Knowledge and Growth in the Very Long Run

Holger Strulik*

First Version: February 2009. This Version: September 2011.

Abstract. This paper proposes a theory for the gradual evolution of knowledge diffusion and growth over the very long run. A feedback mechanism between capital accumulation and the ease of knowledge diffusion explains a long epoch of (quasi-) stasis and an epoch of high growth linked by a gradual economic take-off. It is shown how the feedback mechanism can explain the Great Divergence, the failure of less developed countries to attract capital from abroad, and a productivity slowdown in fully developed countries. An extension towards a two-region world economy shows robustness of the gradual take-off and other interesting interaction between forerunners and followers of the Industrial Revolution.

Keywords: Industrial Revolution; Endogenous Growth; Knowledge Diffusion; Productivity Slowdown; Convergence; Divergence.
JEL: O10, O30, O40, E22.

---

* University of Hannover, Wirtschaftswissenschaftliche Fakultaet, 30167 Hannover, Germany; email: strulik@vwl.uni-hannover.de. I would like to thank Thomas Andersen, Matteo Cervellati, Carl-Johan Dalgaard, David de la Croix, Karl Dietrich, Oded Galor, Christian Groth, Peter Howitt, Andreas Irmen, Ines Lindner, Fidel Perez Sebastian, Karl-Gunnar Persson, and Thomas Steger for their helpful comments and Jakob Madsen and Rajabrata Banerjee for sharing their data. An earlier version of this paper circulated as University Hannover, Working Paper 414 (Strulik, 2009). The earlier version contains some further extensions of the theory and is still available.

1. Introduction

From the perspective of the very long-run, economic growth is a quite recent phenomenon. During most of world history economies were not (visibly) growing in income per capita terms. It took until the time of the so called Industrial Revolution that income per capita set off to grow permanently and at unprecedent rates. How large income and productivity growth exactly were during the Industrial Revolution is heavily debated among economic historians. But they unanimously emphasize one point: the take-off to modern growth was *gradual*.[1]

In economic language, a gradual take-off of income growth and productivity growth means that there were (temporarily) increasing returns to scale at work. This is a troublesome observation for conventional models of economic growth. They are based either on constant returns (if they are in reduced form of the a linear Ak structure) or on decreasing returns (the neoclassical growth model and all its derivatives). Consequently they display either no adjustment dynamics or, in case of the neoclassical growth model, they predict that the rate of economic growth is highest when the endowment with physical (and human) capital is lowest and that growth is subsequently falling with economic development. That is, they generate the wrong adjustment dynamics with respect to the long-run historical record.

One conclusion from the poor performance of conventional growth theory in motivating a gradual take-off could be that these simple models are just unsuitable for the discussion of historical economic development. This is the view of unified growth theory, which emphasizes the interaction between demographic and economic forces as essential for an understanding of economic growth in the very long-run.[2] Here, I propose a complementing view: a modification of conventional growth theory by a feedback mechanism between capital accumulation and knowledge diffusion such that it gets the historical adjustment dynamics right. Of course, I am not denying the importance of demographic change for long-run economic development. But disregarding the demographic channel helps to keep the model simple, to disentangle effects, and to establish, in theory, the accumulation-diffusion feedback as a stand-alone mechanism, which, of course, in practice interacts with demographic forces.

The proposed theory of dynamic knowledge diffusion is built upon the Arrow (1962)-Romer (1986)

[1] For example, according Crafts and Harley (1992) productivity growth in Britain was 0.1 percent from 1760-1800, 0.35 percent from 1801 to 1831 and 0.8 percent from 1831 to 1860. For comparison, from 1947 to 1973 TFP growth in the U.K. was 1.93 percent according Barro and Sala-i-Martin (2004). For evidence on gradual take-off of income and TFP see also Temin (1997), Mokyr (2002, 2005), Antras and Voth (2003), and Clark (2007).

[2] See Galor and Weil (2000), Kögel and Prskawetz (2001), Jones, 2001, Lucas (2002), Galor and Moav (2002), Boucekkine et al. (2002), Doepke (2004), Galor and Mountford (2008), Strulik and Weisdorf (2008) and many others. See Galor (2005) for a survey.

learning-by-doing mechanism. It is simple enough to be represented in reduced form by one equation and it encompasses the neoclassical growth model and the Romer (1986) model as corner solutions. The basic idea consists of a distinction between *existing* knowledge and *accessible* knowledge and a double role for capital accumulation. As in the Arrow-Romer setup the aggregate capital stock approximates existing knowledge. But instead of assuming a constant – and in the Romer case complete – degree of knowledge diffusion the new theory, inspired by Mokyr (2002), views knowledge diffusion as a process that can be alleviated by capital accumulation. More capital (more horses, trains, cars, airplanes etc.) improves the movement of people and ideas and thus improves the degree to which an individual firm has access to existing knowledge. The resulting higher factor productivity in turn triggers even more investment and higher growth, which further alleviates the movement of ideas and so forth.

The mutually enforcing power of capital accumulation and knowledge diffusion generates a long-period of (quasi-) stasis when both forces are low. Over time, however, the implied socially increasing returns on investment manifest themselves in gradually increasing rates of growth of factor accumulation and TFP. The phase were these forces for the first time became visible in the data is then called the Industrial Revolution. But increasing returns are also inherently temporary. The phase of gradual adjustment comes to an end when knowledge diffuses completely, i.e. when the degree of knowledge diffusion approaches unity and the economy converges towards a Romer (1986) economy.

The notion of a gradual evolution of knowledge diffusion, both over time and over space, is supported by empirical evidence. Clark (2007) compiles historical data on the speed of information travel from ancient times until the 20th century and concludes that "in the Malthusian era people lived in a world where information spread so slowly that many died fighting over issues that had already been decided". Of course, the mere flow information is not equal to the speed of knowledge diffusion. As emphasized by Mokyr (2002) most knowledge is not fully codified. It is embodied in people such that the speed of knowledge diffusion depends crucially on the speed and availability of means of (mass) transportation. For the time around the Industrial Revolution Clark (2007) compiled diffusion lags for the steam engine, the cotton mill and the railway of on average thirteen years for Western European countries, 22 years for southern and eastern Europe, thirty-five years for India, and fifty-two years for Latin America. Both, Clark and Mokyr continue and argue that the productivity gains achieved during the Industrial Revolution (and onwards) were to a large extent caused by development and accumulation of means of mass transportation.

For modern times, Keller (2002) and Bottazzi and Peri (2003) present evidence that knowledge spillovers are spatially localized and decay strongly with geographic distance. Keller also demonstrates that the degree of localization has shrunk substantially over time. Similarly, Jaffe et al. (1993) find localization effects for the links between patent creation and patent citation on the level of country, state, and metropolitan area, and that localization fades gradually over time. For DRAM production in fully developed countries Irwin and Klenow (1994) show evidence for international knowledge spillovers. At the same time, Foster and Rosenzweig (1995) present evidence that farmers in LDCs are still learning mostly from their neighbors. Andersen and Dalgaard (2011) show that across countries aggregate TFP and GDP per worker are strongly associated with the intensity of international travel.

Summarizing, there exists corroborating evidence that the diffusion of ideas is associated with the movement of people and that capital accumulation eases the movement of people by shortening the effective distance between them. This way capital accumulation is both consequence and cause of the gradual rise of TFP growth during and after the Industrial Revolution. Still, there exists a another force tying capital accumulation to productivity, which is the factory itself. The idea of the factory became popular around the time of the Industrial Revolution when the costs of moving people became sufficiently low to gather the workforce at one place (Mokyr, 2002). It was through capital accumulation at the factory-level that entrepreneurs elicited to an increasing degree scale effects from knowledge exchange and learning-by-doing at the workplace.

In principle, scale effects through learning-by-doing could not only work through capital but also through population size as in Kremer (1993). A recent study by Comin et al. (2010), however, provides indirect evidence in support of the here proposed channel. Comin et al. investigate technology adoption across countries over the very long-run and show that technological differences are very persistent, an indication for increasing returns to scale. They arrive at the conclusion that their findings are generally supportive of models based on a feedback mechanism from technology to lower access costs for knowledge and learning-by-doing. They then continue by testing whether the feedback channel is through population size. Controlling for historical population size they find that it is significantly negatively associated with knowledge adoption, a result that is “inconsistent with an important role for historical population size determining future technology”.

Naturally the ease of knowledge diffusion is also affected, at any given capital stock, by institutions (as, for example, barriers to travel and trade). Keller and Shiue (2008) provide evidence that early European economic integration was to a larger degree affected by capital accumulation (the expansion of the railway network) than by institutional change (customs liberalization and currency agreements). However, they also document an indirect effect of institutions on economic performance in that better institutions improved the rate of adoption of steam trains. For the modern world Coe et al. (2008) provide evidence that institutional differences are an important determinant for the national appropriation of international R&D spillovers. The present paper includes a parameter representing the capital-independent and potentially institution-based determinants of knowledge diffusion. Comparative statics show that low institutional efficiency can prevent or delay industrialization and convergence towards modern growth.

After developing the basic model in Section 2 and discussing its qualitative implications in Section 3, a network-based foundation of knowledge diffusion is introduced in Section 4. For that purpose I utilize the famous Small World model (Watts and Strogatz, 1998) and replace the original assumption that "nature" establishes long-distance links by the assumption that long-distance links are men-made and built from capital. I show that a mild assumption about the "production function" of long-distance links suffices to motivate the theory of knowledge diffusion and growth. In Section 5 I set up a numerical version of the model and compare the long-run adjustment dynamics predicted by theory with real data. It will be shown that the model time series for income per worker and TFP approximate the historical evolution of England 1620-2006 reasonably well. In particular, the theory gets the gradual historical take-off to growth about right.

The closed economy model is extended in Section 6 towards a two-country version in which knowledge (and capital) is allowed to move across borders. This part is strongly related to a large literature on international knowledge diffusion and catch-up growth.[3] In contrast to the present work, this literature focusses on adjustment dynamics of the *followers* of the industrial revolution. Mostly, it is assumed that the leading "world technology frontier" is exogenous and growing at a constant rate. In any case, to my best knowledge, the gradual take-off of TFP at the "world frontier" itself has not been investigated. Explaining the gradual take off of the *leaders* of the industrial revolution is the main purpose of the present study.

The learning-by-doing mechanism is appropriate to investigate technological and economic development for most of human history because technological advances were not (much) brought

[3]See among many others Grossman and Helpman (1991), Parente and Prescott (1994, 2005), Barro and Sala-i-Martin (1997), Basu and Weil (1998), Eaton and Kortum, 1999, Howitt (2000), Lucas (2009). Klenow and Rodriguez-Clare (2005) provide a survey on the role of knowledge diffusion in explaining international TFP differences.

forward by formally trained scientists before the mid 19th century (Mokyr, 2002). Since then, however, knowledge production has increasingly become a market activity, rendering the Arrow-Romer approach less appropriate. In an earlier version of this paper (Strulik, 2009) I investigated the accumulation-diffusion mechanism also within a Romer (1990)-style, R&D-driven growth model. I briefly discuss results from there and other potential extensions of the present model in the Conclusion.

2. Model Setup

2.1. **Households.** Consider an economy populated by two overlapping generations. The concept of overlapping generations is useful in order to investigate adjustment dynamics analytically but it is not driving the results. Members of the young generation supply one unit of labor, earn wages w_t, and divide their labor income on current consumption c_t^1 and on savings for the second period of life. Members of the old generation do not work and live off the returns on their savings. More specifically, we assume that the young individuals of period t maximize utility $u_t = \log(c_t^1) + \beta \log(c_{t+1}^2)$ where β is the discount factor. They face the current period's budget constraint $c_t^1 = w_t - s_t$ and the next period's budget constraint $c_{t+1}^2 = R_{t+1} s_t$ where R_{t+1} is the expected gross interest rate and s_t are savings. This standard OLG setup provides the well-known solution for savings (1).

$$s_t = \frac{\beta}{1+\beta} \cdot w_t. \tag{1}$$

There is no population growth. The size (mass) of a generation is normalized to one.

2.2. **Firms.** There exists a continuum of size one of competitive firms. Firms produce a homogenous output using a Cobb-Douglas production function and employing capital and labor. In period t a firm i employs capital $k_t(i)$ and labor $\ell_t(i)$ and produces output $y_t(i) = k_t(i)^\alpha [A_t(i)\ell_t(i)]^{1-\alpha}$ where total factor productivity $A_t(i)$ is exogenous to the single firm and $0 < \alpha < 1$. For simplicity and without loss of generality we assume that capital depreciates fully within one generation. Profit maximization implies that production factors are demanded such that factor prices equal the (private) marginal product, i.e. $w_t = (1-\alpha) k_t(i)^\alpha A_t(i)^{1-\alpha}$, $r_t = \alpha k_t(i)^{\alpha-1} A_t(i)^{1-\alpha}$ where w_t denotes wages and r_t denotes the interest rate. Aggregate (i.e. average) employment is denoted by $k_t = \int_0^1 k_t(i) \mathrm{d}i$ and $\ell_t = \int_0^1 \ell_t(i) \mathrm{d}i = 1$.

2.3. **Knowledge Diffusion.** As proposed by Arrow (1962) and Romer (1986) we think of knowledge embodied in capital goods such that the aggregate capital stock approximates the existing knowledge.

In the studies of Arrow and Romer the degree of knowledge diffusion σ was constant; the association between $\log A_t(i)$ and $\log k_t$ was assumed to be linear, $\log A_t(i) = \sigma \cdot \log k_t$. Whereas Arrow focussed on the case of incomplete knowledge diffusion by assuming a degree $\sigma < 1$, the Ak growth model was based on the assumption of complete knowledge diffusion, $\sigma = 1$, such that the knowledge of each firm changes in proportion to the aggregate capital stock, $\mathrm{d}A_t(i)/A_t(i) = \mathrm{d}k/k$. In any case the degree of knowledge diffusion was assumed to be independent from the state of the economy.

Here, we explicitly distinguish between *existing* knowledge, as before, approximated by the aggregate capital stock, and *accessible* knowledge. The degree to which a single firm has access to existing knowledge is identified by σ_t. It is no longer assumed to be a constant. How easily knowledge diffuses through the economy depends on the state of the economy at a given time. Equation (2) is the simplest way to formalize this fact.

$$A_t(i) = \bar{A}{k_t}^{\sigma_t}. \tag{2}$$

In a completely localized economy, knowledge spillovers between firms are at their minimum, the single firm has no access to knowledge created elsewhere, $\sigma_t = 0$, and the model is isomorph to the neoclassical growth model. In a fully integrated economy $\sigma_t = 1$, and firms are capable to access all existing knowledge. For $\sigma_t = 1$ the model is isomorph to (an overlapping generations version of) the original Romer (1986) setup and learning-by-doing and knowledge diffusion generate perpetual economic growth at a constant rate. In between these bounds we have a developing economy at an intermediate degree of knowledge diffusion. The time-invariant efficiency parameter $\bar{A}$ controls the diffusion of knowledge independently from the current state of the economy. It can be thought of capturing the effect of persistent determinants of access to knowledge as, for example, institutional barriers to travel and trade.

2.4. **Static equilibrium.** In equilibrium all firms make the same choices, $\ell_t(i) = \ell_t$, $k_t(i) = k_t$ for all i. Inserting this fact and (2) into wages we get (3).

$$w_t = (1-\alpha)\bar{A}^{1-\alpha}k_t^{\alpha+(1-\alpha)\sigma_t}. \tag{3}$$

Wages are increasing in the capital endowment of the workplace and this effect is in turn increasing in σ_t because a high degree of knowledge diffusion implies that workers have access to much of the existing knowledge, a fact that amplifies worker productivity for any given capital stock k_t. For later reference it is also useful to define GDP per capita y_t, $y_t \equiv \bar{A}^{1-\alpha}k_t^{\alpha+(1-\alpha)\sigma_t}$.

2.5. **Knowledge Diffusion and Growth.** According to the OLG setup the capital stock with which the next period's young generation is working is determined by the savings decision of this period's young generation, $k_{t+1} = s_t$. Inserting (1) – (3) we get the equation of motion (4).

$$k_{t+1} = ak_t^{\alpha+(1-\alpha)\cdot\sigma_t} \tag{4}$$

where $a \equiv (1-\alpha)\bar{A}^{1-\alpha}\beta/(1+\beta)$.

The final element is a positive feedback effect from the size of the capital stock to the degree of knowledge diffusion because capital goods alleviate travel and information exchange, $\sigma_t = \sigma(k_t)$. A rich economy, ceteris paribus, has a higher endowment of capital per capita (e.g. more rails, more cars, more airports, and more miles of telephone lines) than a poor economy, such that people and firms are better economically integrated and knowledge and ideas diffuse more easily through the rich economy. The following assumption about knowledge diffusion and capital stock seems to be reasonable.

Assumption 1 (Knowledge Diffusion). *The degree of knowledge diffusion is continuous in* $[0,1]$ *and monotonously increasing in* k, $\sigma'(k) \geq 0$, *with* $\sigma(0) = 0$ *and* $\lim_{k\to\infty}\sigma(k) = 1$.

Applying Assumption 1 in (4) we get the economy represented by a single difference equation.

$$k_{t+1} = f(k_t) = ak_t^{\alpha+(1-\alpha)\cdot\sigma(k_t)}. \tag{5}$$

Inspection of (5) shows that at the origin – when there is no capital to transport people and knowledge – the model coincides with the neoclassical growth model and that when the capital stock approaches infinity knowledge is completely accessible such that the model coincides with the Ak-growth model, $\lim_{k\to\infty} k_{t+1} = ak_t$. In between these bounds, only some knowledge is accessible to the single firm and the degree of knowledge diffusion is increasing in the existing capital stock.[4] Since the economy is capable of long-run growth only if $a > 1$ we make the following parameter restriction in order to have an interesting problem.

Assumption 2 (Feasibility of Growth). *Technology* (A,α) *and preferences* (β) *support long-run growth at a positive rate, i.e.* $a \equiv (1-\alpha)\beta/(1+\beta)\bar{A}^{1-\alpha} > 1$.

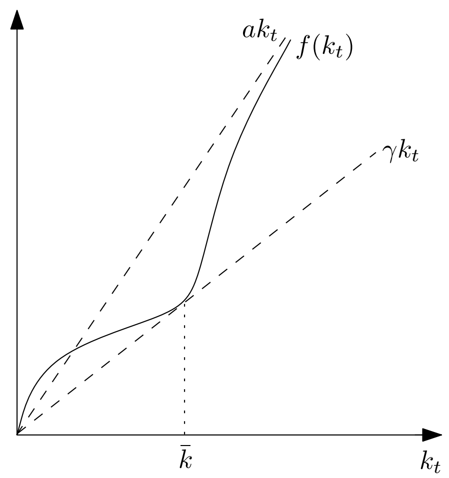
[4]Aside from the definition of σ as a degree there exists a further reason for its upper bound of unity. For $\sigma > 1$ the economy would arrive at infinite capital stock in finite time, a phenomenon that has to be excluded from any meaningful model for logical reasons. Of course, the upper bound could be characterized by less than complete knowledge diffusion. I discuss this possibility in the Conclusion.

If a were smaller than one, we would be back in the neoclassical world without sustainable long-run growth. The interesting question here is not whether a balanced growth path exists but if and how an economy arrives there in the long-run. Finally we impose a mild assumption on the curvature of $\sigma(k)$.

ASSUMPTION 3 (Curvature). *The term* $b(k) \equiv \sigma'(k) \cdot k \cdot \log(k) + \sigma(k) - 1$ *changes its sign exactly once for* $k \in [0, \infty)$.

Intuitively we assume that the degree of knowledge diffusion increases not too rapidly in k. To see this, note that b is negative for small k since $\log(k)$ is negative for $k < 1$ and $\sigma(k) - 1$ is negative for all k. For rising k the term $k \log(k)$ approaches infinity and $\sigma(k) - 1$ approaches zero from below. A change of sign of $b(k)$ thus requires that $\sigma'(k)$ goes less quickly to zero than $k \log(k)$ goes to infinity. This assumption is a very mild one. It is even fulfilled for some functional forms of σ for which the approach to complete knowledge diffusion is implausibly fast already at low capital stocks as, for example, a quadratic logistic ($\sigma(k) = 1 - e^{k^2}$) or a concave ($\sigma(k) = 1 - 1/(1 + k)$) function. Actually, it is hard to come up with a functional form for σ that does not support Assumption 3.

Figure 1: The Shape of $f(k)$



The interaction between capital accumulation and knowledge diffusion introduces an interesting non-linearity into the evolution of capital (5):

LEMMA 1. *There exists a capital stock* $\bar{k}$ *where* $f(k)$ *has a supporting tangent* γk.

The proof is in the Appendix. Intuitively, when the capital stock is low ($k < \bar{k}$) there is little economic integration and little flow of goods and people. As a result, the degree of knowledge diffusion $\sigma(k)$ is small and the neoclassical part of technology, characterized by decreasing returns to physical factor accumulation, is dominating. On the other hand, for $k > \bar{k}$ the modern part of production, characterized by increasing (social) returns through learning-by-doing and knowledge diffusion, is dominating.

Figure 1 visualizes the result. When capital stock is low and there is little knowledge diffusion, the neoclassical channel of decreasing individual returns to investment dominates and $f(k_t)$ is concave, as suggested by the neoclassical approach. At capital stock $\bar{k}$ the modern channel capturing the expansive power of socially increasing returns to investment through knowledge spillovers becomes dominating and the $f(k_t)$ curve becomes convex. In the limit, when k approaches infinity, and aggregate knowledge becomes completely accessible by every firm, $f(k_t)$ becomes linear (with slope a) reflecting constant social returns to investment. Overall the $f(k)$ curve assumes a concave-convex or "hyperbolic" shape.

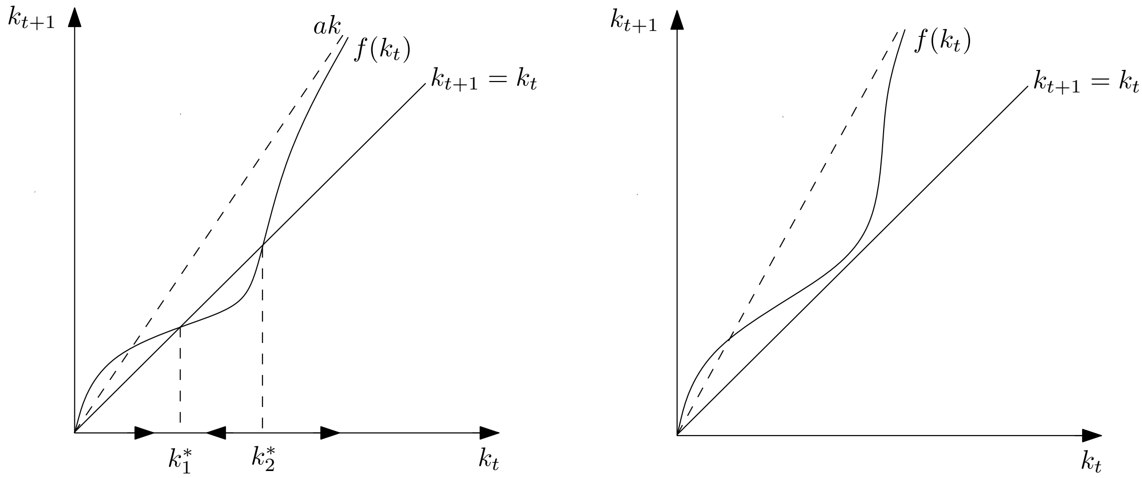
3. Steady-States and Adjustment Dynamics

3.1. **Steady-States.** The hyperbolic shape of $f(k)$ provides some plausible steady-state features and adjustment dynamics which cannot be captured by the simple neoclassical model or by the Ak growth model. In a phase diagram the concave-convex shape implies that the $f(k)$ curve cuts the identity line either twice or never. From Figure 1 it is obvious that an intersection exist if and only if the slope of the supporting tangent is smaller than unity. If two intersections exist, then there exist two equilibria of stagnation (we neglect the trivial steady-state at the origin and the degenerate case where $\gamma = 1$ and the supporting tangent coincides with the identity line).

Proposition 1. *If $\gamma < 1$, then $f(k)$ has two fixed points, k_1^*, k_2^*. The fixed point at $k_1^* < k_2^*$ is a locally stable equilibrium of stagnation, the fixed point at k_2^* is unstable. Otherwise, if $\gamma > 1$, there exists no fixed point. In that case and in the case of $k_0 > k_2^*$ for $\gamma < 1$ the economy converges towards balanced growth. The growth rates of capital and income are increasing during the adjustment towards the balanced growth path and are eventually converging towards the constant $a - 1$.*

The proof is given in the Appendix. Figure 2 explains the result. The figure on the right hand side shows the outcome for $\gamma < 1$, i.e. when the supporting tangent of $f(k_t)$ lies below the identity line such that there exist two equilibria of stagnation. The first equilibrium at k_1^* is locally stable. It is situated along the concave part of $f(k)$, i.e. in the "neoclassical domain" of production when

Figure 2: Phase Diagram: Dynamics of the Arrow-Romer Economy



decreasing returns to physical factor accumulation dominate. In line with conventional neoclassical growth theory the fact of decreasing returns renders the equilibrium stable. Here, however, stability is only a local property. The second equilibrium at $k_2^{*}*$ is unstable. It is situated along the convex branch of $f(k)$, i.e. in the "modern domain" of production when increasing (social) returns through increasing economic integration and diffusion of knowledge dominate.

3.2. **Comparative Statics.** Observe that the compound constant $a \equiv (1-\alpha)\bar{A}^{1-\alpha}\beta/(1+\beta)$ operates like a re-scaling of $f(k_t)$ along the k_{t+1} axis: with rising a the supporting tangent γk and the the asymptote ak get steeper and $f(k)$ is pulled upwards. This observation proves the following result.

Proposition 2. *There exists a parameter constellation $\{\bar{A}, \beta\}$ for which stagnation exists (does not exist). Income per capita at stagnation is increasing in $\bar{A}$ and β. Stagnation does not exist if $\bar{A}$ or β are sufficiently large.*

This means that stagnation is less likely to exist if there are good, knowledge-flow promoting institutions ($\bar{A}$ is high) and if individuals put sufficient weight on consumption in old age (β is high) because, for example. the disease environment is sufficiently mild and medical knowledge is sufficiently advanced. Given an economy trapped in stagnation an exogenous improvement of institutions or survival prospects can initiate an escape from poverty and convergence towards balanced growth.

3.3. **The Poverty Trap.** The Arrow-Romer equilibrium of stagnation differs qualitatively from the usual poverty trap. Stagnation occurs regardless of subsistence needs at an income level which – depending on parameter choice – may exceed by far the income level usually associated with subsistence. This way, the equilibrium of stagnation has more potential to explain actually observed poor growth performance. In particular, stagnation may occur in midst of a process of increasing economic integration at an intermediate degree of knowledge diffusion σ, a result that may help to explain the poor growth performance of countries which are less appropriately characterized as traditional societies.[5]

3.4. **Why Doesn't Capital Flow to Poor Countries?** With contrast to the subsistence argument for stagnation, which relies on too low national savings rates in a closed economy, the present model can provide also an intuition for why capital is not flowing from rich to poor countries (Lucas, 1990). The argument is based on the *dilemma* originating from the fact that knowledge spillovers are external to the individual firm. In the neighborhood of the equilibrium of stagnation k_1^*, capital would not flow from a rich country (where $k > k_2^*$) to the poor country because capital productivity is low. In turn, capital productivity is low because capital endowment per workplace is low such that learning-by-doing effects and knowledge diffusion are small.

However, if the two countries are connected such that at least some knowledge created in the rich country spills over to the poor one, stagnation and missing capital flows are not sustainable in the long-run. We take up issues of catch-up growth and international flows of knowledge and capital in Section 4.

3.5. **The Gradual Transition to Balanced Growth.** If the equilibrium of stagnation does not exist, as shown on the right hand side of Figure 2, the economy converges towards perpetual growth. Along the transition path, the rate of economic growth is perpetually increasing. In the figure this can be seen by the fact that the distance between the $f(k_t)$ curve and the identity line gets larger with rising k_t, implying that the growth rate of k_t, $f(k_t)/k_t - 1$, increases. Intuitively, the interaction and mutually enforcing power of capital accumulation and knowledge diffusion explains why the take-off to balanced growth is gradual. In the "beginning" when the aggregate capital stock is relatively low, knowledge diffuses relatively little such that productivity is relatively low, and income, and investment (savings) are relatively low. With ongoing capital accumulation, firms

[5]See Kraay and Raddatz, 2007, for the difficulties incurred by calibrating actual economies to get support of stagnation at subsistence level.

become better integrated and can access more of the existing knowledge, such that productivity and thus income and investment (savings) rises, which in turn leads to higher aggregate capital stock and even better knowledge diffusion next period etc. In the limit the economy converges towards full integration and completely accessible knowledge.

In order to motivate poor growth performance over a very long stretch of time the model does not rely on *literal* stagnation. In the diagram on the right hand side of Figure 2, very low growth occurs when the $f(k)$ curve is close to but still above the identity line. The fact that economies can spend millennia in the funnel between $f(k)$ and the identity line makes growth at glacier speed observationally equivalent to actual stagnation. Qualitatively, however, it makes a big difference whether an economy stagnates at k_1^* on the left hand side in Figure 2 or whether it develops very slowly through the funnel on the right hand side of Figure 2. In the latter case the economy develops endogenously such that sooner or later growth at a positive rate becomes visible. Since it was invisible before, one may speak of an *Industrial Revolution.* From then on the growth rate of income is visibly increasing over time and approaches a high constant level.

3.6. **Productivity Slowdown.** Along the transition towards balanced growth, capital and income per capita are growing at a monotonously increasing rate. Productivity growth, however, may adjust non-monotonously in an inverted u-shaped way, implying that the economy experiences a *productivity slowdown* along the transition. For this phenomenon I cannot provide a proof but some economic intuition. From (2) we obtain TFP growth as follows.

$$\gamma_{A_t} \equiv \frac{A_{t+1} - A_t}{A_t} = \frac{(k_{t+1})^{\sigma_{t+1}}}{(k_t)^{\sigma_t}} - 1.$$

Since σ is small in comparison with k (which goes to infinity), we set $\sigma_{t+1} \approx \sigma_t$ such that $\gamma_{A_t} \approx (k_{t+1}/k_t)^{\sigma_t} - 1$. Inserting (4) we obtain the the following expression.

$$\gamma_{A_t} \approx a^{\sigma_t} \cdot k_t^{\sigma_t(1-\alpha)(\sigma_t - 1)} - 1.$$

We see that the rate of TFP growth approaches the growth rate of income, $a - 1$, when the degree of knowledge diffusion approaches unity. Along the way, the first term rises monotonously and approaches a as σ_t approaches unity. The k_t term, however, has a non-monotonous, hump-shaped, influence on productivity growth. The exponent vanishes for $\sigma_t = 0$ and for $\sigma_t = 1$ and assumes a maximum for $\sigma_t = 1/2$. Accordingly the k_t term has no influence on TFP growth for $k_t = 0$ and for $k_t \to \infty$. In between, for an intermediate value of k_t the term reaches a maximum. This effect dominates TFP growth if k_t is sufficiently large at the time when the exponent is largest.

Intuitively, TFP growth is not maximal when the level of capital is highest but when the *momentum* of capital-induced improvement of knowledge diffusion is largest. Historically, we probably associate the highest momentum of the improvement of knowledge diffusion with the onset of the IT age in the 1970s. The present model can then explain why this event is – seemingly counterfactually – associated with a productivity slowdown. According to this view, there is nothing alarming or frightening about the productivity slowdown. With increasing speed of knowledge diffusion TFP growth rates were "just" overshooting and are subsequently converging towards the balanced growth level from above.

4. A Network Theoretic Foundation of Knowledge Diffusion

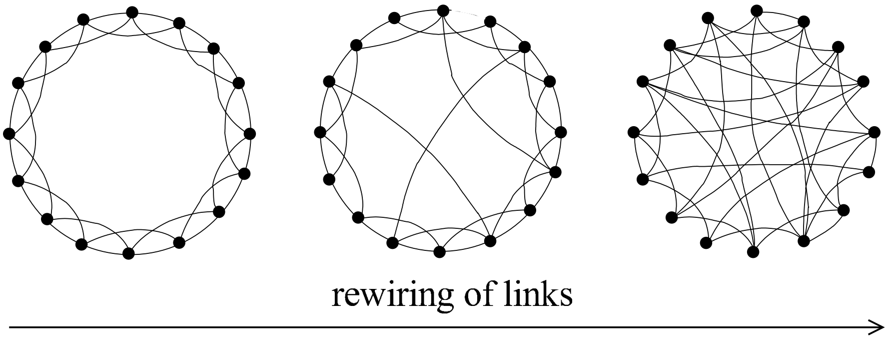
This section illustrates how the knowledge diffusion function $\sigma(k)$ can be derived from an evolving network, the Small World model. The Small World model was originally developed by Watts and Strogatz (1998) and has since been empirically verified for a plethora of biological and social phenomena.[6] The Small World model is particularly suited for the analysis of economic integration because a measure of localization, the clustering coefficient, can be calculated analytically. This way, the impact of network evolution on knowledge diffusion can be conveniently analyzed without any concrete specification of the network itself.

Consider a network in which vertices (nodes) are persons and edges (links) indicate whether any two persons know each other. A stylized version of such a network, a network on a one-dimensional lattice, is shown in Figure 3. The Small World model interpolates smoothly between a regular network, i.e. a network with the highest degree of local connections (left hand side of Figure 3), and a random network, i.e. a network with the highest degree of global connections. This is done by rewiring edges with other edges that are randomly chosen from the net. The process of network evolution and increasing globalization is approximated by continuously increasing from zero to unity the fraction p of rewired edges (movement from the left to the right in Figure 3).

The clustering coefficient is an indicator for the connectivity of local communities. It gives the average probability that two persons which are connected to any person are also directly connected with each other, that is the probability that someone's friends (trading partners) are also friends (trading partners) with each other. A high clustering coefficient indicates a collectivistic society,

[6]See Newmann (2003), for an overview. Serrano and Beguna (2003) show that the world trade trade-web is well represented by the Small World network characteristics

Figure 3: Small World Model



Left hand side: one-dimensional lattice with connections between all vertex pairs by two or fewer lattice spaces away. A Small World is created by randomly rewiring edges.

in which information is exchanged mainly between neighbors, whereas a low clustering coefficient indicates an individualistic society, with long-distance information exchange. Based on insight from other research we assume that the degree of knowledge diffusion σ is high when the clustering coefficient is low.[7]

Generally, the clustering coefficient depends on network size and the number of links per node. Normalizing the clustering coefficient by measuring it relative to the highest possible clustering, however, eliminates all network specifics and yields the clustering coefficient $\tilde{c}$ as a simple function of the fraction of long distance links, $\tilde{c}(p) = (1-p)^3$, see Newman (2003). The inverse relation between the degree of knowledge diffusion and the clustering coefficient can thus be written as $\sigma = 1 - \tilde{c}(p) = 1 - (1-p)^3$. In contrast to Watts and Strogatz' original approach, in which nature randomly created long-distance links, we assume that the formation of long-distance link depends on economic activity. Specifically, it seems reasonable to assume that the relative number of long-distance links depends positively on aggregate capital stock because capital consists partly of effective distance reducing devices (like horses, ships, cars, or airports). A larger capital stock thus increases the probability that a particular long-distance link between two nodes exists. Because $p \in (0,1)$, the most parsimonious and general formal notation of this notion is given by (6).

$$p = \frac{g(k)}{\omega + g(k)}, \;\; g(0) = 0, \;\; g'(k) > 0. \tag{6}$$

This generates a monotonous mapping from k to $p \in (0,1)$ in which the constant ω controls for

[7] Hofstede (2001) and Fogli and Feldkamp (2011). This notion is consistent with the finding of predominantly local knowledge diffusion in LDCs (Foster and Rosenzweig, 1995) and global knowledge diffusion in fully developed countries (Irwin and Klenow, 1994).

capital-independent factors (for example, spatial size of the economy and other geographic factors). Inserting the probability of a long-distance link into the formula for the clustering coefficient and then into the degree of knowledge diffusion we get

$$\sigma(k) = 1 - \left(\frac{\omega}{\omega + g(k)}\right)^3. \tag{7}$$

The final step is to check whether the network-based notion of knowledge diffusion supports the proposed theory of knowledge and growth, that is to check whether it is consistent with Assumption 3. Inserting $\sigma(k)$ and its derivative into $b(k)$ and simplifying we obtain

$$b(k) = \left(\frac{\omega}{\omega + g(k)}\right)^3 \cdot \left(\frac{3g'(k)}{\omega + g(k)} k \log k - 1\right).$$

For Assumption 3 to be fulfilled, $b(k)$ has to change its sign exactly once. Since the first term is always positive and the second term is negative for $k < 1$, it is sufficient to show that the term

$$\frac{g'(k)k}{\omega + g(k)} \log k$$

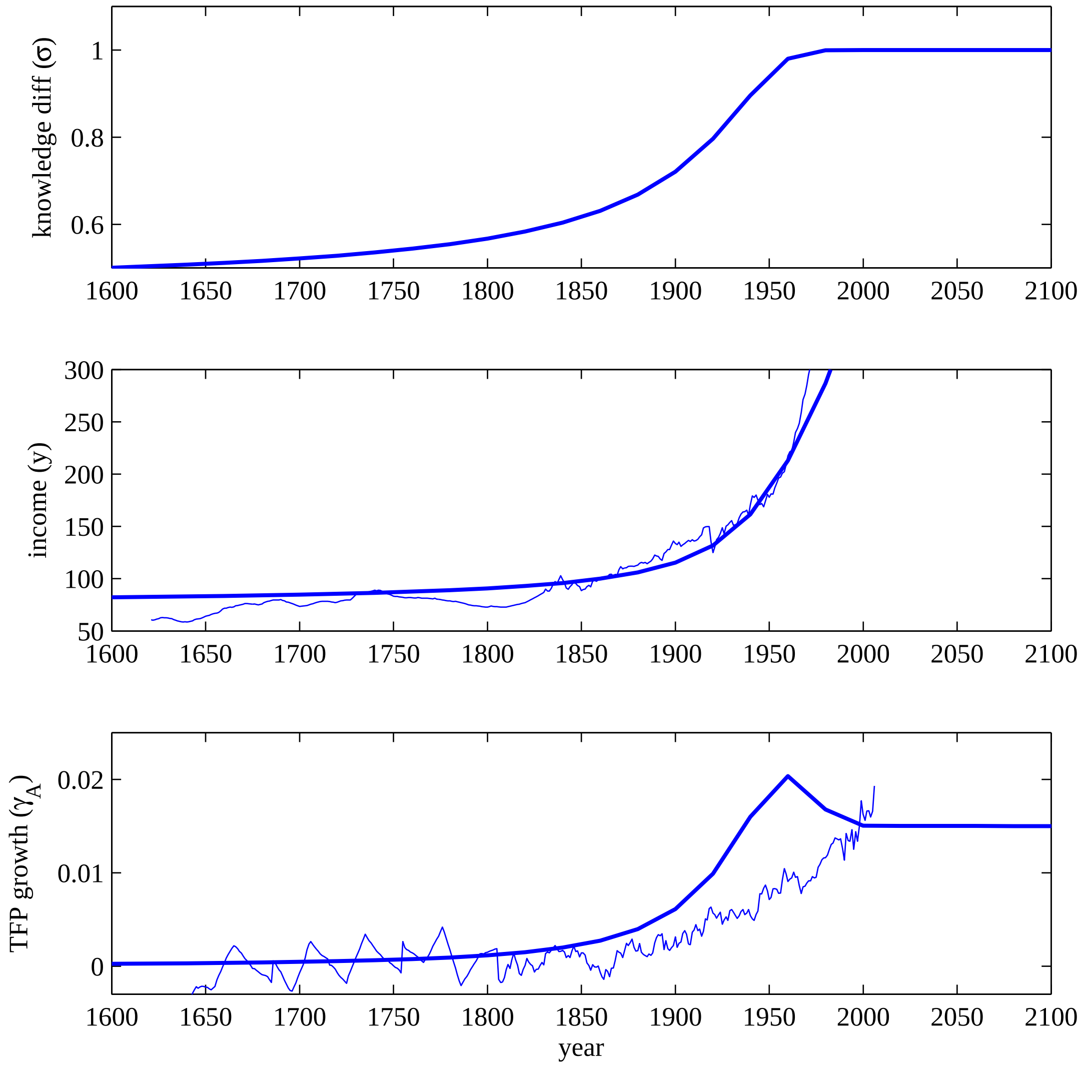
rises monotonously for $k \geq 1$. This is a very mild requirement. For example, it is straightforward to show that it is fulfilled for linear ($g(k) = Bk$), exponential ($g(k) = \exp(k)$) and iso-elastic ($g(k) = k^\theta, \theta > 0$) "production functions". It is actually hard to come up with a function $g(k)$ not fulfilling Assumption 3. The theory is thus largely independent from structure. The essential assumption is that more capital is good for long-distance link formation, $g'(k) > 0$. This seems to be a relatively mild constraint and intuitively plausible.

5. Long-Run Adjustment Dynamics: Numerical Exercise

We next check how well the model explains the actually observed transition to modern growth. Given that the model in its reduced form consists only of one equation and neglects several important aspects of the historical transition to balanced growth (as, for example, the demographic transition, the transition towards mass education, structural change, the emergence of market R&D), it is clear that we should not expect a perfect fit with the historical data. The appropriate yardstick here is whether the model improves quantitatively the adjustment dynamics suggested by its competitors, the neoclassical growth model and the Ak growth model.

We begin with setting $\alpha = 0.33$ as in many related studies. We then fix $\bar{A}$ and β such that along the balanced growth path the savings rate (investment rate) is 0.25 and income per capita grows at 1.5 percent annually. This implies $a = 1.347$. We use the network-based foundation of knowledge diffusion (7) and set $g(k) = \exp(k)$. We set the initial capital stock in year 0 to 1 unit and calibrate the remaining parameter ω such that the industrial revolution kicks off in the 18th century. Specifically we try to match the actual evolution of income and productivity in England 1600-2000. After solving the model numerically we convert periods measured in generations into periods measured in years by assuming that a generation takes 20 years. This alleviates the comparison with real data.

Figure 4 Adjustment Dynamics – The Take-off to Modern Growth



Parameters: $\alpha = 0.33$, $\beta = 0.25$, $a = 1.35$, $\omega = 42.4$, $k_0 = 1$ at year 0. Income has been normalized such that income $y_{1860} = 100$.

The predicted adjustment dynamics for knowledge diffusion, income, and TFP growth are shown by wide lines in Figure 4. The adjustment process is very gradual. For better visibility the period from year 1 to year 1600 is not shown in the figure. Basically, it looks very much like the time between 1500 and 1600. Although the middle ages seemingly look like a world of stasis, this is in fact not the case. To begin with, there is a lot of learning-by-doing going on; the degree of knowledge diffusion is around 0.5. This allows income levels larger than predicted by the neoclassical model without learning-by-doing. Secondly, the economy is effectively growing. The parameter values do not support a poverty trap. Growth, however, is exceedingly slow such that it looks like stagnation for the uninformed spectator.

Because income levels are so low from modern perspective the uninformed spectator could also have difficulties in distinguishing the growing economy from one resting at the poverty trap. Only when the growing economy takes off to the Industrial Revolution a stagnating economy is visibly left behind and we observe what has been called the Great Divergence (Pomeranz, 2000).

Income and TFP growth become first visible around the late 18th century. Until the end of 20th century growth rates are sharply on the rise and reach levels unseen so far in history. While income growth adjusts monotonously towards its balanced growth value, TFP growth exhibits a maximum.

In order to better compare with the real economic evolution I take the time series for real GDP, employment, and TFP as compiled by Madsen at al. (2010) for England for the years 1620-2006. From the data for GDP and employment I compute GDP per worker, which is shown by thin lines in the middle panel of Figure 4. For better comparability I have normalized the data such that in 1860, at the onset of the second industrial revolution, GDP per worker in the data and predicted y by the model equal 100. Given that the model provides only one degree of freedom to control for adjustment dynamics (ω), the fit with the real time series is remarkably well. The model overestimates income before the industrial revolution somewhat and underestimates it somewhat at the end of the 20th century. But the crucial era of industrialization and economic take-off is approximated quite well.

The lower panel of Figure 4 compares the model's prediction of TFP growth with actual data for England. The TFP time series has been computed by Madsen et al. (2010) for England on an annual basis for the years 1620-2006. Naturally, because TFP is computed as a residual, there is huge fluctuation in annual TFP growth. In order to identify the trend I have computed 50-year moving averages from the Madsen et al. data. The obtained time series is shown by the thin line in Figure 4. The model gets the gradual long-run evolution of TFP growth about right. It only predicts a somewhat "too early" take-off of TFP growth. This (mis-) prediction is in line with observations from economic historians who emphasize that – from today's perspective – TFP growth contributed relatively little to early industrialization in England (see Crafts and Harley, 1992, and the literature cited in footnote 1). Another interesting aspect is that the model predicts a productivity slowdown after the 1960s, a phenomenon debated in the literature (e.g. Crafts, 2003) but not visible in Madsen et al.'s (smoothed) TFP times series.

6. A Two-Region Model of Knowledge Diffusion

6.1. **Model Setup.** The closed economy model of knowledge diffusion and growth displayed multiply equilibria and suggested that low capital productivity associated low capital stock would prevent capital to flow from rich to poor countries. This view of a completely dichotomized world, however, was derived from the strong assumption of absent international knowledge flows. The assumption of isolation from international knowledge may be a good approximation for some traditional societies, but in general it seems to be more reasonable to allow (some) knowledge to flow between countries or regions.

The here proposed model is a useful tool to analyze international knowledge diffusion because it explicitly distinguishes between *existing* and *accessible* knowledge. In particular is seems reasonable to continue the Arrow-Romer argument and assume that world-wide existing knowledge is a function of the world-wide existing capital stock. The knowledge accessible by a single firm, however, is a local variable; it is a function of region-specific economic integration, which in turn depends on the region-specific capital stock.[8] Knowledge diffuses less easily through a region with little capital endowment and any firm of that region has inferior access to the world-wide existing knowledge.

To simplify the exposition we assume that the world consists of two regions, A and B. A region is defined as a set of households and firms facing the same fundamental parameters capturing institutions and preferences. Formalizing the ideas from above, productivity of a typical firm in region A and B is given by (8).

$$A_t^A(i) = \bar{A} \cdot \left(k_t^A + k_t^B\right)^{\sigma(k_t^A)}, \qquad A_t^B(i) = \bar{B} \cdot \left(k_t^A + k_t^B\right)^{\sigma(k_t^B)} \tag{8}$$

where $\bar{A}$ and $\bar{B}$ capture the influence of knowledge-flow promoting institutions in region A and B. World capital stock, $k_t^A + k_t^B$, approximates the existing knowledge whereas region specific capital

[8]There are two alternative assumptions possible. Both do not lead to an interesting reformulation of the original problem. If accessible knowledge were also a function of world-wide capital, the model would boil down to the one-region model discussed earlier. If both existing knowledge and accessible knowledge were just a function of regional capital, we would have two isolated economies without interaction, which, separately, could be again analyzed within the available one-region framework.

stocks (k_t^A or k_t^B) determine the degree of knowledge diffusion within the region, such that A_t^A and A_t^B reflect the accessible knowledge in region A and B. Each region is populated by a measure one of firms such that in equilibrium aggregate productivity $A_t^A = A_t^A(i)$ and likewise for region B.

We assume that both regions are of the same size. As before, population in a region is given by a mass of unity of young and old adults. For simplicity we assume that both regions share the same α. Aggregating across firms we get regional GDP $y_t^A = \left(A_t^A\right)^{1-\alpha}\left(k_t^A\right)^{\alpha}$ and $y_t^B = \left(A_t^B\right)^{1-\alpha}\left(k_t^B\right)^{\alpha}$. As before the young save for consumption in old age. Let β^A and β^B denote the region-specific weights on old age-consumption in utility (capturing, for example, the region-specific disease environment). Individual maximization then yields region-specific aggregate savings.

$$s_t^A = \frac{\beta^A}{1+\beta^A}(1-\alpha)\left(A_t^A\right)^{1-\alpha}\left(k_t^A\right)^{\alpha}, \qquad s_t^B = \frac{\beta^B}{1+\beta^B}(1-\alpha)\left(A_t^B\right)^{1-\alpha}\left(k_t^B\right)^{\alpha}. \tag{9}$$

In order to avoid case differentiation about international capital mobility we assume that at least some capital is flowing internationally such that $r_t^A = \eta r_t^B$. Here, η is an institutional parameter capturing regional-specific property rights (or, more generally regional-specific institutional barriers to investment and international capital movements). If property rights are secure everywhere and capital is fully mobile, $\eta = 1$. If property rights are less well respected in region A, $\eta > 1$, reflecting a risk premium (exceeding the risk premium in region B). Despite this relatively crude treatment of capital flows, the model is capable to display very rich and plausible dynamics of regional capital and productivity growth.[9]

Inserting the regional rates of return, $r_t^A = \alpha(A_t^A)^{1-\alpha}(k_t^A)^{\alpha-1}$ and $r_t^B = \alpha(A_t^B)^{1-\alpha}(k_t^B)^{\alpha-1}$ into the interest-parity, $r_t^A = \eta r_t^B$ we get the correlation between regional capital stocks.

$$k_t^B = \eta^{1-\alpha}\frac{A_t^B}{A_t^A}\cdot k_t^A. \tag{10}$$

Naturally, there is relatively more capital allocated to region B if in region B property rights are relatively well respected and TFP is relatively high. Using the definition $k_t = k_t^A + k_t^B$ we can express the regional capital stocks as functions of the world capital stock and regional TFP:

$$k_t^A = \left(1+\eta^{1-\alpha}\frac{A_t^B}{A_t^A}\right)^{-1}\cdot k_t, \qquad k_t^B = \left(1+\eta^{\alpha-1}\frac{A_t^A}{A_t^B}\right)^{-1}\cdot k_t. \tag{11}$$

The model is closed by the fact that regional-specific savings of the young generations must add up

[9]Caselli and Feyrer (2007) have shown that there exist relatively little variation of the marginal product of capital across countries. In the context of the present model this would imply an η close to unity.

to next period's world capital stock.

$$k_{t+1} = s_t^A + s_t^B. \tag{12}$$

6.2. **Steady-State.** By focussing on regional capital stocks, i.e. by focussing on equation (10) it seems to be possible that the two regions behave qualitatively differently at the steady-state. Seemingly, the Great Divergence could continue forever if one country, say A, grows perpetually (k_t^A and A_t^A grow in sync) while the other region stagnates at the poverty trap identified in Section 3 (k_t^B and A_t^B are constant). This view, however, is ill-informed. It ignores the power of international knowledge diffusion.

Proposition 3. *At a steady-state it is impossible that one regions grows at a positive rate while the other region stagnates in poverty.*

The proof by contradiction is straightforward. Assume that one region, say A, is growing at constant rate, while the other region, B, stagnates. Since k^A is perpetually growing and k^B stagnating, world capital stock k is growing at the same rate as k^A, implying that k_t^A/k_t is constant along the steady-state. From (11) this implies that A^B/A^A stays constant along the steady-state. But A^A is growing because k^A is growing, which implies that A^B is growing. But then k^B is growing, a contradiction to the initial assumption that k^B is constant.

This leaves open two long-run possibilities, world-wide stagnation and world-wide growth. Henceforth we focus on the case of growth. If there is growth, $\sigma(k_t^A) \to \infty$ and $\sigma(k_t^B) \to \infty$ for $t \to \infty$ and thus from (8) along the steady-state $\lim_{t\to\infty}(A_{t+1}^B/A_{t+1}^A) = (\bar{B}k_t)/(\bar{A}k_t) = \bar{B}/\bar{A}$. Inserting this into (11) and applying $\lim_{t\to\infty}\sigma(k_t^A) = 1$, we see that, along the steady-state, capital and productivity in country A are growing in proportion to the world capital stock.

$$k_t = \left(1 + \eta^{1-\alpha}\frac{\bar{B}}{\bar{A}}\right) k_t^A \qquad \Rightarrow \qquad A_t^A = \bar{A}\left(1 + \eta^{1-\alpha}\frac{\bar{B}}{\bar{A}}\right) k_t^A. \tag{13}$$

Inserting this information into (9) we get (14a). Proceeding analogously for region B we get (14b).

$$s_t^A = (1-\alpha)\bar{A}^{1-\alpha}\left(1 + \eta^{1-\alpha}\frac{\bar{B}}{\bar{A}}\right)^{-\alpha} \frac{\beta^A}{1+\beta^A} \cdot k_t \tag{14a}$$

$$s_t^B = (1-\alpha)\bar{B}^{1-\alpha}\left(1 + \eta^{\alpha-1}\frac{\bar{A}}{\bar{B}}\right)^{-\alpha} \frac{\beta^A}{1+\beta^A} \cdot k_t. \tag{14b}$$

Having obtained the region-specific savings rates as a function world-wide capital stock it is straightforward to obtain world-wide growth. Inserting (14) into (12) and solving for $g \equiv k_{t+1}/k_t - 1$ we get the rate of economic growth for the world.

$$g = (1-\alpha)\left[\frac{\beta^A}{1+\beta^A}\bar{A}^{1-\alpha}\left(1+\eta^{1-\alpha}\frac{\bar{B}}{\bar{A}}\right)^{-\alpha} + \frac{\beta^B}{1+\beta^B}\bar{B}^{1-\alpha}\left(1+\eta^{\alpha-1}\frac{\bar{A}}{\bar{B}}\right)^{-\alpha}\right] - 1.$$

From (13) and Proposition 3 it is then obvious that both regions grow also at rate g.

If both regions are symmetric, i.e. $\beta^A = \beta^B = \beta$ and $\bar{A} = \bar{B}$ world growth simplifies to $g = 2^{1-\alpha}a - 1$. This result is very intuitive. In Section 3 we have obtained the growth rate in isolation as $g = a - 1$. Thus there is a scale effect from economic integration. Integrating two symmetric regions, however, does not double economic growth because the regions share "only" knowledge, i.e. non-physical ideas whereas physical capital remains to be rivalrous and unshared. Recalling that $1-\alpha$ is the share of knowledge in production it is clear that by doubling the number of countries (regions) world growth increases by factor $2^{1-\alpha}$. By calculating the comparative statics for the general case we get the following result.

Proposition 4. *A higher weight on future consumption increases world-wide growth (that is $\partial g/\partial \beta^A > 0$, $\partial g/\partial \beta^B > 0$). An improvement of regional institutions ($\bar{A}$ or $\bar{B}$ or η) does not generally increase world-wide growth.*

To see why improving institutions can deteriorate growth, assume an extreme case of $\beta^A = 0$ and take the derivative of growth with respect to $\bar{A}$ and obtain $\partial g/\partial \bar{A} < 0$. In this extreme case only the citizens of B are saving. If knowledge-flow improving institutions in region A improve, savings are withdrawn from B to A. As a consequence productivity and income of the young (worker-) generation in A rises while it falls in B. However, the young in A are not saving while for the young in B, who save at a positive rate, income is lower. Facing lower income, they save less and, consequently, world capital accumulation and thus economic growth deteriorates.

A similar argument holds for property rights (or institutional barriers to investment). Taking the derivative with of g with respect to η we see that $\partial g/\partial \eta > 0$ if and only if

$$\eta^{2(1-\alpha)} > \left(\frac{\bar{A}}{\bar{B}}\right)^{1+\alpha}\left(\frac{\beta^A}{1+\beta^A}\right)\left(\frac{\beta^B}{1+\beta^B}\right)^{-1}.$$

This means that if knowledge-promoting institutions and weight on future consumption in region A are sufficiently low compared to B, then increasing distortions on investment in A effectively raise world economic growth and growth in both regions. Increasing distortions prevent the young generation in A and B from investing in a region where a relatively small part of the fruits of investment is saved as capital stock available for the next generation.

In contrast, a higher weight on future consumption (capturing for example an improvement of the disease environment or higher life-expectancy through medical technological progress) always increases world-wide growth. While the simple model is certainly too crude to elicit explicit policy conclusions it highlights a so far overlooked channel in the geography vs. institutions debate and lets it appear less obvious how "institutions rule" long-run economic behavior from a world-wide perspective.

6.3. **West – East Adjustment Dynamics.** In an integrated world differences in institutional parameters are no longer reflected in multiple equilibria but in a differentiated take-off to growth. Naturally, countries or regions with more favorable conditions to growth will take off first thereby generating a *temporary* divergence of regional income levels and growth rates. Later on, when sufficient knowledge diffuses into the world, the gap in growth rates is closed. As mentioned in the Introduction this part of the analysis relates to the rich literature on catch-up growth. In contrast to that literature (which mostly keeps growth of the world technological frontier exogenous and constant) we focus here particular on the gradual take-off of the technological frontier.[10] Besides clarifying the robustness of the results obtained so far we get also interesting insights about the interaction between forerunners, followers, and trailers of the Industrial Revolution.

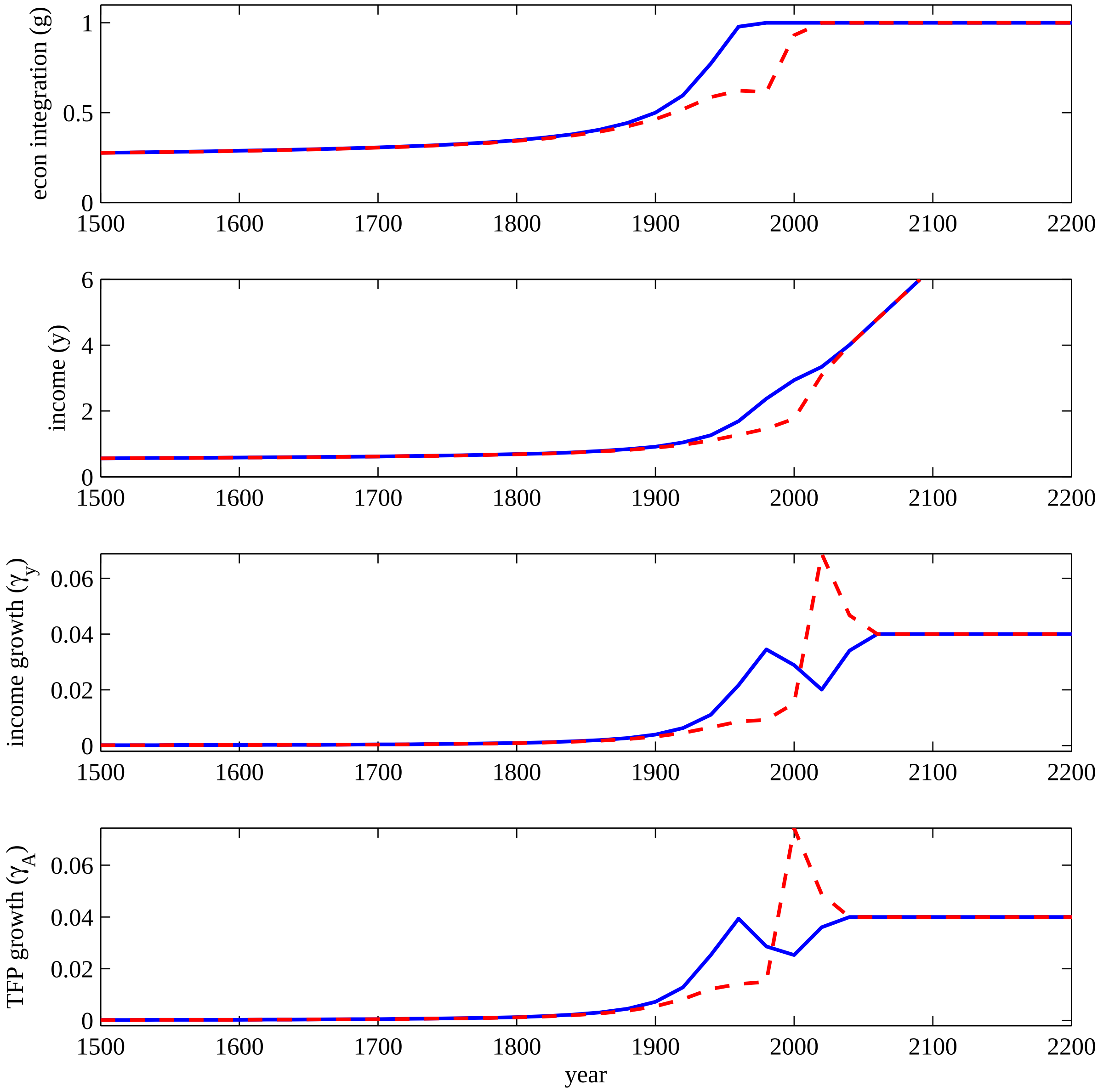
For the numerical exercise we assume that both regions deviate only in their knowledge-flow promoting institutions. They share the same weight on future consumption and η is set to unity. Keeping the analogy from the simple model we conceptualize the forerunners of the Industrial Revolution as the "West" and imagine – for the current experiment – the followers as the "East". We maintain the assumption of a savings rate of 0.25 and fix the remaining parameters, $\bar{A}$, $\bar{B}$, and ω such that

- the West takes off in the late 18th century
- during the 20th century the West grows by factor 6.5, i.e. experiences an average annual growth rate of 1.89 percent (which is the average growth rate of Maddison's 12 Western Eurpean countries).
- the West experiences a productivity slowdown in the second half of the 20th century.

[10]One notable exception is Howitt (2000) where the world technology frontier is also endogenous. But there the focus is on the speed of convergence of contemporary countries and not on the gradual "take off of the frontier".

- income growth in the East begins to catch up with the West about in 1980.

Figure 5: West–East Adjustment Dynamics



Parameters: $\alpha = 0.33$, $\beta = 0.25$, $\omega = 106.4$, $k_0 = 1$ for both countries. $\bar{A} = 32.380$, $\bar{B} = 32.374$.

Figure 5 shows the obtained time paths. Blue solid lines identify the West and red dashed lines the East. As before, the take-off to growth is gradual. This is true for the forerunner region and the follower region of the Industrial Revolution. During the take-off phase of the West there is also income growth in the East but the better integrated West displays temporarily higher productivity and attracts more investment such that the East falls behind.

At the time when growth loses momentum in the West because firms are almost completely integrated and knowledge diffusion is close to its maximum degree, the East takes off. Now the East experiences a phase of temporarily increasing returns to scale (convex shape of $f(k)$) and TFP growth surpasses that of the West in the 1970s. Consequently the East attracts capital from the West, which has the effects of (i) amplifying growth of income and TFP further in the East and (ii) slowing down growth of TFP and income in the West. As a consequence the productivity slowdown in the West is more pronounced than predicted by the closed-economy model of Section 3. The temporarily higher growth rate in the East and the slowdown of growth in the West work together to allow for an almost complete catch up of income levels until about the year 2030.[11]

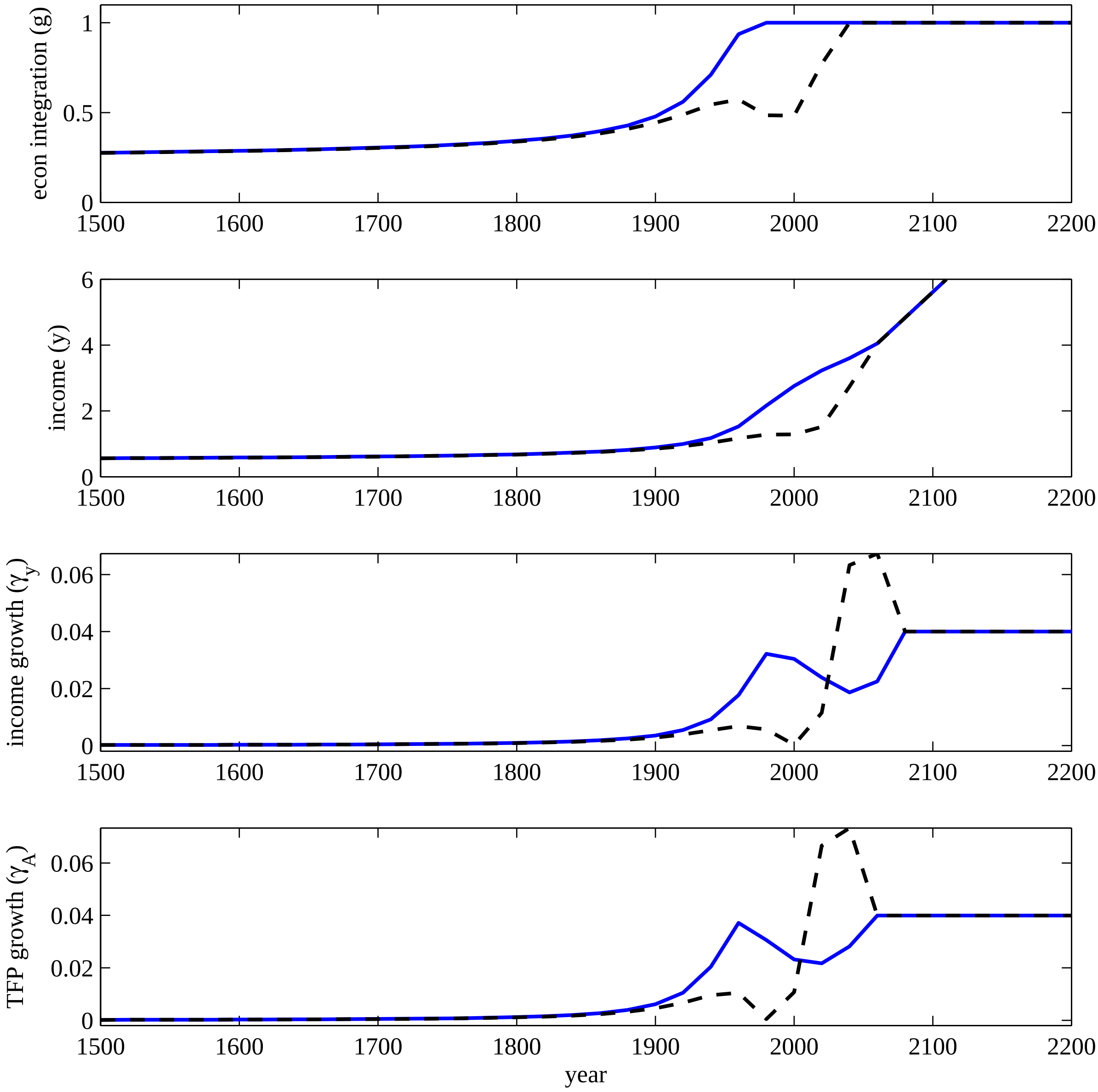
6.4. **North – South Adjustment Dynamics.** The previous example has shown a case of *relative* divergence during the take-off phase: Both regions start to grow but the forerunner grows much faster thereby generating an income gap. But the model is also capable of generating *absolute* divergence in the sense of temporarily deteriorating income levels for a latecomer of the Industrial Revolution. Figure 6 shows such a case. Here, we kept all parameters from the previous case but reduced $\bar{B}$ further. As a consequence, take off of the follower country is further delayed. Keeping the regional analogy, this case may be called North-South adjustment dynamics.

The main difference to the West-East scenario is that during the take-off phase of the North (blue solid lines) the South (black dashed lines) remains in its neoclassical domain of production. Diagrammatically, capital stock in the South remains in the concave part of the $f(k)$ curve. For a while it looks like as if income and TFP would be stagnating in the South and then, in the second half of the 20th century when growth in the North reaches unprecedented levels, income in the South deteriorates.

The reasons for divergence are capital flight and deteriorating capital stock. It is worthwhile for South to allocate capital to the North because the North is already in its modern phase where production is characterized by increasing returns through perpetually improving knowledge diffusion whereas the less well integrated South is still stuck in the neoclassical phase. As a consequence of capital flight, the South experiences a period of absolute decline in income and productivity and, as a consequence, economic integration declines as well (because, for example, the capital stock of railroads is not completely maintained from one generation to the next). Ultimately, however, even the inferior conditions for knowledge diffusion cannot prevent that sufficient knowledge reaches the South such that the region escapes from the neoclassical domain and starts growing and catching

[11]See Phillips and Sul (2009) for evidence on the gradual take-off of followers of the Industrial Revolution.

Figure 6: North-South Adjustment Dynamics



Parameters: $\alpha = 0.33$, $\beta = 0.25$, $\omega = 106.4$, $k_0 = 1$ for both countries. $\bar{A} = 32.380$, $\bar{B} = 32.372$.

up with the North.

Once the take-off has been initiated catch-up is much faster than in the West-East case, visible in Figure 6 by the high growth rates of TFP and income predicted for the South during the 21th century. The reason is that – compared to the West-East case – the forerunner region has experienced a much longer phase of high growth before the take-off of the latecomer region. Consequently, the North is much richer and more capital flows from the forerunner to the latecomer during the catch-up phase.

7. Final Remarks

This paper has proposed a feedback mechanism between knowledge diffusion and capital accumulation. The feedback mechanism explains why the take-off to growth has to be gradual. Capital embodied technological progress (incorporated, for example, in ships, trains, and planes) alleviates the travel of people and ideas. More capital accumulation leads to better diffusion of knowledge, which raises factor productivity, which in turn leads to even more accumulation and better diffusion of knowledge etc. Unlike growth, the process of improving knowledge diffusion has a certain end, the fully integrated economy, a fact that generates convergence towards balanced growth.

An extension towards a two-region world economy has shown the robustness of the gradual take-off with respect to leaders and followers of the Industrial Revolution. Unfavorable conditions, which would have caused stagnation in poverty within the closed-economy framework, are causing a delay of the take-off to modern growth within the world-economy framework. The interaction between knowledge flows and capital flows was also helpful to explain why economic integration of markets leads to higher growth, how a phase of absolute divergence between leader and follower region may emergence and why catch-up growth will be higher for countries that take off later.

There are several extension possible. Two of these, market R&D and income dependent savings rates, I have addressed in Strulik (2009). When the feedback effect between accumulation and knowledge diffusion is integrated into a model with market R&D it can been explained why R&D effort, TFP growth, and income growth are jointly rising during the Industrial Revolution. A combination between the learning-by-doing setup and the market R&D setup can explain how a long phase of growth driven exclusively by learning-by-doing eventually triggers a transition towards market R&D activities. Within such a double TFP-driven growth model it is also possible to abandon the assumption of complete knowledge diffusion through learning-by-doing ($\lim_{k\to\infty} \sigma(k) = 1$ in the current setup). Intuitively, spillovers in learning-by-doing have only be strong enough to generate a certain *level* of TFP from which onwards market-based R&D becomes profitable.

When the model is extended toward non-homothetic utility such that the savings rate increases with income, there exists a phase during which income growth and knowledge diffusion are further amplified through increasing investment rates such that overshooting behavior occurs. This mechanism is helpful to explain a phase of extraordinary high income and TFP growth (the Roaring Twenties). As shown by Steger (2000), Carroll et al. (2000), and Strulik (2010) increasing savings rates are also a stand-alone mechanism that generates gradually increasing income during the transition to balanced growth. But, of course, the savings-rate approach cannot motivate the gradual rise of TFP growth observable in association with the take-off to modern growth.

Recently, Jones and Romer (2010) have set up the "new Kaldor facts", i.e. stylized facts about growth that cannot be addressed within the neo-classical growth paradigm. Here, I have shown that a simple extension of the neo-classical growth model can actually address the following facts from their list:

- increasing extent of the market (increasing flow of ideas)
- accelerating growth
- variation in modern growth rates (depending from distance to the technological frontier)
- large income and TFP differences across countries.

But these are only 4 out of the 6-item list. The simple model does not speak to the evolution of human capital and to labor income differentials. In order to show the explanatory power of the diffusion-accumulation feedback as a stand-alone mechanism of growth over the very long run, it has not been unified with a theory of fertility and education. This remains a challenging task for future research.

Appendix

Proof or Lemma 1. Denote by $\bar{k}$ the capital stock where the term $\equiv \sigma'(k) \cdot k \cdot \log(k) + \sigma(k) - 1$ changes its sign. Obtain $f(k)/k = ak^{(1-\alpha)(\sigma(k)-1)}$ from (4). This expression has an extremum where

$$\left(\frac{f(k)}{k}\right)' = ak^{(1-\alpha)(\sigma(k)-1)} \left[\sigma'(k)\log(k) + \frac{(1-\alpha)(\sigma(k)-1)}{k}\right]$$
$$= \frac{1-\alpha}{k} f(k) \left[\sigma'(k) k \log(k) + \sigma(k) - 1\right] = 0$$

i.e. where $k = \bar{k}$. Observe that the right hand side is negative for $k < \bar{k}$ and positive for $k > \bar{k}$. Thus $f(k)/k$ is monotonously falling for $k > \bar{k}$, monotonously increasing for $k > \bar{k}$, and assumes a global minimum at $\bar{k}$. Let $\gamma \equiv f(\bar{k})/\bar{k}$ and thus $f(k) > \gamma k$ for $k \neq \gamma$. Compute

$$f'(k) = ak^{\alpha+(1-\alpha)\sigma(k)} \left[\sigma'(k)\log(k) + \frac{\alpha + (1-\alpha)\sigma(k))}{k}\right]$$
$$= f(k) \left[\sigma'(k)\log(k) + \frac{\alpha + (1-\alpha)\sigma(k))}{k}\right].$$

And thus at $\bar{k}$

$$f'(\bar{k}) = \gamma\bar{k} \left[\sigma'(\bar{k})\log(\bar{k}) + \frac{\alpha + (1-\alpha)\sigma(\bar{k})}{\bar{k}}\right] = \gamma \left[(1-\alpha)\left\{\sigma'(\bar{k})\bar{k}\log(\bar{k}) + \sigma(\bar{k})\right\} + \alpha\right].$$

Recall that $\sigma'(\bar{k})\bar{k}\log(\bar{k}) + \sigma(\bar{k}) = 1$ to conclude $f'(\bar{k}) = \gamma$. q.e.d.

Proof of Proposition 1. Recall from the proof of Lemma 1 that $f(k)/k$ assumes a global minimum at $\bar{k}$. This implies that for $\gamma < 1$ we have $f(\bar{k}) < \bar{k}$, implying that $f(k) < k$ in a neighborhood of $\bar{k}$. For small k, $0 < k < 1$, $f(k) > k$ whereas for large k, $k \to \infty$, $f(k)$ approaches asymptotically $ak > k$. Applying the Intermediate Value Theorem for continuous functions, we know that $f(k)/k = 1$ at least once in the interval $(0, \bar{k})$. And since $f(k)/k$ is monotonously falling in $(0, \bar{k})$, we know that $f(k)/k = 1$ at most once in $(0, \bar{k})$. Thus there exists exactly one equilibrium in $(0, \bar{k})$. We call the associated capital stock k_1^*. Analogous reasoning verifies that there exists exactly one equilibrium in in $(\bar{k}, \infty)$, we call the associated capital stock k_2^*, $k_1^* < k_2^*$.

Since $f(k)/k$ is falling in $(0, \bar{k})$, we have $k_{t+1} = f(k_t) > k_t$ for $k_t < k_1^*$ and $k_{t+1} = f(k_t) < k_t$ for $k_2^* > k_t > k_1^*$. Thus the equilibrium at k_1^* is locally stable. Analogous reasoning verifies that the equilibrium at k_2 is unstable.

If $\gamma > 1$, $f(k)/k > 1$ everywhere and there exists no fixed point. In that case, and for $k > k_2^*$

if $\gamma < 1$, the capital stock is perpetually growing. For $k \to \infty$, $\sigma \to 1$ and $f(k) \to ak$. Thus the economy approaches a balanced growth path where $k_{t+1}/k_t - 1 = a - 1$. Along the transition the growth rate is perpetually rising since $f(k)/k$ is monotonously increasing in k for $k > \bar{k}$.

References

Andersen, T.B. and Dalgaard, C.-J., 2011, Flows of People, Flows of Ideas, and the Inequality of Nations, *Journal of Economic Growth* 16, 1-32.

Antras, P. and H.J. Voth, 2003, Factor prices and productivity growth during the English Industrial Revolution, *Explorations in Economic History* 40, 52-77.

Arrow, K.J., 1962, The economic implications of learning-by-doing, *Review of Economic Studies* 29, 155-173.

Barro, R.J. and X. Sala-i-Martin, 1997, Technological diffusion, convergence, and growth, *Journal of Economic Growth* 2, 1-6.

Barro, R.J. and X. Sala-i-Martin, 2004, *Economic Growth* (2nd ed.), MIT Press, Cambridge, MA.

Basu, S. and D.N. Weil, 1998, Appropriate technology and growth, *Quarterly Journal of Economics* 113, 1025-54.

Bottazi, L. and G. Peri, 2003, Innovation, demand, and knowledge spillovers: Evidence form European patent data, *European Economic Review* 47, 687-710.

Boucekkine, R., D.de la Croix and O. Licandro, 2002, Vintage human capital, demographic trends, and endogenous growth, *Journal of Economic Theory* 104: 340-375.

Carroll, C.D. and Weil, D.N., 2000, Saving and growth with habit formation, *American Economic Review* 90, 341-355.

Clark, G., 2007, *A Farewell to Alms*, Princeton University Press, Princeton.

Caselli, F. and J. Feyrer, 2007, The marginal product of capital, *Quarterly Journal of Economics* 122, 535–568.

Coe, D.T., E. Helpman, and A.W. Hoffmaister, 2008, International R&D spillovers and institutions, NBER Working Paper 14069.

Comin, D., Easterly, W., and Gong, E., 2010, Was the wealth of nations determined in 1000 BC? *American Economic Journal: Macroeconomics* 2, 65-97.

Crafts, N.F.R. and Harley, C.K, 1992, Output growth and the British Industrial Revolution: A Restatement of the Crafts-Harley view, *Economic History Review* 45, 703-730

Crafts, N.F.R., 2003, Quantifying the contribution of technological change to economic growth in different eras: a review of the evidence, Economic History, London School of Economics and Political Science.

Doepke, M., 2004, Accounting for fertility decline during the transition to growth. *Journal of Economic Growth* 9: 347–383

Eaton, J. and S. Kortum, 1999, International patenting and technology diffusion: Theory and measurement, *International Economic Review* 40, 537-70.

Fogli, A. and Veldkamp, L., 2011, Germs, social networks and growth, mimeo, University of Minnesota.

Foster, A.D. and Rosenzweig, M.R., 1995, Learning by doing and learning from others: Human capital and technical change in agriculture, *Journal of Political Economy* 103, 1176-1209.

Galor, O., 2005, From stagnation to growth: unified growth theory, in: P. Aghion and S. Durlauf, *Handbook of Economic Growth* Vol 1A, Elsevier, Amsterdam.

Galor, O. and D.N. Weil, 2000, Population,technology and growth: From the Malthusian regime to the demographic transition and beyond, *American Economic Review* 110, 806–828.

Galor, O. and O. Moav, 2002, Natural selection and the origin of economic growth, *Quarterly Journal of Economics* 117, 1133-1192.

Galor, O. and A. Mountford, 2008, Trading population for productivity: theory and evidence, *Review of Economic Studies* 75, 1143-1179.

Grossman, G.M. and E. Helpman, 1991, *Innovation and Growth in the Global Economy*, MIT Press, Cambridge, MA.

Hofstede, G., 2001, *Culture's Consequences : Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, Sage Publications.

Howitt, P., 2000, Endogenous growth and cross-country income differences, *American Economic Review* 90, 829-846.

Irwin, D.A. and Klenow, P.J., 1994, Learning-by-doing spillovers in the semiconductor industry, *Journal of Political Economy* 102, 1200-1227.

Jaffe, A. M. Tratjenberg, and R. Henderson, 1993, Geographic localization of knowledge spillovers as evidenced by patent citations *Quarterly Journal of Economics* 63, 577-598.

Jones, C.I., 2001, Was an industrial revolution inevitable? Economic growth over the very long run, *Advances in Macroeconomics* 1, 1-43.

Jones, C.I. and Romer, P.M., 2010, The new Kaldor facts: ideas, institutions, population, and human capital, *American Economic Journal: Macroeconomics* 2, 224-245.

Keller, W., 2002, Geographic localization of international technology diffusion, *American Economic Review* 92, 120-142.

Keller, W. and C.H. Shiue, 2008, Institutions, technology, and trade, Discussion Paper, University of Colorado.

Klenow, P.J. and Rodriguez-Clare, A., 2005, Externalities and growth, in: Aghion, P. and S.N. Durlauf (eds.) *Handbook of Economic Growth* Vol. 1, 817-861, Elsevier, Amsterdam.

Kögel, T. and Prskawetz, A., 2001, Agricultural productivity growth and escape from the Malthusian trap, *Journal of Economic Growth* 6, 337-357.

Kraay, A. and C. Raddatz, 2007, Poverty traps, aid, and growth, *Journal of Development Economics* 82, 315-347.

Kremer, M., 1993, Population growth and technological change: one million B.C. to 1990, *Quarterly Journal of Economics* 108, 681-716.

Lucas, R.E., 1990, Why doesn't capital flow from rich to poor countries?, *American Economic Review* 80, 92-96.

Lucas, R.E., 2002, *The Industrial Revolution: Past and Future*, Cambridge: Harvard University Press.

Lucas, R.E., 2009, Trade and the industrial revolution, *American Economic Journal: Macroeconomics* 1, 1-25.

Maddison, A., 2001, *The World Economy – A Millennial Perspective*, Development Center Studies, OECD, Paris.

Madsen, J.B. and Ang, J.B. and Banerjee, R., 2010, Four centuries of British economic growth: the roles of technology and population, *Journal of Economic Growth* 15, 1–28.

Mankiw, G., D. Romer and D. Weil, 1992, A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107,407-37.

Mokyr, J., 2002, *The Gifts of Athena*, Princeton University Press, Princeton.

Mokyr, J. 2005, Long-term economic growth and the history of technology, in: Aghion, P. and S.N. Durlauf (eds.) *Handbook of Economic Growth*, Elsevier, Amsterdam, Chapter 17.

Newman, M.E.J., 2003, The structure and function of complex networks, *SIAM Review* 45, 167-256.

Parente, S.L. and Prescott, E.D., 1994, Barriers to technology adoption and development, *Journal of Political Economy* 102, 298–32.

Parente, S.L. and Prescott, E.D., 2005, A unified theory of the evolution of international income levels, in: Aghion, P. and S.N. Durlauf (eds.) *Handbook of Economic Growth*, Elsevier, Amsterdam, Chapter 21.

Phillips, P.C.B. and D. Sul, 2009, Economic transition and growth, *Journal of Applied Econometrics* 24, 1153-1185.

Pomeranz, K., 2000, *The Great Divergence: China, Europe, and the Making of the Modern World*, Princeton, Princeton University Press.

Romer, P.M., 1986, Increasing returns and long-run growth, *Journal of Political Economy* 94, 1002-1037.

Romer, P.M., 1990, Endogenous technological change, *Journal of Political Economy* 98, S71-S102.

Steger, T.M., 2000, Economic growth with subsistence consumption, *Journal of Development Economics* 62, 343-361.

Strulik, H., 2009, Knowledge and growth in the very long run, Hannover University, Working Paper 414.

Strulik, H., 2010, A note on economic growth with subsistence consumption, *Macroeconomic Dy-*

namics, forthcoming.

Strulik, H. and J. Weisdorf, 2008, Population, food, and knowledge: A simple unified growth model, *Journal of Economic Growth*

Temin, P., 1997, Two views of the British Industrial Revolution, *Journal of Economic History* 57, 63-82.

Watts, D.J., and S.H. Strogatz, 1998, Collective dynamics of 'small-world" networks, *Nature* 393, 440-442.