Economics Division
School of Social Sciences
University of Southampton
Southampton SO17 1BJ UK

Discussion Papers in Economics and Econometrics

The Long Wave of Conditional Convergence

Jian Tong

No. 0614

ISSN 0966-4246

The Long Wave of Conditional Convergence[1]

Jian Tong[2]

University of Southampton

October 3, 2006

[1]The author wishes to thank Ákos Valentinyi, Jean-Yves Pitarakis, John Bluedorn, Sevi Rodríguez-Mora, Juuso Välimäki and seminar participants at Southampton for helpful comments and suggestions. Any remaining errors are the author's own.

[2]Correspondence address: Economics Division, School of Social Sciences, University of Southampton, Highfield, Southampton SO17 1BJ, United Kingdom. Tel: +44-23-8059 6801, Fax: +44-23-8059 3858, E-mail: j.tong@soton.ac.uk

Abstract

We calculate the time series of the speed of convergence for 21 high-income countries over the period: 1953-1996, using low-pass filtered time series of per-capita GDP which are thus isolated from the influence of the short-run business cycle components. The observed patterns contradict the conventional 'time-invariant speed of convergence' hypothesis. Furthermore, dynamic panel data analysis provides strong evidence of the existence of stationary long cycles in the per capita GDP time series. We develop and estimate a technology-diffusion-based endogenous growth model, which shows that the endogenous growth of the domestic knowledge stock can account for the long cycles observed in the data.

Key words: trend reversion, speed of convergence, growth cycles

1 Introduction

The study of the speed of convergence is basically concerned with understanding how the time series of per capita GDP converges to its (conditional) steady-state trend. By empirically measuring the speed of convergence, growth economists aim to obtain identification of the growth mechanism that generates the data. In this regard, the existing literature on conditional convergence has appeared to be a bit single-minded in that, it focuses on achieving an accurate estimate of a time-invariant speed of convergence. This narrow focus, i.e., the 'time-invariant speed of convergence' hypothesis, originated in a broad class of neoclassical growth models, e.g., the Solow-Swan model, the Ramsey-Cass-Koopmans model (See Solow (1956), Swan (1956), Ramsey (1928), Cass (1965) and Koopmans (1965)) and their variations[1]. The equilibrium transition paths in these models all follow a first order autoregressive process[2], e.g., a first-order differential or difference equation.

Under the 'time-invariant speed of convergence' hypothesis, one can derive the speed of convergence from the estimate of the average speed of convergence. This hypothesis therefore provides a theoretical underpinning for the cross-sections estimations, which form the bulk of the vast literature on conditional convergence (see for example, Mankiw et al. (1992), Barro and Sala-I-Martin (1992)). The consensus that emerges from the cross-sections literature is that the logarithm of per capita income converges to its steady-state trend at the speed of about 2% per year, i.e., each year, the deviation from the trend shrinks by about 2%. However, if the 'time-invariant speed of convergence' hypothesis does not hold, then the estimate of the average speed of convergence will become less meaningful, and the speed of convergence must be understood in radically different light.

In this paper, we critically reexamine the 'time-invariant speed of convergence' hypothesis on both empirical and theoretical grounds. Empirically, we calculate the speed of convergence (of log per capita GDP toward the linear trend) for 21 high-income countries over the period 1953-1996, using low-pass filtered time series which are isolated from the short-run business cycle components. The observed patterns show that the speed of convergence is time-dependent in an important way. First, even after the removal of short-run business cycle components from the data, the time series of the speed of convergence have both positive and negative values, implying both convergence toward and transitory divergence from the long-run trend. Second, the time series of the speed of convergence fluctuate around

[1] For example, the augmented Solow model with human capital (Mankiw, Romer and Weil (1992)), and the Ramsey model with international capital flows (Barro, Mankiw and Sala-I-Martin (1995)).

[2] In these models, even if the dynamic system has multiple independent state variables, e.g., as in the Ramsey-Cass-Koopmans model, there happens to be only one stable autoregressive root (eigen value); then the unique saddle-path-stable transition dynamics becomes a first-order dynamic system.

zero at some low frequencies, indicating low-frequency cyclical components in the per capita GDP time series. We then use autoregressive models with linear time trend and country-specific intercepts to fit the low-pass filtered/smoothed time series of per capita GDP. Model selection tests indicate that the best model to describe the data is the fourth order autoregressive - AR(4) - model, the complex autoregressive roots of which identify two bands of frequencies for the cyclical components. The first has a cyclical period of around 10 years, attenuating around 10% per year in amplitude. The second has a much longer cyclical period, which is above 50 years, and is more persistent with a rate of attenuation around 5% per year, hence "a long wave" in the processes of conditional convergence. The empirical results hence contradict the conventional 'time-invariant speed of convergence' hypothesis. Consequently, they reject the theoretical models which predict time-invariant speed of convergence, including a broad class of neoclassical growth models as explanations for long-term economic growth.

In order to explain these newly found empirical regularities, we develop a technology diffusion-based endogenous growth model in this study. In our setting, an economy's domestic knowledge stock can deviate from its steady-state trend. The growth of domestic knowledge stock is a function of intentional R&D investments and the stock of opportunities of emulating the world technological frontier. Each country's steady-state trend of domestic knowledge stock is determined by its country fixed effects. The convergence of each country's per capita income towards its steady-state trend, is primarily influenced by the trend-reversion process of its domestic knowledge stock. In the market economy, the intensity of investments in R&D and progress in technology is affected by the prospect of asset value appreciation. International technological diffusion is found to ultimately drive long-term conditional convergence, and the long-run appreciation (depreciation) of asset values. Sufficient strength of asset value appreciation, however, can complicate the long-run trend reversion with overshootings, hence create long waves in conditional convergence.

The transition dynamics of the model economy is determined by its characteristic polynomial. By matching the theoretical characteristic polynomial to its empirical counterparts, we can estimate the structure model.[3] With empirically plausible parameters, the model can accurately reproduce the frequencies and the rates of attenuation of the cycles empirically established. The close match between the theory and the empirical regularities suggests that the long term growth (transition dynamics) of per capita income is primarily determined by the trend-reversion process of the domestic knowledge stock and total factor productivity (TFP). The neoclassical mechanism of capital deepening only plays a secondary role.

[3] In doing so, we need to calibrate three of the parameters to the values commonly agreed in the empirical growth literature.

2 Related Literature

Our research contributes to the empirical literature on conditional convergence. It differs from both the cross sections literature (e.g., Mankiw et al. (1992), Barro and Sala-I-Martin (1992)) and the existing panel data literature (e.g., Islam (1995), Caselli, Esquivel and Lefort (1996)) in that we abandon the restriction that the long-term growth should be an AR(1) process, or, equivalently, the 'time-invariant speed of convergence' hypothesis.[4] Our finding of long waves in conditional convergence suggests that future empirical research on growth convergence should take into account the low-frequency cyclicality of the long-term trend-reversion processes of per capita incomes.

The theoretical model extends those developed by Tong and Xu (2004) and Tong and Xu (2006), and is related to the work by Barro and Sala-I-Martin (1997) and Howitt (2000). All these models study the influence of international technological diffusion on domestic R&D investments, and they all predict that the growth rate of every country converges to a common steady-state rate. Tong and Xu (2004) and Tong and Xu (2006) study the joint determination of steady-state trend of per capita income and financial institutions, which is absent in the current paper. There is a close comparison between the current model and Barro and Sala-I-Martin (1997) in that both of them extend the variety expansion model due to Romer (1990). The key departure of our model from Barro and Sala-I-Martin (1997) is that we use a discrete time setting while their model is of continuous time. One justification for preferring the discrete-time setting to the continuous one is that the former can conveniently capture the lumpiness of the time scale of R&D projects.[5] This difference generates an interesting difference in the transition dynamics between the two settings: while Barro and Sala-I-Martin (1997) predicts an AR(1) autoregressive conditional convergence process and hence a time-invariant speed of convergence, as do the neoclassical growth models, our setting can generate stationary long cycles in the trend-reversion process. Since these model all predict convergence of growth rate, they do not address the issue of long-run growth rate divergence of the poorest countries relative to the rich countries. As an exception, a very recent contribution by Aghion, Howitt and Mayer-Foulkes (2005), which also models the influence of international technological diffusion on domestic investments in innovations, predicts a 'poverty trap' and growth rate divergence for countries which suffer from severe credit constraints.

As a study of long waves, our finding inevitably relates to the notion of Kondratieff cycles, Schumpeter's theory of business cycles, and Kuznets' critique. Kondratieff (1935) identified cycles of about 50 years' length in a number of price, consumption and production series between the period of 1780-1921.

[4] Another recent study on growth convergence, which departs from the 'time-invariant speed of convergence' hypothesis is Phillips and Sul (2003).

[5] In contrast, in the Barro and Sala-I-Martin (1997) model, the time scale of an R&D project can be infinitely short.

Schumpeter (1939) was an attempt to link the Kondratieff cycles and the movement in innovations. Kuznets (1940) pointed two difficulties with Schumpeter's analysis. The first is the lack of convincing explanation for the "bunching" of innovations over time. The second is the difficulty of subjecting the theoretical claims to statistical time series analysis; particularly, the lack of reliable statistical means to differentiate the long cycles from the much more clearly marked shorter cyclical swings called into question the validity of the Kondratieff cycles. The lack of (empirical) regular recurrence is the main reason why Kondratieff cycles are not recognized by modern economics. The long waves of conditional convergence which we identify in the current study differ from the Kondratieff cycles as conventionally understood. First, the former are identified by much more rigorous statistical analysis of time series. Second, we find that the amplitudes of the long cycles attenuate over time at non-negligible rate, therefore they do not recur with the same the amplitudes, and will die out eventually if not 'renewed' by exogenous shocks. The lack of regular recurrence of the long cycles is a prediction of our findings, rather than a contradiction.

The remainder of the paper is organized as follows: Section 2 documents a set of new empirical regularities about conditional convergence. Section 3 constructs and analyzes a discrete-time technology diffusion-based endogenous growth. Section4 calibrates and estimates the model. Section 5 concludes.

3 New Empirical Regularities about Conditional Convergence

In the existing literature on conditional convergence, the speed of convergence β is featured in the following equation:[6]

$$\ln \frac{y_{it}}{y^*_{it}} = (1-\beta)^t \ln \frac{y_{i0}}{y^*_{i0}}, \tag{1}$$

which implies[7]

$$\ln \frac{y_{it+1}}{y^*_{it+1}} - \ln \frac{y_{it}}{y^*_{it}} = -\beta \ln \frac{y_{it}}{y^*_{it}} \tag{2}$$

where y_{it} is the per capita income of country i at time t, y^*_{it} is the time-invariant steady-state log-linear trend of per capita income and $y^*_{it} = y^*_{i0}e^{g_y t} \approx y^*_{i0}\left(1+g_y\right)^t$, g_y is the common trend rate of growth, y^*_{i0}

[6] In the literature the following alternative specification is often used:

$$\ln \frac{y_{it}}{y^*_{it}} = e^{-\beta t} \ln \frac{y_{i0}}{y^*_{i0}},$$

which is a close approximation when β is close to 0.

[7] The respective alternative formulation is:

$$\ln \frac{y_{it+1}}{y^*_{it+1}} - \ln \frac{y_{it}}{y^*_{it}} = -\left(1-e^{-\beta}\right) \ln \frac{y_{it}}{y^*_{it}}.$$

is the country-specific intercept.

Without loss of generality, we can define the speed of convergence as

$$\beta_{it} \triangleq -\frac{\ln \frac{y_{it+1}}{y^*_{it+1}} - \ln \frac{y_{it}}{y^*_{it}}}{\ln \frac{y_{it}}{y^*_{it}}}. \tag{3}$$

Studying the behavior of the time series of β_{it} can reveal some properties of the trend-reversion process of per capita GDP. Using panel data can increase the signal-noise ratio by adding the cross-sectional variation. However, one has to be cautious when pooling the cross-country time series together because uncontrolled heterogeneity may violate the assumption of common trend growth rate. To reduce the probability of this problem one can confine the pooling to a group of richest countries, which are less likely to diverge in long-run growth rate. Since the speed of convergence is a property of long-run growth, its measurement needs to be isolated from the influence of the short-run business cycles components. To that end, one can use the highpass or bandpass filters to remove the short-run cyclical components from the time series of per capita GDP. In this investigation, we employ the Hodrick-Prescott filter[8] (Hodrick and Prescott (1997)) to smooth the 21 per capita GDP time series (from year 1950 to 2000) of richest countries from the Penn World Table Mark 6.1 (Heston, Summer and Aten). Using the approximate bandpass filter due to Baxter and King (1999) gives a very similar outcome. The deviation of $\ln y_{it}$ from the linear trend, $\ln \frac{y_{it}}{y^*_{it}}$, can be measured as the error term ν_{it} in following regression:

$$\ln y_{it} = \ln y^*_{i0} + g_y t + \nu_{it}. \tag{4}$$

Figure 1 plots the entire panels of $\ln \frac{y_{it}}{y^*_{it}}$ against time[9]. Conditional convergence relies on that $\ln \frac{y_{it}}{y^*_{it}}$ being stationary. The standard deviation does appear to be bounded. Furthermore, the Im-Pesaran-Shin and Maddala-Wu tests for unit root in panel data (see Im, Pesaran and Shin (2003) and Maddala and Wu (1999)) can reject the unit root hypothesis[10]. Figures 2 displays the time series of speed of convergence and the deviation of $\ln y_{it}$ from the linear trend for all the countries in the sample. To eliminate extreme values in the speed of convergence, any values above 1 or below -1 have been truncated at 1 and -1 respectively in the figure. It is now plainly evident that the speed of convergence is both country-specific and time-dependent. Two related novel empirical patterns emerge from the figure[11] as follow:

[8]The smooth parameter is set to $\lambda = 7$ for the annual data, which according to Ravan and Uhlig (2002), is consistent with setting $\lambda = 1600$ for the quarterly data in the business cycles literature. The effect is to remove the usual business cycle components with periods of 2-8 years.

[9]The results from the OLS regression are shown in Figure 10 in Appendix A. The results reported here are based on a two-step estimation, which first obtains a more efficient estimate of the trend growth rate by taking into account the serial correlation in the error term. More details about this are provided in later part of this section.

[10]More details about unit root tests will be provided later in this section.

[11]Figure 11 in Appendix A shows the results based on the OLS estimation. They confirm what is reported here.

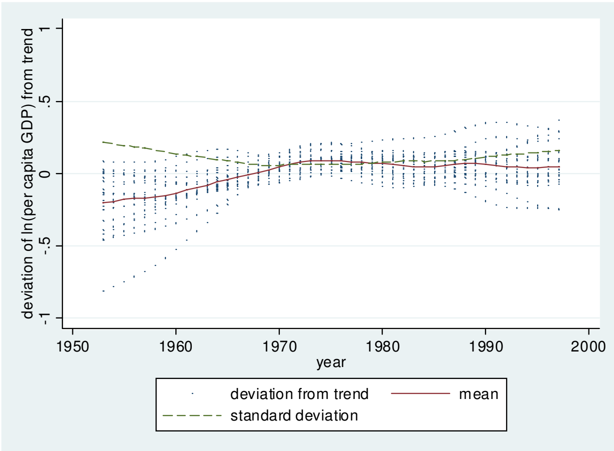

Figure 1: The deviation of $\ln y_{it}$ from linear trend

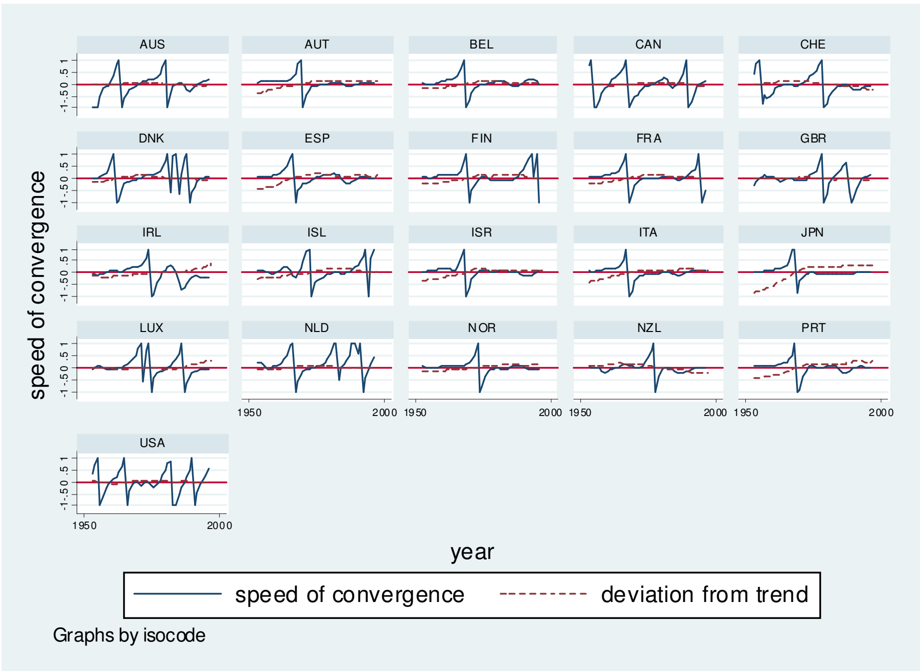

Figure 2: The time series of speed of convergence and the deviation of $\ln y_{it}$ from the linear trend for 21 countries

Stylized fact A The speed of convergence has both positive and negative values; accordingly the time series of per capita GDP can both converge toward and (transitorily) diverge from the long-run trend.

This pattern is consistent with Phillips and Sul (2003)'s emphasis on the role of time-dependence of the speed of convergence in reconciling long-run convergence and transitory divergence among economies which share the same steady-state trend. Closer examination of the alternation between convergence and divergence reveals the following pattern.

Stylized fact B Each time series of the speed of convergence fluctuates around 0 with below the usual short-run business cycle frequencies; accordingly the time series of per capita GDP has some long-run cyclical components.

Although the overall mean of the speed of convergence among the entire panels is .0244658, indicating an overall conditional convergence, the richness of the panel data information reveals that the devil is in the details. The novel finding here is the cyclical nature of the trend-reversion process of per capita GDP, which we try to understand in the remainder of the paper.

Figure 3 plots the time series of the speed of convergence and the deviation of $\ln y_{it}$ from the linear trend for nine example countries: Austria, Belgium, Spain, Israel, Italy, Japan, Norway, New Zealand and Portugal. One commonality of these nine countries is that their time series of per capita GDP all overshot (i.e., crossed) their long-run trend once within the sample period. All except New Zealand overshoot the trend from below. Each of the nine panels displays a clear pattern of cyclicality. Figure 4 shows the periodograms (spectral density functions) of $\ln \frac{y_{it}}{y_{it}^*}$ for these nine countries, which provide a crude indication of the major cyclical components of $\ln \frac{y_{it}}{y_{it}^*}$. In all the nine panels it can be seen that a lot of variance of $\ln \frac{y_{it}}{y_{it}^*}$ is at the low frequency end (between 0-0.125, i.e., cyclical period longer than 8 years); particularly, there is a peak value between 0-0.02 (i.e., longer than 50 years cyclical period). These frequencies (or cyclical periods) are characteristic of the dynamic system of long-run growth. They imply that the time series of the logarithm per capita GDP has corresponding complex regressive roots. For each different characteristic frequency, we need a complex conjugate pair regressive roots: so for a single characteristic frequency, we need at least an AR(2) model to describe $\ln y_{it}$; for two characteristic frequencies, we need at least an AR(4) model.

We conjecture that the behavior of $\ln y_{it}$ is best described by an AR(n) model ($n \geq 2$) with common time trend and individual intercept as follows:

$$\ln y_{it} = u_i + \phi_0 t + \sum_{s=1}^{n} \phi_s \ln y_{it-s} + \varepsilon_{it}, \tag{5}$$

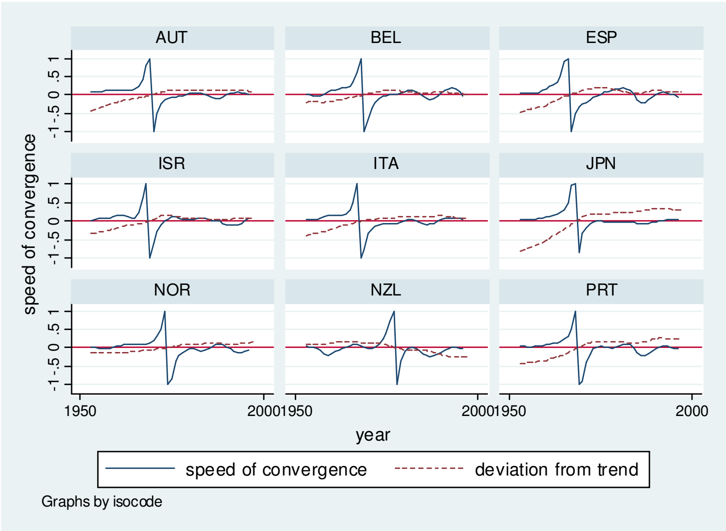

Figure 3: The time series of speed of convergence and the deviation of $\ln y_{it}$ from the linear trend for 9 countries

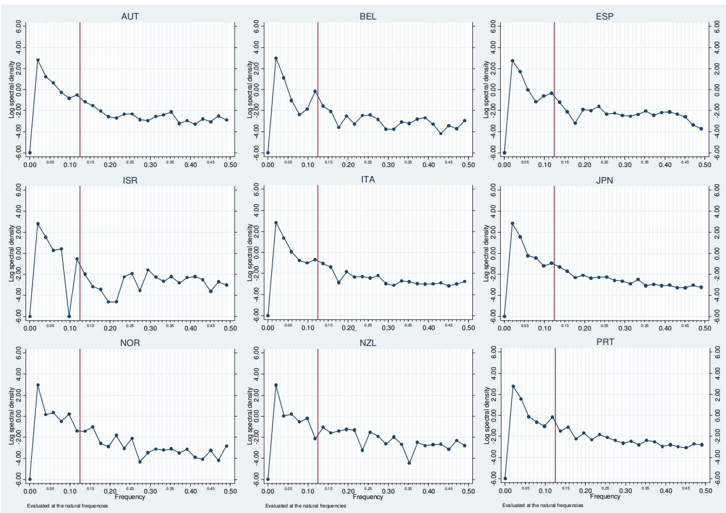

Figure 4: The periodogram (log spectral density function) of $\ln \frac{y_{it}}{y_{it}^*}$ for 9 countries, year: 1950-2000, the frequency unit is one cycle period per year. (The data is non-filtered, the vertical line indicates cyclical period of 8 years, which is a usual cutoff point for high/low-pass filters.)

where ε_{it} is i.i.d. with zero mean. It is essential to assume that $\ln y_{it}$ does not have a unit root, i.e., formally, $\sum_{s=1}^{n}\phi_s \neq 1$. The above equation can be reformulated as the following:

$$\ln\frac{y_{it}}{y_{it-1}} = u_i + \phi_0 t + \theta_0 \ln y_{it-1} + \sum_{j=1}^{n-1}\theta_j \ln\frac{y_{it-j}}{y_{it-j-1}} + \varepsilon_{it}, \tag{6}$$

where $\theta_0 = \sum_{s=1}^{n}\phi_s - 1$, $\theta_j = -\sum_{s=j+1}^{n}\phi_s$. Therefore the null hypothesis for the unit root test is H_0: $\theta_0 = 0$.

We use the Im-Pesaran-Shin and Maddala-Wu tests (see Im et al. (2003) and Maddala and Wu (1999)) to test the unit root (null) hypothesis against the alternative hypothesis H_1: $\theta_0 < 0$ for a variety of values of n ranging from 2 to 8. It turns out that for $n = 2$ and 4, the null hypothesis is rejected at 0.05; for other values of n the null hypothesis can not be robustly rejected. These results have two important implications. First, they confirm the stationarity assumption. Second, they provide a useful guidance on how n should be selected. Therefore we should use the AR(2) and AR(4) models. Furthermore, the likelihood ratio tests also strongly reject the AR(1) model against the AR(2) and AR(4) models respectively, hence strongly rejecting the conventional hypothesis of time-invariant speed of convergence.

We start with the AR(2) model, for which the characteristic polynomial is:

$$X^2 - \phi_1 X - \phi_2, \tag{7}$$

which has complex roots if and only if

$$\phi_1^2 + 4\phi_2 < 0. \tag{8}$$

The existence of complex roots of (7) would imply cyclical transition path. We can test a second null hypothesis that the trend-reversion process is not cyclical, which can be formulated as H_2: $\phi_1^2 + 4\phi_2 \geq 0$, against the alternative hypothesis H_3: $\phi_1^2 + 4\phi_2 < 0$.

The regression and test results are reported in Table 1. The test of H_2 against H_3 is by the Wald test of nonlinear restriction. The four sets of estimations (1)-(4) use different data filters or band widths. For regressions (1) and (2) we remove the usual business cycle components which have cyclical periods in the range of 2-8 years. For regressions (3) and (4) the cycles removed range between 2-11 and 2-10 years respectively. And we alternate between the HP filter and BK filter for similar bandwidth. The variation in the data filtering procedures and the bandwidths allows us to check the robustness of the results.

The null hypothesis H_2 is rejected at the confidence level of 1%. The test result is not sensitive to the data filtering technique we use. The rejection of hypothesis H_2 is confirmative of a cyclical trend-reversion process.

	(1) HP filter ($\lambda = 7$)	(2) BK filter (p_l, p_h= 2, 8 $k = 3$)	(3) HP filter ($\lambda = 25$)	(4) BK filter (p_l, p_h= 2, 10 $k = 4$)
$\ln y_{it-1}$ (ϕ_1)	1.924 ***	1.902122***	1.952687***	1.935067***
	(.0127003)	(.0088477)	(.0093619)	(.0072537)
$\ln y_{it-2}$ (ϕ_2)	-.9374352 ***	-.9165936***	-.963349***	-.9449827***
	(.0122747)	(.0133463)	(.0090374)	(.0072845)
t (ϕ_0)	.0003591 ***	.0003834***	.0002761***	.0001812***
	(.0000389)	(.0000452)	(.0000245)	(.0000183)
observation	903	903	861	861
R^2	0.9999	0.9999	0.9999	0.9999
hypothesis H_2 vs. H_3	rejected***	rejected***	rejected***	rejected***
cyclical period τ	55.438	54.620	61.249	55.705
rate of attenuation γ	.031788	0.04261 2	0.018 5	.03070 1

Table 1: Regression of logarithm of per capita GDP - AR(2)

(Note: *** indicates statistical significance at 1% level.

From the estimated coefficients we can retrieve the cyclical period τ and the rate of attenuation of amplitude γ, using the following formula:

(9)
$$\tau = \begin{cases} \frac{2\pi}{\arcsin \frac{|\mathrm{Im}(\chi)|}{|\chi|}} & \text{if } \mathrm{Re}(\chi) \geq 0 \\ \frac{2\pi}{\arcsin \frac{|\mathrm{Im}(\chi)|}{|\chi|} + \pi} & \text{if } \mathrm{Re}(\chi) < 0 \end{cases},$$

and

(10)
$$\gamma = 1 - |\chi|,$$

where χ is a complex root of (7). The estimations of the $AR(2)$ model suggest that the dominant cyclical component in long-term growth should have a period of oscillation of above 50 years, and an attenuation rate of up to 4.3% per year. These imply a "long wave" in the process of conditional convergence.

Figure 5 plots the error terms from the AR(2) regression (1). The time series of the mean error shows clear cyclical pattern with the cyclical period around 10 years. This is suggestive of the restrictiveness of the AR(2) model in accounting for some significant cyclical component. To address this problem,

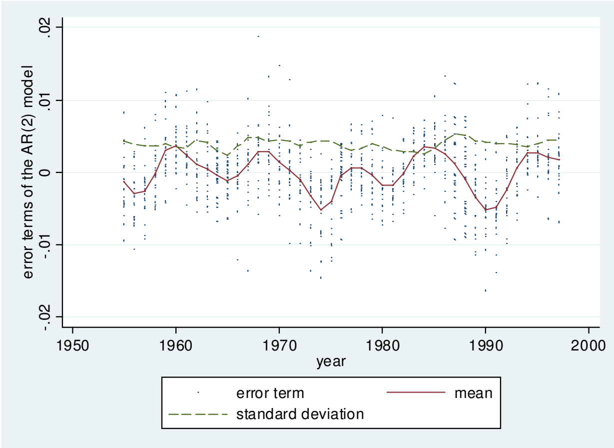

Figure 5: The error term of the AR(2) regression (1)

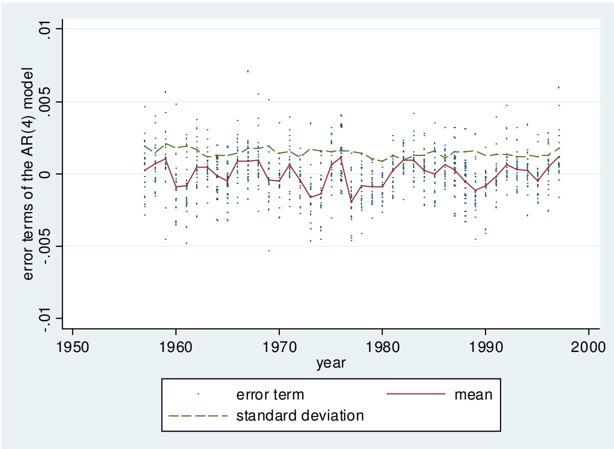

Figure 6: The error term of the AR(4) regression (5)

next we estimate the AR(4) model, for which the characteristic polynomial model is:

$$X^4 - \phi_1 X^3 - \phi_2 X^2 - \phi_3 X - \phi_4, \tag{11}$$

the complex roots of which determine the cyclical periods and the rates of amplitude attenuation. Also, the trend grow rate can be inferred as follows:

$$g_y = \frac{\phi_0}{1 - \sum_{s=1}^{n} \phi_s} \text{ for AR(n) model.} \tag{12}$$

Table 2 summarizes the results of estimating the AR(4) model. Again four sets of estimations are obtained using different data filters or band widths. The R^2 for each regression here is extremely large, as is for each AR(2) regression. This is due to the fact that the predominant cross-section variance is well explained by the lagged cross-section variances; and the time series have been smoothed by the low-pass filters, which have removed the high-frequency components from the variance of the original time-series. Figure 6 plots the error terms of the AR(4) regression (5), which can be compared to Figure 5. Besides a significant reduction (roughly halving) in the variance of the errors, it is also apparent that the cyclical component with cyclical period of around 10 years disappears from the mean error time series. The likelihood ratio test also indicates that the AR(4) model is significantly better than the AR(2) model.

Using the AR(4) model, the estimated cyclical periods of the long waves are longer than using the AR(2) model, they are in the range between 60-70 years. The cause for this difference relates to the fact that the AR(2) model is mis-specified and fails to account for the around 10-year-period cyclical components. As a result the error term is serially correlated and also correlated with the lagged dependent variables. The resultant bias in the OLS estimator then induces a bias in the estimation of the long cycle period. The AR(4) results confirm the existence of a type of shorter cycles with periods between 9-14 years in the conditional convergence process. These cycles match what have been known as the Juglars in the old business cycle literature[12] (see for example, Schumpeter (1939) and Kuznets (1940)). To check the robustness of results of the AR(4), we also estimate an AR(6) model. The results, which are summarized in Table 3, are close to the AR(4) results for the first two cyclical components and the trend growth rate g_y.[13] Besides, they suggest that some high-frequency business cycle components (around 4-5 year cyclical period) do "leak" through the HP filter; but have little effect on the AR(4) estimator.

The current results suggest that the long waves with sizeable amplitudes are unlikely to have regular recurrences because their amplitudes attenuate over time. If not "renewed" or "reenforced" by exogenous shocks, they tend to die out eventually. The rate of attenuation of the long waves are in the range between 3-6% per year, implying 'half lives'–the time that it takes for half of initial amplitude to be eliminated–from 11 to 23 years. It is unlikely that exogenous shocks could sustain a sizeable amplitude of the long wave.

The lack of regular recurrence of the long waves in a single time series is illustrated by the example of the UK. Figure 7 shows the logarithm of per capita income of the UK from year 1830 to 2001 (Data source: Maddison (2003)) and its linear trends. The spectral density functions for the pre-1914 era and post-1945 era are presented in Figures 8 and 9. It can be seen that a long wave (i.e., with cyclical period longer than 50 year) was more strongly present in the pre-1914 era (hence Figure 8) than the post-1945

[12] In that literature, cycles were usually called by the names of economists who identified them. Therefore cycles with periods between 50-60 years are Kondratieffs; between 15-25 years: Kuznets'; between 7-11 years: Juglars; around 40 months: Kitchins.

[13] Actually, we use the AR(6) result of the trend growth rate in the 2-step estimation of equation (4).

	(5) HP filter ($\lambda = 7$)	(6) BK filter (p_l, p_h= 2, 8 $k = 3$)	(7) HP filter ($\lambda = 25$)	(8) BK filter (p_l, p_h= 2, 10 $k = 4$)
$\ln y_{it-1}$ (ϕ_1)	3.312033***	3.279736***	3.568417***	3.394918***
	(.0316516)	(.0278954)	(.0276212)	(.0271036)
$\ln y_{it-2}$ (ϕ_2)	-4.369273***	-4.339645***	-4.957507***	-4.575488***
	(.0880009)	(.0761624)	(.0780728)	(.0748587)
$\ln y_{it-3}$ (ϕ_3)	2.752199***	2.782082***	3.189778***	2.926294***
	(.0877513)	(.0766304)	(.0767245)	(.074257)
$\ln y_{it-4}$ (ϕ_4)	-.6993725***	-.7269854***	-.8026889***	-.7489372***
	(.0312118)	(.0282277)	(.0261588)	(.0263697)
t (ϕ_0)	.0000976***	.0001035***	.0000459***	.0000697***
	(.0000155)	(.0000183)	(.00000659)	(.0000131)
observation	861	861	819	819
R^2	1.0000	1.0000	1.0000	1.0000
cyclical period τ_1	63.094	66. 946	68.655	71.537
rate of attenuation γ_1	.056757	.057331	.034 69	.049817
cyclical period τ_2	10.005	9.1935	13.071	10.446
rate of attenuation γ_2	.113 39	.095509	.07187 4	.08922 4
growth rate g_y	.02211933	.02150609	.02296034	.02169331

Table 2: Regression of logarithm of per capita GDP - AR(4)

era (hence Figure 9). If the rate of attenuation is 4.5% per year, then the amplitude of a long wave will shrink to 0.0048 of its original size after 116 years (i.e., from 1830 to 1946). The lack of its recurrence in the post-1945 era is hardly surprising. This example illustrates the point that one cannot interpret this kind of lack of regular recurrence as evidence of non-existence of long waves.

In summary, the empirical results presented in this section are at odds with the conventional assumption that the long-term conditional convergence has a time-invariant speed of convergence. Consequently, they are not consistent with growth models which predict a time-invariant speed of convergence, including a broad class of neoclassical growth models. It challenges growth theories to reproduce the long cyclical components in the trend-reversion processes of per capita incomes. In the next section, we develop a theoretical technology-diffusion-based endogenous growth model to explain the new empirical

HP filter	τ_1	γ_1	τ_2	γ_2	τ_3	γ_3	g_y
$(\lambda = 7)$	64.491	.066664	10.749	0.13782	4.8035	0.4893	.02236098

Table 3: The results of an AR(6) regression

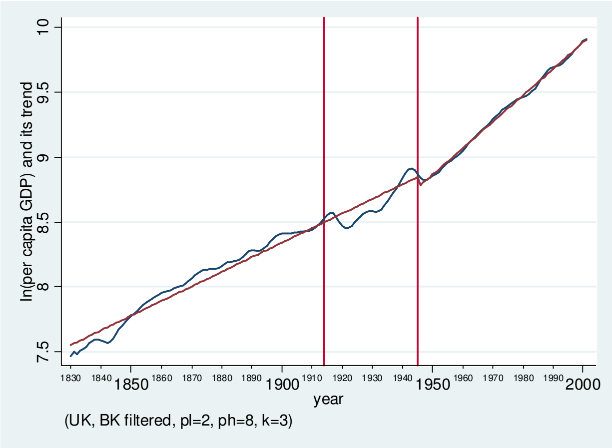

Figure 7: Logarithm of per capita GDP of the UK and its linear trends. The vertical lines indicate year 1914 and 1945 respectively. (A structural change in the linear trend is assume for the post-1945 era.)

regularities.

4 The Model

The model economy exists for an infinite number of periods labelled $t = 0, 1, 2, \cdots, \infty$. There are L identical consumers who live forever, each has one unit of labor supply per period. Each consumer's utility maximization problem is:

$$\begin{aligned} &\max \textstyle\sum_{s=t}^{\infty} \frac{1}{(1+\rho)^{s-t}} \left(\frac{(c_s)^{1-\theta}}{1-\theta} - 1 \right) \\ &s.t. : b_{s+1} = w_s + b_s (1 + r_s) - c_s, \end{aligned} \tag{13}$$

where c_s is consumption, ρ is the utility discount rate, θ is the parameter of preference over smoothness of consumption, w_s is wage income, b_s is the holding of risk-free bond or bank deposit, with interest rate r_s. The *Euler* equation for optimal consumption is

$$\frac{c_{s+1}}{c_s} = \left(\frac{1 + r_s}{1 + \rho} \right)^{\frac{1}{\theta}}. \tag{14}$$

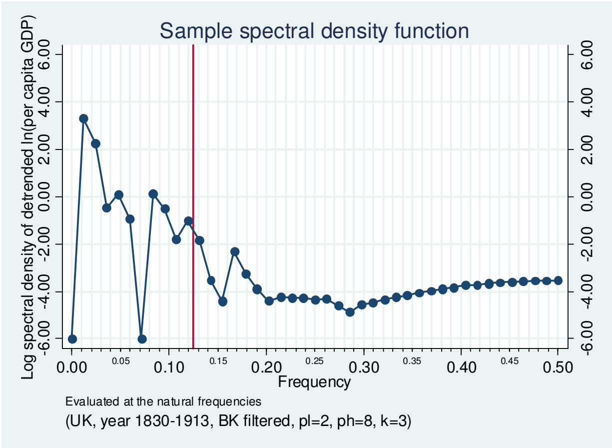

Figure 8: Spectral density function of detrended logarithm of per capita GDP of the UK for 1830-1913, the frequency unit is one cycle period per year

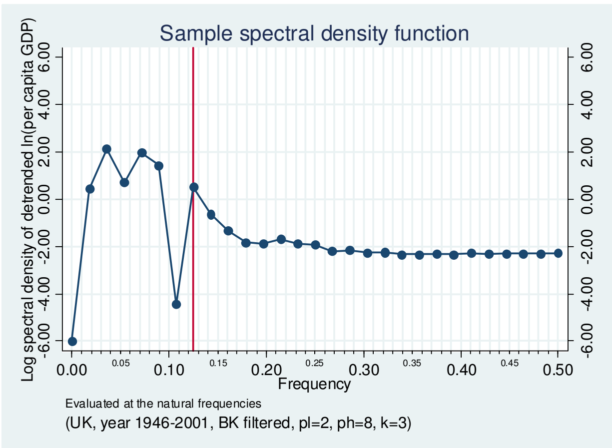

Figure 9: Spectral density function of detrended logarithm of per capita GDP of the UK for 1946-2001, the frequency unit is one cycle period per year

Denote by r the steady-state (balanced growth path) value of r_s, and g_y the steady-state growth rate of consumption; we have the following results:

$$\left(\frac{1+r}{1+\rho}\right)^{\frac{1}{\theta}} = 1 + g_y,$$

and

$$r = (1+\rho)(1+g_y)^{\theta} - 1. \tag{15}$$

Production in this economy comprises a final good sector and an intermediate good sector. The final good sector is perfectly competitive, and it has the following Cobb-Douglas/CES production function[14] with intermediate inputs x_{it}, labor input L_{1t}, and output:

$$Y_t = L_{1t}^{1-\alpha}\left(\sum_{i=1}^{A_t} x_{it}^{\frac{\alpha}{\sigma}}\right)^{\sigma},\ 0 < \alpha < 1 \text{ and } \sigma > \alpha, \tag{16}$$

where A_t is the number of varieties of intermediate goods, which is also a measure of domestic knowledge stock. Parameter σ determines whether the different varieties of intermediate goods are direct complements or direct substitutes. If $\sigma > 1$ (respectively $\sigma < 1$) then a new variety is a direct complement (respectively a direct substitute) to the existing varieties because it increases (respectively decreases) their marginal products.[15] The assumption: $\sigma > \alpha$ rules out the possibility of $\sigma = \alpha$, which is the case that all the intermediate goods are perfect substitutes. The firm's maximization program is

$$\max_{x_{it}, L_{1t}} \left(L_{1t}^{1-\alpha}\left(\sum_{i=1}^{A_t} x_{it}^{\frac{\alpha}{\sigma}}\right)^{\sigma} - (1+r_t)\sum_{i=1}^{A_t} p_{it}x_{it} - L_{1t}w_t\right), \tag{17}$$

where p_{it} is the rental price of intermediate good x_{it}, and w_t is the wage of labor in period t. The final good producers pay the intermediate goods producers at the beginning of each period to get the inputs, and sell their own products and pay their workers at the end of each period. The inverse demand functions for intermediate goods and labor are:

$$p_{it} = \frac{L_{1t}^{1-\alpha}\alpha\left(\sum_{i=1}^{A_t} x_{ti}^{\frac{\alpha}{\sigma}}\right)^{\sigma-1} x_{it}^{\frac{\alpha}{\sigma}-1}}{(1+r_t)} \tag{18}$$

[14] This type of production function has been originally used by Spence (1976) and Dixit and Stiglitz (1977) in studies on monopolistic competition and product selection/diversification, and is now standard in the growth literature.

[15] To see this, note the cross-partial derivative

$$\frac{\partial^2 Y_t}{\partial x_i \partial x_j} = \frac{\sigma-1}{\sigma}\alpha^2 L_{1t}^{1-\alpha}\left(\sum_{s=1}^{A_t} x_{st}^{\frac{\alpha}{\sigma}}\right)^{\sigma-1} x_{it}^{\frac{\alpha}{\sigma}-1}x_{jt}^{\frac{\alpha}{\sigma}-1} \gtreqless 0 \text{ if } \sigma \gtreqless 1.$$

and

$$w_t = \frac{(1-\alpha) Y_t}{L_{1t}}. \tag{19}$$

The producer of intermediate good i is a monopolist with the following profit maximization program:

$$\begin{aligned} \max_{p_{it}, x_{it}} \pi_{it} &= \max_{p_{it}, x_{it}} (p_{it} x_{it} - d x_{it}), \\ s.t. &: \quad p_{it} = \frac{L_{1t}^{1-\alpha} \alpha \left(\sum_{i=1}^{A_t} x_{it}^{\frac{\alpha}{\sigma}} \right)^{\sigma-1} x_{it}^{\frac{\alpha}{\sigma}-1}}{(1+r_t)} \end{aligned} \tag{20}$$

where d is the depreciation rate of the intermediate good stock, which can be produced from the final good on a one-to-one basis. It can be shown that the Nash equilibrium of the game is symmetric, so the subscript i can be dropped. Using the following approximation based on the assumption that A_t is sufficiently large whenever appropriate

$$\left(\frac{\sigma - 1}{A_t} + 1 \right) \approx 1,$$

then it can be shown that equilibrium levels of output, price and profit in each intermediate good sector are given by,

$$x_t = \frac{\alpha^{\frac{2}{1-\alpha}} (A_t)^{\frac{\sigma-1}{1-\alpha}}}{\sigma^{\frac{1}{1-\alpha}} d^{\frac{1}{1-\alpha}} (1+r_t)^{\frac{1}{1-\alpha}}} L_{1t}, \tag{21}$$

$$p_{it} = \frac{\sigma d}{\alpha}, \tag{22}$$

and

$$\pi_t = \left(\frac{\sigma}{\alpha} - 1 \right) d x_t = \frac{\alpha^{\frac{1+\alpha}{1-\alpha}} (\sigma - \alpha) (A_t)^{\frac{\sigma-1}{1-\alpha}}}{\sigma^{\frac{1}{1-\alpha}} d^{\frac{\alpha}{1-\alpha}} (1+r_t)^{\frac{1}{1-\alpha}}} L_{1t}. \tag{23}$$

Both x_t and π_t are proportional to L_{1t} and decreasing in $(1+r_t)^{\frac{1}{1-\alpha}}$. They both are proportional to $(A_t)^{\frac{\sigma-1}{1-\alpha}}$. So given that A_t has a time trend in steady state, x_t and π_t both have a positive time trend in steady state if $\sigma > 1$.

The equilibrium relative price of labor is given by

$$\frac{w_t}{(1+r_t)} = \frac{(1-\alpha) \alpha^{\frac{2\alpha}{1-\alpha}}}{\sigma^{\frac{\alpha}{1-\alpha}} d^{\frac{\alpha}{1-\alpha}} (1+r_t)^{\frac{1}{1-\alpha}}} (A_t)^{\frac{\sigma-\alpha}{1-\alpha}}. \tag{24}$$

Denote by K_t the aggregate capital stock (employed in the final good sector), and define

$$K_t \triangleq A_t x_t = \frac{\alpha^{\frac{2}{1-\alpha}} (A_t)^{\frac{\sigma-\alpha}{1-\alpha}}}{\sigma^{\frac{1}{1-\alpha}} d^{\frac{1}{1-\alpha}} (1+r_t)^{\frac{1}{1-\alpha}}} L_{1t},$$

thereby, the final good sector output is

$$Y_t = (A_t)^{\sigma-\alpha} L_{1t}^{1-\alpha} (K_t)^{\alpha},$$

and the aggregate (final good) production function is

$$Y_t = (N_t)^{1-\alpha} L^{1-\alpha} (K_t)^{\alpha}, \tag{25}$$

where $N_t \triangleq (A_t)^{\frac{\sigma-\alpha}{1-\alpha}} \left(\frac{L_{1t}}{L}\right)$ corresponds to the labor augmenting factor, or $(N_t)^{1-\alpha}$ can be seen as what is known as the total factor productivity (TFP).

Denote by $\hat{k}_t$ and y_t the capital stock per effective unit of labor and (final good) output per capita respectively, then we have

$$\hat{k}_t \triangleq \frac{K_t}{N_t L} = \frac{x_t}{(A_t)^{\frac{\sigma-1}{1-\alpha}} L_{1t}} = \frac{\alpha^{\frac{2}{1-\alpha}}}{\sigma^{\frac{1}{1-\alpha}} d^{\frac{1}{1-\alpha}} (1+r_t)^{\frac{1}{1-\alpha}}}, \tag{26}$$

and

$$y_t \triangleq \frac{Y_t}{L} = N_t \hat{k}_t^{\alpha}. \tag{27}$$

It is clear from the above expressions that $\hat{k}_t$ is stationary in steady state given that r_t is stationary, and y_t has a positive time trend since $\sigma > \alpha$. It follows that

$$\ln \frac{y_t}{y_t^*} = \ln \frac{N_t}{N_t^*} + \alpha \ln \frac{\hat{k}_t}{\hat{k}_t^*} = \ln \frac{N_t}{N_t^*} - \frac{\alpha}{1-\alpha} \ln \frac{1+r_t}{1+r}. \tag{28}$$

where y_t^* and N_t^* are the steady-state trends of y_t and N_t. The implication is that the deviation of $\ln y_t$ from its steady-state trend $\ln y_t^*$ is determined by the deviation of $\ln N_t$ from its steady-state trend $\ln N_t^*$ and the deviation of $\ln \hat{k}_t$ (or $\ln(1+r_t)$ respectively) from its steady-state level $\ln \hat{k}_t^*$ (or $\ln(1+r)$ respectively).

Define the relative technological development position $a_t \triangleq \frac{A_t}{A_{ft}}$, where $A_{ft} \triangleq A_{f0}(1+g_f)^t$ is the knowledge stock of the world frontier and g_f is the constant growth rate of knowledge stock at the world frontier. The number of new intermediate products introduced in period $t+1$ as results of R&D activities at t is determined by the productivity of the R&D sector, $\frac{\delta}{(a_t)^{\eta}}$, the labor input in the R&D sector, L_{2t}, and the domestic knowledge stock at the time, A_t, i.e.,

$$A_{t+1} - A_t = \frac{\delta}{(a_t)^{\eta}} L_{2t} A_t. \tag{29}$$

L_{2t} is determined by the labor market clearing condition:

$$L_{2t} = L - L_{1t}. \tag{30}$$

The aggregate R&D productivity $\frac{\delta}{(a_t)^\eta}$ decreases with a_t. This is a standard feature of the technology-diffusion-based endogenous growth model. This feature is the driving force of trend-reversion of A_t, and parameter η measures the strength of the trend reversion.[16]

The aggregate R&D productivity is affected by some country fixed effects which are parameterized by δ. It covers a broad range of factors, including what Parente and Prescott (1994) call the "barriers to technology adoption".

Define the growth rate of domestic knowledge stock as

$$g_t \triangleq \frac{A_{t+1} - A_t}{A_t}. \tag{31}$$

It follows from (29) that

$$g_t = \delta \left(a_t\right)^{-\eta} L_{2t}, \tag{32}$$

which implies that in our model the growth domestic knowledge stock is determined by the allocation of labor force between production and R&D.

Denote by V_t the value of the ownership of one intermediate good firm. With free entry into the R&D sector, equilibrium entails the following zero-profit condition:

$$V_t \delta \left(a_t\right)^{-\eta} L_{2t} A_t - \frac{w_t}{1+r_t} L_{2t} = 0 \tag{33}$$

where $V_t \delta \left(a_t\right)^{-\eta} L_{2t} A_t$ is the expected present value of an R&D project that employs L_{2t} units of labor, $\frac{w_t}{1+r_t} L_{2t}$ is the present value of its labor cost. Reorganizing using eq. (24), therefore

$$V_t = \frac{(1-\alpha)\, \alpha^{\frac{2\alpha}{1-\alpha}} \left(A_t\right)^{\frac{\sigma-1}{1-\alpha}} \left(a_t\right)^{\eta}}{\delta \sigma^{\frac{\alpha}{1-\alpha}} d^{\frac{\alpha}{1-\alpha}} \left(1+r_t\right)^{\frac{1}{1-\alpha}}}. \tag{34}$$

The asset value of a firm, V_t, is proportional to $\left(A_t\right)^{\frac{\sigma-1}{1-\alpha}}$. It thus has a positive time trend in steady state if $\sigma > 1$, which implies that the stock price should have a positive time trend in steady state. Then factor $(\sigma - 1)$ determines the strength of the asset value appreciation. If A_t grows faster (slower) than its trend rate, *ceteris paribus*, then the appreciation of asset value will be faster (slower) than the corresponding trend rate.

[16] To see this, note that the R&D productivity $\delta \left(a_t\right)^{-\eta}$ can be rewritten as

$$\delta \left(\frac{A_t}{A_{ft}}\right)^{-\eta} = \delta \left(\frac{A_t^*}{A_{ft}} \frac{A_t}{A_t^*}\right)^{-\eta} = \delta \left(a^* \frac{A_t}{A_t^*}\right)^{-\eta} = \delta \left(a^*\right)^{-\eta} e^{-\eta \ln \frac{A_t}{A_t^*}},$$

where a^* is the steady-state value of a_t and $A_t^* = a^* A_{ft}$ is the steady-state trend of A_t. Hence $\delta \left(a_t\right)^{-\eta}$ is a decreasing function of $\ln \frac{A_t}{A_t^*}$, which is a measurement of the deviation of A_t from its steady-state trend A_t^*. Parameter η determines the effect of this deviation on R&D productivity.

The size of R&D labor force is given by

$$L_{2t} = \frac{g_t\,(a_t)^{\eta}}{\delta},$$

and hence from eq. (23) it follows that the dividend income from the ownership of an intermediate firm is

$$\pi_t = \frac{\alpha^{\frac{1+\alpha}{1-\alpha}}\,(\sigma-\alpha)\,(A_t)^{\frac{\sigma-1}{1-\alpha}}\left(L - \frac{g_t(a_t)^{\eta}}{\delta}\right)}{\sigma^{\frac{1}{1-\alpha}}\,d^{\frac{\alpha}{1-\alpha}}\,(1+r_t)^{\frac{1}{1-\alpha}}}. \tag{35}$$

The non-arbitrage condition of an equilibrium

$$\pi_{t+1} + V_{t+1} = (1+r_t)\,V_t, \tag{36}$$

implies

$$g_{t+1} = \frac{\delta L}{(a_{t+1})^{\eta}} - \frac{\sigma\,(1-\alpha)\left(\frac{(1+r_{t+1})^{\frac{1}{1-\alpha}}\,(a_t)^{\eta}}{(1+r_t)^{\frac{\alpha}{1-\alpha}}\,(a_{t+1})^{\eta}}\right)}{\alpha\,(\sigma-\alpha)\,(1+g_t)^{\frac{\sigma-1}{1-\alpha}}} + \frac{\sigma\,(1-\alpha)}{\alpha\,(\sigma-\alpha)}. \tag{37}$$

When the domestic knowledge stock of an economy grows at the rate of g_t, its position of relative development changes according to the following identity:

$$a_{t+1} \equiv a_t \frac{1+g_t}{1+g_f}, \tag{38}$$

thereby, it catches up if $g_t > g_f$; it lags behind if $g_t < g_f$; in the steady state, $g_t = g^* = g_f$ and $a_{t+1} = a_t = a^*$.

The final good market clearing condition is given by

$$Y_t = C_{t+1} + x_{t+1}A_{t+1} - (1-d)\,x_t A_t \tag{39}$$

which implies

$$\frac{\left(L - \frac{g_{t+1}(a_{t+1})^{\eta}}{\delta}\right)}{\sigma^{\frac{1}{1-\alpha}}\,d^{\frac{1}{1-\alpha}}\,(1+r_{t+1})^{\frac{1}{1-\alpha}}} = \frac{\frac{\alpha^{\frac{2\alpha}{1-\alpha}}\left(\frac{1}{1+g_t}\right)^{\frac{\sigma-\alpha}{1-\alpha}}\left(L - \frac{g_t(a_t)^{\eta}}{\delta}\right)}{\sigma^{\frac{\alpha}{1-\alpha}}\,d^{\frac{\alpha}{1-\alpha}}\,(1+r_t)^{\frac{\alpha}{1-\alpha}}}\left(1 + \frac{\alpha^2(1-d)}{\sigma d(1+r_t)}\right) - \hat{c}_{t+1}}{\alpha^{\frac{2}{1-\alpha}}}, \tag{40}$$

where $\hat{c}_{t+1} \triangleq \frac{C_{t+1}}{(A_{t+1})^{\frac{\sigma-\alpha}{1-\alpha}}}$ is the normalized level of total consumption. The *Euler* condition (14) can now be rewritten as

$$\hat{c}_{t+1} = \left(\frac{1+r_t}{1+\rho}\right)^{\frac{1}{\theta}} \frac{\hat{c}_t}{(1+g_t)^{\frac{\sigma-\alpha}{1-\alpha}}}. \tag{41}$$

Equations (38), (37), (40) and (41) can be used to construct a recursive system of difference equations (54), which is shown in appendix B.

The complete set of steady-state values of the four state variables are given by (55) in Appendix B. Here we only show the following result:

$$a^* = \left(\frac{\alpha\delta L}{r + \alpha g_f}\right)^{\frac{1}{\eta}}. \tag{42}$$

Clearly, the steady-state level of a_t, hence, the steady-state trend of A_t, depends on parameter δ, which summarizes all the information about the country fixed effects.

To analyze the transition dynamics of the model, we log-linearize the difference equations system (54) as follow:

$$\begin{cases} \ln \frac{a_{t+1}}{a^*} = \ln \frac{a_t}{a^*} + \ln \frac{1+g_t}{1+g^*} \\ \ln \frac{1+g_{t+1}}{1+g^*} = B_1 \ln \frac{a_t}{a^*} + B_2 \ln \frac{1+g_t}{1+g^*} + B_3 \ln \frac{1+r_t}{1+r^*} + B_4 \ln \frac{\hat{c}_t}{\hat{c}^*} \\ \ln \frac{1+r_{t+1}}{1+r^*} = G_1 \ln \frac{a_t}{a^*} + G_2 \ln \frac{1+g_t}{1+g^*} + G_3 \ln \frac{1+r_t}{1+r^*} + G_4 \ln \frac{\hat{c}_t}{\hat{c}^*} \\ \ln \frac{\hat{c}_{t+1}}{\hat{c}^*} = D_2 \ln \frac{1+g_t}{1+g^*} + D_3 \ln \frac{1+r_t}{1+r^*} + \ln \frac{\hat{c}_t}{\hat{c}^*} \end{cases} \tag{43}$$

where the coefficients B_i, G_i for $i = 1, 2, 3$ and D_j for $j = 2, 3$ are defined in Appendix B. The characteristic polynomial of the above linear difference equations system is:

$$X^4 + \psi_1 X^3 + \psi_2 X^2 + \psi_3 X + \psi_4, \tag{44}$$

where the coefficients ψ_i for $i = 1, 2, 3, 4$ are also are defined in Appendix B. Let χ_j for $j = 1, 2, 3, 4$ be the roots of eq. (44), i.e., the eigen values of the system (43). Then the time series of $\ln \frac{a_t}{a^*}$, $\ln \frac{1+g_t}{1+g^*}$, $\ln \frac{1+r_t}{1+r^*}$ and $\ln \frac{\hat{c}_t}{\hat{c}^*}$ are all solutions to the same fourth-order linear difference equation as follows:

$$z_t = -\psi_1 z_{t-1} - \psi_2 z_{t-2} - \psi_3 z_{t-3} - \psi_4 z_{t-4}, \tag{45}$$

i.e., they all follow the same AR(4) process, and have the same general solution:

$$z_t = \tilde{C}_1^k \chi_1 + \tilde{C}_2^k \chi_2 + \tilde{C}_3^k \chi_3 + \tilde{C}_4^k \chi_4, \tag{46}$$

where $\tilde{C}_j^k$, for $j = 1, 2, 3, 4$, and $k = \ln \frac{a_t}{a^*}$, $\ln \frac{1+g_t}{1+g^*}$, $\ln \frac{1+r_t}{1+r^*}$ and $\ln \frac{\hat{c}_t}{\hat{c}^*}$, are arbitrary constant coefficients.

5 Calibration and Estimation of the Model

In what follows we show that the time series of $\ln \frac{y_t}{y_t^*}$ should (approximately) follow the AR(4) process given by eq. (45). From (28) it follows that

$$\ln \frac{y_t}{y_t^*} = \frac{\sigma - \alpha}{1 - \alpha} \ln \frac{A_t}{A_t^*} + \ln \frac{L_{1t}}{L_1^*} - \frac{\alpha}{1 - \alpha} \ln \frac{1 + r_t}{1 + r},$$

where L_1^* is the steady-state level of labor input in the final good sector. It can be shown that

$$\ln \frac{A_t}{A_t^*} = \ln \frac{a_t}{a^*}$$

and

$$\ln \frac{L_{1t}}{L_1^*} = \ln \frac{L - L_{2t}}{L - L_2^*} = \ln \frac{L - \frac{g_t (a_t)^\eta}{\delta}}{L - \frac{g_f (a^*)^\eta}{\delta}} \approx -\frac{\eta \alpha g_f}{r} \ln \frac{a_t}{a^*} - \frac{\alpha (1+g_f)}{r} \ln \frac{1+g_t}{1+g_f}.$$

Consequently

$$\ln \frac{y_t}{y_t^*} \approx \left(\frac{\sigma - \alpha}{1-\alpha} - \frac{\eta \alpha g_f}{r} \right) \ln \frac{a_t}{a^*} - \frac{\alpha (1+g_f)}{r} \ln \frac{1+g_t}{1+g_f} - \frac{\alpha}{1-\alpha} \ln \frac{1+r_t}{1+r}, \tag{47}$$

i.e., $\ln \frac{y_t}{y_t^*}$ is a linear combination of the time series of $\ln \frac{a_t}{a^*}$, $\ln \frac{1+r_t}{1+r}$ and $\ln \frac{1+r_t}{1+r}$, each of which follows the AR(4) process described by eq. (45). As a result, $\ln \frac{y_t}{y_t^*}$ must also follow the same AR(4) process described by eq. (45). Given that L_{1t} and $\hat{k}_t$ are stationary in steady state, from (27) it follows that the steady-state growth rates of y_t and A_t have the following relationship:

$$g_f = (1+g_y)^{\frac{1-\alpha}{\sigma-\alpha}} - 1. \tag{48}$$

Matching the theoretical and the empirical characteristic polynomials (44) and (11), we arrive at the following four equations.

$$\begin{cases} \psi_1 = -\phi_1 \\ \psi_2 = -\phi_2 \\ \psi_3 = -\phi_3 \\ \psi_4 = -\phi_4 \end{cases} \tag{49}$$

5.1 Calibration

Since parameters L and δ do not feature in the theoretical characteristic polynomial (44), we have seven independent parameters to be determined. They are α, g_y, ρ, η, σ, θ and d. We calibrate α, ρ and g_y, and then retrieve the values of θ, σ, η and d using the equations system (49) and the empirical values of ϕ_1, ϕ_2, ϕ_3 and ϕ_4 reported in the section 2.

We set $\alpha = 0.3$, which is the commonly agreed the value in the empirical growth literature. We set $g_y = 0.02$, which is very close to our own empirical result, and is also the commonly agreed value. We choose two alternative values for ρ, namely, $\rho = 0.03$ and $\rho = 0.05$, which are by no means unusual. The values of ϕ_1, ϕ_2, ϕ_3 and ϕ_4 come from the four alternative estimations (5)-(8) reported in Table 2. Therefore we have eight alternative sets of estimations of θ, σ, η and d, as reported in Table 4.

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
α	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
g_y	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
ρ	0.03	0.03	0.03	0.03	0.05	0.05	0.05	0.05
η	0.031177	0.031093	0.013545	0.02156	0.020311	0.020297	0.0089202	0.009573
σ	1.2179	1.2166	1.2417	1.2287	1.2348	1.233	1.2579	1.2443
θ	0.025605	0.02.4765	0.030704	0.027182	0.038722	0.037616	0.047414	0.041473
d	0.091092	0.097593	0.11254	0.10153	0.079232	0.08501	0.0009864	0.088709
g_f	0.015216	0.01.5238	0.014829	0.015038	0.014939	0.014968	0.014576	0.014788
r	0.030522	0.03.0505	0.030626	0.030555	0.050805	0.050782	0.050986	0.050863
γ_r	0.53161	0.55415	0.54730	0.54581	0.47678	0.49508	0.47907	0.48442

Table 4: Calibration and estimation of the model

The estimated values of σ are in the range between 1.21-1.26, which is significantly above 1. This implies that future new products on average are direct complements rather than direct substitutes to the existing products. Their introduction is likely to enhance of the value of existing products, and cause capital gains to the existing stock shares. This result is consistent with the observation that the stock indices, such as the Dow Jones, or the FTSE, have positive time trends.

The estimated values of η are rather small, in the range between 0.008-0.032. This suggests that the strength of the trend-reversion of the domestic knowledge stock is rather weak. The parameters ρ and d are to some extent substitutes. Larger ρ is offset by smaller d. The estimated values of θ are very small, which implies that the consumers do not care too much about smoothing consumptions. The implied supply elasticity of capital with respect to risk-free interest rate, consequently, is very high.

In Appendix C, a discrete-time Ramsey growth model is presented as a benchmark. Thereby, the speed of convergence in the unique saddle-path-stable equilibrium, γ_r, is given by eq. (72). With $\alpha = 0.3$, $g_y = 0.02$, $\rho \in [0.03, 0.05]$, $\theta \in (0.02, 0.05)$, $\sigma \in [1.21, 1.26]$ and $d \in (0.07, 0.12)$, the Ramsey growth model would predict a time-invariant speed of convergence in the range of 47-56% per year. This indicates that the neoclassical capital deepening process *per se* would be a very fast non-cyclical trend-reversion process. When interacting with the process of TFP growth, it should respond to the TFP growth very swiftly. The slowness of the trend-reversion of per capita income, therefore, is primarily due to the rather weak trend-reversion of the domestic knowledge stock, as is measured by η. To confirm this intuition, we do the following exercise to highlight the role of the primary economic mechanism.

5.2 Explaining the long cycles

By considering the limiting case: $\theta \to 0$, we can completely trivialize the neoclassical capital deepening mechanism. Now we have

$$r_t = r = \rho, \tag{50}$$

therefore the channel whereby capital deepening can interact with TFP growth through interest rate variation is shut. Then eq. (37) becomes

$$g_{t+1} = \frac{\delta L}{(a_t)^\eta \left(\frac{1+g_t}{1+g_f}\right)^\eta} - \frac{\sigma(1-\alpha)(1+\rho)}{\alpha(\sigma-\alpha)(1+g_t)^{\frac{\sigma-1}{1-\alpha}}\left(\frac{1+g_t}{1+g_f}\right)^\eta} + \frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)} \tag{51}$$

When η is small, we can use the following approximation

$$\left(\frac{1+g_t}{1+g_f}\right)^\eta \approx 1$$

and hence

$$g_{t+1} \approx \frac{\delta L}{(a_t)^\eta} - \frac{\sigma(1-\alpha)(1+\rho)}{\alpha(\sigma-\alpha)(1+g_t)^{\frac{\sigma-1}{1-\alpha}}} + \frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}. \tag{52}$$

It is more convenient to use the following the log-linearized approximation:

$$\ln\frac{1+g_{t+1}}{1+g_f} \approx -\eta\frac{(\rho+\alpha g_f)}{\alpha(1+g_f)}\ln\frac{a_t}{a^*} + (\sigma-1)\frac{\sigma(1+\rho)}{\alpha(\sigma-\alpha)(1+g_f)^{\frac{\sigma-\alpha}{1-\alpha}}}\ln\frac{1+g_t}{1+g_f}. \tag{53}$$

The first term in the right-hand side of the above equation shows that a positive value of η tends to lead to a trend-reversion of a_t to a^* (or A_t to A_t^*). If A_t is below the trend A_t^*, then A_t tends grow faster than the trend. When η is small eq. (34) approximately becomes

$$V_t \approx \frac{(1-\alpha)\,\alpha^{\frac{2\alpha}{1-\alpha}}(A_t)^{\frac{\sigma-1}{1-\alpha}}}{\delta\sigma^{\frac{\alpha}{1-\alpha}}d^{\frac{\alpha}{1-\alpha}}(1+r_t)^{\frac{1}{1-\alpha}}},$$

which implies that the expected reversion of A_t toward A_t^* from below enhances the asset price appreciation in the time series of V_t. The expected increase in capital gain accelerates the investments in R&D and the growth in A_t, which then feed back positively on asset price appreciation in the time series of V_t. This positive feedback loop is captured by the second term in the right-hand of eq. (53). The factor $(\sigma-1)$ which affects the strength of this positive feedback mechanism is quite large in our empirical result. Since parameter η which measures the strength of trend-reversion of a_t to a^* (or A_t to A_t^*) is small in the empirical result, the trend-reversion force is not strong enough to stop the time series of A_t from overshooting its trend A_t^* and then (transitorily) diverging from it. The smaller η is, the longer it takes for the (transitory) divergence to end and for A_t to start reversing to A_t^* from above. This kind of repeated sequence of trend-reversion, over-shooting, (transitory) divergence, and then trend-reversion again forms the long cycles in conditional convergence.

5.3 Discussion

The purpose of the growth current model is to explain long term growth, i.e., the very low frequency components in the time-series variance of per capita incomes. It is not designed to explain the usual business cycle components in range of 2-8 years of period of oscillation. The fact that the model can match the two separate empirical bands of frequencies indicates that the cycles in the range of 9-14 years cyclical periods may be due to the interaction between the neoclassical capital accumulation/deepening mechanism and the endogenous technological progress. Without the mechanism of endogenous technological progress, there would have been no long cycles; without the influence of the neoclassical capital accumulation mechanism, all the cycles would have been in the very low frequency band. The above analysis suggests that the shorter cycles are linked to the channel of interest rate variation through which the two mechanisms interact.

In the current study, the steady-state trends of per capita GDP, comprising both the trend growth rates and the intercepts, have been treated as time-invariant, as is common in the conditional convergence literature. Presumably, the steady-state trends may be subject to random shocks that have permanent effects, i.e., stochastic trend shifts, therefore it is important to know how restrictive this assumption is when confronted to the data. The unit root tests results from this study suggest that when coupled with a suitably-specified cyclical growth model, this assumption appears to be appropriate for our sample as a first order approximation.

6 Conclusion

In this study we find strong evidence that the long term conditional convergence of per capita income possesses low frequency cyclical components. This finding contradicts the conventional assumption that the speed of convergence is time-invariant. Consequently, the study rejects growth models that predict time-invariant speed of convergence, including a broad class of neoclassical growth models, as explanations for long-term economic growth. We propose that the long cycles in the trend-reversion process of per capita income can be explained by the endogenous growth of total factor productivity under the influence of international technological diffusion.

With international technological diffusion, the opportunities for technologically backward economies to emulate the more advanced economies tend to lead all economies to converge to parallel steady-state trends of per capita incomes. The disparities in the steady-state trends are caused by country-specific fixed effects, such as the quality of institutions. The focus of the current paper is to understand the process of per capita incomes to revert to their steady-state trends. We analyze two economic mechanisms underlying this trend-reversion process, in their joint presence, and in isolation. The first is the neoclassical capital deepening mechanism, the second is the mechanism of endogenous investments in technological progress and total factor productivity. Our quantitative analysis shows that on the one hand, the reversion of capital intensity to its steady-state trend *per se* would be a very fast non-cyclical process; on the other hand, the reversion of total factor productivity to its steady-state trend is a slow and cyclical process. The long-term conditional convergence of per capita incomes is primarily explained by the trend-reversion process of the total factor productivity.

The focus of the current study has been on understanding the primary economic mechanism that generates long growth cycles. This inevitably leaves the details about how TFP growth and capital deepening interact, and the determinants of the shorter cycles under-explored. These remain interesting open questions to be addressed by future research.

References

Aghion, Philippe, Peter Howitt, and David Mayer-Foulkes, "The Effect of Financial Development on Convergence: Theory and Evidence," *Quarterly Journal of Economics*, 2005, *120*, 173–222.

Barro, Robert and Xavier Sala-I-Martin, "Convergence," *Journal of Political Economics*, 1992, *100*, 223–51.

Barro, Robert J. and Xavier Sala-I-Martin, "Technological Diffusion, Convergence, and Growth," *Journal of Economic Growth*, 1997, *2*, 1–26. NBER Working Paper No. 5151.

Barro, Robert, N. Gregory Mankiw, and Xavier Sala-I-Martin, "Capital Mobility in Neoclassical Models of Growth," *American Economic Review*, 1995, *85*, 103–115.

Baxter, M. and R. G. King, "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, 1999, *81*, 5–37.

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort, "Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics," *Journal of Economic Growth*, 1996.

Cass, David, "Optimal Growth in an Aggregative Model of Capital Accumulation," *Review of Economic Studies*, 1965, *32*, 233–240.

Dixit, Avinash K. and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 1977, *67*, 297–308.

Hodrick, Robert J. and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 1997, *29*, 1–16.

Howitt, Peter, "Endogenous Growth and Gross-Country Income Differences," *American Economic Review*, 2000, *90*, 829–46.

Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin, "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 2003, *115*, 53–74.

Islam, Nazrul, "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 1995, *110*, 1127–70.

Kondratieff, N. D., "The Long Waves in Economic Life," *Review of Economic Statistics*, 1935, *17*, 105–115.

Koopmans, Tjalling C., "On the Concept of Optimal Economic Growth," in "The Econometric Approach to Development Planning," Amsterdam: North Holland, 1965.

Kuznets, Simon, "Schumpeter's Business Cycles," *American Economic Review*, 1940, *30*, 257–271.

Maddala, G. S. and Shaowen Wu, "A Comparative Study of Unit Root Tests with Panel Data and A New Simple Test," *Oxford Bulletin of Economics and Statistics*, 1999, *Special Issue*, 631–652.

Maddison, Angus, *The World Economy: Historical Statistics*, OECD Publishing, 2003.

Mankiw, Gregory, David Romer, and David Weil, "A Contribution to the Empirics of Economic Growth," *Quartely Journal of Economics*, 1992, *107*, 407–37.

Parente, Stephen and Edward Prescott, "Barriers to Technology Adoption and Development," *Journal of Political Economy*, 1994, *102*, 298–321.

Phillips, Peter C. B. and Donggyu Sul, "The Elusive Empirical Shadow of Growth Convergence," February 2003. Cowles Foundation Disscussion Paper No. 1398.

Ramsey, Frank, "A Mathematical Theory of Saving," *Economic Journal*, 1928, *38*, 543–559.

Ravan, Morten O. and Harald Uhlig, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, 2002, *84*, 371–380.

Romer, Paul, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98, Part II*, S71–S102.

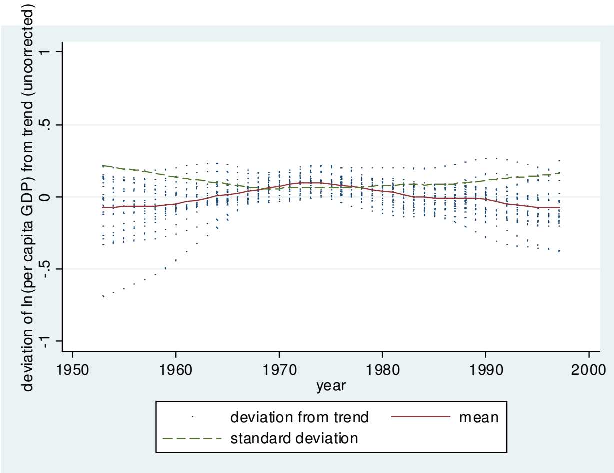

Figure 10: The deviation of $\ln y_{it}$ from linear trend (uncorrected)

Schumpeter, Joseph A., *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, New York: McGraw-Hill, 1939.

Solow, Robert M., "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 1956, *1956*, 65–94.

Spence, Michael, "Product Selection, Fixed Costs, and Monopolistic Competition," *Review of Economic Studies*, 1976, *43*, 217–235.

Swan, Trevor W., "Economic Growth and Capital Accumulation," *Economic Record*, 1956, *32*, 334–361.

Tong, Jian and Chenggang Xu, "Financial Institutions and Wealth of Nations: Tales of Development," 2004. CEPR Discussion Paper Series, No. 4348.

____ **and** ____ , "Emulation-based Growth: Technology and Institutions," 2006. (unpublished).

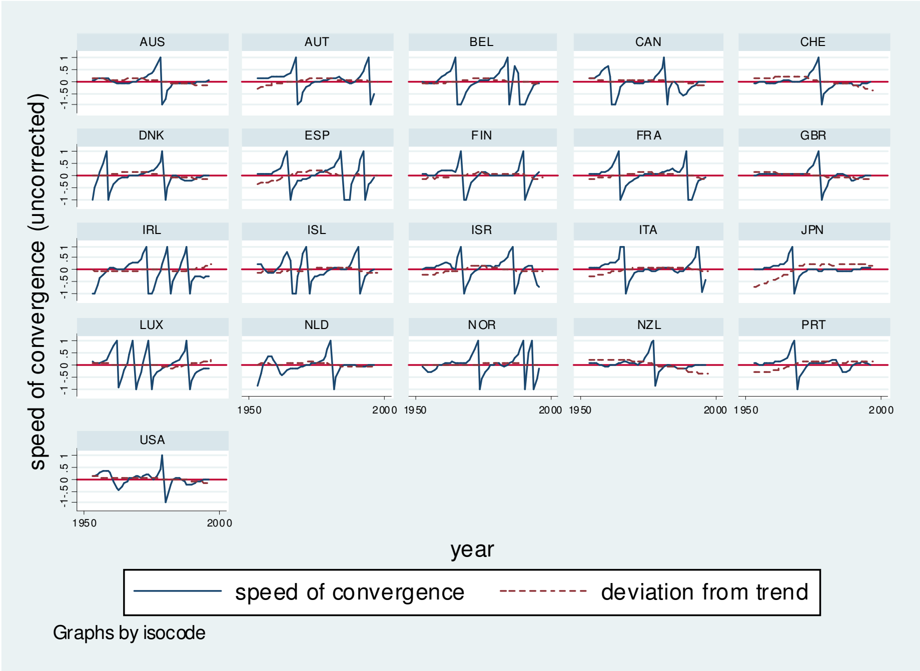

Figure 11: The time series of speed of convergence and the deviation of $\ln y_{it}$ from the linear trend for 21 countries (uncorrected)

A Deviation from Trend and Speed of Convergence (Uncorrected Resulted)

B Difference Equations

The recursive difference equations system is:

(54)

$$\begin{cases} a_{t+1} = a_t \frac{1+g_t}{1+g_f} \\ g_{t+1} = \frac{\delta L}{\left(a_t \frac{1+g_t}{1+g_f}\right)^{\eta}} - \frac{\frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}(a_t)^{\eta}}{(a_t)^{\eta} - \frac{\delta(\sigma-\alpha)d}{\alpha(1-\alpha)} \frac{\left(L - \frac{g_t (a_t)^{\eta}}{\delta}\right)}{(1+g_t)} \left(1 + \frac{\alpha^2(1-d)}{\sigma d(1+r_t)}\right) + \frac{\delta(\sigma-\alpha)\sigma^{\frac{\alpha}{1-\alpha}} d^{\frac{1}{1-\alpha}} (1+r_t)^{\frac{\alpha}{1-\alpha}} \left(\frac{1+r_t}{1+\rho}\right)^{\frac{1}{\theta}} \hat{c}_t}{\alpha^{\frac{1+\alpha}{1-\alpha}} (1-\alpha)(1+g_t)}} + \frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)} \\ 1 + r_{t+1} = \frac{\left(a_t \frac{1+g_t}{1+g_f}\right)^{\eta(1-\alpha)} (1+r_t)^{\alpha} (1+g_t)^{\sigma-1}}{\left((a_t)^{\eta} - \frac{\delta(\sigma-\alpha)d}{\alpha(1-\alpha)} \frac{\left(L - \frac{g_t (a_t)^{\eta}}{\delta}\right)}{(1+g_t)} \left(1 + \frac{\alpha^2(1-d)}{\sigma d(1+r_t)}\right) + \frac{\delta(\sigma-\alpha)\sigma^{\frac{\alpha}{1-\alpha}} d^{\frac{1}{1-\alpha}} (1+r_t)^{\frac{\alpha}{1-\alpha}} \left(\frac{1+r_t}{1+\rho}\right)^{\frac{1}{\theta}} \hat{c}_t}{\alpha^{\frac{1+\alpha}{1-\alpha}} (1-\alpha)(1+g_t)}\right)^{(1-\alpha)}} \\ \hat{c}_{t+1} = \left(\frac{1+r_t}{1+\rho}\right)^{\frac{1}{\theta}} \frac{\hat{c}_t}{(1+g_t)^{\frac{\sigma-\alpha}{1-\alpha}}}. \end{cases}$$

The steady-state values of the four state variables are given by

$$
\left\{
\begin{array}{c}
a^* = \left(\frac{\alpha\delta L}{r+\alpha g_f}\right)^{\frac{1}{\eta}} \\
g^* = g_f \\
r^* = r = (1+\rho)(1+g_f)^{\frac{\theta(\sigma-\alpha)}{1-\alpha}} - 1 \\
\hat{c}^* = \dfrac{\alpha^{\frac{2\alpha}{1-\alpha}}(1-\alpha)\left(\frac{(\sigma-\alpha)\left(\sigma d(1+r)+\alpha^2(1-d)\right)}{(1-\alpha)(1+g)} - \frac{\sigma(1+r)(\sigma-\alpha)\alpha^2}{r(\sigma-\alpha)+\sigma(1-\alpha)}\right)Lr}{(\sigma-\alpha)\sigma^{\frac{1}{1-\alpha}} d^{\frac{1}{1-\alpha}}(1+r)^{\frac{1}{1-\alpha}}(1+g)^{\frac{\sigma-1}{1-\alpha}}(r+\alpha g)}
\end{array}
\right. \tag{55}
$$

The log-linearized difference equations are:

$$
\left\{
\begin{array}{c}
\ln \frac{a_{t+1}}{a^*} = \ln \frac{a_t}{a^*} + \ln \frac{1+g_t}{1+g^*} \\
\ln \frac{1+g_{t+1}}{1+g^*} = B_1 \ln \frac{a_t}{a^*} + B_2 \ln \frac{1+g_t}{1+g^*} + B_3 \ln \frac{1+r_t}{1+r^*} + B_4 \ln \frac{\hat{c}_t}{\hat{c}^*} \\
\ln \frac{1+r_{t+1}}{1+r^*} = G_1 \ln \frac{a_t}{a^*} + G_2 \ln \frac{1+g_t}{1+g^*} + G_3 \ln \frac{1+r_t}{1+r^*} + G_4 \ln \frac{\hat{c}_t}{\hat{c}^*} \\
\ln \frac{\hat{c}_{t+1}}{\hat{c}^*} = D_2 \ln \frac{1+g_t}{1+g^*} + D_3 \ln \frac{1+r_t}{1+r^*} + \ln \frac{\hat{c}_t}{\hat{c}^*}
\end{array}
\right.
$$

where

$$
B_1 = -\eta\frac{r+\alpha g_f}{\alpha\left(1+g_f\right)} - \frac{\eta\frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}(1+r)}{\left(1+g_f\right)\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}} + \frac{\frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}(1+r)^2\left(\eta+\eta\frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{g_f}{\left(1+g_f\right)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right)\right)}{\left(1+g_f\right)\left(\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}\right)^2}
$$

$$
B_2 = -\eta\frac{r+\alpha g_f}{\alpha\left(1+g_f\right)} + \frac{\frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}(1+r)^2\left(\frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right) - \frac{\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}{(1+r)} + 1\right)}{\left(1+g_f\right)\left(\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}\right)^2}
$$

$$
B_3 = \frac{\frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}(1+r)^2\left(\frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{r}{\left(1+g_f\right)}\left(\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right) + \left(\frac{\alpha}{1-\alpha}+\frac{1}{\theta}\right)\left(\frac{\alpha\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}{(1+r)} - \alpha + \frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{r}{\left(1+g_f\right)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right)\right)\right)}{\alpha\left(1+g_f\right)\left(\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}\right)^2}
$$

$$
B_4 = \frac{\frac{\sigma(1-\alpha)}{\alpha(\sigma-\alpha)}(1+r)^2\left(\frac{\alpha\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}{(1+r)} - \alpha + \frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{r}{\left(1+g_f\right)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right)\right)}{\alpha\left(1+g_f\right)\left(\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}\right)^2}
$$

$$
G_1 = \eta(1-\alpha) - \frac{(1-\alpha)(1+r)\left(\eta+\eta\frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{g_f}{\left(1+g_f\right)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right)\right)}{\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}
$$

$$
G_2 = (\eta(1-\alpha)+(\sigma-1)) - \frac{(1-\alpha)(1+r)\left(\frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right) - \frac{\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}{(1+r)} + 1\right)}{\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}
$$

$$
G_3 = \alpha - \frac{(1-\alpha)(1+r)\left(\frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{r}{\left(1+g_f\right)}\left(\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right) + \left(\frac{\alpha}{1-\alpha}+\frac{1}{\theta}\right)\left(\frac{\alpha\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}{(1+r)} - \alpha + \frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{r}{\left(1+g_f\right)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right)\right)\right)}{\alpha\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}
$$

$$
G_4 = -\frac{(1-\alpha)(1+r)\left(\frac{\alpha\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}{(1+r)} - \alpha + \frac{(\sigma-\alpha)d}{\alpha(1-\alpha)}\frac{r}{\left(1+g_f\right)}\left(1+\frac{\alpha^2(1-d)}{\sigma d(1+r)}\right)\right)}{\alpha\left(1+g_f\right)^{\frac{\sigma-1}{1-\alpha}}}
$$

$$
D_2 = -\frac{\sigma-\alpha}{1-\alpha}
$$

$$
D_3 = \frac{1}{\theta}
$$

The linear difference equations system has the following characteristic polynomial:

$$X^4 + \psi_1 X^3 + \psi_2 X^2 + \psi_3 X + \psi_4$$

where

$\psi_1 = (-B_2 - G_3 - 2)$

$\psi_2 = (2B_2 - B_1 + G_3 + G_3 (B_2 + 1) - D_2 B_4 - D_3 G_4 - B_3 G_2 + 1)$

$\psi_3 = B_1 - B_2 - G_3 (B_2 + 1) + (D_2 B_4 + D_3 G_4)(B_2 + G_3 + 1)$

$-B_3 G_1 + B_3 G_2 - D_2 (B_2 B_4 + B_3 G_4) - D_3 (B_4 G_2 + G_3 G_4)$

$+G_3 (B_1 - B_2) + B_3 G_2 (B_2 + 1) - B_2 B_3 G_2$

$\psi_4 = (D_2 B_4 + D_3 G_4)(B_1 - B_2 - G_3 (B_2 + 1) + B_3 G_2)$

$-D_3 (B_4 G_1 + G_2 (B_2 B_4 + B_3 G_4) + G_3 (B_4 G_2 + G_3 G_4))$

$-D_2 (B_1 B_4 + B_2 (B_2 B_4 + B_3 G_4) + B_3 (B_4 G_2 + G_3 G_4)) + B_3 G_1 - G_3 (B_1 - B_2)$

$+ (D_2 (B_2 B_4 + B_3 G_4) + D_3 (B_4 G_2 + G_3 G_4))(B_2 + G_3 + 1)$

$-B_3 G_2 (B_2 + 1) + B_2 B_3 G_2$

C Ramsey Growth Model

We consider a Ramsey model with exogenous technological progress. The aggregate production function is given by

$$Y_t = L^{1-\alpha} (A_t)^{\sigma} (x_t)^{\alpha}, \tag{56}$$

and can be rewritten as

$$Y_t = L^{1-\alpha} (A_t)^{\sigma-\alpha} (K_t)^{\alpha}, \tag{57}$$

where $K_t \equiv A_t x_t$, $A_t = A_0 (1 + g_f)^t$. Define $N_t \triangleq (A_t)^{\frac{\sigma-\alpha}{1-\alpha}}$ as the labor augmenting factor, hence the steady-state growth rate of N_t is

$$g_N = (1 + g_f)^{\frac{\sigma-\alpha}{1-\alpha}} - 1. \tag{58}$$

and the production function can be rewritten as:

$$Y_t = L^{1-\alpha} (N_t)^{1-\alpha} (K_t)^{\alpha}. \tag{59}$$

Define $\hat{y}_t \triangleq \frac{Y_t}{LN_t}$ and $\hat{k}_t \triangleq \frac{K_t}{LN_t}$, we have

$$\hat{y}_t = \hat{k}_t^{\alpha}. \tag{60}$$

The *Euler* condition for optimal consumption is

$$\frac{C_{t+1}}{C_t}=\left(\frac{1+r_t}{1+\rho}\right)^{\frac{1}{\theta}}. \tag{61}$$

The law of motion for capital stock is

$$K_{t+1}=K_t\left(1-d\right)+I_{t+1} \tag{62}$$

where d is the depreciation rate, I_{t+1} is investment. Define $\hat{\imath}_{t+1}\triangleq\frac{I_{t+1}}{LN_{t+1}}$. Hence we have

$$\hat{k}_{t+1}=\hat{k}_t\frac{1-d}{1+g_N}+\hat{\imath}_{t+1}. \tag{63}$$

The market clearing condition is:

$$Y_t=C_{t+1}+I_{t+1} \tag{64}$$

which implies

$$\frac{\hat{y}_t}{1+g_N}=\hat{c}_{t+1}+\hat{\imath}_{t+1} \tag{65}$$

and

$$\hat{k}_{t+1}=\hat{k}_t\frac{1-d}{1+g_N}+\frac{\hat{y}_t}{1+g_N}-\hat{c}_{t+1} \tag{66}$$

where $\hat{c}_{t+1}\equiv\frac{C_{t+1}}{LN_t}$.

The first order condition for profit maximization entails that

$$r_t+d=\alpha\hat{k}_t^{\alpha-1}. \tag{67}$$

The laws of motion for $\hat{k}_t$, $\hat{c}_t$ are given by

$$\begin{cases}\hat{k}_{t+1}=\frac{1-d}{1+g_N}\hat{k}_t+\frac{\hat{k}_t^{\alpha}}{1+g_N}-\left(\frac{1+\alpha\hat{k}_t^{\alpha-1}-d}{1+\rho}\right)^{\frac{1}{\theta}}\frac{\hat{c}_t}{1+g_N}\\ \hat{c}_{t+1}=\left(\frac{1+\alpha\hat{k}_t^{\alpha-1}-d}{1+\rho}\right)^{\frac{1}{\theta}}\frac{\hat{c}_t}{1+g_N}\end{cases} \tag{68}$$

The steady state is characterized by

$$\begin{cases}g_N=g_y=\left(1+g_f\right)^{\frac{\sigma-\alpha}{1-\alpha}}-1\\ r^*=\left(1+\rho\right)\left(1+g_y\right)^{\theta}-1\\ \hat{k}^*=\frac{\alpha^{\frac{1}{1-\alpha}}}{\left((1+\rho)(1+g_y)^{\theta}-(1-d)\right)^{\frac{1}{1-\alpha}}}\\ \hat{c}^*=\frac{\alpha^{\frac{1}{1-\alpha}}\left((1+\rho)(1+g_y)^{\theta}-(1-d)-\alpha(d+g_y)\right)}{\alpha(1+g_y)\left((1+\rho)(1+g_y)^{\theta}-(1-d)\right)^{\frac{1}{1-\alpha}}}\end{cases}. \tag{69}$$

Log-linearizing the system around the steady state, we arrive at the following linear equations:

$$
\begin{cases} \ln \frac{\hat{k}_{t+1}}{\hat{k}^*} = J_1 \ln \frac{\hat{k}_t}{\hat{k}^*} + J_2 \ln \frac{\hat{c}_t}{\hat{c}^*} \\ \ln \frac{\hat{c}_{t+1}}{\hat{c}^*} = M \ln \frac{\hat{k}_t}{\hat{k}^*} + \ln \frac{\hat{c}_t}{\hat{c}^*} \end{cases} \tag{70}
$$

where

$J_1 = \frac{\left((1-d)+\frac{\left((1+\rho)(1+g_y)^{\theta}-(1-d)\right)}{\alpha}+\frac{(1-\alpha)\left((1+\rho)(1+g_y)^{\theta}-(1-d)\right)\left((1+\rho)(1+g_y)^{\theta}-(1-d)-\alpha(d+g_y)\right)}{\theta\alpha(1+\rho)(1+g_y)^{\theta}}\right)}{1+g_y} > 0$

$J_2 = -\frac{\left((1+\rho)(1+g_y)^{\theta}-(1-d)-\alpha(d+g_y)\right)}{\alpha(1+g_y)} < 0$

$M = -\frac{(1-\alpha)\left((1+\rho)(1+g_y)^{\theta}-(1-d)\right)}{\theta(1+\rho)(1+g_y)^{\theta}} < 0$

The characteristic polynomial is

$$
\chi^2 - (1 + J_1)\,\chi + (J_1 - J_2 M)\,. \tag{71}
$$

There is a unique root that is below 1, related to the saddle-path-stable equilibrium, and the speed of convergence thereof is given by

$$
\gamma_r = 1 - \frac{(1 + J_1) - \sqrt{(1 - J_1)^2 + 4J_2 M}}{2}. \tag{72}
$$