Development Discussion Papers

Empirical Evidence for the Endogenous Growth Generated by Evolution in Division of Labor

Been-Lon Chen, Chien-fu Jeff Lin, and Xiaokai Yang

Development Discussion Paper No. 671
January 1999

Harvard Institute for
International Development

HARVARD UNIVERSITY

VE RI TAS

Empirical Evidence for the Endogenous Growth Generated by Evolution in Division of Labor

Been-Lon Chen, Chien-fu Jeff Lin, and Xiaokai Yang*

Abstract

Following up the empirical works of Jones (1995a,b) that reject two classes of the major endogenous growth models: the AK models and R&D based models, the paper tests the third class of endogenous growth models that generate endogenous evolution in division of labor against empirical data. It is shown that this class of endogenous growth models not only avoid scale effects, but also accommodate both convergence and divergence phenomena. One of the hypotheses generated by this class of endogenous growth models is that the divergence phenomenon takes place first between each pair of economies that enter the take off stage at different points in time, then the convergence phenomenon follows. This implies that the difference in per capita real income between the two economies is an inverse U shape curve. The empirical data strongly support the hypothesis.

Xiaokai Yang is a Reader in Department of Economics at Monash University, Australia and a Development Associate at the Harvard Institute for International Development. He received his Ph.D. from Princeton University. **Been-Lon Chen** is an Associate Research Fellow at the Institute of Economics, Academia Sinica, Taiwan. **Chien-fu Jeff Lin** is affiliated with the Department of Economics at the National Taiwan University.

---

* We are grateful for comments and criticisms from the participants of the seminar at Yale University, our sessions in 1997 Far East Econometric Society Meeting, 1997 Australian Econometric Society Meeting, and International Symposium of Dynamic Models. Special thanks go to Don Snodgrass for comments and suggestions. We are solely responsible for the remaining errors.

Empirical Evidence for the Endogenous Growth Generated by Evolution in Division of Labor

Been-Lon Chen, Chien-fu Jeff Lin, and Xiaokai Yang

I. Introduction

The paper is motivated by two recent developments in the literature of endogenous growth: the puzzle of scale effects and the debate over convergence. Two classes of major endogenous growth models, the AK model and R&D based model, have recently been tested against empirical data (see Jones, 1995a, b). The AK model predicts a positive relationship between growth rates in per capita GDP and investment rates[1]. The relationship is referred to as type-I scale effect which is rejected by the data. This suggests that "the AK models do not provide a good description of the driving forces behind growth" (Jones, 1995a, pp. 508-509). The R&D based model generates a positive relationship between the growth rates in per capita GDP and the level of resources devoted to R&D, referred to as type-II scale effect[2]. Type-II scale effect is also rejected by the empirical observations (Jones, 1995a, b). Jones suggests a way to salvage the R&D based model by adding some feature of the AK model to it, but the modified model still has a type-III scale effect, a positive relationship between the growth rates in per capita GDP and the growth rates of population, which is also wildly at odds with empirical evidence as shown in National Research Council (1986). As Jones (1995b) indicates, endogenous growth cannot be preserved if the scale effect in the R&D based model is eliminated.

Let us now turn to the debate over convergence. The new theory of endogenous growth was motivated in part by the criticism (Lucas, 1988 and Romer, 1986) of the convergence theory based on the neoclassical growth model (Solow, 1956). The absolute convergence hypothesis - per capita incomes of countries converge to one another in the long-run independently of their initial conditions - has been rejected by empirical data (Barro, 1991). The empirical evidence thus calls for a new endogenous growth model that generates a divergence hypothesis. It also

[1] According to Jones (1995a,b) and Barro and Sala-i-Martin (1995), the Romer model (1987), the Rebelo model (1987, 1991), the Barro model (1991), and the Benhabib and Jovanovic model (1991) can be considered as the AK model since their reduced forms are the same as the AK model.

[2] Judd (1985), Romer (1990), Grossman and Helpman (1990, 1991), and Aghion and Howitt (1992), among others,

calls for more empirical works that try to salvage the convergence conjecture. The new theoretical models are not only disturbed by the puzzle of scale effects, but also challenged by the new evidence for convergence. Sala-i-Martin (1996) defines absolute β-convergence as the case in which poor economies tend to grow faster than rich ones and defines σ-convergence as the case in which the dispersion of real per capita GDP levels of a group of economies tends to decrease over time. His empirical works show that there was neither σ-convergence nor absolute β-convergence in the cross-country distribution of world GDP between 1960 and 1990, although the sample of OECD economies displays σ-convergence and so do the samples of regions within a country, such as the USA, Japan, Germany, the UK, France, Italy, or Spain[3]. Since the new evidence is not conclusive, many refined concepts have been proposed to interpret the new evidence as supporting convergence theory. Conditional convergence (Sala-i-Martin, 1996) is defined as the case in which per capita incomes of countries with the same structural characteristics converges to their own steady state in the long-run independently of their initial conditions. Hence, conditional divergence may take place if their structural characteristics are not the same. Club convergence (Galor, 1996) is defined as the case that per capita incomes of countries that are identical in their structural characteristics converge to one another in the long-run provided that their initial conditions are similar as well. The latter concept relates to multiple steady states to which different groups of countries with different initial conditions converge. Hence, the divergence between clubs takes place as each club converges to a steady state different from the steady state to which the other club converges if different clubs of countries with the same structural characteristics have different initial states.

All the new concepts of convergence generate more confusion than clarity. Their invention cannot be used to reject the facts that convergence and divergence, no matter how we define them, may coexist and that such coexistence has not received the attention it deserves. Quah (1996) and Bernard and Jones (1996) complain that the debate of convergence is misguided. They suggest the use of alternative frameworks to address the debate.

The conclusive empirical results in Jones (1995 a, b) that reject scale effect call for a shift of attention to the third class of endogenous growth models proposed by Yang and Borland (YB

---

are along this line.

[3] Tamura (1991) shows that convergence phenomenon can be generated by an endogenous growth model.

model, 1991)[4]. This model explains growth by spontaneous evolution in division of labor which has no scale effect. Moreover, one of the empirical implications of the YB model is the sequential divergence and convergence phenomena. This implication seems to echo Quah, Bernard and Jones' concerns.

The YB model begins from the premise that the distinguishing feature of learning by doing in an economic system is specialized learning by doing through the division of labor[5]. In their model, a mechanism of spontaneous evolution of the division of labor can generate endogenous growth. The YB model has the following attractive features that can be used to address the puzzle of scale effects. The driving force of economic growth in the YB model is economies of division of labor rather than economies of scale. Economies of division of labor differ from economies of scale as discussed in Allyn Young (1928)[6]. According to him, the notion of economies of scale is misleading. The endogenous evolution of the number of traded goods in the YB model is associated with endogenous evolution in the size of the network of division of labor. As the size of the network increases, many separate local communities merge into an increasingly more integrated market. This can take place in the absence of an increase in population size and of other scale effects. Hence, the speed at which new traded goods emerge is determined by the level of division of labor rather than by the population size or the size of the R&D sector. In other words, in the R&D based models, there is a "if and only if" relationship between investment in R&D and the number of new goods and related technology. But in the YB model, there is no such relationship. Endogenous technical progress is associated with the social learning capacity created by a sufficiently large network of division which makes new traded goods commercially viable. In the YB model, the blend of individuals' specialized learning by doing and an increase in the size of the network of division of labor can generate network effects of social learning without scale effect. In contrast, the learning by doing in the other models (Arrow, 1962, Lucas,

---

[4]Other papers along this line of research include Yang and Ng (1993), Borland and Yang (1995), Wen (1997), and Zhang (1997).

[5] Becker and Murphy (1992), Lucas (1988, 1993), Rosen (1977, 1983), and Schultz (1993) all emphasize the implications of interactions between learning by doing and specialization for economic growth.

[6] Young (1928, p. 539) stated "The mechanism of increasing returns is not to be discerned adequately by observing the effects of variations in the size of an individual firm or of a particular industry, for the progressive division of labor and specialization of industries is an essential part of the process by which increasing returns are realized." He argued (p. 533) that the use of the notion of large-scale-production misses the phenomenon of economies of division of labor.

1988; Stocky, 1991; Young, 1993, and others) is independent of the evolution of division of labor and generates scale effect.

The YB model not only formalizes Young's insights into the interdependency between the level of division of labor and the extent of the market and his point that demand and supply are two sides of the division of labor[7] but also accommodates convergence and divergence phenomena of growth rates between developed and less developed economies. The YB model predicts that if the transaction cost coefficient falls, the evolution of division of labor will be speeded up although the change in the parameter is not necessary for the spontaneous evolution of division of labor. Therefore, if an economy started the evolution process earlier than other economies because of a favorable transaction condition and entered the takeoff stage when other economies were still in autarky (with decelerated growth), then the growth rates between the two kinds of economies will diverge. As the developed economy has finally exhausted the potential for further evolution of division of labor, and other economies have eventually entered the takeoff stage, then the growth rates of the two types of economies will converge. The YB model thus generates the following empirical implications: the divergence and convergence phenomena take place sequentially between each pair of economies that enter the take off stage at different points in time and the difference in per capita real income between them is thus an inverse U shape curve.

In addition, the YB model generates the following empirical implications. (1) The income share of the transaction sector increases as division of labor evolves and per capita real income increases. (2) Growth performance and speed of evolution of division of labor critically depend on transaction conditions. Hypothesis (1) is verified by North's empirical work (1986) which shows that the employment share of the transaction sector increased with economic development in the USA in the past century. Hypothesis (2) is verified by historical evidences documented in North (1958) and North and Weingast (1989) and by empirical evidence provided in Barro (1997), Easton and Walker (1997), Frye and Shleifer (1997), Sachs and Warner (1997), and Yang, Wang, and Wills (1992). North shows that the continuous fall of ocean freight rates contributed significantly to early economic development in Europe. He and Weingast show that it was the institution (constitutional monarch and parliamentary democracy) established after England's

[7] Allyn Young (1928) argued that not only the division of labor depends upon the extent of the market, but the extent of the market also depends upon the division of labor (p.534).

Glorious Revolution (1688) that ensured a credible commitment of the British government to the constitutional order and significantly reduced endogenous transaction costs caused by rent seeking, corruption, and state opportunism. Hence, economic development could take off in the UK in the 18^{th} and 19^{th} century.

Barro uses a data set of one hundred countries in 1965-1990 to establish a positive effect of an index of rule of law, which affects transaction conditions, on growth rates of per capita real GDP. Easton and Walker use cross country data to show that growth performance is positively affected by an index of economic freedom. Frye and Shleifer use survey data from Poland and Russia to establish a negative correlation between growth performance and governments' violation of private property rights. Sachs and Warner use a data set of 83 countries in 1965-1990 to establish a positive correlation between growth performance and indices of openness and the quality of institutions that affect transaction conditions. Yang, Wang, and Wills use a Chinese data set in 1979-1988 to establish positive correlations among per capita real income, the degree of commercialization (commercialized consumption divided by total consumption including self-provided consumption, one aspect of division of labor), and an index of institutions and policies that affect transaction conditions in specifying and enforcing private property rights.

But the empirical evidence, except that in North (1986) and Yang, Wang, and Wills (1992), can support both the YB model and neoclassical growth model. In this paper, the hypothesis of sequential divergence and convergence will be tested in connection to the puzzle of scale effects and the debate of convergence versus divergence. We shall show that the new evidence rejects the neoclassical hypothesis of monotonic convergence or divergence. Two types of previous works relate closely to this paper. Kelly (1997) has developed a dynamic equilibrium model with endogenous evolution of the degree of integration of the market network. In his model the trade offs between economies of scale and economies of variety of goods, between economies of scale and fixed transaction costs in establishing trade connections, and between current and future consumption are specified to generate spontaneous evolution of the number of goods and degree of market integration.

The growth rates may increase and take-off may occur when the degree of market integration evolves, while stagnation occurs when the potential for further expansion of market network has been exhausted. Hence, the Kelly model can predict two sequential growth patterns: accelerated growth and stagnation. But the Kelly model generates the type-III scale effect: growth performance improves with the size of the economy since economies of scale instead of economies of specialization are the driving forces in the model. The scale effect distinguishes the Kelly model from the YB model empirically. Kelly provides historical evidence from Song Dynasty China for the relationship between two patterns of growth and transaction conditions. The evidence can support both the Kelly model and the YB model, while the empirical evidence against the type-III scale effect in National Research Council (1986) supports the YB model and is incompatible with the Kelly model. Also, the empirical evidence for evolution of the degree of commercialization in Yang, Wang, and Wills (1992) supports the YB model, in which the degree of commercialization evolves from 0 to 1 as division of labor evolves, but does not support the Kelly model, where the degree of commercialization is always 1.

Empirical evidence for catch up provided by, for instance, Dowrick and Nguyen (1989) and Barro (1997), relate to our paper too. However, the extended neoclassical growth models that predict catch up of the poor countries with the rich ones cannot generate endogenous divergence phenomenon. The initial differences between poor and rich countries are determined by some parameters and do not endogenously increase over time in the models. Hence, there is an empirical distinction between the YB model and these catch up models. In other words, evidence for monotonic convergence rejects the YB model and the Kelly model, while an inverted U curve of difference in per capita real income supports the two models and rejects the neoclassical catch up models. Also, the empirical evidences, mentioned above, for the relationship between evolution of the degree of commercialization, which relates to individuals' levels of specialization, growth, increasing income share of the transaction sector, and transaction conditions distinguish the YB model from those neoclassical catch up models with no evolution of division of labor.

We now consider econometric approaches that relate to our paper. Numerous researchers have recently examined the tendency of convergence and divergence via cross-section regressions of measured growth rates on the initial level of real income per capita[8]. Quah (1993) called them Barro regressions and argued that using a negative coefficient on the initial level of per capita real income to infer convergence has a spurious effect which turns out to be plagued by

[8] See, for example, Barro (1991), Barro and Sala-i-Martin (1991, 1992), Baumol (1986), Delong (1988), Dowrick and Nguyen(1989), Mankiw, Romer and Weil (1992).

Galton's fallacy of regression towards the mean. Quah (1993) proposes a direct test on the cross-section distribution of output per worker over time using the four or five-state fractile Markov chain models and provides the evidence against the convergence. This method, however, is not popular. The existing empirical studies still calculate the average growth rate within a period across economies and then run the regression of these average growth rates on per capita income in the initial year. The focus statistics are the R^2, estimated coefficients and their t-values.

This existing approach is appealing to criticisms. First, it ignores the dynamic path of the growth as all the information in an economy is compressed into a point. Second, the regression can only be acclaimed as in the framework of the absolute convergence in which the varied economic conditions in different countries are not taken into account. Third, the heteroskedasiticity problem might emerge in the regression, creating the inefficiency of the estimator and lowering the t-value. Fourth, the examination of s -convergence is always conducted using the dispersion of income across different economies over time. However, this analysis ignores the new shocks which might increase the dispersion between countries. To overcome the second problem mentioned above, a modification is made in regressions to add regional dummies and structural variables. This can also alleviate the problem of heteroscedasticity. However, the first and last problems mentioned above are never touched upon. To study the pattern of economic growth, a data set for a long period of time, which includes information of time path of economic growth, is needed. A method will be devised in this paper to test the hypothesis of sequential divergence and convergence against a long-run empirical data over more than 120 years.

The rest of the paper is organized as follows. In section II, the empirical implications of the YB model are outlined. Section III rigorously specifies the hypotheses to test. Then the hypotheses are tested against empirical data in section IV. We have found that empirical data strongly support the hypothesis of sequential divergence and convergence.

II. The Empirical Implications of the Yang-Borland Model

The algebra of the YB model can be found from their paper or from Yang and Ng (1993, chapter 7). In this section, we shall present the results that are not worked out in that paper and then discuss the empirical implication.

Consider an economy with M ex ante identical consumer-producers and m consumption goods, where M>m. An individual selling good i and buying n-1 goods from the market has the following decision problem.

(1) $$\text{Max:}U_i = \int_0^\infty u_{it} e^{-\rho t} \mathrm{d}t ,$$

s.t. $$u_{it} = x_{it} \left[\Pi_{r\in R_t} \left(K_t x_{rt}{}^d\right)\right]\left(\Pi_{j\in J_t} x_{jt}\right),$$ (utility at t)

$$x_{it} + x_{it}^s = L_{it}{}^a , x_{jt} = L_{jt}^a \;, j \in J_t , a \geq 1 ,$$ (production function)

$$L_{st} = \int_0^t l_{st}\, \mathrm{d}t \;\text{ and }\; l_{st} = \mathrm{d}L_{st} / \mathrm{d}t \;\text{ if }\; l_{st} > 0 ,$$ (state equation)

$$L_{st} = 0 \;\text{ if }\; l_{st} = 0 ,$$ (current labor input is necessity)

$$l_{it} + \Sigma_{j\in J_t} l_{jt} = 1 ,$$ (endowment constraint of time)

$$K_t = k / n_t , \quad k \in (0,1),$$ (transaction condition)

$$p_{it} x_{it}^s = \Sigma_{r\in R_t} p_{rt} x_{rt}^d ,$$ (budget constraint)

$$n_t \big|_{t=0} = 1,\; L_{st} \big|_{t=0} = 0 ,\; s = i, j , \quad j \in J_t ,$$ (initial conditions)

where ρ is the subjective discount rate, x_{it} is the quantity of traded good i that is self-provided at time t, x_{rt}^d is the quantity of good r that is purchased from the market at t and x_{jt} is the quantity of non-traded good that is self-provided. A fraction 1-K_t of a unit of goods purchased disappears in transit because of transaction costs. A buyer receives K_t from a unit of goods purchased and therefore the quantity of good r that is consumed at t is $K_t\, x_{rt}^d$. Instantaneous utility at t, u_{it}, is a product of quantities of all goods consumed. While k is a parameter of transaction efficiency that relates to institutional, geographical, and technical conditions, K_t is assumed to be a decreasing function of the number of the person's traded goods since her traveling distance increases as the number of her traded goods and thereby the average distance between her and the trade partners increase.

Assume that the number of consumption goods m is much smaller than the population size M. The individual trades n_t goods and therefore self-provides m-n_t goods at t. R_t denotes the set of n_t-1 goods purchased, whereas J_t denotes the set of m-n_t non-traded goods. $x_{it}{}^s$ is the quantity of good i sold to the market at t and therefore x_{it} + $x_{it}{}^s$ is the individual's output level of good i at t. Each individual is endowed with a unit of labor at any point in time and a fraction l_{st} is allocated to produce good *s* at *t*, referred to as the individual's level of specialization in producing good *s*. L_{st} is the amount of labor accumulated in producing good *s* up to *t*. The conditions associated with the definition of L_{st} imply that all previous experience in producing a good will be forgotten if a person ceases to produce this good at *t*[9]. In other words, the labor input in producing good *s* at *t*, l_{st}, is essential for the production of good *s* at *t*. L_{st} - l_{st} can thus be considered as the existing stock of human capital in producing good *s* at time *t*. Parameter p_{st} is the price of good *s* at time *t*. A Walrasian regime prevails since economies of specialized learning by doing are individual specific (increasing returns are localized) and all terms of trade are assumed to be sorted out at $t = 0$ when individuals are yet to specialize. The distinction between individual specific economies of specialized learning by doing and economies of scale ensures that there is no scale effect in this model.

Because of the symmetry, the market equilibrium conditions yield equal prices of all traded goods. The dynamic general equilibrium for $n_t \in (1, m)$ is characterized by the following equations[10].

(2a) $$r_n \equiv \frac{\dot{n}_t}{n_t} = \frac{a r_i}{2 + 1/n_t} = \frac{a n_t^{2(1-1/a)} k^{1/a} e^{(1-2n_t)/a n_t}}{(2n_t + 1)m - a n_t (m - n_t)},$$

(2b) $$r_u \equiv \frac{\dot{u}_t}{u_t} = a[n_t r_i + (m + n_t) r_j] = [(2n_t + 1)m - a n_t (m - n_t)] \frac{r_n}{n_t} = a n_t^{1-\frac{2}{a}} k^{1/a} e^{\frac{1-2n_t}{a n_t}},$$

(2c) $$l_{it} = \frac{n_t (2n_t + 1)}{m(2n_t + 1) - a n_t (m - n_t)}, \quad L_{it} \quad [n^2 e^{2-(1/n_t)} / k^{1/},$$

(2d) $$\frac{\dot{r}_u}{r_u} = \frac{[(a-2)n_t - 1] r_n}{a n_t},$$

where ρ_n is the growth rates of the number of each individual's traded goods, $r_i \equiv l_{it} / L_{it}$ is the growth rates of a person's human capital in producing the good that is sold, $\rho_j \equiv l_{jt} / L_{jt}$ is the growth rates of per capita human capital in producing a non-traded good, and ρ_u is the growth rates

[9] This assumption is too strong to be realistic, but it is necessary for keeping tractable the model that needs the control theory to solve. In Borland and Yang (1995), the assumption is relaxed, so that the control theory does not work and dynamic programming is needed.

[10] There is a typo of the counterpart of equation (2a) in Yang and Borland (1991) and Yang and Ng (1993, Chapter 7). We thank Jack Zhang for pointing out the typo.

of per capita real income (utility). $\dot{r}_u/r_u$ is the growth rates of the growth rates of per capita real income.

The distinguishing feature of the model relates to n_t. For $n_t = 1$, the economy is in autarky where each individual self-provides m consumption goods and there is no market since the number of goods purchased is 0 for each individual. For $n_t = m$, the economy is associated with the complete division of labor that implies that each individual is completely specialized in producing a good and sells the good to and buys a good from each of m - 1 trade partners. If $\rho_n > 0$, the economy evolves from $n_0 = 1$ to $n_T = m$ for some t=T. As the number of each individual's traded goods, n_t, increases from 1 to m, the number of her non-traded goods, m-n_t, is reduced so that her level of specialization in producing some non-traded goods l_{jt} discontinuously jumps from a positive value to 0. This implies that the dynamic optimum solution of l_{jt} jumps from an interior value to a corner solution. As n_t approaches to m, set J_t becomes empty and l_{it} discontinuously jumps from an interior solution to a corner solution $l_{it} = 1$. Such discontinuous jumps of decision rules are referred to as bang-bang control. This feature is essential not only for endogenizing an individual's number of traded goods, but also for endogenizing the size of the network of division of labor.

In the YB model, when an individual chooses her number of traded goods n_t, she chooses the number of her trade partners, which is n_t - 1 in a symmetric model. The aggregate outcome of all individuals' decisions in choosing their n_t determines the size of the network of division of labor and the degree of connectedness of the network. Hence, in the YB model, the degree to which learning by producing new traded goods can promote growth is determined by the level of division of labor and related size of the market network.[11].

(2a) is a non-linear differential equation of n that determines dynamics of the general equilibrium since demand, supply, and other endogenous variables are determined by the level of division of labor n_t. From (2b), it is clear that there is no steady state of per capita real income, a

[11] Yang and Borland's 1991 model has not endogenized the number of all goods though it endogenizes each individual's number of traded goods. Borland and Yang (1995) has introduced the CES production function and intermediate goods into the Yang-Borland model (1991) to endogenize the number of intermediate goods, the number of traded goods, the emergence of the institution of the firm from evolution of division of labor in roundabout production process. The growth rate of per capita real income in the YB model is always positive for $t<\infty$. This differs from the neoclassical growth model with 0 growth rate for a particular $t<\infty$. When growth of population size is introduced and m is an endogenous variable the growth rate in the YB model will be positive even if $t=\infty$.

feature shared with other endogenous growth models. Also, there is no balanced-growth rate of per capita real income. But rather, the growth rate of per capita real income keeps changing over time.

When $n_t \notin (1, \mathrm{m})$, the dynamic general equilibrium is different from (2). If n_t = 1 for all t, then the equilibrium is autarky and is characterized by:

(3) $r_u = amr_j = aml_j / L_j = -am(1/m)/(t/m) = am/t$ and r_n = 0.

For $n_t = m$, all goods are traded and each individual is completely specialized in equilibrium, namely l_{it}=1 from (2c). The potential for further evolution in division of labor has therefore been exhausted. The growth rates of per capita real income are determined solely by the growth rates of human capital in producing traded good (ρ_i) which equals $1/L_i$ for l_{it}=1. Though r_i is always positive, it monotonously declines as human capital L_i increases over time.

The intuition behind the model is quite straightforward. Suppose that there are transaction costs and productivity gains from specialized learning by doing and that consumer-producers prefer current diverse consumption. At t=0 each person does not have much experience in producing each and every goods, so that her productivity is low and she cannot afford the transaction costs caused by specialization and division of labor. Autarky is thus chosen. As time goes by, each person builds up some experience (or so called human capital) in producing each and every goods, so that her productivity goes up slightly and she can afford a slightly higher transaction cost and therefore will choose a slightly higher level of specialization. The specialized learning by doing will speed up the accumulation of professional experience. Consequently, each person's productivity in her professional activity increases further and therefore she can afford an even higher transaction cost and will choose an even higher level of specialization, and so on, until the potential for further evolution of division of labor has been exhausted. In the process, the growth rate of per capita real income declines initially due to a low level of division of labor, then increases (takeoff) as the evolution in division of labor is speeded up by compounded effects of learning over time and increasingly more specialized learning at each point in time, and finally declines again (but is always positive) as the potential for further evolution of division of labor has been exhausted and learning over time becomes the sole driving force of growth.

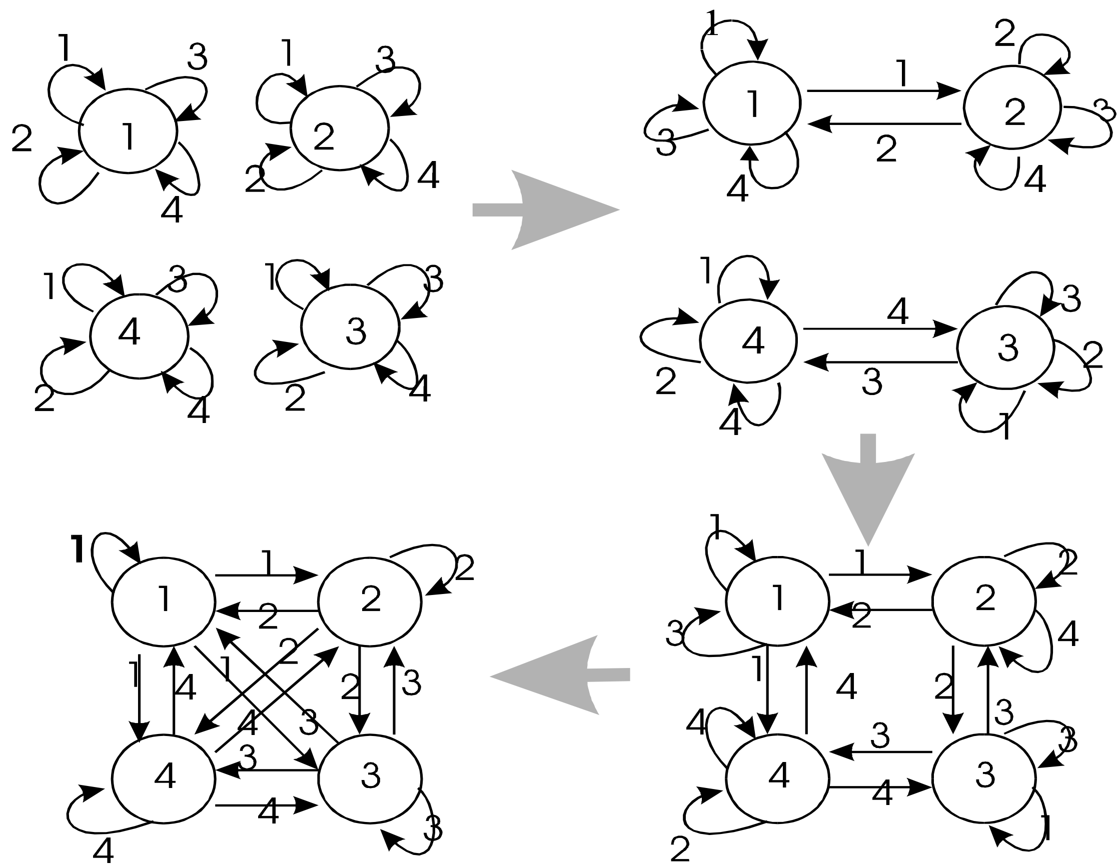
The trade off between economies of specialized learning by doing, which increase future productivity through speeding up accumulation of human capital, and related transaction costs, which reduce current consumption is the driving force of the story. As time goes by, the scope for individuals to trade off economies of specialized learning by doing against transaction costs and to trade off current consumption for future consumption is enlarged, so that the efficient trade off evolves over time, generating evolution of division of labor and other concurrent phenomena. The evolution of division of labor will increase the extent of the market (per capita effective demand), production concentration of each traded good, the diversity of different professions, the extent of endogenous comparative advantage, the degree of market integration, each person's level of specialization, an income share of transaction cost, each person's productivity in her profession, and so on.

The evolution of division of labor features expansion of the market network and merge of separated local communities into an increasingly integrated market. The network can be well defined by a directed graph where each individual is a point (or vertex, node) and each flow of goods is a directed line (or edge, arc) that connects a pair of points. The following graph provides a better illustration of the topological and graphical properties of the evolution in division of labor than above intuition and algebra.[12]

[12] There are multiple dynamic equilibria in the YB model since it is indeterminate who is specialized in producing which good in the model with ex ante identical individuals. Also, both of the dynamic equilibrium shown in Fig. 1b and the dynamic equilibrium shown in Fig. 2b of Yang and Borland (1991) may exist when $n_t = 2$. But if the assumption is made that there is neither central clearing house nor money and related credit systems, then all of the structures, such as the one in Fig.1b of Yang and Borland (1992) with one way trade (without double coincidence of demand and supply) cannot take place in equilibrium. All of the multiple dynamic equilibria generate the same total present value of utility for each individual and have similar dynamics and comparative dynamics.

Figure 1: Topological Properties of Evolution in Division of Labor

(a) Autarky $n_t = 1$ (b) Partial division of labor $n_t = 2$



(d) Complete division of labor $n_t = 4$ (c) Partial division of labor $n_t = 3$

Fig. 1 shows an economy with 4 individuals and 4 goods (M = m = 4). The lines denote goods flows, whereas the small arrows indicate direction of goods flows. The numbers beside the lines signify goods involved. A circle with number i denotes a person selling good i. Panel a denotes autarky where each person self-provides 4 goods. There is no market and trade. The economy is divided into 4 separate local "communities" and there are 4 producers of each good. Panel b denotes partial specialization where each person sells one good, buys one good, trades two goods, and self-provides three goods. There are two separate local communities for the pattern of partial division of labor. Panel c denotes the partial division of labor where each person sells and self-provides one good, buys 2 goods, and trades 3 goods. Note that if the population size is, say, 300 instead of 4, the partial division of labor would be organized in 100 separate local communities. In each of them three individuals trade with each others. The integer problem for the existence of equilibrium with n_t = 3 and with the particular population size M = 4 yields the pattern of organization in panel c. Panel d denotes the complete division of labor for the society where each person is completely specialized. There is a single integrated market. Large arrows indicate evolution in division of labor that features merge of separate local communities into a few of larger separate communities, followed by a single integrated market. The merging process of separate total communities into an increasingly integrated network generates endogenous growth without scale effect.

Let us consider empirical implications of the model in connection with the puzzle of the scale effect and the debate of convergence versus divergence. Certainly, there is no scale effect in the YB model despite the endogenous evolution in the number of traded goods that generates endogenous growth. This is because the endogenous growth in the model comes from network effects of division of labor associated with interpersonal complementarity. The growth rates of level of division of labor and the growth rates of per capita real income in (2) are independent of population size M. The population size plays a very passive role in this model. It sets up a limit for the evolution in division of labor: the number of professions cannot be larger than population size.

To see the empirical implications of the model, let us examine (2d) which indicates that ρ_u increases over time for $n_t \in (1, m)$ when r_n >0 and a > 2. Together with our discussion of the corner cases where n_t = 1 (autarky) and $n_t = m$ (complete division of labor), this implies that three stages will take place sequentially: preindustrialization growth, accelerating growth and take off, and mature growth. Growth rates first decline, then increase, and finally decline again.

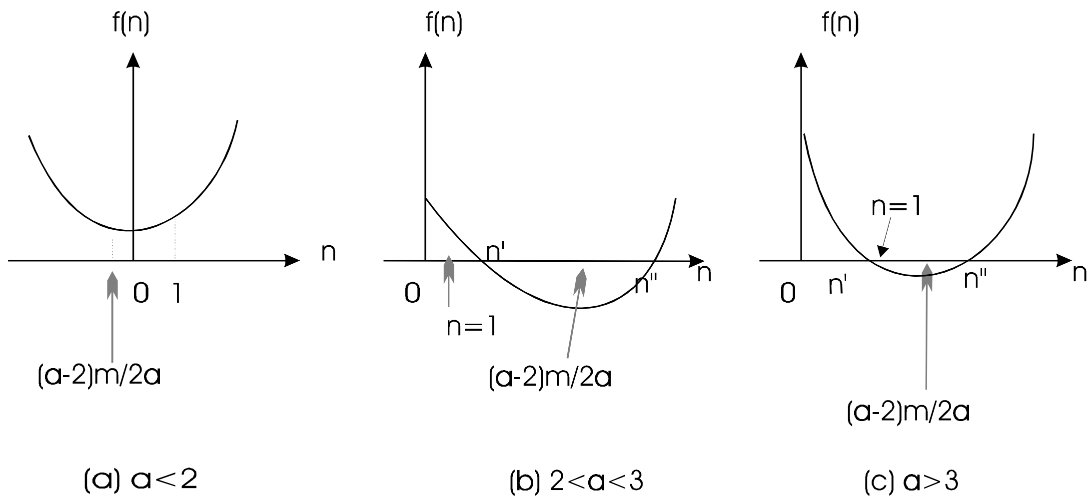
This is consistent with Rostow's (1960) description of three stages of economic growth. (2b) gives comparative dynamics of the general equilibrium. As transaction-condition parameter k increases, the growth rates of per capita real income rise despite the fact that the change in k is not necessary for the accelerated growth and for endogenous evolution in division of labor. The result of this comparative dynamics can be used to justify cross-country growth differentials since different tariff regimes and degree of openness, different institutional arrangements, different legal systems and related property rights regimes, and different geographical conditions across countries all imply different values of k across countries. Hence, those countries with a larger k enter the take off stage earlier.

The UK, for example, entered the take off stage earlier than other countries since it is an island country which implied a higher shipment efficiency in the UK than in inland countries, such as Germany and China, when automobiles and trains were not available. The UK was the first country to have Statute of Monopolies (patent laws) (1624) which significantly improved transaction efficiency in trading intellectual properties. Also in 1846, the UK stopped the Corn Law and three years later, in 1849, abolished Navigation Act and adopted a free trade policy.

Indeed there are more growth patterns and sub-patterns that can be generated by the model. For instance, there are two sub-patterns under the pattern of take off: big push industrialization (Nurske, 1953; Murphy, et al., 1989) and smooth take off (Matsuyama, 1991; Chen and Shimomura, 1997). In order to work out the intervals of parameter values that divide between the patterns and between the sub-patterns, we shall have a close look at (2). (2a)-(2c) indicate that the growth rates of the level of division of labor, ρ_n, the growth rates of per capita real income, ρ_u, and each individual's level of specialization l_{it} are positive if and only if

(4) $$f(n_t) = (2n_t + 1)m - an_t(m - n_t) > 0 .$$

Figure 2: Intervals of *a* that Divide Between Various Growth Patterns



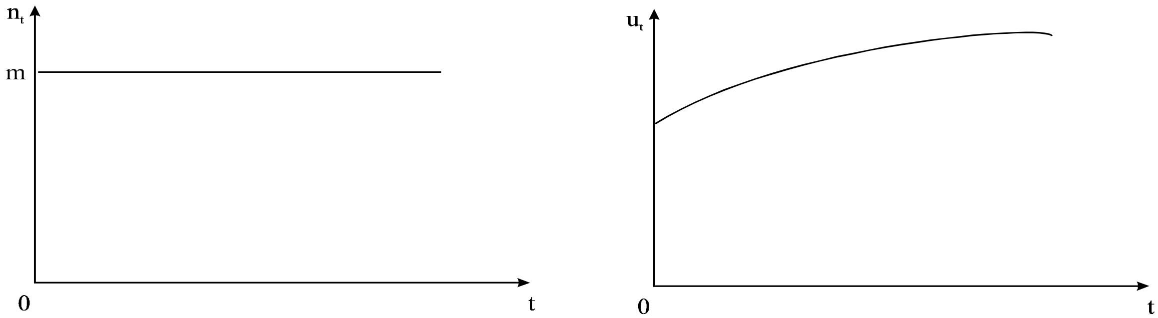
Note that for an interior solution, $f(n_t)$ must be positive. The maximum value of l_{it} is 1 due to the endowment constraint and its minimum value is 0 due to nonnegative constraint. (2c) indicates l_{it} takes on its maximum value, 1, when $f(n_t)$ = 0 and takes on its minimum value, 0, when $f(n_t)$<0. $f(n_t)$ can be used to identify three patterns of path of n_t. Having differentiated (4) with respect to n twice, we can show that $f(n_t)$ is a non-monotonous convex function with the minimum point $n \equiv \frac{(a-2)m}{2a}$, as shown in Fig. 2. It is not difficult to see that for $a \leq 2$, the minimum value of $f(n_t)$ is positive, so that $f(n_t)$ is always positive. If a>2, the curve associated with $f(n_t)$ cuts the horizontal axis at two points (a single point for $m = [2/(a-2)]^2 a$). From $f(n_t)$ =0, the two cutting points can be solved as follows.

(5a) $\quad n' = \{(a-2)m - [(a-2)^2 m^2 - 4am]^{½}\} / 2a$,

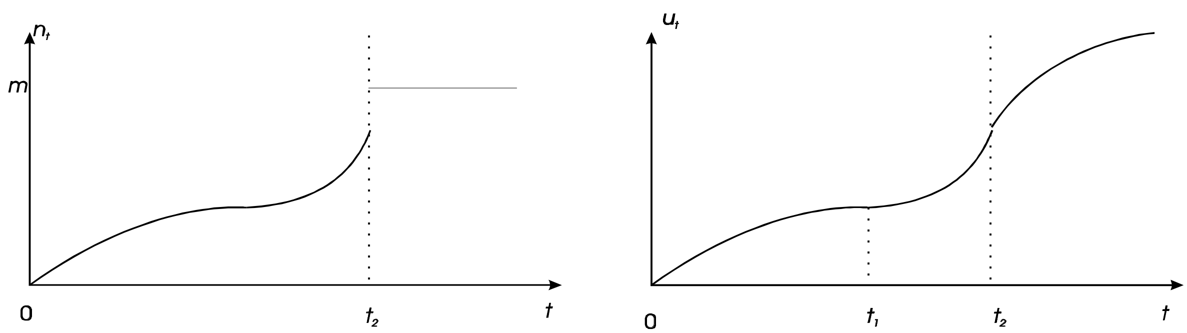
(5b) $\quad n'' = \{(a-2)m + [(a-2)^2 m^2 - 4am]^{½}\} / 2a$,

where $n'' \geq n'$. It is not difficult to show that $n' \geq 1$ if $a \in (2,3)$ and $n' < 1$ if a>3. Since the minimum value of n is 1, which means autarky, $n' < 1$ is impossible . Hence, for a>3, n' is irrelevant. This discussion generates three intervals of a that are associated with three growth patterns, as shown in Fig. 3.

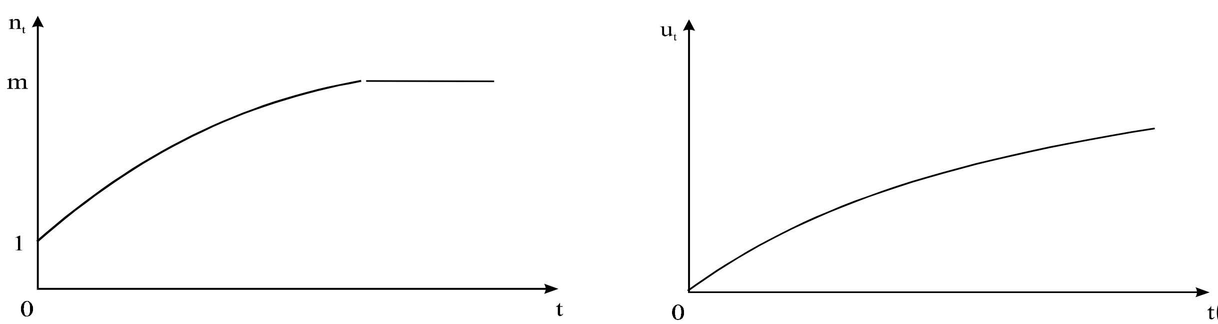
Figure 3: Various Patterns of Evolution of Division of Labor and Growth of Per capita Real Income



(a) *a*>3 and *k* is sufficiently large

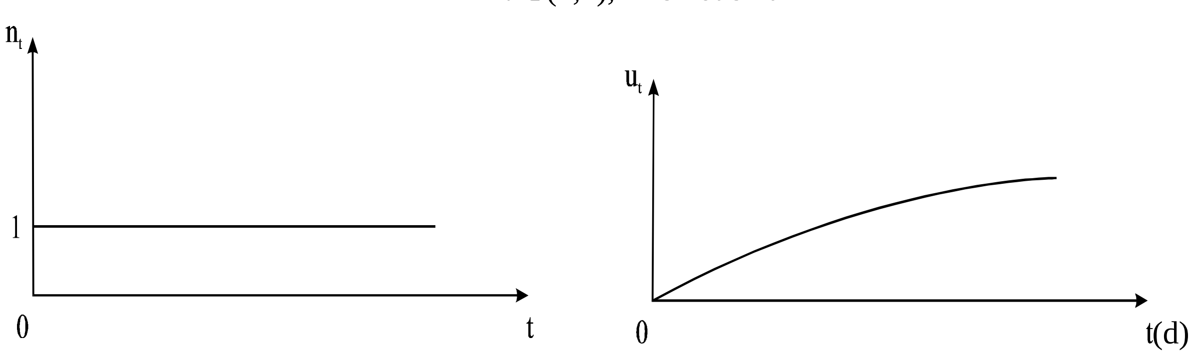


(b) $a \in (2,3)$, k is sufficiently large



(c)

$a \in (1,2)$, *k* is not small



(d) *a*=1 and/or $k \to 0$

First we consider interval a>3 in Fig. 2(c). Since the initial condition is n_t=1 for t = 0, the system starts from the point n_t=1 which is on the right hand side of the cutting point n'. At point n', $f(n_t)$ = 0 implies that l_{it} tends to be as large as possible according to (2c), but the maximum value of l_{it} is 1 because of the endowment constraint of time. The application of the Kuhn-Tucker condition to the individual's decision of level of specialization yields l_{it} = 1. This implies that at t = 0, each individual jumps from autarky (n_t=1) to complete specialization (l_{it} =1 and n_t= m). Afterward, the level of division of labor stays there and the growth rates of per capita real income decline although they are always positive. The paths of the level of division of labor and per capita real income for this growth pattern are shown in Fig. 3(a). This pattern features a single big push industrialization in the absence of a smooth take off and smooth evolution in division of labor. We cannot observe the phenomenon of human society jumping to a completely commercialized state from autarky as soon as the society comes to existence. The lack of empirical evidence for this pattern of growth can be attributed to the fact that the degree of economies of specialization, a, in the real world is not larger than 3. Of course, the threshold value of a would be much more complicated if the model were not symmetric.

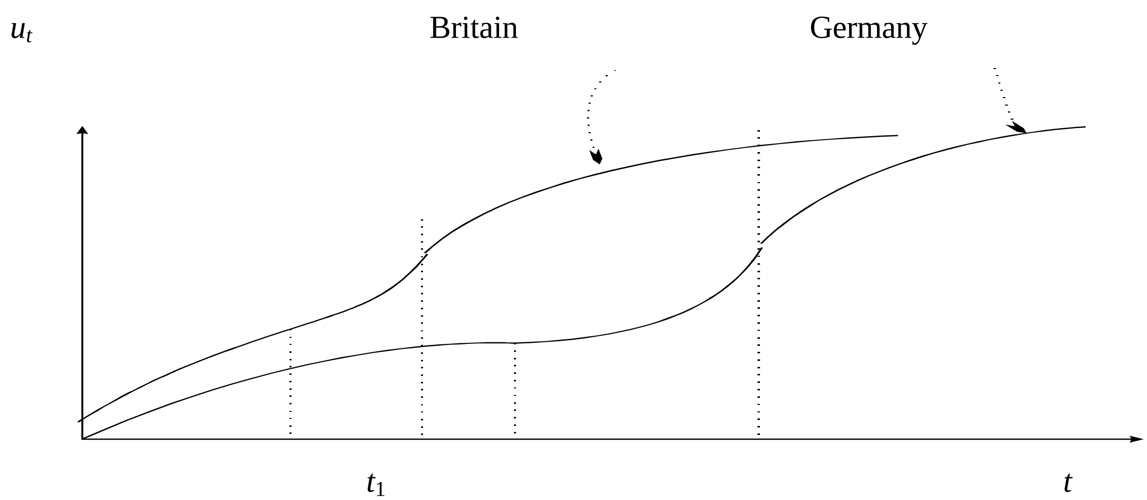
Let us consider the case of $a \in (2,3)$ in Fig. 2(b). From (5), it can be shown that $1 < n' < n'' < m$ for this case. The dynamic equilibrium value of n_t increases over time before $t = t_2$ or before $n_t = n'$ since $n_t < n'$ for $f(n_t) > 0$ which implies $r_n > 0$ according to equation (2a). At $t = t_2$ or at $n_t = n'$, $f(n_t)$ = 0, which implies l_{it} =1 according to equation (2c). Hence, the level of division of labor jumps from $n_t = n'$ to n_t= m (complete division of labor) at this point in time. From equation (2d), it is obvious that there is a point t_1 before t_2. The growth rates of per capita real income decrease (although they are positive) before t_1 but increase between t_1 and t_2, as shown in Fig. 3(b). The sub-growth pattern before t_1 is referred to as pre-industrialization growth which features a low level of division of labor and decreasing growth rates. The sub-growth pattern between t_1 and t_2 is referred to as smooth take off which features smoothly increasing growth rates of per capita real income and increasing n_t. The sub-growth pattern at t_2 is referred to as big push industrialization, which features a discontinuous jump of the level of division of labor (or the level of commercialization). At this point, many new professions and new traded goods simultaneously emerge and many separate local business communities suddenly merge into the integrated market.

After the big jump, the economy has reached the complete division of labor and the potential for further evolution in division of labor has been exhausted. The economy enters the mature growth stage which features decreasing growth rates of per capita real income and with zero growth in the level of division of labor.

For $a < 2$, equations (2b) and (2d) indicate that the growth rates of per capita real income are always positive and declining if $n_t \in (1, m)$. Also, equations (2a) and (5) imply that the growth rates of the level of division of labor are always positive for $n_t \in (1, m)$ and $a < 2$. The pattern of evolution in division of labor and the growth pattern of per capita real income for this case are illustrated by Fig. 2(c) which features decreasing growth rates of both per capita real income and level of division of labor. Finally, equations (2a) and (2b) imply that the growth rates of level of division of labor are 0 if parameter k is close to 0. The application of the Kuhn-Tucker condition to each individual's decision of the level of specialization (see Yang and Borland, 1991) indicates that the optimum value of n_t is always at a corner $n_t = 1$ if k is sufficiently close to 0. The pattern of growth is illustrated in Fig. 3(d) which features autarky and decelerated growth forever. This pattern of growth is similar to so called growth trap. However, in the YB model such a growth trap is not an accidental outcome caused by multiple steady states for given parameters. But rather, it is based on comparative dynamics (changes in the dynamic equilibrium in response to changes in parameters) and generated by a small k resulting from deficient institutional arrangements, a low degree of openness, high tariffs, underdeveloped transportation infrastructures, deficient banking and legal systems, and other unfavorable transaction conditions.

It is obvious that for an economy that has experienced mature growth the time path of structure of division of labor and per capita real income are like those in Fig 3(b). As t increases from its initial value to infinity, the curve representing per capita real income is first concave, then convex, finally concave again, if short-run fluctuations are eliminated. However, economies may differ in their parameter k that relates to transaction conditions. We should therefore see that different countries enter the take off stage at different points in time. Between each pair of economies that enter the take off stage at different points in time, the difference in per capita real incomes should be an inverse U curve in the coordinates of time and per capita real income. That is, when the country that enters the take off stage more early starts accelerated growth, the latecomer is still in the decelerated growth stage. Their per capita real incomes and growth rates will diverge. As the latecomer ultimately enters the take off stage and the leader reaches the mature growth stage, the difference in their per capita real incomes and growth rates converge. The sequential divergence and convergence is illustrated in Fig. 4 in which per capita real incomes between two economies diverge before t_1 and converge afterwards. The convergence may be conditional and different economies may not end up with the same income level and growth rates due to the difference in k between them and to possible changes in k over time[13].

Figure 4: Sequential Divergence and Convergence Phenomena



III. The Hypothesis to Test

The maintained hypothesis to test is that for a pair of economies that have experienced mature growth, the time path of per capita income differential between them should be an inverted U curve. We may use a second-order polynomial time trend of per capita real income differential between each pair of countries to fit the data. However, this setting lacks flexibility if the inverted U curve is not symmetric around the extreme value point of the curve or the number of points on one side of the extreme value point is much larger than that of other side. In those cases, the pattern of

[13] For instance, the invention of trains, automobiles, and airplanes raised the relative size of k of the inland economies vs. the island or coastal economies.

quadratic polynomial fitted will reduce to a straight line. To increase the flexibility, a cubic polynomial time trend is chosen to fit the data. The equation is therefore

(6) $y_t = a_0 + a_1 t + a_2 t^2 + a_3 t^3$.

where y_t is the long-run time trend of per capita real income differential between a pair of countries, which excludes short-run stochastic fluctuation, t is time, and α_i are parameters to be estimated. This functional form allows for various shapes of a curve. Before identifying parameter subspaces that demarcate the four hypotheses, we first figure out the features of (6).

A differentiation of (6) yields

(7a) $dy_t(0)/dt > 0$ iff $\alpha_1 > 0$

(7b) $dy_t(\bar{t})/dt = 0$ yields two interior extreme points:

$$\bar{t}_1, \bar{t}_2 = \left(-2a_2 \pm \sqrt{4a_2^2 - 12a_1a_3}\right) \Big/ \left(6a_3\right)$$

(7c) $d^2y_t(0)/dt^2 > 0$ iff $\alpha_2 > 0$

(7d) $d^2y_t(t^*)/dt^2 = 0$ yields the inflection point $t^* = -\alpha_2/3\alpha_3$.

(7b) and (7d) imply that the curve is non-monotone and has a maximum point and a minimum point and the inflection point t^* is in between the two extreme points, if $\alpha_2^2 > 3\alpha_1\alpha_3$. Without loss of generality, we assume $\bar{t}_1$ is the minimum point and $\bar{t}_2$ is the maximum point. The curve is concave on one side of t^* and is convex on the other side of t^*. (7) can be used to identify the condition under which a particular side is concave.

For $\alpha_2^2 \leq 3\alpha_1\alpha_3$, the curve is monotonically increasing (consistent with the hypothesis of divergence based on the Romer model) or decreasing (consistent with the hypothesis of convergence based on the Solow model) because it has no interior extreme point. With the knowledge about the features of the curve given by (6), we can identify the parameter subspace within which one of the four hypotheses holds. We consider only $t \in [1, 123]$, where $t = 1$ (year 1870) is the initial year and $t = 123$ (year 1992) is the final year for available data.

(0) Hypothesis H_0: An inverse U curve within $t \in [1, 123]$ or sequential divergence and convergence. This is our maintained hypothesis.[14] There are six parameter subspaces that might be consistent with H_0. Each of them requires $\alpha_2^2 > 3\alpha_1\alpha_3$. The first four of them are conclusively consistent with H_0 and the other two are partially consistent with H_0.

(0a) ($\bar{t}_1 < 1,\ \bar{t}_2 \in (1,\ 123)$, $\alpha_1<0$, $\alpha_2>0$, $\alpha_3<0$, $\alpha_2^2>3\alpha_1\alpha_3$), as shown in Fig. 7(a).

According to (7), this parameter subspace implies $t^* > 0$, $dy_t(0)/dt < 0$, and $d^2y_t(0)/dt^2 > 0$. This implies that the curve is convex on the left hand side of t^* and is concave on the right hand side, as shown in Fig. 7(a). Fib. 7(a) also shows that the curve is of inverse U-shape on the right hand side of $\bar{t}_1$. Having noted the fact that the initial time for the data set is $t = 1$, this implies for $\bar{t}_1 <$ 1, all data fall in the parameter subspace that ensures an inverse U curve. This is true even if we extend data from the final time $t = 123$ of the available data. Hence, hypothesis H_0 is conclusively accepted in both senses of fitting and forecasting if estimated parameter values fall in the subspace.

(0b) ($\bar{t}_1 < 1,\ \bar{t}_2 \in (1,\ 123)$, $\alpha_1>0$, $\alpha_2<0$, $\alpha_3<0$, $\alpha_2^2>3\alpha_1\alpha_3$), as shown in Fig. 7(b).

If $a_2<0$ and $a_3<0$, (7b) implies $\bar{t}_1<1$. Here, $\bar{t}_2>0$ is the maximum point, $\bar{t}_1<0$ is the minimum point, and the inflection point t^* is negative and is on the right hand side of $\bar{t}_1$. Since the curve of the polynomial function is concave (inverse-U shape) for $t > t^*$, and the initial year for the data set is $t = 1 > t^*$ (because of $t^* < 0$). Therefore, the hypothesis H_0 is accepted in both senses of fitting and forecasting within the parameter subspace (0b).

(0c) ($\bar{t}_1 < 1,\ \bar{t}_2 \in (1,\ 123)$, $\alpha_1>0$, $\alpha_2>0$, $\alpha_3<0$, $\alpha_2^2>3\alpha_1\alpha_3$).

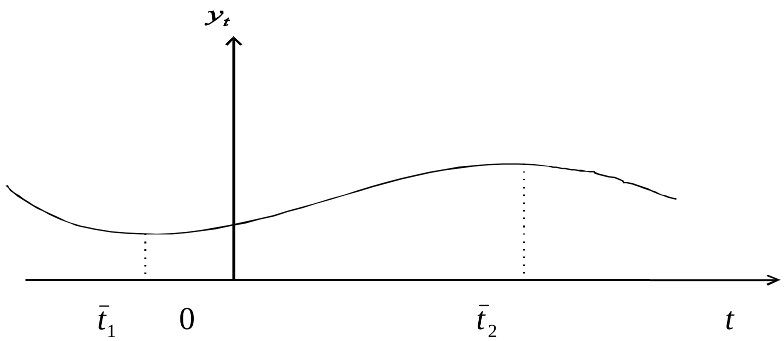
Following the discussion for parameter subspaces (0a) and (0b), we can show that hypothesis H_0 is conclusively accepted if estimated parameters fall in this subspace.

(0d) ($\bar{t}_1 > 1,\ \bar{t}_2 > 123$, $\alpha_1>0$, $\alpha_2<0$, $\alpha_3>0$, $\alpha_2^2>3\alpha_1\alpha_3$).

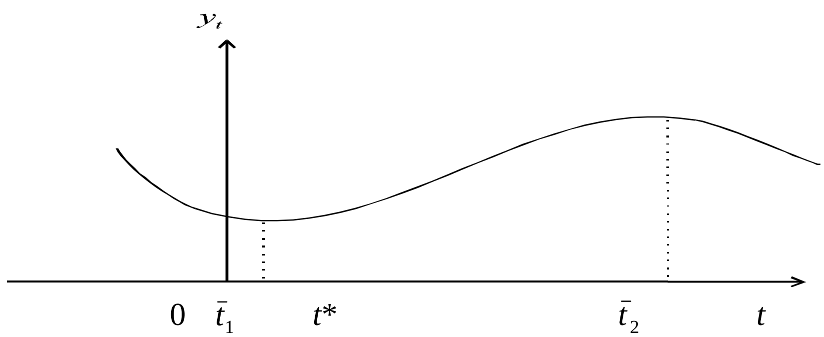
For this parameter subspace, (7) implies $t^* > 0$, $dy_t(0)/dt > 0$, and $d^2y_t(0)/dt^2 < 0$, so that the curve is concave on the left hand side of t^* and is convex on the right hand side, as shown in Fig. 7(c). But for this subspace, we have $\bar{t}_2 > 123$. This, together with the fact that for all available data $t \leq$ 123, implies that $t < \bar{t}_2$, that is, all data fall in the interval of t within which the curve is of inverse U-shape. Though H_0 is accepted if estimated parameters fall in this parameter subspace, the acceptance is in the sense of fitting rather than in the sense of forecasting since as t increases such that $t > \bar{t}_2$, the curve becomes U shape.

[14] Hypothesis H_0 allows per capita real income in an economy to take over that in the benchmark economy.

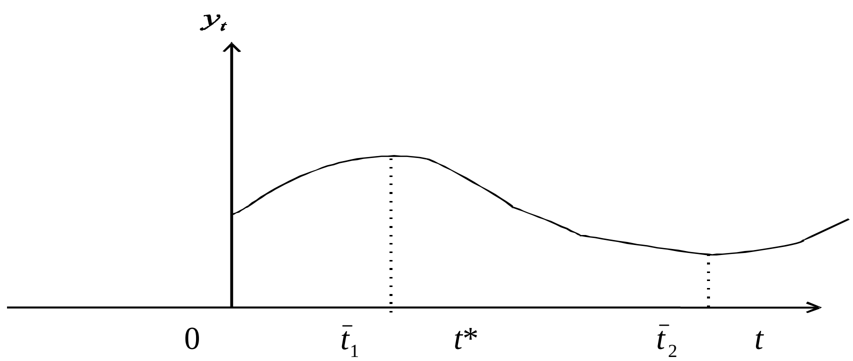
Fig 5: Possible Shape of A Cubic Polynomial Time-trend Function



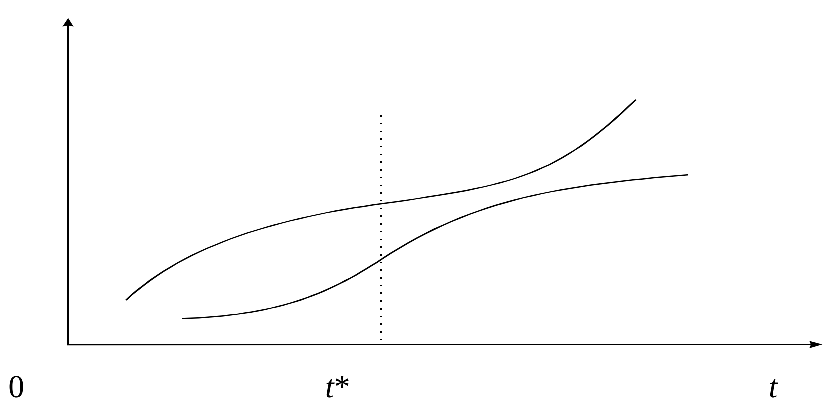
(a) (i) α_2 <0 and α_3 <0, (ii) α_2 >0, α_3 <0 and $\bar{t}_1$ <0



(b) α_2 >0, α_3 <0 and $\bar{t}_1$ >0



(c) α_1 >0, α_2 <0 and α_3 >0



(d) Parameter subspace (1a): α_1 >0, $\alpha_2^2 < 3\alpha_1\alpha_3$

If estimated parameter values fall in any of the four parameter subspaces, then data conclusively support the hypothesis of an inverse U curve. Two other parameter subspaces are partially consistent with H_0 since both U and inverse U shape are possible within each of them for $t \in [1, 123]$.

(0e) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 < 0$, $\alpha_2 > 0$, $\alpha_3 < 0$, $\bar{t}_1 > 1, \bar{t}_2 > 123$), as shown in Fig 7(b).

This parameter subspace is the same as subspace (0a) except $\bar{t}_1 > 1$. Hence, for the data in the interval $t \in (\bar{t}_1, 123]$, we have an inverse U curve and H_0 is accepted. But for the data in the interval $t \in [1, \bar{t}_1]$, we have a U curve and H_0 is rejected.

(0f) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 > 0$, $\alpha_2 < 0$, $\alpha_3 > 0$, $\bar{t}_1 < 1$, $\bar{t}_2 < 123$), as shown in Fig. 7(c).

This parameter subspace is the same as subspace (0a) except $\bar{t}_2 < 123$. Hence, for the data over $t \in [1, \bar{t}_2]$, we have an inverse U curve and H_0 is accepted. But for the data over $t > \bar{t}_2$, we have a U curve, so that H_0 is rejected. But even if H_0 is accepted within this parameter subspace, it is accepted in the sense of fitting rather than in the sense of forecasting.

For parameter subspaces (0e) and (0f), if (1-*x*)% data are inconsistent with inverse-U shape curve, then the hypothesis H_0 is rejected by (1-*x*)% of data or accepted by *x*% of data.

(1) Hypothesis H_1: A monotonically increasing per capita real income differentials within $t \in [1, 123]$. Following the approach to establishing the connection between hypothesis H_0 and subspaces (0a) and (0b), we can show that there are six parameter subspaces that are consistent with H_1:

(1a) ($\alpha_2^2 < 3\alpha_1\alpha_3$, $\alpha_1 > 0$), as shown in Fig. 5(d);

(1b) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 < 0$, $\alpha_2 > 0$, $\alpha_3 < 0$, $\bar{t}_1 < 1$ and $\bar{t}_2 > 123$), a rejection of (0a);

(1c) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 > 0$, $\alpha_2 < 0$, $\alpha_3 < 0$, $\bar{t}_2 > 123$), a rejection of (0b);

(1d) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 > 0$);

(1e) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 > 0$, $\alpha_2 > 0$, $\alpha_3 < 0$, $\bar{t}_2 > 123$);

(1f) ($\alpha_2^2 = 3\alpha_1\alpha_3$, $\alpha_1 \geq 0$).

(2) Hypothesis H_2: A monotonically decreasing per capita real income differentials within $t \in [1, 123]$. There are five parameter subspaces that are consistent with H_2:

(2a) ($\alpha_2^2 < 3\alpha_1\alpha_3$, $\alpha_1 < 0$);

(2b) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 > 0$, $\alpha_2 < 0$, $\alpha_3 > 0$, $\bar{t}_1 < 1$ and $\bar{t}_2 < 123$), a rejection of (0d);

(2c) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 < 0$, $\alpha_2 < 0$, $\alpha_3 < 0$, $\bar{t}_1 < \bar{t}_2 < 0$), a rejection of (0a);

(2d) ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 < 0$, $\alpha_2 > 0$, $\alpha_3 > 0$);

(2f) ($\alpha_2^2 = 3\alpha_1\alpha_3$, $\alpha_2 \leq 0$).

(3) Hypothesis H_3: A U curve of the difference in per capita real income.

Following the approach to establishing the connection between hypothesis H_0 and the parameter subspace (0a), we can show that hypothesis H_3 will be conclusively accepted if estimated parameters fall in the subspace ($\alpha_2^2 > 3\alpha_1\alpha_3$, $\alpha_1 < 0$, $\alpha_2 > 0$, $\alpha_3 > 0$). Acceptance of either of H_1, H_2, H_3 implies a conclusive rejection of H_0.

In the next section, we test the four hypotheses against empirical observations.

IV. Empirical Evidence

In this section, fifteen OECD countries included in Maddison (1991) are employed for our empirical study[15]. The countries are Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, Switzerland, UK and US. They are used because they all have experienced the stages of preindustrialization decelerated growth, take off, and mature growth with declining growth rates, which render our empirical study possible. The data are annual real GDP per capita at 1985 international prices, ranged from 1870 to 1992, except for Japan (1885-1992), Netherlands (1902-1992) and Switzerland (1899-1992). The data sources before 1950 come from Madison (1991) , whereas those after 1951 come from Penn World Table 5.6[16]. With year 1913 serving as the base year, Maddison provides indexes of annual aggregate real GDP (adjusted to exclude the impact of boundary changes) in terms of 1985 US relative prices that are constructed using OECD purchasing power parity units of national currency per US dollar. Levels of aggregate GDP at 1985 international prices are calculated by multiplying the indexes of annual aggregate real GDP and the level of GDP in 1913 at 1985 international prices. Maddison also provides annual mid-year population levels. Annual real GDP per capita of each country was calculated by dividing its aggregate GDP by its mid-year population level. Although the data in

[15] Australia is excluded for its real GDP per capita was higher than the UK in 1870.
[16] There are some data points available prior to 1870 and far back to 1770 in Madison (1991). These points are useful if we want to study the evolution of an economy. They are, however, not recorded annually and not available for most of

Maddison end in 1989, we use data in Penn World Table 5.6 for the years after 1950 as Penn World Table 5.6 is now more popular and ends in 1992. Nevertheless, our examination reveals that the annual real GDP per capita after 1950 have a similar trend for both data sets. The data variable we obtain from Penn World Table 5.6 is RGDPL.

Figure 5 illustrates per capita real incomes in the fifteen countries from 1870 or later to 1992. Overall, the curves are concave first, then convex and finally concave again. The convex stages are very impressive for most of countries, whereas the initially concave and finally concave stages are not that clear. The latter result may be because the data series in preindustrialization decelerated-growth stage is not long enough due to the lack of data series prior to 1870 and the shortness of the data series for the mature stage as the stage has just arrived. As can be seen, the take off growth stage, roughly from 1940 to 1990, dominates the graph. In the analysis below, we take the UK economy as the benchmark country to examine the sequential divergence and convergence phenomena as the UK is the first country to experience a take off.

Figure 6 portrays the differences in real GDP per capita by subtracting annual real GDP per capita of each of the other fourteen countries in the period under study from that of the UK. From Figure 6, it is obvious to see that most of the curves roughly show an inverse-U shape. The UK was a leading economy long before 1870 and grew faster than the other economies before World War II. Its income gaps with the other countries therefore widened before around 1940 and narrowed afterward except with the US and Canada, whose gaps started to narrow earlier and occurred in the late 19 century. Sequential divergence and convergence between the UK and each of the other countries seem to be evident. In what follows, we test the maintained hypothesis statistically.

---

the countries. They therefore do not serve our purpose.

Figure 6: Per capita Real Incomes in the Fifteen Countries from 1870 to 1992

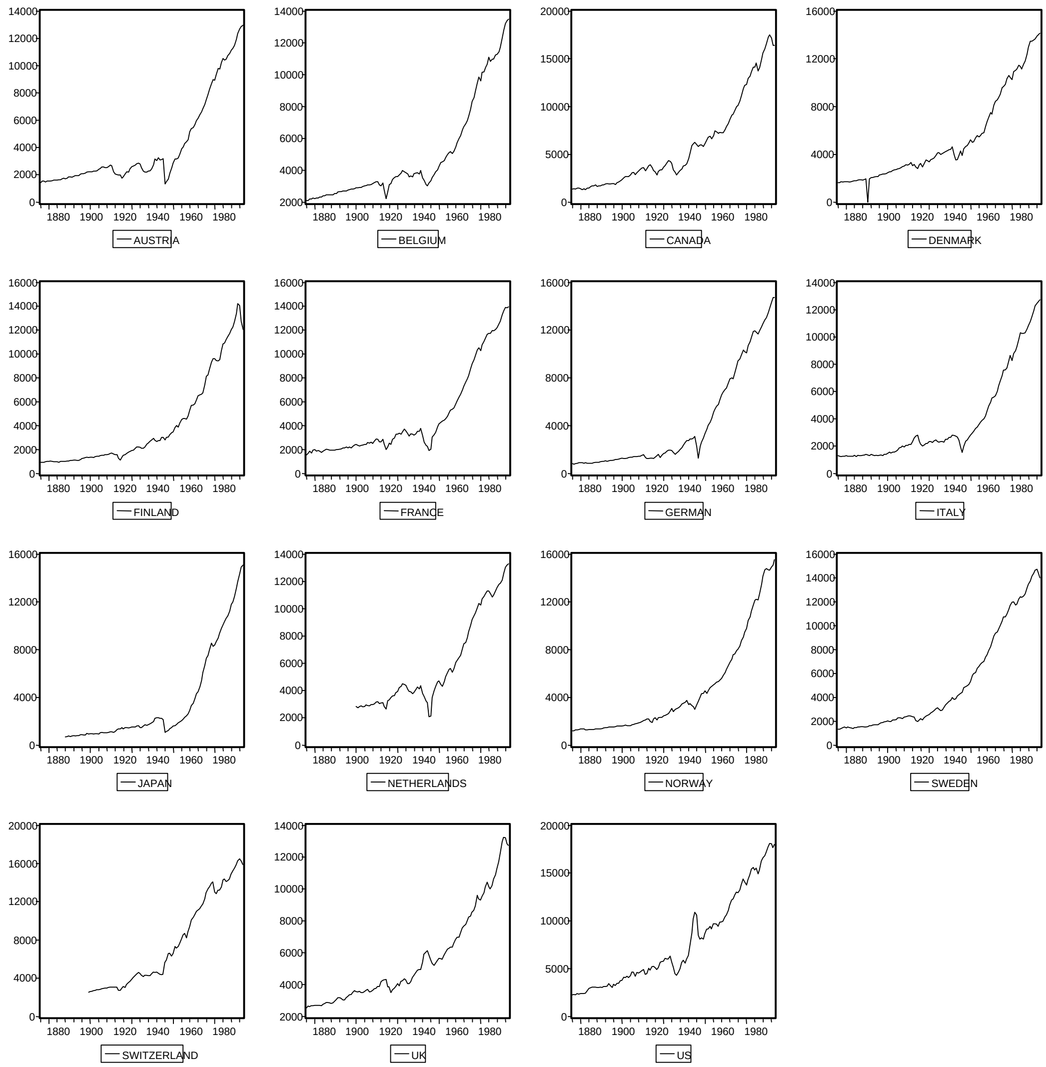


Figure 7: Differences in Real GDP Per capita Between UK and the Other Fourteen Countries from 1870 to 1992

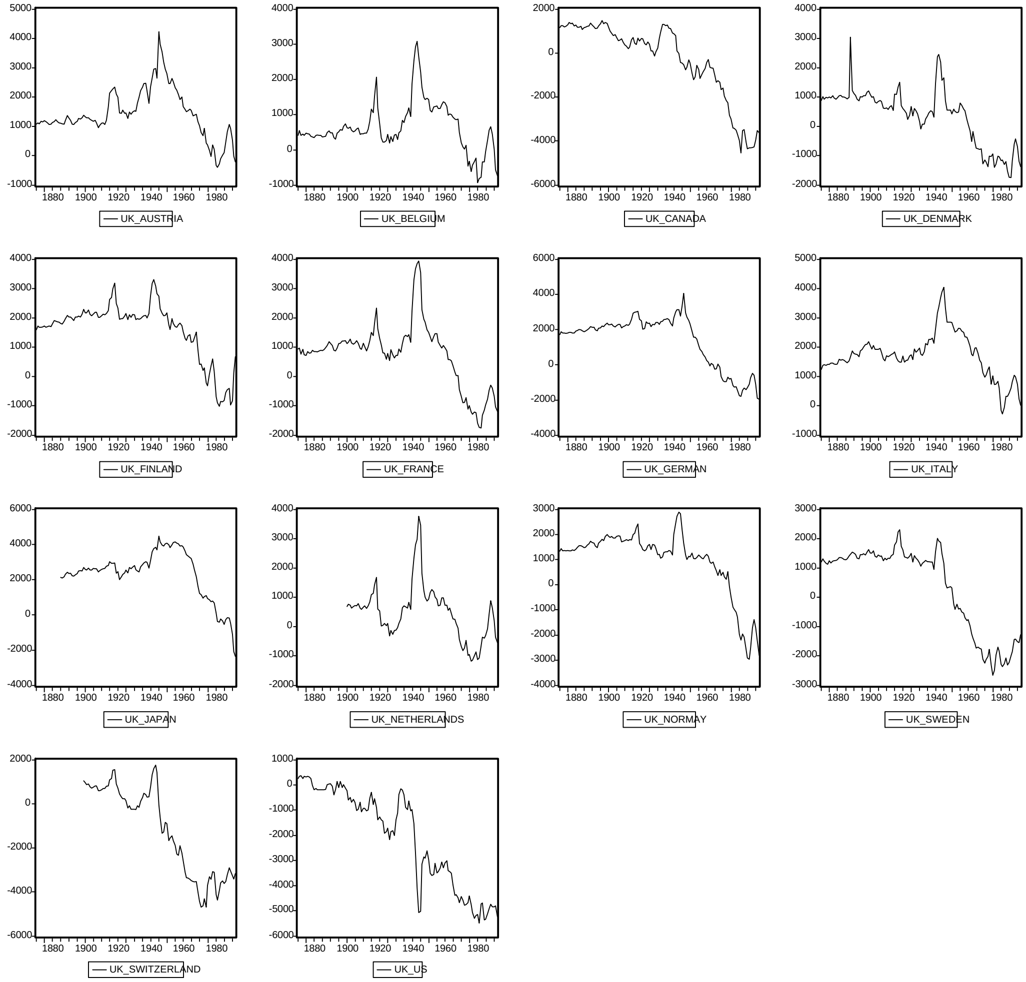
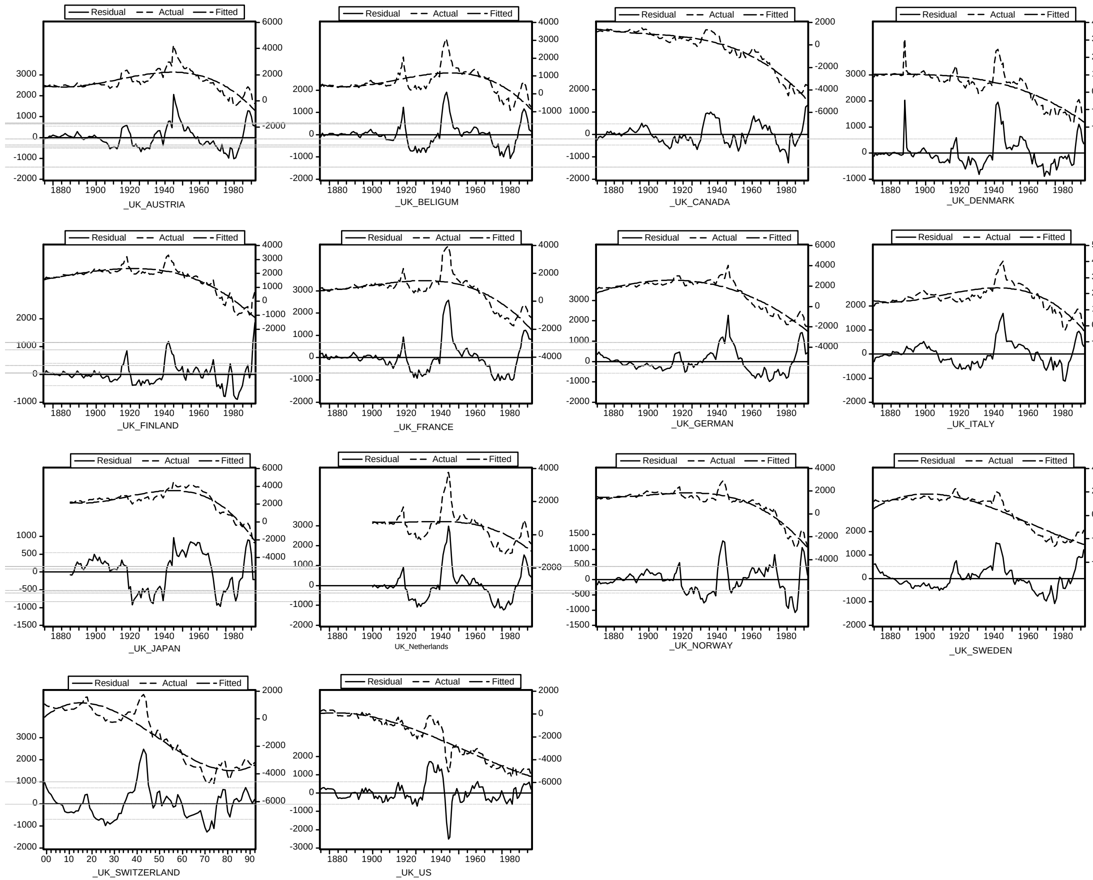


Figure 8. Actual, Fitted and Residual Values of the Difference of Real Income per capita in UK and the Other Fourteen Countries in Equation 8.



To proceed, we follow the recent macro-econometrics to process the data. The observed time-series income differentials can be decomposed into the long-run growth part and short-run fluctuation part. The model can be expressed as follows.

(8) $y_t = y_t^g + y_t^c$,

where y_t^g is the long-run growth part of the series which helps us to identify the long run tendency or pattern of the series, whereas y_t^c is the short-run fluctuation part which is expected to be an irregular and cyclical type series with some unexpected shocks or exogenous stochastic process. The research in the literature of business cycle focuses on the fluctuation part of the data. For example, Kydland and Prescott (1982) used the HP filter developed in Hodrick and Prescott (1980) to study the real business cycle theory. King and Rebelo (1993) compared the HP filter with exponential smoothing filter and questioned the measurement of the filters. Baxter and King (1995) constructed the band pass filters. Here we shall emphasize the long-run growth component. First we set y_t^g as a polynomial trend function as shown in (6).

Note that a cubic polynomial time trend is not an exact representation of the difference in per capita real income between a pair of countries, but rather it is an approximation. There is also another advantage to apply such a trend polynomial. The model can have structural changes embedded in the trend function over a long time period. The setting can absorb the effect of shocks and technological innovation in the economy. In a similar vein, by fitting the model on the difference of the pairs can also reduce the simultaneous effect of shocks and technological innovation spilled over the world.

The model to test is

(9) $y_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + u_t$

The estimated coefficients are distributed as asymptotically normal. However, we do not separate the irregular and cyclical part, y_t^c, from y_y^g, so that the residuals of the fitted time trend polynomial from our data are expected to be heteroskedastic and serial correlated. White (1980) once proposed a heteroskedasticity consistent covariance matrix estimator. However, the more important issue is the serial correlation problem. This will cause the unit root in the process. From our not reported study here, all the process cannot reject an unit root model with a second order time trend. Although this can help us to justify the use of a third order time trend polynomial function, the unit root model is hard to aid us to identify the pattern of the growth. We should stick to equation (8). Newey and West (1987) proposed a nonparametric method to calculate the heteroskedasticity and autocorrelation consistent standard errors. These estimators are robust to a variety of forms of heteroskedasticity and autocorrelation of the residuals in the regression. Newey and West (1987) suggested using the Bartlette's kernel to choose an optimum value of the lag truncation or bandwith parameter for the estimator. Andrews (1991) replaced Bartlette's kernel by the quadratic spectral kernel in the procedure. But it is more complicated and the effect is similar. Hence, we will apply Bartlette's kernel to obtain the robust standard errors which can also help us to see the significance of the estimated coefficients. Because the data set covers a long period of time, we might doubt the stability of the estimated equation. Therefore, we apply the test of cumulative sum of recursive residuals (CUSUM) in Brown, Durbin and Evans (1976) to check the stability of the estimated equations.

Here, year 1870 is replaced with 1 and year 1992 is then replaced with 123. The estimated statistics from both models are listed in Table 1, whereas the data fitted and the residuals are shown in Figure 8. In the table, we report the truncation parameters and the *t*-value of the coefficients by dividing their robust standard errors in the parenthesis. We also report the quality of the fit by using adjusted R^2, the p-value of overall coefficients significance F-test.

The p-value of overall coefficients significance F-test are zeros in all series in table 1, showing the importance of our explanatory variables. The estimated result shows that all of the cubic time trend coefficients are statistically significant at less than 10% statistical level except the coefficient of *t* for the UK_German pair. The graphs of the CUSUM test for all the pairs are shown in Figure 9. We can see there is no instability for most of the pairs except UK_Italy, UK_Japan, UK_Switzerland, which go out of bound at the end of samples at 5% significance level. If we allow the 10% significance level, they all stay in bounds. Then we can accept the estimated equations.

Figure 9. The Instability Test Results of Equation (9): the Plot of Cumulative Sum of Recursive Residuals.

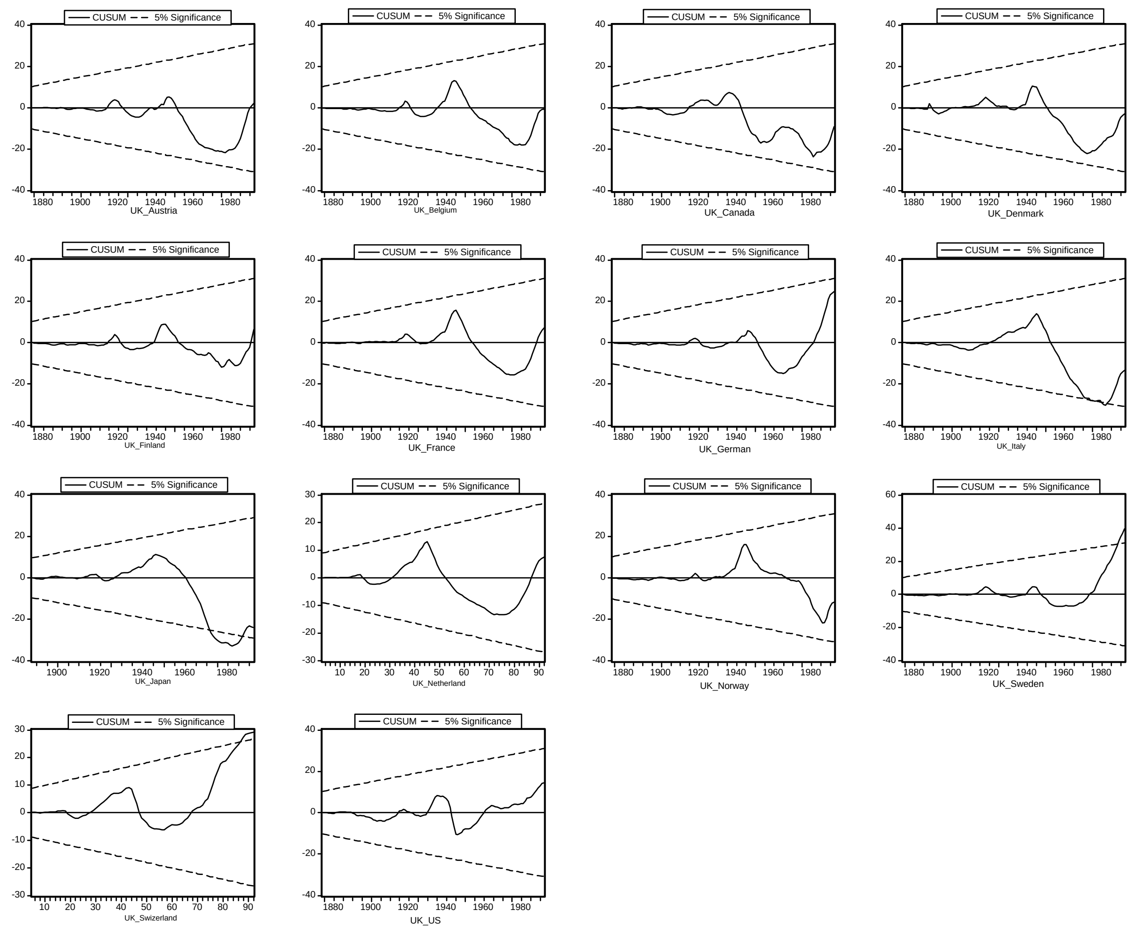


Table 1. The Result of Regressions from Equation (9).

	T	C	α_1	α_2	α_3	truncation lag	$\bar{R}^2$	F-test	$\bar{t}_1$	reject H_0 %	$\bar{t}_2$
UK_Denmark	123	1010.47 (7.84)	0.81 (0.05)	0.01 (0.04)	-0.02 (-0.74)	4	0.69	0.00	-3.5	0	3.8
UK_Finland	123	1558.14 (14.09)	21.41 (1.60)	0.056 (0.17)	-0.003 (-1.79)	4	0.83	0.00	-42.9	0	55.4
UK_France	123	769.61 (4.69)	2.50 (0.12)	0.50 (1.01)	-0.006 (-1.83)	4	0.59	0.00	-2.4	0	58
UK_Germany	123	1358.11 (6.08)	55. 02 (2.50)	-0.57 (-1.08)	-0.001 (-0.35)	4	0.85	0.00	- 423.3	0	43.3
UK_Austria	123	1164.81 (8.20)	-21.67 (-1.19)	1.14 (2.61)	-8.95E-3 (-3.28)	4	0.59	0.00	10.9	8.9	74
UK_Belgium	123	500.99 (4.00)	-19.72 (-1.31)	0.88 (2.53)	-6.66E-3 (-3.07)	4	0.41	0.00	13.2	9.9	74.9
UK_Italy	123	1549.02 (10.66)	-14.94 (-0.90)	0. 89 (2.35)	-0.007 (-3.18)	4	0.60	0.00	0.9	6.9	75
UK_Japan	108	3106.44 (4.51)	-97.97 (-2.35)	2.86 (4.04)	-0.01 (-5.60)	4	0.85	0.00	19	17.6	172
UK_Netherland	93	1534.95 (2.54)	-50.36 (-1.40)	1. 01 (1.55)	-0.006 (1.74)	3	0.24	0.00	37.4	29.8	74.8
UK_Norway	123	1490.11 (12.75)	-8.41 (-0.67)	0.69 (2.37)	-0.008 (-4.14)	4	0.89	0.00	7	5.7	51
UK_Canada	123	1399.98 (9.68)	-24.78 (-1.48)	0.52 (1.40)	-6.12E-3 (-2.72)	4	0.93	0.00	NA	100	NA
UK_Sweden	123	594.28 (2.16)	84.04 (3.59)	-1.62 (-3.23)	0. 006 (2.04)	4	0.88	0.00	31	0	149
UK_Swizerland	93	-8821.69 (-3.18)	508.10 (4.04)	-7.87 (-4.63)	0.033 (4.68)	3	0.87	0.00	45.04	7.4	114
UK_US	123	42.96 (0.23)	15.18 (0.99)	-0.97 (-2.94)	0. 004 (1.92)	4	0.90	0.00	8	0	153

Note: NA: complex root

Figure 8 shows the shapes of the cubic trend polynomial function and their residuals. We can see the evident long run trend well fitted by the curve. However, the residuals are not close to white noise and have a strong tendency of serial correlation. Since our focus is only on the long-run growth part rather than on the short-run fluctuation, the serial correlation problem of this sort is not of our main concern. Therefore we will use the estimated coefficients of long-run growth part in model (9) to test the hypothesis of the inverse U-shaped curve.

Let us start with the first four pairs. The UK_German pair has $\varnothing_2$<0 and $\varnothing_3$<0 that necessarily imply $\bar{t}_1 < 1$. Also, estimated $\bar{t}_2$ is between 1 and 123. Hence, estimated parameter values fall in parameter subspace (0b). The estimated coefficients of the UK_Denmark, UK_Finland, UK_France pairs are $\varnothing_2$ >0 and $\varnothing_3$<0 that also imply $\bar{t}_1 < 1$. Also, estimated $\bar{t}_2$ is between 1 and 123. They fall in subspace (0a), as illustrated in Fig. 7(a). The long-run difference in per capita real income for these four pairs are conclusively confirmed by the data as of an inverse-U shape for all $t > 1$.

For the next seven pairs of economies in Table 1, we obtain $\varnothing_2$ >0 and $\varnothing_3$<0. Since the first six of them have $\bar{t}_1$ >1, it follows that estimated parameters fall in subspace (0e), as shown in Fig. 7(b). In this case, the curve is of an inverse-U shape only for $t>\bar{t}_1$ which is greater than 1. It can be computed that 8.1% of the observations are inconsistent with hypothesis of inverse U shape curve for UK_Austria, 9.9% for UK_Belgium, 6.9% for UK_Italy, 14.8% for UK_Japan, 29.8% for UK_Netherlands and 4.9% for UK_Norway.

For UK_Canada, all data fall in parameter subspace (2a), meaning that the curve of real income differentials, monotonically decreasing, is not an inverse U curve in the period under consideration. The maintained hypothesis of inverse-U shape curve is thus rejected. A major reason is the lack of data prior to 1870 for which the divergence occurred.

For the last three pairs in Table 1, the UK_Sweden, UK_Switzerland and UK_US pairs, we obtain $\varnothing_1$>0, $\varnothing_2$ <0 and $\varnothing_3$>0. The estimated parameters fall in parameter subspace (0d), as shown in Fig. 7(c). Since 123< $\bar{t}_2$ for Sweden and US, 100% observations of the countries are consistent with the hypothesis of an inverse-U shape in the period under study. For Switzerland, $\bar{t}_2 = 144 < 123$. However, 92.6% observations for this country and UK are consistent with H_0. But hypothesis H_0 is accepted in the sense of fitting rather than in the sense of forecasting.

In conclusion, our analysis of data conclusively rejects the hypotheses of monotonically increasing difference in per capita real income (H_1) and of monotonically decreasing differences in per capita real income (H_2), except for Canada's data. Canada's data conclusively rejects the maintained hypothesis H_0 and accepts hypothesis H_2. Also, the data of four countries (Denmark, Finland, France and German) conclusively confirm the maintained hypothesis in both senses of fitting and forecasting. For other five countries, most of the data are consistent with H_0 in both senses of fitting and forecasting. Among them, 91.9% of Austria's data, 90.1% of Belgium's data, 93.1% of Italy's data, 85.2% of Japan's data, 70.2% of Netherlands' data, 95.1% of Norway's data and 92.6% of Switzerland's data are consistent with H_0. 100% of Sweden and the USA's data fall in the interval for H_0 to be accepted in the sense of fitting, but the data cannot be used to forecast an inverse U curve.

Overall, the data set strongly supports the maintained hypothesis H_0: an inverse-U relation exists for long-run per capita real income differentials between the UK and thirteen out of fourteen countries. It conclusively rejects hypothesis H_1. Also, it conclusively rejects hypothesis H_2 except for Canada.[17]

V. Conclusions

Following up the empirical works of Jones (1995a, b) that reject two classes of the major endogenous growth models: the AK models and R&D based models, this paper tests the third class of endogenous growth models that generate endogenous evolution in division of labor against empirical data. It is shown that this class of models not only avoids scale effects, but also accommodates both divergence and convergence phenomena. One of the hypotheses generated by this class of models is that the divergence phenomenon takes place first between each pair of economies that enter the take off stage at different points in time, and then the convergence phenomenon follows. This implies that the difference in per capita real income between the two economies is an inverse-U shape curve.

Annual long-run data of fifteen developed countries are used to conduct the empirical analysis as they all have experienced the stages of preindustrialization decelerated growth, take

[17] Those countries experimenting with a state socialist system, for instance, Argentina and Russia, may have a growth pattern with sequential divergence-convergence-divergence in comparison to the UK. This relates to implications of experiments with state socialism for growth pattern, which is beyond the scope of the current paper.

off, and mature growth which render the investigation possible. The time-series data cover a period from 1870 or later to 1992, a long-run perspective of over 120 years. Our data show that the graph of per capita real income for each country in the period under study is concave initially, then convex and finally concave again. Using the UK as the benchmark country, we have found that the difference in per capita real income between the UK and any of the other countries except Canada is an inverse-U curve, showing the phenomenon of sequential divergence and convergence. In our unreported experiments, using Germany and the US as the benchmark countries, with some other OECD countries and four little dragons (Hong Kong, Singapore, South Korea, Taiwan), which entered the take off stage later than the benchmark countries, we got similar results. The model we have employed is the cubic time trend function which allows a variety of shapes. In general, an inverse-U curve is confirmed. This indicates that both sides of the current debate concerning convergence versus divergence in per capita real income look only at a portion of the whole picture.

A promising further push of this research is to process data to push the initial year from 1870 back to 1780 and push the final year from 1992 to 1997. This will significantly improve the quality of the empirical test and more strongly support our maintained hypothesis since missing data for divergence between UK and other countries, in particular Australia and Canada, before 1870 and missing data for mature growth of Japan after 1992 are extremely unfavorable to our maintained hypothesis.

References

Aghion, Philippe and Peter Howitt (1992): "A Model of Growth Through Creative Destruction", *Econometrica*, 60, 323-351.

Andrews, D. W. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation", *Econometrica* 59, 817-858.

Arrow, Kenneth J. (1962): "The Economic Implications of Learning-by-Doing", *Review of Economic Studies*, 29, 155-173.

Barro, R. (1991): "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics*, 106(2), 407-443.

Barro, R. (1997): *Determinants of Economic Growth: A Cross-Country Empirical Study*, Cambridge, MA, MIT Press.

Barro, R. and Sala-i-Martin, X. (1991): "Convergence Across States and Regions", *Brookings Papers on Economic Activities*, # 1, 159-82.

Barro, R. and Sala-i-Martin, X. (1992): "Convergence", *Journal of Political Economy*, 100, 223-51.

Barro, R. and Sala-i-Martin, X. (1995): *Economic Growth*, New York, McGraw-Hill.

Baumol, W.J. (1986): "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show", *American Economic Review*, 76(5), 1072-1085.

Baxer, M. and King, R.G. (1995): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series", NBER working paper No. 5022.

Becker, G. and Murphy, K. (1992): "The Division of Labor, Coordination Costs, and Knowledge", *Quarterly Journal of Economics*, 107: 1137-60.

Benhabib, J. and Joranovic, B. (1991): "Externalities and Growth Accounting", *American Economic Review*, 81, 82-113.

Bernard, A. and Jones, C.I. (1996): "Technology and Convergence", *Economic Journal*, 106, 1037-1044.

Borland, J. and Yang, X. (1995): "Specialization, Product Development, Evolution of the Institution of the Firm, and Economic Growth", *Journal of Evolutionary Economics*, 5, 19-42.

Brown, R.L., J. Durbin and J.M. Evans (1975), "Techniques for Testing the Constancy of Regression Relationships Over Time" (with discussion), *Journal of the Royal Statistical Society*, Series B, 37, 149-92.

Chen, B-L. and K. Shomomura (1997): "Self-fulfilling Expectations and Economic Growth: A Model of Technology Adoption and Industrialization", *International Economic Review* (forthcoming).

Delong, B. (1988): "Productivity Growth, Convergence and Welfare: Comments", *American Economic Review*, 78, 1138-1154.

Dowrick, S. and Nguyen, D.T. (1989): "OECD Comparative Economic Growth 1950-85: Catch-up and Convergence", *American Economic Review*, 79, 1010-1030.

Easton, Stephen and Michael, Walker (1997): "Income, Growth, and Economic Freedom", *American Economic Review, Papers and Proceedings*, 87, 328-32.

Frye, Timothy and Shleifer, Andrei (1997): "The Invisible Hand and the Grabbing Hand." *American Economic Review, Papers and Proceedings*, 87, 354-58.

Galor, O. (1996): "Convergence ? Inferences from Theoretical Models", *Economic Journal*, 106, 1056-1069.

Grossman, G. and Helpman, E. (1990): "Comparative Advantage and Long-Run Growth", *American Economic Review*, 80, 796-815.

Grossman, G. and Helpman, E. (1991), "Quality Ladders and Product Cycles", *Quarterly Journal of Economics*, 106, 557-86.

Hodrick, R. J. and Prescott, E. (1980): "Post-war U.S. Business Cycles: An Empirical Investigation," working paper, Carnegie-Mellon.

Jones, C.I. (1995a): “Time Series Tests of Endogenous Growth Models”, *Quarterly Journal of Economics*, 110, 695-525.

Jones, C.I. (1995b): “R & D-Based Models of Economic Growth”, *Journal of Political Economy*, 103, 759-784.

Judd, K. (1985): "On the Performance of Patents”, *Econometrica*, 53, 579-85.

King, R. G. and Rebelo, S. T.(1993): "Low Frequency Filtering and Real Business Cycles," *Journal of Economic Dynamics and Control*, Vol.17(1), 207-231.

Kelly, Morgan (1997): “The Dynamic of Smithian Growth”. *Quarterly Journal of Economics*, 112, 939-64.

Kydland, F. and Prescott, E. (1982): Time to Build and Aggregate Fluctuations. *Econometrica*, 50, 1345-1370.

Lucas, R. E. Jr. (1988): "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22, 3-42.

Lucas, R. E. Jr. (1993): "Making a Miracle", *Econometrica*, 61, 251-72.

Madison, A. (1991): *Dynamic Forces in Capitalist Development: A Long Run View*, New York: Oxford University Press.

Mankiw, N.G., Romer, D. and Weil, D. N., (1992): “A Contribution to the Empirical of Economic Growth”, *Quarterly Journal of Economics*, 107(2), 407-437.

Matsuyama, K. (1991): “Increasing Returns, Industrialization, and Indeterminancy of Equilibrium”, *Quarterly Journal of Economics*, 106, 617-650.

Murphy, K., Shleifer, A., and Vishny, R. (1989a): "Industrialization and the Big Push", *Journal of Political Economy*, 97: 1003-26.

National Research Council (1986), *Population Growth and Economic Development: Policy Questions.* Washington, DC, National Academy of Sciences Press.

Newey, W. and K. West (1987): “A Simple Positive Definite, Heterskedasticity and Autocorrelation Consistent Covariance Matrix”, *Econometrica*, 55, 703-708

North, D. (1958), "Ocean Freight Rates and Economic Development", *Journal of Economic History*, ??, 537-55.

North, D. (1986), "Measuring the Transaction Sector in the American Economy", S. Eugerman and R. Gallman, eds,. *Long Term Trends in the American Economy*, Chicago, University of Chicargo Press.

North, Douglass and Weingast, Barry (1989), “Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England,” *Journal of Economic History*, XLIX, pp 803-32.

Nurske, R. (1953): *Problems of Capital Formation in Underdeveloped Countries*, New York: Oxford University Press.

Outliaris, S., Park, J.Y. and Phillips., P.C.B. (1989): “Testing for a Unit Root in the Presence of a Maintained Trend,” in Raj, B. edit *Advance in Econometrics and Modeling*, Kluwer Academic Publishers, Boston.

Quah, D. (1993): “Galton‘s Fallacy and Tests of The Convergence Hypothesis”, *Scandivarian Journal of Economics*, 95(4), 427-443.

Quah, D. (1996): “Twin Peaks: Growth and Convergence in Models of Distribution Dynamics”, *Economic Journal*, 100, 1045-1055.

Rebelo, S. (1991): “Long-Run Policy Analysis and Long-Run Growth”, *Journal of Political Economy*, 99: 500-521.

Romer, P. (1986): "Increasing Returns and Long Run Growth", *Journal of Political Economy*, 94, 1002-37.

Romer, P. (1987): "Increasing Returns, Specialization, and External Economies: Growth as Described by Allyn Young", *American Economic Review, Papers and Proceedings*.

Romer, P. (1990): "Endogenous Technological Change", *Journal of Political Economy*, 98, S71-S102.

Romer, P. (1993): "The Origins of Endogenous Growth", *Journal of Economic Perspectives*, 8, 3-22.

Rosen, S. (1978): "Substitution and the Division of Labor", *Economica*, 45: 235-50.

Rosen, S. (1983): "Specialization and Human Capital", *Journal of Labor Economics*, 1: 43-49.

Rostow,W.W. (1960): *The Stages of Economic Growth: A Non-Communist Manifesto*, Cambridge: Cambridge University Press.

Sachs, Jeffrey and Warner, Andrew (1997): "Fundamental Sources of Long-Run Growth." *American Economic Review, Papers and Proceedings*, 87, 184-88.

Sala-i-Martin, (1996): "The Classical Approach to Convergence Analysis", *Economic Journal*, 106, 1019-1036.

Schultz, T. (1993): *Origins of Increasing Returns*, Cambridge, MA, Blackwell.

Smith, Adam (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*. Reprint, edited by E. Cannan. Chicago: University of Chicago Press, 1976.

Solow (1956): "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70: 65-94.

Stoky, Nancy (1991): "Human Capital, Product Quality and Growth", *Quarterly Journal of Economics*, 106, 587-616.

Tamura, R. (1991): "Income Convergence in an Endogenous Growth Model", *Journal of Political Economy*, 99, 522-40.

Wen, M. (1997): "Infrastructure and Evolution in Division of Labor", *Review of Development Economics*, No. 2.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48, 817-838.

Yang, X. and Borland, J. (1991): "A Microeconomic Mechanism for Economic Growth", *Journal of Political Economy*, 99: 460-82.

Yang, X. and Ng, Y-K. (1993): *Specialization and Economic Organization, a New Classical Microeconomic Framework*, Amsterdam, North-Holland.

Yang, X., Wang, J., and Wills, I. (1992): "Economic Growth, Commercialization, and Institutional Changes in Rural China, 1979-1987", *China Economic Review*, 3, 1-37.

Young, Allyn (1928): "Increasing Returns and Economic Progress", *The Economic Journal*, 152, 527-542.

Zhang, J. (1997): "Evolution in Division of Labor and Macroeconomic Policies", *Review of Development Economics*, No. 2.