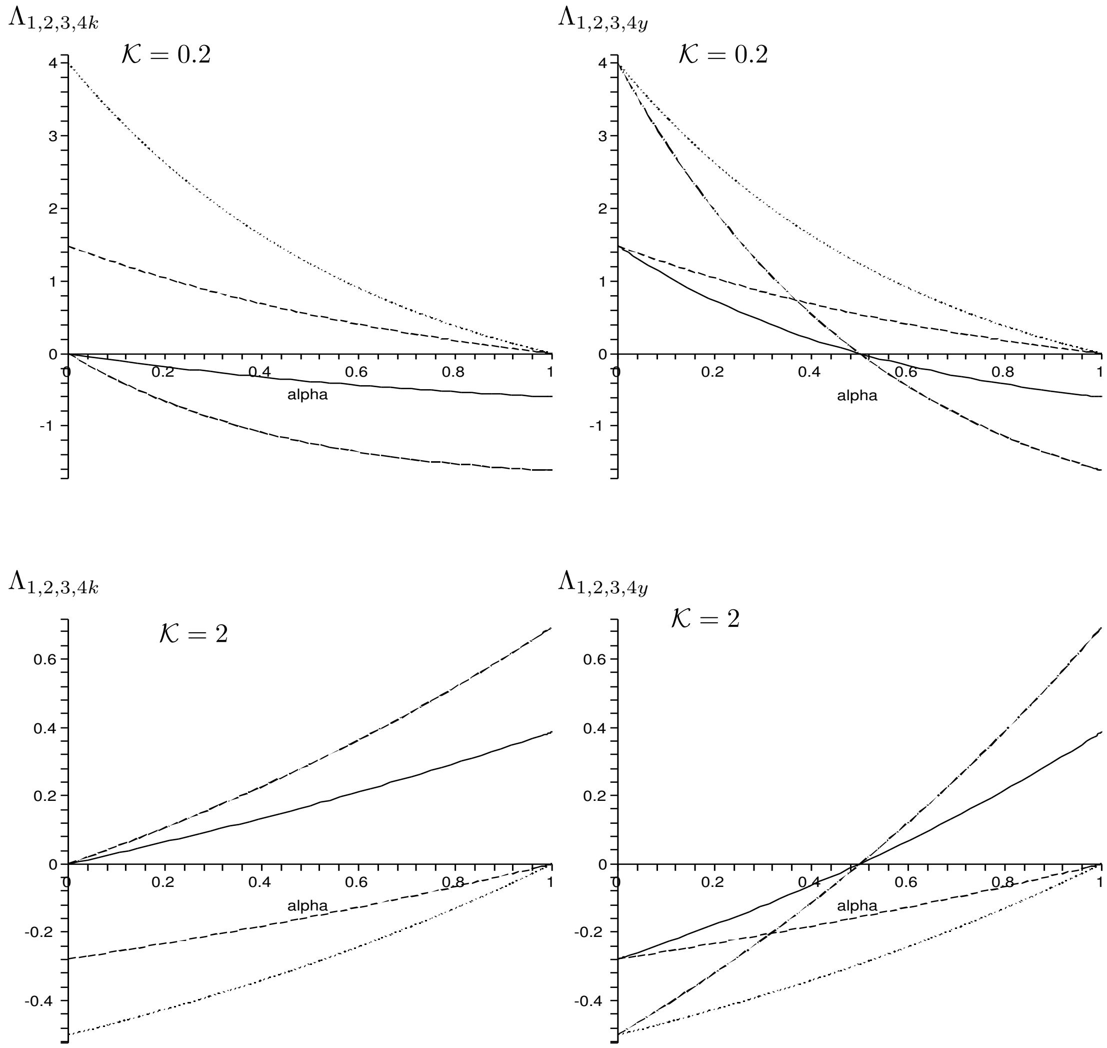

Figure 7: Proportional error — $\Lambda_1/\lambda - 1$, $-\cdot-\cdot$ $\Lambda_2/\lambda - 1$, - - -$\Lambda_3/\lambda - 1$ and $\cdots$ $\Lambda_4/\lambda - 1$ plotted against α for $\mathcal{K} = 0.2$ and $\mathcal{K} = 2$. In all cases $n = 0.01, g = 0.02$ and $\delta = 0.05$.

differences are essentially caused by changes in the concavity of the function (that occur as a result of expressing in terms of different variables) that determines the time evolution of the system in the phase plane. Whenever this function is concave up, speeds of convergence are higher than λ to the left of $\mathcal{K} = 1$, and slower than λ to the right of $\mathcal{K} = 1$. This can be explained using the graphical illustrations of the speeds of convergence, shown in figures 3 and 4. Consider the line whose (negative) slope gives Λ_{3k} in figure 4, for example. Call this line $\ell\Lambda_{3k}$, say. Since the evolution function is concave up, $\ell\Lambda_{3k}$ will always lie above the curve. This means, to the left of $\mathcal{K} = 1$, $\ell\Lambda_{3k}$ will be steeper than the gradient of the curve at $\mathcal{K} = 1$, while to the left, $\ell\Lambda_{3k}$ will be flatter. Since the (negative) slope of the gradient at $\mathcal{K} = 1$ equals λ, whenever $\ell\Lambda_{3k}$ is steeper then $\Lambda_{3k} > \lambda$ and if $\ell\Lambda_{3k}$ is flatter then $\Lambda_{3k} < \lambda$. The reverse happens if the curve is concave down, and then speeds of convergence are lower than λ to the left of $\mathcal{K} = 1$, and higher than λ to the right of $\mathcal{K} = 1$.

4.2 Half-life times

Since the theoretically predicted length of time required by an economy to fully attain its balanced growth equilibrium is infinity, it is conventional to use the notion of half-life to compare convergence speeds. Short half-life times indicate high speeds of convergence while long half-life times imply low speeds of convergence. There are two commonly used definitions of the half-life, an OV-based definition and an LV-based one. We derive expressions for half-life times of $y(t)$ and $k(t)$ based on these definitions, and then discuss their properties.

In terms of OVs, the half-life is defined as the time it takes for the gap $\left(k-\hat{k}\right)$ to decrease by half. Finding the half-life T_k involves solving the equation

$$\tfrac{1}{2}(k_0-\hat{k}) = k(T_k) - \hat{k},$$

which can be written as

$$k(T_k) = \tfrac{1}{2}(\hat{k}+k_0), \tag{48}$$

or

$$\mathcal{K}(T_k) = \tfrac{1}{2}(1+\mathcal{K}_0). \tag{49}$$

The equation (48) says that T_k marks the time at which $k(t)$ equals the arithmetic mean of k_0 and $\hat{k}$.

In terms of logarithmic variables, the half-life is defined as the time it takes for the log distance $\left(\ln k - \ln \hat{k}\right)$ to be halved. We use $\mathcal{T}_k$ to denote half-life times derived in this way. The condition satisfied by $\mathcal{T}_k$ is

$$\tfrac{1}{2}(\ln k_0 - \ln \hat{k}) = \ln k(\mathcal{T}_k) - \ln \hat{k},$$

which yields

$$k(\mathcal{T}_k) = \sqrt{\hat{k}k_0}, \tag{50}$$

and hence

$$\mathcal{K}(\mathcal{T}_k) = \sqrt{\mathcal{K}_0}. \tag{51}$$

Equation (50) indicates that the log version of half-life corresponds to the time at which $k(t)$ equals the geometric mean of k_0 and $\hat{k}$.

Consequently, the two approaches will generally give different values of the half-life. However, noting that $\sqrt{\mathcal{K}_0} \simeq \frac{1}{2}(1+\mathcal{K}_0)$ in the neighbourhood of $\mathcal{K}_0 = 1$ shows that, in the vicinity of the steady-state, half-life times obtained using the two approaches will be comparable. From the equations (31) and (36), the evolution paths of economies in the vicinity of $k=\hat{k}$, are given by

$$\begin{aligned} k(t)-\hat{k} &= e^{-\lambda t}(k_0-\hat{k}), \quad \text{and} \qquad (52)\\ \ln k(t) - \ln \hat{k} &= e^{-\lambda t}(\ln k_0 - \ln \hat{k}), \qquad (53) \end{aligned}$$

in the two frames, respectively. Substituting (52) and (53) into the equations (49) and (51), respectively, yields

$$\tilde{T} = \tfrac{1}{\lambda}\ln 2 \tag{54}$$

in both cases. Hence, for economies that evolve from within the neighbourhood of $k = \hat{k}$ (that is $\mathcal{K}_0 \sim 1$), the half-life is independent of the initial distance from steady-state, but depends only on λ.

For economies evolving from outside the neighbourhood of the steady-state, the expressions (52) and (53) are no longer applicable and we use the exact solutions (16) and (17) to compute the half-life times.

First, substituting (16) into the OV-based definition (49) gives

$$\left[\left(1 - e^{-\lambda T_k}\right) + \mathcal{K}_0^{1-\alpha} e^{-\lambda T_k}\right]^{\frac{1}{1-\alpha}} = \tfrac{1}{2}(1 + \mathcal{K}_0),$$

from which the half-life time is given by[9]

$$T_k = \frac{1}{\lambda}\left\{\ln\left|\mathcal{K}_0^{1-\alpha} - 1\right| - \ln\left|\left[\tfrac{1}{2}(\mathcal{K}_0 + 1)\right]^{1-\alpha} - 1\right|\right\}. \tag{55}$$

Similar calculations for $y(t)$ give

$$T_y = \frac{1}{\lambda}\left\{\ln\left|\mathcal{K}_0^{1-\alpha} - 1\right| - \ln\left|\left[\tfrac{1}{2}(\mathcal{K}_0^{\alpha} + 1)\right]^{\frac{1}{\alpha}-1} - 1\right|\right\}, \tag{56}$$

which is generally unequal to (55) except when $\mathcal{K}_0 = 1$. The half-life time expression (56) has previously been established by Reiss (2000).

Substituting (16) into the LV-based definition (51) and solving for $\mathcal{T}_k$ gives

$$\mathcal{T}_k = \frac{1}{\lambda}\ln\left(\mathcal{K}_0^{\frac{1}{2}(1-\alpha)} + 1\right), \tag{57}$$

and we have $\mathcal{T}_k = \mathcal{T}_y$ in this case.[10]

Again, along the balanced growth path where $\mathcal{K}_0 = 1$, all formulas (55)–(57) converge to the half-life of $\tilde{T} = \frac{1}{\lambda}\ln 2$ as predicted by the linear Taylor expansions.

Away from $\mathcal{K}_0 = 1$, the half-life times depend on both λ and $\mathcal{K}_0$, and the times given by (55), (56), and (57) are then significantly different. The variations of T_k, T_y and $\mathcal{T}_k$ with $\mathcal{K}_0$ and α are illustrated graphically in figures 8 and 9. We compare half-life times obtained in the two frames for $k(t)$ and $y(t)$.

OV-based half-life times: For economies that start below their steady-states, the half-life times T_k are longer than $\tilde{T}$, irrespective of the value of α – and vice versa is true for $\mathcal{K}_0 > 1$.

For output, provided $0 < \alpha < \frac{1}{2}$, the half-life times T_y are shorter than $\tilde{T}$ if $\mathcal{K}_0 < 1$ and longer if $\mathcal{K}_0 > 1$. When $\alpha = \frac{1}{2}$, the half-life $T_y = \tilde{T}$ irrespective of the value of $\mathcal{K}_0$. When $\frac{1}{2} < \alpha < 1$, the situation is the reverse of the case when $0 < \alpha < \frac{1}{2}$.

LV-based half-life times: In this case, the half-life times $\mathcal{T}_y = \mathcal{T}_k$ are shorter than $\tilde{T}$ whenever $\mathcal{K}_0 < 1$ and longer if $\mathcal{K}_0 > 1$.

[9] See appendix for the complete solution.

[10] See the appendix for complete solution.

We remark that the OV-based and LV-based half-times give results that are consistent with the OV-based and LV-based definitions of speed of convergence, respectively. For example, OV-based definitions gives slower speeds of convergence $\Lambda_{1,2k}$ than λ whenever $\mathcal{K}_0 < 1$. This means, in the OV-based frame of reference, an economy approaching steady-state from below actually takes longer than predicted by λ – and hence $T_k > \tilde{T}$. In the LV-based frame, however, $\Lambda_{3,4k} > \lambda$ whenever $\mathcal{K}_0 < 1$ implying that the economy takes a shorter time to converge than predicted by λ. Hence the half-life times $\mathcal{T}_k < \tilde{T}$.

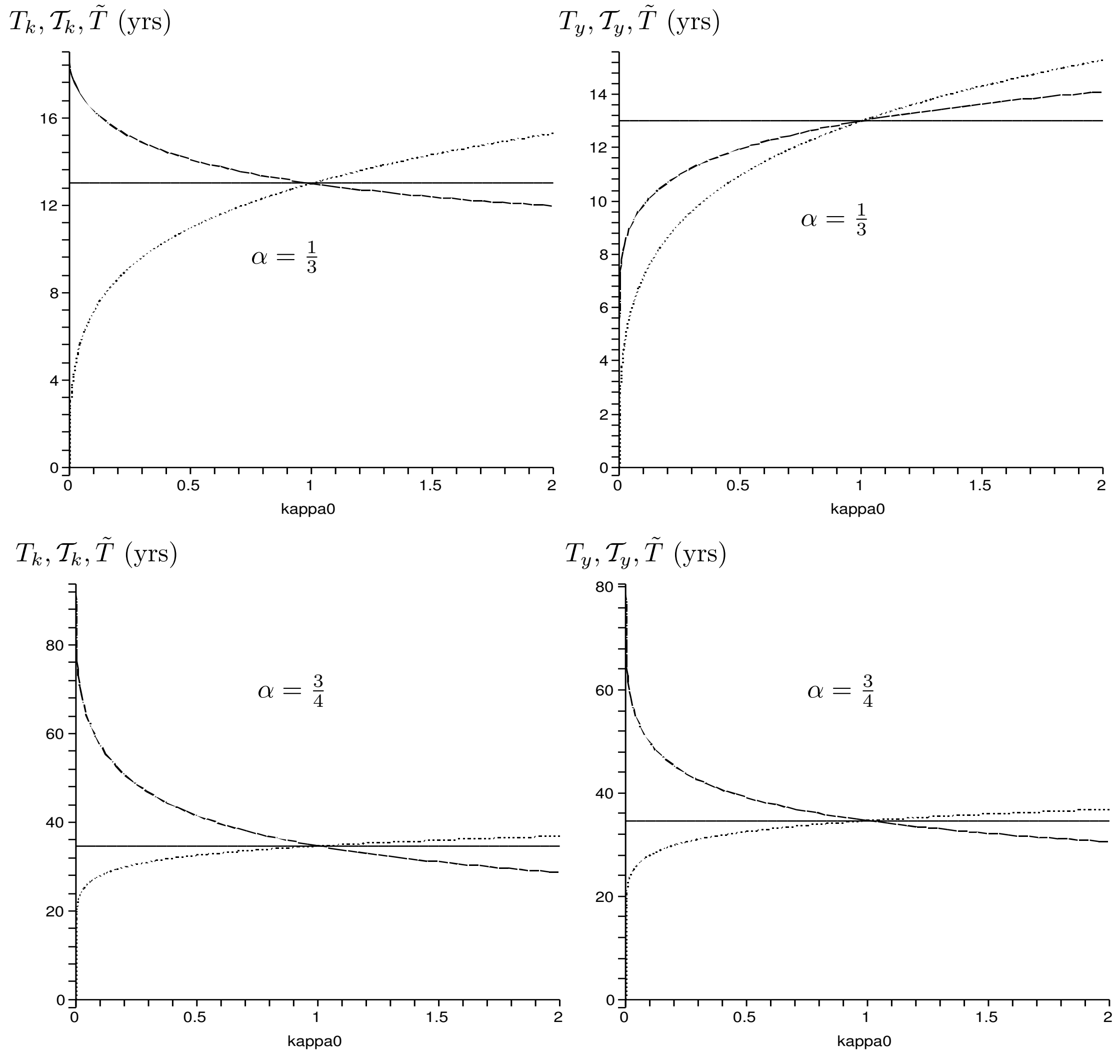

Figure 8: Half-life times for k (on the left) and y ($\cdots \mathcal{T}_y = \mathcal{T}_k$, –·–·– T_k, T_y) plotted as functions of $\mathcal{K}_0$ for $\alpha = \frac{1}{3}$ and $\frac{3}{4}$. The curve of $\tilde{T}$, is also shown (solid) in each case for comparison. In all cases $n = 0.01, g = 0.02$ and $\delta = 0.05$.

4.3 Discussion

Our analysis has demonstrated that definitions of the speed of convergence and half-life times can be derived in either an OV-based frame or an LV-based frame. It has been shown that all definitions give the same speed of convergence $\lambda = (1-\alpha)(n+g+\delta)$ and half-life $\tilde{T} = \frac{1}{\lambda}\ln 2$ in the neighbourhood of the steady-state.

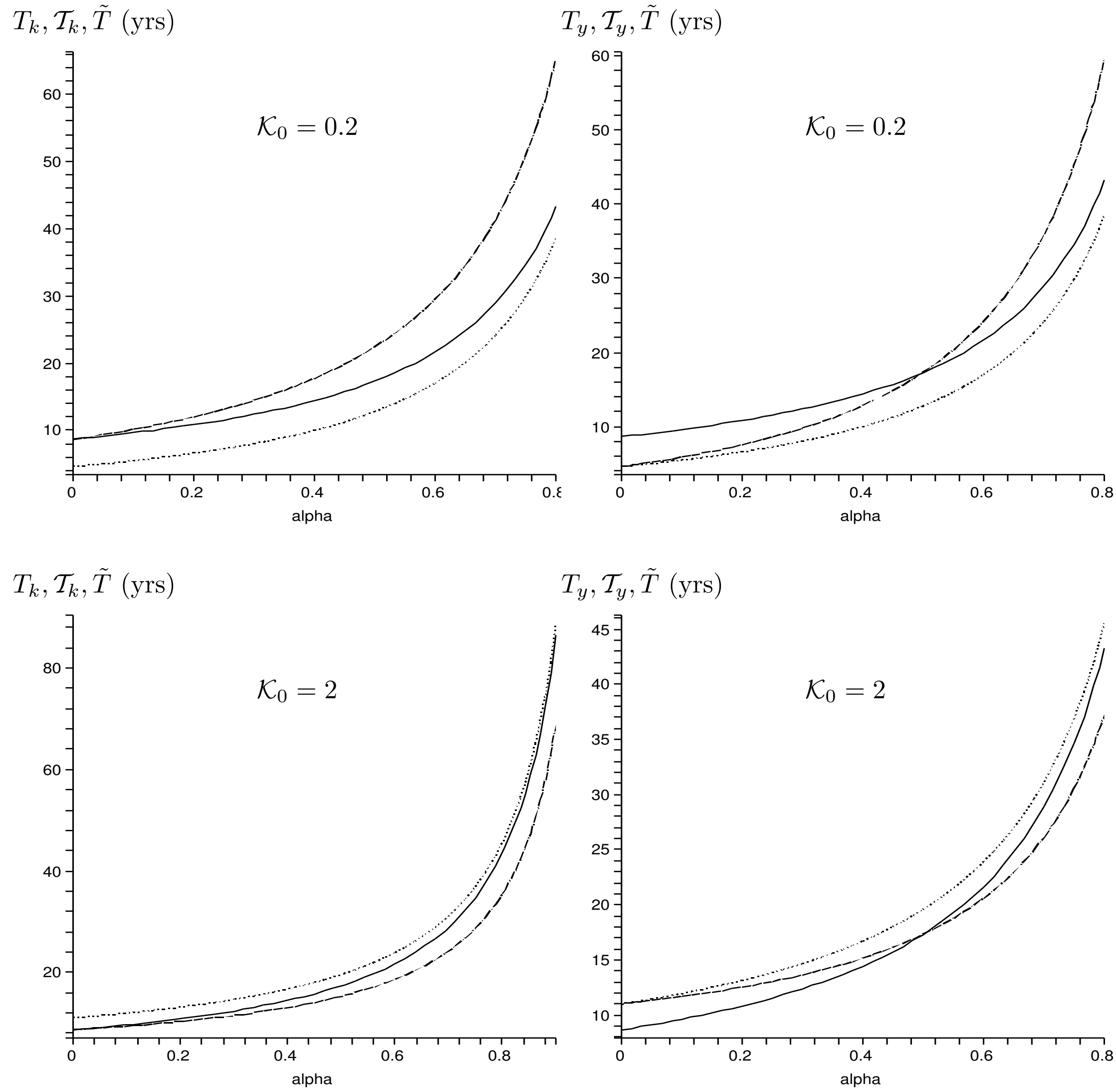

Figure 9: Half-life times for k (on the left) and y ($\cdots\, \mathcal{T}_y = \mathcal{T}_k$, $-\cdot-\cdot-\ T_k, T_y$) plotted as functions of α for $\mathcal{K}_0 = 0.2$ and $\mathcal{K}_0 = 2$. The curve of $\tilde{T}$, is also shown (solid) in each case for comparison. In all cases $n = 0.01, g = 0.02$ and $\delta = 0.05$.

Far from the steady-state however, convergence properties derived in the two frames are generally inconsistent. We have used the case of a Cobb-Douglas production function to demonstrate that OV-based measures generally yield different speeds of convergence and half-life times for output and capital. Provided $\frac{1}{2} < \alpha < 1$, speeds of convergence for output and capital are slower than, and increase towards λ if the economy is evolving from below the steady-state, and are faster than, and decrease towards λ if the economy evolves from above. When $0 < \alpha < \frac{1}{2}$, the convergence measures for output and capital are shown to exhibit opposite evolution trends as the economy approaches the steady-state.

An advantage of the LV-based frame is that, for any value of α, the convergence measures for output and capital are always equal. For economies below steady-state, speeds of convergence are shown to be faster than λ, and decrease towards λ as the economy approaches steady-state. Above steady-state, speeds of convergence are slower than λ, and increase towards λ as the economy approaches steady-state.

Therefore, far from steady-state, measures of convergence derived in the OV-based frame generally yield different results to those derived in the LV-based frame.

5 Empirical implications

A key implication of exogenous growth models is conditional convergence. There have been numerous empirical studies aimed at testing the hypothesis of conditional convergence against the data. Based on equations derived by log-linearising about the steady-state, these studies typically assume that the speed of convergence exhibited by economies is constant and independent of the distance from steady-state. The theoretical analysis of the preceding sections, however, has indicated that, except in the close vicinity of steady-state, the speed of convergence depends on both capital's share of output and the economy's distance from its steady-state. Moreover, economies approaching steady-state from below have been predicted to have faster (than λ) LV-based speeds of convergence while economies approaching from above have slower speeds of convergence.

In this section, we consider the empirical implication of the preceding analysis. In the next section, we will test these implications. Our empirical work builds upon the approach introduced by Mankiw, Romer and Weil (1992) (referred to as MRW henceforth) and extended by Cho and Graham (1996) and Okada (2006). We test the hypothesis that economies converging from below have higher speeds of convergence than those converging from above. Using the data set of MRW, we split the sample into two groups based on whether an economy was above or below their predicted steady-state in 1960. We then perform a nonlinear regression that allows us to test whether speeds of convergence are significantly different between the two groups. Our results show that the hypothesis is supported by the data.

5.1 The Augmented Solow model

5.1.1 The model

Following MRW, we adopt a Cobb-Douglas production function which incorporates human capital and takes the form

$$Y = K^{\alpha} H^{\beta} (AL)^{1-\alpha-\beta}, \quad 0 < \alpha + \beta < 1, \tag{58}$$

where H is the human stock of capital and all other variables are defined as before. If s_k and s_h are the fraction of output invested in physical and human capital, respectively, and human capital is assumed to depreciate at the same rate as physical capital, then the economy evolves according to

$$\begin{aligned} y(t) &= \big[k(t)\big]^{\alpha}\big[h(t)\big]^{\beta}, &(59)\\ \dot{k}(t) &= s_k y(t) - (n+g+\delta)k(t), &(60)\\ \dot{h}(t) &= s_h y(t) - (n+g+\delta)h(t), &(61) \end{aligned}$$

where $h = H/(AL)$. Eliminating y between the equations (60) and (61) shows that human capital and physical capital satisfy the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{k(t)}{s_k} - \frac{h(t)}{s_h}\right) = -(n+g+\delta)\left(\frac{k(t)}{s_k} - \frac{h(t)}{s_h}\right) \tag{62}$$

the solution of which, is

$$\frac{k(t)}{s_k} - \frac{h(t)}{s_h} = e^{-(n+g+\delta)t}\left(\frac{k_0}{s_k} - \frac{h_0}{s_h}\right). \tag{63}$$

This expression indicates that, at steady-state, the condition

$$k(t)/s_k = h(t)/s_h \tag{64}$$

is satisfied irrespective of the initial levels of human capital (h_0) and physical capital (k_0).

Under the assumption that the initial levels of human and physical capital are such that is $k_0/h_0 = s_k/s_h$ from the outset, it is possible to find an exact solution of the system (59)–(61).[11] It is instructive to study this case because exact expressions of quantities can be found even outside the vicinity of the steady-state. The time evolution of output per effective unit of labour is given by

$$y(t) = \left(\frac{s_h}{s_k}\right)^{\beta}\left\{\hat{k}^{1-\alpha-\beta}\left(1-e^{-\lambda_h t}\right) + k_0^{1-\alpha-\beta}e^{-\lambda_h t}\right\}^{\frac{\alpha+\beta}{1-\alpha-\beta}}, \tag{65}$$

where $\lambda_h = (1-\alpha-\beta)(n+g+\delta)$.[12] The economy therefore evolves towards the steady-state

$$\hat{y} = \left(\frac{s_h}{s_k}\right)^{\beta}\hat{k}^{\alpha+\beta} = \left\{\frac{s_k^{\alpha}s_h^{\beta}}{(n+g+\delta)^{\alpha+\beta}}\right\}^{\frac{1}{1-\alpha-\beta}}. \tag{66}$$

The steady-state output per capita can be obtained from this expression by taking logs and is given by

$$\ln\frac{\hat{Y}(t)}{L(t)} = \ln A_0 + gt - \tfrac{\alpha+\beta}{1-\alpha-\beta}\ln(n+g+\delta) + \tfrac{\alpha}{1-\alpha-\beta}\ln s_k + \tfrac{\beta}{1-\alpha-\beta}\ln s_h. \tag{67}$$

Another equation involving steady-state output per capita can be derived by log-linearising the differential equation for $y(t)$ about the steady-state.[13] The solution of the linear log expansion can be expressed in the form

$$\ln y(t) - \ln y_0 = \theta\big(\ln\hat{y} - \ln y_0\big), \tag{68}$$

where $\theta = 1 - e^{-\lambda_h t}$ and λ_h is the speed of convergence near the steady-state. Expressing in terms of per capita variables and combining with (67) gives

$$\ln\frac{Y_t}{L_t} - \ln\frac{Y_0}{L_0} = gt + \theta\ln A_0 - \tfrac{\alpha+\beta}{1-\alpha-\beta}\theta\ln(n+g+\delta) + \tfrac{\alpha}{1-\alpha-\beta}\theta\ln s_k + \tfrac{\beta}{1-\alpha-\beta}\theta\ln s_h - \theta\ln\frac{Y_0}{L_0}, \tag{69}$$

[11] See appendix for complete solution.

[12] Note that we use the notation $\lambda_h = (1-\alpha-\beta)(n+g+\delta)$ to denote speed of convergence in the augmented Solow framework, compared to $\lambda = (1-\alpha)(n+g+\delta)$ in the framework of the traditional Solow model.

[13] See appendix for derivation.

where the left hand side is the growth of output per capita over the period (not annualized). This framework, based on log-linearising about the steady-state, takes the speed of convergence to be constant irrespective of the distance from, and the direction of approach to, the steady-state. However, as our analysis in section 4 has shown, the speed of convergence for economies converging from outside the vicinity of steady-state is dependent on both the distance from and the direction of approach to steady-state.

For, consider the speed of convergence parameter Λ_3 (introduced in section 3) which can be expressed in the form

$$\Lambda_{3y}(t) = -\frac{\mathrm{d}\Big[\ln\big(y/\hat{y}\big)\Big]\Big/\mathrm{d}t}{\ln\big(y/\hat{y}\big)} = -\frac{\mathrm{d}}{\mathrm{d}t}\Big\{\ln\Big(\ln\big(y/\hat{y}\big)\Big)\Big\}. \tag{70}$$

Since the right hand side is an exact differential, (70) can be readily integrated and rearranged to give

$$\ln y(t) - \ln y_0 = \Theta\big(\ln\hat{y} - \ln y_0\big), \tag{71}$$

where $\Theta = 1 - \exp\left\{-\int_0^t \Lambda_{3y}(\tau)\mathrm{d}\tau\right\}$. This expression (71) is to be compared with (68) in which the speed of convergence is constant.

For the case under consideration (a Cobb-Douglas production function under the assumption that $k_0/h_0 = s_k/s_h$ from the outset), we can substitute the exact solution for $y(t)$ into (70) to obtain an analytical expression for $\Lambda_{3y}(t)$, namely

$$\Lambda_{3y}(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\left\{\ln\left(\ln\Big[1 + e^{-\lambda_h t}\big(\mathcal{K}_0^{1-\alpha-\beta} - 1\big)\Big]^{\frac{\alpha+\beta}{1-\alpha-\beta}}\right)\right\}. \tag{72}$$

The equations (70) and (72) indicate that $\Lambda_{3y}(t)$ (and hence Θ), is generally a function of time, approaching λ_h in the limit $t \to \infty$. Figures 10 and 11 show the time evolution of $\Lambda_{3y}(t)$ and the relative error $(\Theta - \theta)/\theta$, respectively, for economies with $\alpha = \frac{1}{3}$, starting from different initial levels.[14] The value of $\alpha = \frac{1}{3}$ is conventionally used within the framework of the traditional Solow model. For economies that begin below/above their steady-states, the speeds of convergence $\Lambda_{3y}(t)$ are shown to be initially faster/slower than, and decrease/increase towards λ in time. The rate at which $\Lambda_{3y}(t)$ converges onto λ is shown to be higher for economies that start below their steady-states compared to those that start from above.

Figures 12 and 13 show the time evolution of $\Lambda_{3y}(t)$ and the relative error $(\Theta - \theta)/\theta$, respectively, for economies with $\alpha = \frac{2}{3}$. This value of the capital share is typical in the framework of the augmented Solow model. The graphs show the behaviour of the quantities $\Lambda_{3y}(t)$ and $(\Theta - \theta)/\theta$ in time to be qualitatively similar to that of the $\alpha = \frac{1}{3}$ case. The main differences are that, in this case, the relative errors are comparatively lower[15] and the development of Λ_{3y} on both sides of steady-state appears to be fairly symmetric.

[14]Note: The initial positions relate to the predicted steady-state levels in 1960, and correspond to 10-, 25-, 50-, 75-, and 90-percentiles computed for the MRW sample.

[15]While the low relative errors are partly due to smaller initial deviations from the steady-state predicted in this model, our calculations show that higher values of α, keeping everything else fixed, generally lead to lower relative errors (see Appendix 6.2).

These results mean that λ generally over-estimates the actual speeds of convergence and hence the coefficient Θ for economies that approach steady-state from above. Equally, the coefficient $\theta = 1 - e^{\lambda t}$ will generally under-estimate its counterpart Θ for economies that converge from below.

6 Convergence from both sides

In this section, we investigate whether economies that approach a steady-state growth path from below have higher speeds of convergence than those that converge from above. Our starting point will be the papers by Mankiw, Romer and Weil (1992) and Cho and Graham (1996). We use the method of Cho and Graham to identify countries that may be converging from above, and then estimate the augmented Solow model by nonlinear least squares, allowing a different rate of convergence for the countries that converge from above. Our empirical test complements an alternative, more complex approach developed by Okada (2006).

The group of countries is the main, 'non-oil' sample used in MRW. For this group of countries, we calculate whether output per worker was above or below the steady-state level in 1960. To determine this, we first run the MRW growth regression

$$\ln\frac{Y_{85}}{L_{85}} - \ln\frac{Y_{60}}{L_{60}} = b_0 + b_1(\ln s_k - \ln(n+g+\delta)) + b_2(\ln s_h - \ln(n+g+\delta)) + b_3 \ln\frac{Y_{60}}{L_{60}}, \quad (73)$$

based on equation (69). Here Y_{85} denotes output in 1985. Once the coefficients in (73) have been determined, they are combined with (67) to derive an equation that gives 1960 steady-state per capita output levels.[16] Notice that, since 1960 corresponds to time $t = 0$, equation (67) yields

$$\begin{aligned} \ln\frac{\hat{Y}_{60}}{L_{60}} &= \ln A_{60} + \tfrac{\alpha}{1-\alpha-\beta}(\ln s_k - \ln(n+g+\delta)) + \tfrac{\beta}{1-\alpha-\beta}(\ln s_h - \ln(n+g+\delta)) \\ &= -\tfrac{1}{b_3}\Big[b_0 - gt + b_1(\ln s_k - \ln(n+g+\delta)) + b_2(\ln s_h - \ln(n+g+\delta))\Big], \quad (74) \end{aligned}$$

where the last expression is obtained by comparing (69) with (73), for which the coefficients have already been determined. Note that we have employed the widely used *a priori* value of $g = 0.02$, corresponding to technical progress of 2% per year, in our calculations. We are also imposing the theoretical restriction that the coefficients on the investment, schooling and population growth terms sum to zero.

The steady-state output values, once found, can be compared with the observed level of output per capita in 1960. We compute the ratio $\mathcal{Y}_{60} = Y_{60}/\hat{Y}_{60} = y_{60}/\hat{y}_{60}$, on the basis of which the sample is split into two groups depending on whether $\mathcal{Y}_{60} \leq 1$ or $\mathcal{Y}_{60} > 1$. This is close to the method adopted in Cho and Graham (1996). Okada (2006) uses a related method.

When we adopt this approach, we find that 49 of the 98 countries are classified as converging from above. This may seem surprising, but output per capita grew

[16] Notice that this is the regression presented in table VI of MRW, and also shown in the first column of Table 1 here.

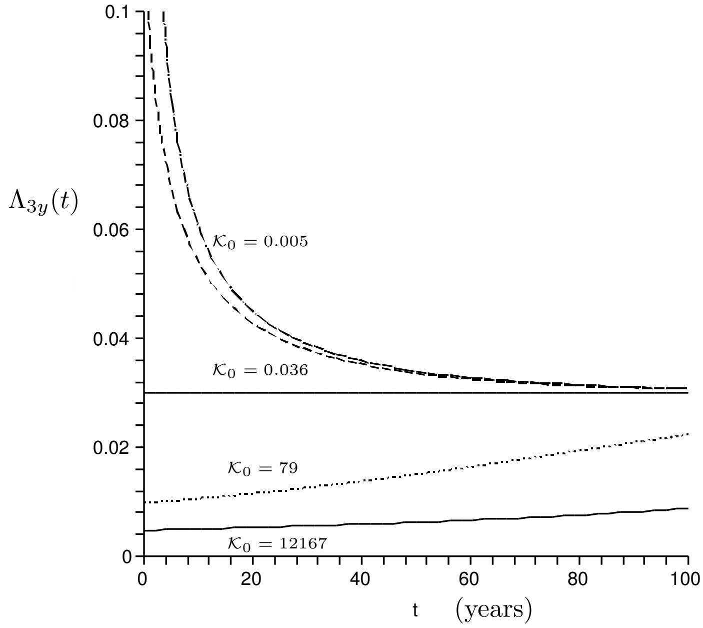

Figure 10: The time evolution of $\Lambda_{3y}(t)$ for economies that start from $\mathcal{K}_0 = 0.005, 0.036, 79$ and $\mathcal{K}_0 = 12167$. In all cases $n = 0.01, g = 0.02, \delta = 0.05, \alpha = \frac{1}{3}$. The graph of $\boldsymbol{\lambda} = 0.03$ is shown as a solid line.

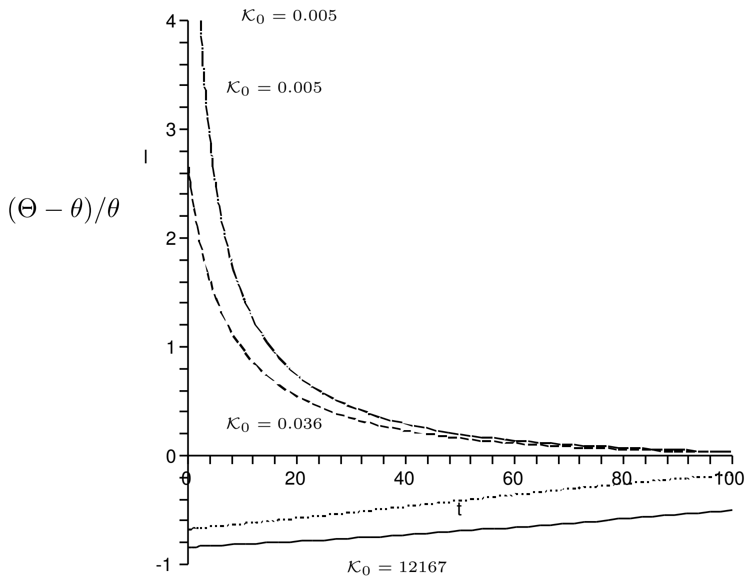

Figure 11: The time evolution of the relative error $(\Theta - \theta)/\theta$ for economies that start from $\mathcal{K}_0 = 0.005, 0.036, 79$ and $\mathcal{K}_0 = 12167$. In all cases $n = 0.01, g = 0.02, \delta = 0.05, \alpha = \frac{1}{3}$.

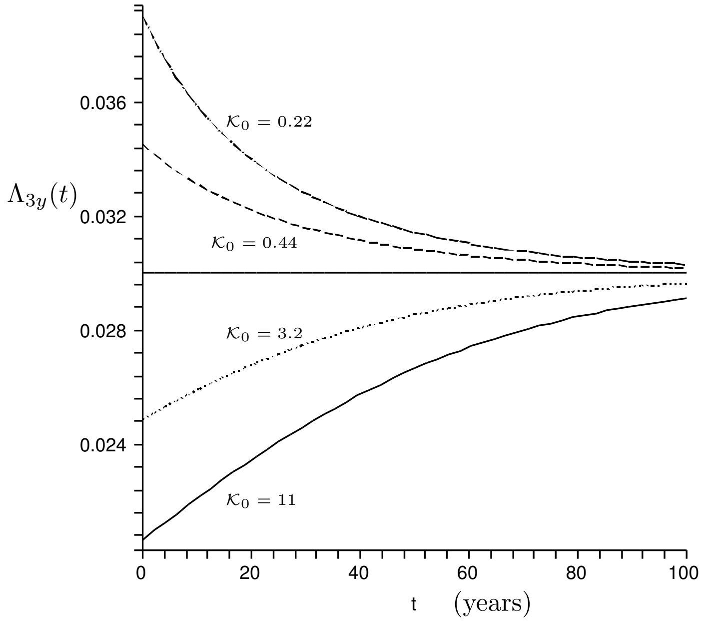

Figure 12: The time evolution of $\Lambda_{3y}(t)$ for economies that start from $\mathcal{K}_0 = 0.22, 0.44, 3.2$ and $\mathcal{K}_0 = 11$. In all cases $n = 0.01, g = 0.02, \delta = 0.05, \alpha = \frac{2}{3}$. The graph of $\lambda_h = 0.03$ is shown as a solid line.

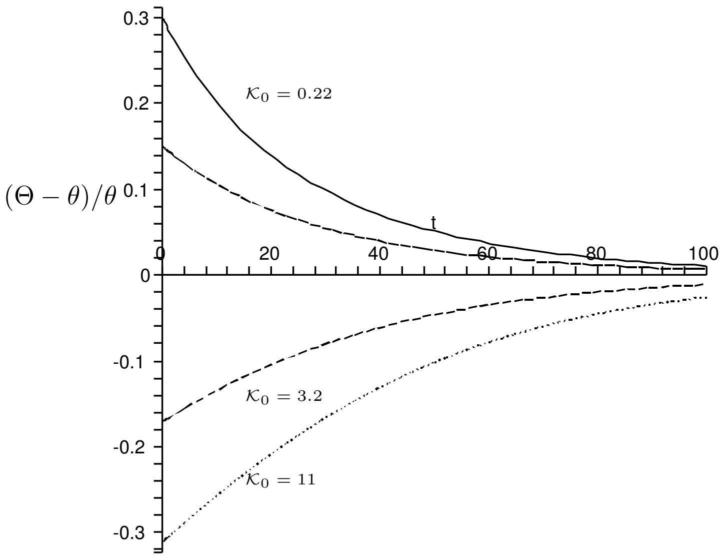

Figure 13: The time evolution of the relative error $(\Theta - \theta)/\theta$ for economies that start from $\mathcal{K}_0 = 0.22, 0.44, 3.2$ and $\mathcal{K}_0 = 11$. In all cases $n = 0.01, g = 0.02, \delta = 0.05, \alpha = \frac{2}{3}$.

at less than 2% a year over 1960-85 in a large number of countries.[17] Under the assumptions of the augmented Solow model, they must have been converging from above, perhaps reflecting declines in investment, or increases in population growth. Note that converging to a growth path from above does not imply strictly negative growth in output per head, given the maintained assumption that the level of efficiency is growing over time.

Now that we have classified the set of countries into 'above' and 'below' we can test the theoretical implication that convergence is slower for countries converging from above. Our starting point will be growth regressions of the form derived by MRW:

$$\begin{aligned} \ln \frac{Y_{85}}{L_{85}} - \ln \frac{Y_{60}}{L_{60}} &= \theta \ln A_{60} + gt + \frac{\alpha}{1-\alpha-\beta}\theta(\ln s_k - \ln(n+g+\delta)) \\ &\quad + \frac{\beta}{1-\alpha-\beta}\theta(\ln s_h - \ln(n+g+\delta)) - \theta \ln \frac{Y_{60}}{L_{60}} \end{aligned}$$

A key prediction of our theoretical analysis is that the parameter θ should be smaller in absolute terms for countries converging from above, reflecting slower convergence. If we create a dummy variable, d_{above}, that is equal to one for countries that are classified as 'above' and zero otherwise, we can rewrite the right-hand-side of the growth regression in the following form:

$$gt + (1+\gamma d_{above})\theta \left[b_0 + b_1(\ln s_k - \ln(n+g+\delta)) + b_2(\ln s_h - \ln(n+g+\delta)) - b_3 \ln \frac{Y_{60}}{L_{60}} \right] \tag{75}$$

where the theoretical prediction is that the new parameter $\gamma < 0$. This prediction is easily tested by estimating (75) by nonlinear least squares.

First of all, we look at estimates for 1960-85. The data and sample are exactly that used by MRW. In the first column of Table 1, we show a replication of the MRW results based on their Table VI. In the second column, we show the outcome obtained by using nonlinear least squares (NLS). The parameter γ is negative, as predicted, but is significantly different from zero only at the 25% level.

It is possible that successfully detecting heterogeneity in convergence rates may require a longer span of data. We therefore carry out a similar analysis for 1960-2000, using data from version 6.1 of the Penn World Table (PWT) due to Heston, Summers and Aten (2002). We first use the Cho and Graham method to classify countries as converging from above, and then construct an updated version of the MRW regression.

Our output measure will be output per adult, using PWT data on output, and data on adult population from the World Bank's *World Development Indicators*. Since we do not have data on the MRW schooling variable for 1985-2000, we use their measure for 1960-85 as a proxy for the whole 1960-2000 period. Due to gaps in the PWT 6.1 data for the 1990s, we are missing data for 11 of the original 98

[17] It might be thought that countries could be classified as converging from above whenever their observed growth in output per capita is less than 2%. But this would create problems for our later statistical analysis, since it amounts to selection based on the dependent variable, and would therefore generate a bias.

observations in MRW.[18]

In column 3 of Table 6, we show the results obtained for MRW's growth regression, estimated for 87 countries over 1960-2000. Then, in the final column, we report the NLS estimates. In this case, γ is not only negatively signed, but also significant at the 5% level. This supports the theoretical prediction that convergence from above is relatively slow. Since the point estimate of γ is -0.436, the effect could be substantial. Taken at face value, the point estimate implies that the value of θ for countries converging from above is less than 60% of the value for countries converging from below. We can reinforce this point by using the estimates to calculate convergence rates for the two groups of countries. For those converging from below, the annual convergence rate is calculated at 2.88%; for countries converging from above, the rate is roughly halved, at 1.37%.

These results are subject to some familiar criticisms. First, equation (69) shows the familiar point that output growth depends on the initial level of technology, A_0. This can lead to omitted variable bias if the unobserved variable A_0 is correlated with any of the regressors. Second, we assume that the rates of technical progress are the same across economies (here, taken to be 2% per year). Third, we have used the standard log-linearized framework, strictly valid only close to the steady-state, to study behaviour away from equilibrium. A fourth and final criticism, especially relevant to our specific empirical tests, is that identifying examples of convergence from above is not straightforward. In particular, our use of the Cho-Graham method assumes that countries have remained above or below their steady-state growth paths throughout the period. This will be unsatisfactory to the extent that some have changed sides. Note, however, that many of these problems might work to hide the effect that we have identified in the data.

[18]These are Angola, the Democratic Republic of Congo, Germany, Haiti, Liberia, Myanmar, Sierra Leone, Singapore, Somalia, Sudan, and Tunisia.

Growth regressions				
Time period	60-85	60-85	60-00	60-00
Estimation	OLS	NLS	OLS	NLS
constant	2.46	2.45	3.29	3.56
	(0.473)	(0.836)	(0.652)	(1.12)
$\log(s_k) - \log(n+g+\delta)$	0.501	0.642	0.520	0.770
	(0.082)	(0.190)	(0.094)	(0.274)
$\log(s_h) - \log(n+g+\delta)$	0.235	0.336	0.332	0.453
	(0.059)	(0.114)	(0.079)	(0.194)
$\log(Y_0/L_0)$	-0.298	-0.397	-0.367	-0.513
	(0.060)	(0.121)	(0.082)	(0.184)
γ		**-0.302**		**-0.436**
		(0.246)		**(0.222)**
g		0.025		0.016
		(0.014)		(0.007)
N	98	98	87	87
R^2	0.48	0.49	0.53	0.54
α	0.48	0.47	0.43	0.44
β	0.23	0.24	0.27	0.26
λ_{below}	1.41%	2.02%	1.83%	2.88%
λ_{above}	1.41%	1.30%	1.83%	1.37%

Table 1: The dependent variable is the log difference of either output per equivalent adult (1960-85) or output per adult (1960-2000). Estimation is by ordinary least squares (OLS) or nonlinear least squares (NLS). Standard errors in parentheses. The last two rows of the Table indicate that convergence from above appears to be slower than convergence from below, as the analysis earlier in the paper predicts.

7 Conclusions

This paper has made a number of contributions to the study of convergence behaviour in exogenous growth models. We have investigated convergence away from the steady-state, and discussed a number of ways of measuring the rate of adjustment, extending the work of Reiss (2000). The analysis reveals that convergence rates are likely to be heterogeneous in systematic ways. In particular, we showed that, for log-linearized models of the kind commonly used in empirical work, rates of convergence are faster for economies that converge from below than for economies that converge from above. Using some straightforward modifications to cross-country growth regressions, we have shown that there is some support for this prediction in the data.

8 Appendix

8.1 Solution of the Solow equation

Starting with the Solow equation $\dot{k} = sk^\alpha - (n+g+\delta)k$, we divide through by k^α to get

$$\frac{\dot{k}}{k^\alpha} = s - (n+g+\delta)k^{1-\alpha}.$$

Then, setting $u = k^{1-\alpha}$ yields $\dot{k}/k^\alpha = \dot{u}/(1-\alpha)$, and the equation becomes

$$\dot{u} + (1-\alpha)(n+g+\delta)u = (1-\alpha)s.$$

This is linear in u and the integrating factor is $e^{\lambda t}$, where $\lambda = (1-\alpha)(n+g+\delta)$. Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(ue^{\lambda t}\Big) = (1-\alpha)se^{\lambda t},$$

which integrates to

$$u = k^{1-\alpha} = \frac{s}{n+g+\delta} + C_0 e^{-\lambda t},$$

where C_0 is a constant of integration. Use of the condition that $k = k(0) = k_0$ at $t = 0$ yields $C_0 = k_0^{1-\alpha} - s/(n+g+\delta)$, and the particular solution is given by

$$k(t) = \Big[\hat{k}^{1-\alpha}\Big(1 - e^{-\lambda t}\Big) + k_0^{1-\alpha}e^{-\lambda t}\Big]^{\frac{1}{1-\alpha}}. \tag{76}$$

8.2 Higher-order Expansions

In this section, we compare linear, quadratic and cubic log expansions and find that linear expansions are the most useful.

A common criticism of linear Taylor expansions is that their validity is strictly limited only to within the close vicinity of the steady-state. One possible response to this criticism is to include more terms in the expansions to increase the region of validity. Using the case of a Cobb-Douglas production function, we derive quadratic and cubic log expansions for the variable $y(t)$, compare their performances with that of the linear log expansion, and then discuss whether there are practical benefits of higher-order expansions.

Starting with

$$\begin{aligned} \dot{y}/y = \mathrm{d}\big[\ln y(t)\big]/\mathrm{d}t = \mathcal{G}(y) &= \alpha s y^{1-\frac{1}{\alpha}} - \alpha(n+g+\delta), \\ &= \alpha(n+g+\delta)\Big(\mathcal{Y}^{1-\frac{1}{\alpha}} - 1\Big), \end{aligned} \tag{77}$$

where, $\mathcal{Y} = y/\hat{y}$, we have

$$\begin{aligned} y\partial_y\mathcal{G} &= (\alpha-1)(n+g+\delta)\mathcal{Y}^{1-1/\alpha}, \\ (y\partial_y)^2\mathcal{G} &= \tfrac{1}{\alpha}(\alpha-1)^2(n+g+\delta)\mathcal{Y}^{1-1/\alpha}, \\ (y\partial_y)^3\mathcal{G} &= \tfrac{1}{\alpha^2}(\alpha-1)(n+g+\delta)\mathcal{Y}^{1-1/\alpha}, \end{aligned}$$

The Taylor log expansion of (77) about $y = \hat{y}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\ln\mathcal{Y} = \mathcal{G}(\hat{y}) + \ln\mathcal{Y}\Big[y\partial_y\mathcal{G}\Big]_{y=\hat{y}} + \tfrac{1}{2}\ln^2\mathcal{Y}\Big[(y\partial_y)^2\mathcal{G}\Big]_{y=\hat{y}} + \tfrac{1}{6}\ln^3\mathcal{Y}\Big[(y\partial_y)^3\mathcal{G}\Big]_{y=\hat{y}} + \cdots. \tag{78}$$

Hence, the linear, quadratic and cubic log expansions are given by

$$\dot{\mathcal{L}}_y(t) = -\lambda \mathcal{L}_y(t), \tag{79}$$

$$\dot{\mathcal{L}}_y(t) = -\lambda \mathcal{L}_y(t) + \tfrac{1}{2}\big(\tfrac{1}{\alpha} - 1\big)\lambda \mathcal{L}_y^2(t), \tag{80}$$

$$\dot{\mathcal{L}}_y(t) = -\lambda \mathcal{L}_y(t) + \tfrac{1}{2}\big(\tfrac{1}{\alpha} - 1\big)\lambda \mathcal{L}_y^2(t) - \tfrac{1}{6}\big(\tfrac{1}{\alpha} - 1\big)^2\lambda \mathcal{L}_y^3(t), \tag{81}$$

respectively, where $\mathcal{L}_y(t) = \ln \mathcal{Y} = \ln y(t) - \ln \hat{y}$, and $\lambda = (1-\alpha)(n+g+\delta)$. The degree to which each of the expansions (79)−(81) can estimate the original equation (77) is measured by how well the left-hand-side expressions approximate $\mathcal{G}(\mathcal{Y})$ in the neighbourhood of $\mathcal{Y} = 1$. Figure 14 shows graphs of $\mathcal{G}(\mathcal{Y})$ and the expansions.[19] All expansions are shown to be very accurate representations of $\mathcal{G}(\mathcal{Y})$ in the very close vicinity of $\mathcal{Y} = 1$. The interval in which the expansions are accurate generally increases with α, and in fact all expansions collapse onto $\mathcal{G}(\mathcal{Y})$ in the limit $\alpha \to 1$.[20]

For $0 < \alpha < 1$, the linear expansion always has the smallest interval in which it accurately reproduces $\mathcal{G}(\mathcal{Y})$. Outside this region, it under-estimates $\mathcal{G}(\mathcal{Y})$ on both sides of $\mathcal{Y} = 1$, but predicts the correct sign of $\mathcal{G}(\mathcal{Y})$ at all points. The usefulness of the linear expansion very close to the steady-state is based on the fact that, when $|\mathcal{L}_y|$ is very small, then $|\mathcal{L}_y| > |\mathcal{L}_y|^2 > |\mathcal{L}_y|^3 > \cdots$ implying that good approximations of (77) can generally be obtained by only considering terms linear in $\mathcal{L}_y$. In this regime, integrating (79) yields[21]

$$\ln y - \ln \hat{y} = e^{-\lambda t}\big(\ln y_0 - \ln \hat{y}\big) \tag{82}$$

which indicates that, in the close vicinity of the steady-state, the speed of convergence of an economy (λ) depends only on α and not on the economy's distance from its steady-state. A form of this equation obtained by subtracting $\ln y_0$ from both sides is widely used as a basis for growth regressions (e.g. Mankiw, Romer and Weil 1992).

As the distance from steady-state increases, $|\mathcal{L}_y|$ becomes large and then $|\mathcal{L}_y| < |\mathcal{L}_y|^2 < |\mathcal{L}_y|^3 < \cdots$. In this case, the higher-order terms of the expansion become significant. Thus, based on (80), the criterion $|\mathcal{L}_y| \gg \frac{1}{2}\big(\frac{1}{\alpha} - 1\big)|\mathcal{L}_y|^2$ can be used as a rough guide to determine when the linear expansion is a reasonable approximation. It leads to

$$\tfrac{1}{2}\big(\tfrac{1}{\alpha} - 1\big)\big|\ln y - \ln \hat{y}\big| \ll 1. \tag{83}$$

This shows that the interval in which the linear log approximation is reasonable is directly proportional to α. For small values of α, the interval is generally small, and as $\alpha \to 1$, the linear approximation provides good estimates in larger intervals around the steady-state.

The quadratic log expansion gives accurate values of $\mathcal{G}(\mathcal{Y})$ in a wider interval than the linear expansion. Moreover, in the region where the linear expansion provides reasonable estimates, the quadratic expansion will yield even better estimates. Figures 15 and 16 show the percentage errors incurred by using the linear, quadratic

[19]We remark that a similar graph appears as figure 13.24 in Carlin and Soskice (2006).

[20]Note that $\lim\limits_{\alpha \to 1} \mathcal{G}(\mathcal{Y}) = 0$ and, in this limit, the model under consideration simplifies to the AK-model which exhibits no transitional dynamics.

[21]See appendix for full solution.

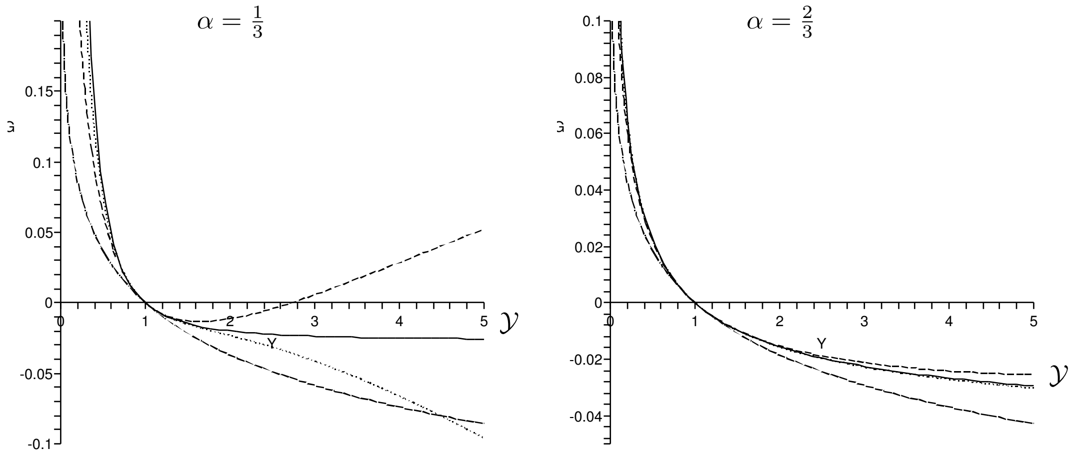

Figure 14: Graphs of the linear (—·—·—), quadratic (— — —), and cubic (· · ·) log expansions of $\mathcal{G}(\mathcal{Y})$ plotted against $\mathcal{Y}$ for $\alpha = \frac{1}{3}$ and $\alpha = \frac{2}{3}$. The curve of $\mathcal{G}(\mathcal{Y})$ is also shown (solid) for comparison. In all cases $(n + g + \delta) = 0.08$.

and cubic expansions to estimate $\mathcal{G}(\mathcal{Y})$ for $\alpha = \frac{1}{3}$ and $\alpha = \frac{2}{3}$, respectively. It is shown that a higher value of α generally leads to lower percentage errors for all values of $\mathcal{Y}$.

The equation (80) can be solved and we obtain[22]

$$\ln y - \ln \hat{y} - \tfrac{1}{2}\left(\tfrac{1}{\alpha} - 1\right)\Big(1 - e^{-\lambda t}\Big)\big(\ln y_0 - \ln \hat{y}\big)\big(\ln y - \ln \hat{y}\big) = e^{-\lambda t}\big(\ln y_0 - \ln \hat{y}\big). \quad (84)$$

This formula indicates that, in addition to α, an economy's speed of convergence also depends on its initial (log) distance from the steady-state. The predictions of $\ln y(t)$ provided by (84) are compared with those provided by (82) in figure 17. For $0 < \mathcal{Y}_0 < 1$, both expansions under-estimate the actual growth rate of $y(t)$, but the values provided by the quadratic expansion are always superior to those given by the linear approximation even for very small values of $\mathcal{Y}_0$. For example, with $\alpha = \frac{1}{3}, y_0 = \frac{1}{10}$ and $\hat{y} = 1$, the quadratic solution converges to the exact solution in about 60 years while the linear solution converges after 100 years.

For values of $\mathcal{Y} > 1$ (above the steady-state) however, the qualitative properties of the quadratic expansion are fundamentally different from those of $\mathcal{G}(\mathcal{Y})$, which it is meant to approximate. For, while $\mathcal{G}(\mathcal{Y})$ is a monotone decreasing function for all $\mathcal{Y} > 0$, the quadratic expansion is decreasing only in $0 < \mathcal{Y} < \mathcal{Y}_{\text{min}} = \exp\left(\frac{\alpha}{1-\alpha}\right)$. Then, for $\mathcal{Y} > \mathcal{Y}_{\text{min}}$, the quadratic expansion is an increasing function, and equals zero when $\mathcal{Y} = \mathcal{Y}_{\text{zero}} = \exp\left(\frac{2\alpha}{1-\alpha}\right)$. Thus, approximating $\mathcal{G}(\mathcal{Y})$ with a quadratic log expansion introduces an extraneous unstable equilibrium at $\mathcal{Y} = \mathcal{Y}_{\text{zero}}$. Economies that start with $\mathcal{Y}_0 > \mathcal{Y}_{\text{zero}}$ evolve in the right-ward direction towards $\mathcal{Y} = \infty$ while those that start in $1 < \mathcal{Y} < \mathcal{Y}_{\text{zero}}$ evolve towards $\mathcal{Y} = 1$. Economies that start in the range $\mathcal{Y}_{\text{min}} < \mathcal{Y} < \mathcal{Y}_{\text{zero}}$ do not satisfy the conditional convergence property because economies farther from $\mathcal{Y} = 1$ are shown to have lower growth rates (absolute values).

[22]The full solution is presented in the appendix.

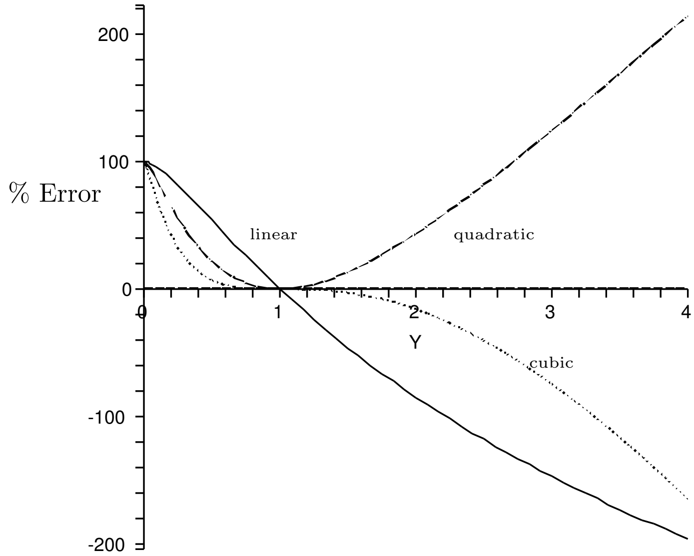

Figure 15: The percentage errors of the growth rate γ_y incurred by using the linear, quadratic and cubic Taylor log expansions, for $\alpha = \frac{1}{3}$. The variable on the horizontal axis is $\mathcal{Y}$.

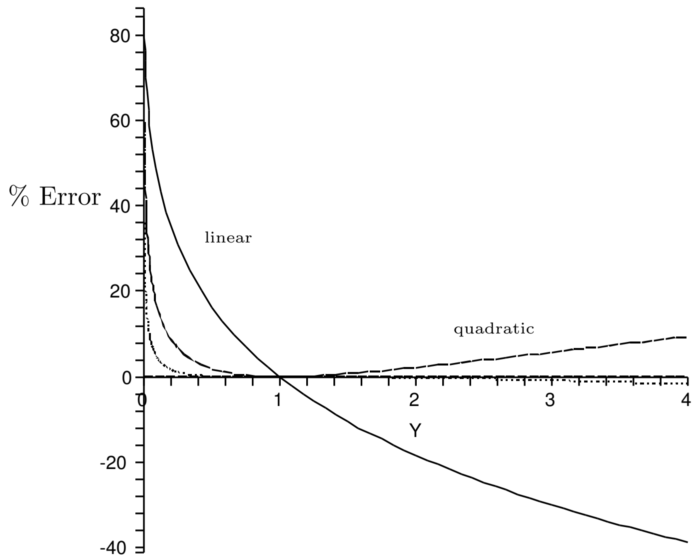

Figure 16: The percentage errors of the growth rate γ_y incurred by using the linear, quadratic and cubic Taylor log expansions, for $\alpha = \frac{2}{3}$.

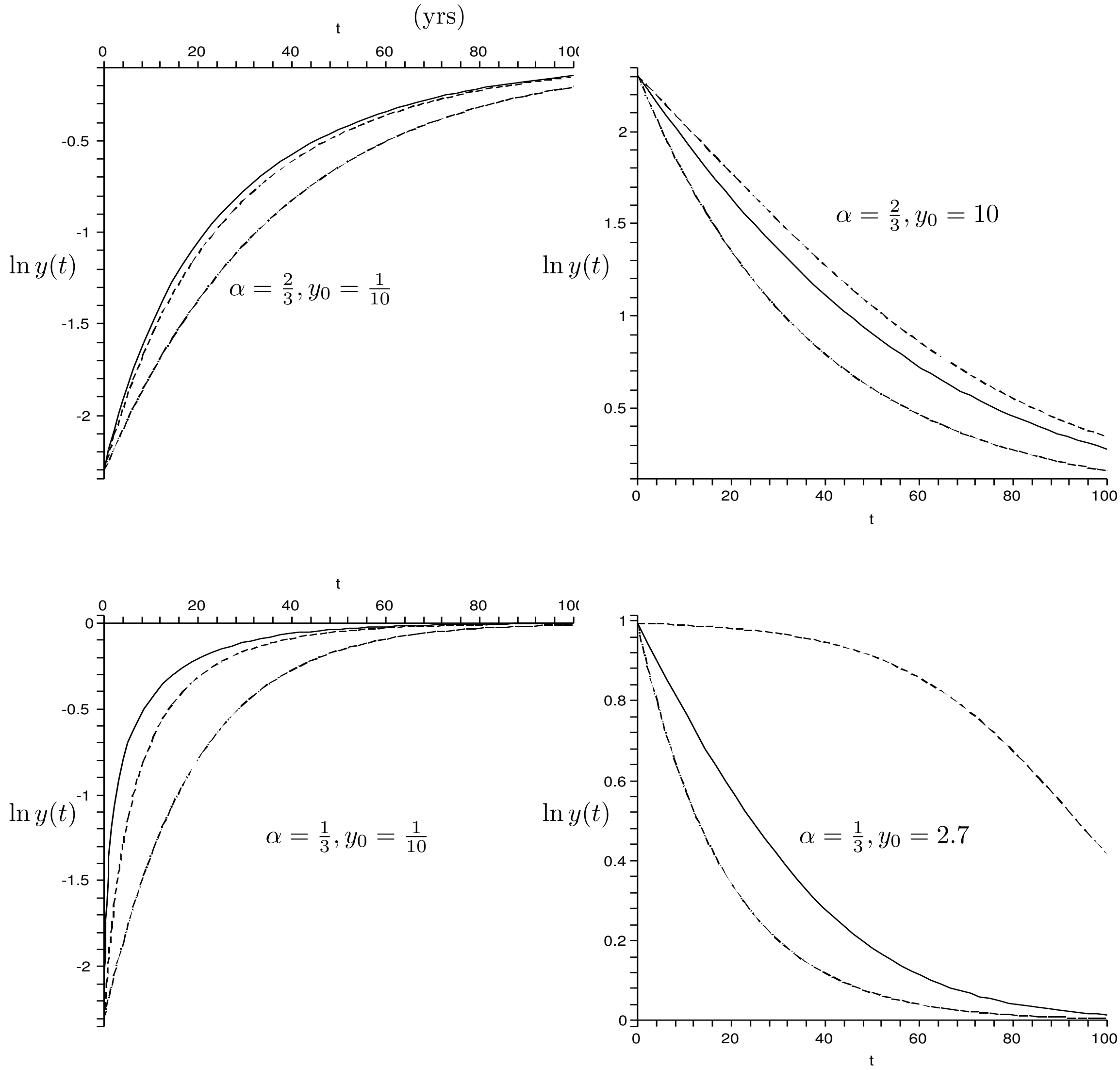

Figure 17: Comparisons of the predicted evolution paths of $\ln y(t)$ provided by the linear $(- \cdot - \cdot -)$ and quadratic $(- - -)$ log expansions of the governing equation. The curve of $\ln y(t)$ given by the exact solution is also shown (solid) in each case for comparison. In all cases $\hat{y} = 1, n = 0.01, g = 0.02$ and $\delta = 0.05$.

Therefore, although the performance of the quadratic expansion is clearly superior to that of the linear formulation in $0 < \mathcal{Y} < \mathcal{Y}_{\min}$, the situation is different when $\mathcal{Y} > \mathcal{Y}_{\min}$. Figure 17 shows, for example, an economy that starts just below $\mathcal{Y} = \mathcal{Y}_{\text{zero}}$ at $\mathcal{Y}_0 = 2.7$ (with $\alpha = \frac{1}{3}, \hat{y} = 1$, we have $\mathcal{Y}_{\text{zero}} = e = 2.718\cdots$). The predictions given by the quadratic expansion are shown to be clearly inferior to those obtained using the linear expansion. The performance of the quadratic formulation is thus unsatisfactory in the regime $\mathcal{Y} \to \mathcal{Y}_{\text{zero}}$ and for $\mathcal{Y} > \mathcal{Y}_{\text{zero}}$.

As a possible response to the foregoing issues, one might want to consider the cubic log expansion (81). Figure 14 shows that the interval in which this expansion accurately represents $\mathcal{G}(\mathcal{Y})$ is even wider than in the quadratic case and the cubic curve is monotone decreasing just like $\mathcal{G}(\mathcal{Y})$. However, the concavity of the cubic expansion is opposite to that of $\mathcal{G}(\mathcal{Y})$ to the right of $\mathcal{Y}_{\min}$. In this interval, the cubic curve is concave down while $\mathcal{G}(\mathcal{Y})$ is concave up. This means that beyond a certain value of $\mathcal{Y}$ $\left(=\exp\left(\frac{2\alpha}{1-\alpha}\right)\right)$, the cubic estimates will be worse than the linear ones. However, the main obstacle with regards to the cubic expansion is the fact that the resulting equations are fairly difficult to solve. The obtained solutions are complicated and hence do not facilitate any meaningful analysis of the transitional dynamics. For example, the solution of (81) is

$$\left[\ln \mathcal{L}_y - \tfrac{1}{2}\ln\left(6 - 3\ell\mathcal{L}_y + \ell^2\mathcal{L}_y^2\right) + \tfrac{1}{5}\sqrt{15}\tan^{-1}\left(\sqrt{15}\left[-\tfrac{1}{5} + \tfrac{2}{15}\ell\mathcal{L}_y\right]\right)\right]_0^t = -\lambda t,$$

where $\ell = \frac{1}{2}\left(\frac{1}{\alpha} - 1\right)$.

In summary, the linear log expansions are the simplest to derive and yield linear exact solutions that are well-suited for use in linear regression empirical tests. Away from the vicinity of the steady-state, although these expansions generally yield the least accurate estimates, their qualitative predictions are always consistent with those of the basic governing equation. The linear evolution equation (82) has been shown to be robust in the sense that, no matter how far from steady-state the economy starts, the path predicted by this equation always converges to the exact solution after sufficiently long times.

The quadratic log expansions improve on the linear approximation in $0 < \mathcal{Y} < \mathcal{Y}_{\min}$, and give exact solutions which, even though nonlinear, are relatively simple enough and can be estimated using nonlinear regression methods. However, a potentially serious drawback of the quadratic formulation is that, for $\mathcal{Y} > \mathcal{Y}_{\min}$, its predictions are inconsistent with those of the basic governing equation.

The main disadvantage of cubic (and higher-order) expansions is that the benefits gained from the solutions of the resulting equations are generally outweighed by the amount of effort required to derive them. When (if) the solutions are found, they are usually too complicated to be of much use.

One assessment is that of Romer (2001), who has stated that "Taylor-series approximations are generally quite reliable ... for the Solow model with conventional production functions."[23] Perhaps surprisingly, the analysis carried out here has demonstrated that, for a Solow model with a Cobb-Douglas production, working in terms of LVs, only the linear Taylor expansion is reliable in qualitative terms.

8.3 Derivations of half-life formulas

Derivation of formula (55)

Starting with

$$\left[\left(1 - e^{-\lambda T_k}\right) + \mathcal{K}_0^{1-\alpha} e^{-\lambda T_k}\right]^{\frac{1}{1-\alpha}} = \tfrac{1}{2}(1 + \mathcal{K}_0),$$

then raising both sides to the power of $(1-\alpha)$ and rearranging gives

$$e^{-\lambda T_k} = \frac{\left[\frac{1}{2}\left(\mathcal{K}_0 + 1\right)\right]^{1-\alpha} - 1}{\mathcal{K}_0^{1-\alpha} - 1}.$$

Taking logs on both sides then yields (55)

[23] Romer (2001), page 25.

Derivation of formula (57)

In this case, we solve the equation

$$\left[\left(1-e^{-\lambda \mathcal{T}_k}\right)+\mathcal{K}_0^{1-\alpha}e^{-\lambda \mathcal{T}_k}\right]^{\frac{1}{1-\alpha}}=\sqrt{\mathcal{K}_0},$$

for $\mathcal{T}_k$. Raising both sides to the power of $(1-\alpha)$ and rearranging gives

$$\begin{aligned} e^{-\lambda T_k} &= \frac{\mathcal{K}_0^{\frac{1}{2}(1-\alpha)}-1}{\mathcal{K}_0^{1-\alpha}-1}=\frac{\mathcal{K}_0^{\frac{1}{2}(1-\alpha)}-1}{\left(\mathcal{K}_0^{\frac{1}{2}(1-\alpha)}-1\right)\left(\mathcal{K}_0^{\frac{1}{2}(1-\alpha)}+1\right)} \\ &= \frac{1}{\left(\mathcal{K}_0^{\frac{1}{2}(1-\alpha)}+1\right)} \end{aligned}$$

Finally taking logs on both sides gives (57).

8.4 Solution of Taylor expansions

8.4.1 Solution of linear expansion

Starting from $\dot{\mathcal{L}}_y(t)=-\lambda\mathcal{L}_y(t)$, we separate variables and obtain

$$\frac{\mathrm{d}\mathcal{L}_y}{\mathcal{L}_y}=-\lambda\mathrm{d}t.$$

Integrate from time 0 to time t to get

$$\ln\left(\frac{\mathcal{L}_y(t)}{\mathcal{L}_y(0)}\right)=-\lambda t.$$

Note that $\mathcal{L}_y(0)=\ln y_0-\ln\hat{y}$. Then, taking exponentials and rearranging yields $\ln y(t)-\ln\hat{y}=e^{-\lambda t}\big(\ln y_0-\ln\hat{y}\big)$.

8.5 Solution of quadratic expansion

Starting from $\dot{\mathcal{L}}_y(t)=-\lambda\mathcal{L}_y(t)+\frac{1}{2}\big(\frac{1}{\alpha}-1\big)\lambda\mathcal{L}_y^2(t)$, we factorise the expression on the right hand side, separate variables, to get

$$\frac{\mathrm{d}\mathcal{L}_y}{\mathcal{L}_y(1-\ell\mathcal{L}_y)}=-\lambda\mathrm{d}t, \tag{85}$$

where $\ell=\frac{1}{2}\big(\frac{1}{\alpha}-1\big)$. The left hand side can now be expressed in terms of partial fractions as

$$\left(\frac{1}{\mathcal{L}_y}+\frac{\ell}{1-\ell\mathcal{L}_y}\right)\mathrm{d}\mathcal{L}_y=-\lambda\mathrm{d}t.$$

Integrating from time 0 to time t then gives

$$\ln\left(\frac{\mathcal{L}_y(t)\big[1-\ell\mathcal{L}_y(0)\big]}{\mathcal{L}_y(0)\big[1-\ell\mathcal{L}_y(t)\big]}\right)=-\lambda t.$$

Finally, taking exponentials and rearranging leads to

$$\mathcal{L}_y(t)-\tfrac{1}{2}\big(\tfrac{1}{\alpha}-1\big)\big(1-e^{-\lambda t}\big)\mathcal{L}_y(0)\mathcal{L}_y(t)=e^{-\lambda t}\mathcal{L}_y(0),$$

from which (84) follows after expressing in terms of $y(t)$.

8.6 Derivations

8.6.1 Solution of Augmented Solow model

Using the equations (59) and (64), it is possible to decouple the differential equations (59) and (64), to get

$$\begin{aligned}\dot{k} &= s_h^{\beta} s_k^{1-\beta} k^{\alpha+\beta} - (n+g+\delta)k, \\ \dot{h} &= s_k^{\alpha} s_h^{1-\alpha} h^{\alpha+\beta} - (n+g+\delta)h.\end{aligned}$$

Then, since k, h and y are related through $h = (s_h/s_k)k$ and $y = k^{\alpha}h^{\beta}$, it is necessary to solve only one of these equations. Following the procedure employed to derive (76), the solution of the capital equation is

$$k(t) = \left\{ \hat{k}^{1-\alpha-\beta}\Big(1 - e^{-\boldsymbol{\lambda} t}\Big) + k_0^{1-\alpha-\beta} e^{-\boldsymbol{\lambda} t} \right\}^{\frac{1}{1-\alpha-\beta}}.$$

8.6.2 Log-linearising the $y(t)$ equation

In the neighbourhood of the steady-state, the condition $k/h = s_k/s_h$ always holds, and the evolution equation for output can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln y(t)\Big] = \mathcal{R}(y) = (\alpha+\beta)s_k\left(\frac{s_k}{s_h}\right)^{\frac{\beta}{\alpha+\beta}} y^{1-\frac{1}{\alpha+\beta}} - (\alpha+\beta)(n+g+\delta).$$

To derive the linear log expansion, first note that

$$\hat{y}^{\frac{1}{\alpha+\beta}-1} = \frac{s_k}{n+g+\delta}\left(\frac{s_k}{s_h}\right)^{\frac{\beta}{\alpha+\beta}}.$$

Then

$$\begin{aligned}\left[\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\ln y}\right]_{y=\hat{y}} = \left[y\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}y}\right]_{y=\hat{y}} &= \left[(\alpha+\beta-1)s_k\left(\frac{s_k}{s_h}\right)^{\frac{\beta}{\alpha+\beta}} y^{1-\frac{1}{\alpha+\beta}}\right]_{y=\hat{y}} \\ &= -(1-\alpha-\beta)(n+g+\delta).\end{aligned}$$

Hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big[\ln\hat{y} + \big(\ln y - \ln\hat{y}\big)\Big] &= \left[\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\ln y}\right]_{y=\hat{y}}\big(\ln y - \ln\hat{y}\big) \\ \frac{\mathrm{d}}{\mathrm{d}t}\big(\ln y - \ln\hat{y}\big) &= -(1-\alpha-\beta)(n+g+\delta)\big(\ln y - \ln\hat{y}\big)\end{aligned}$$

References

[1] BARRO R J (1991) Economic Growth in a Cross-section of Countries. *Quarterly Journal of Economics,* **106**, 407–443.

[2] BARRO R J AND SALA-I-MARTIN X (1992) Convergence. *Journal of Political Economy,* **100**(2), 224–254.

[3] Barro R J and Sala-i-Martin X X (2004) *Economic Growth.* Cambridge MA: MIT Press.

[4] Baumol W J (1986) Productivity growth, convergence, and welfare: what the long-run data show. *Amer. Econ. Review* **96**(5), 1072–1085.

[5] Carlin W and Soskice D (2006) *Macroeconomics: imperfections, institutions and policies.* Oxford University Press, Oxford.

[6] Cho Dongchul & Graham Stephen (1996) The other side of convergence. *Economic Letters* **50**, 285−290.

[7] Heston A, Summers R and Aten B (2002). Penn World Table Version 6.1, Center for International Comparisons at the University of Pennsylvania (CICUP), October.

[8] Mankiw N G, Romer D & Weil D (1992) A contribution to the empirics of economic growth. *The Quarterly Journal of Economics.* **107**(2), 407 – 437.

[9] Okada T (2006) What does the Solow model tell us about economic growth? *Contributions to Macroeconomics,* **6**(1), Article 5.

[10] Reiss J P (2000) On the convergence speed in growth models. FEMM Working Paper No. 22/2000.

[11] Romer David (2001) *Advanced Macroeconomics.* Boston MA: McGraw Hill.

[12] Sala-i-Martin X X (1996) The classical approach to convergence analysis. *The Economic Journal* **106**, 1019 – 1036.

[13] Solow R M (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics* **LXX**, 65−94.

[14] Williams R L & Crouch R L (1972) The adjustment speed of neoclassical growth models. *Journal of Economic Theory* **4**, 552−556.