

Knowledge and Growth in the Very Long Run

Holger Strulik*

First Version: February 2009. This Version: November 2012.

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 Goettingen, Wirtschaftswissenschaftliche Fakultät, Platz der Goettinger Sieben 3, 37073 Goettingen, Germany; email: holger.strulik@wiw.uni-goettingen.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. 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 Industrial Revolution that income per capita set off to grow permanently and at unprecedented 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 on an output per worker function with either constant returns (if they are in reduced form of the a linear Ak structure) or 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 (Galor, 2005; Aghion and Howitt, 2009, Chapter, 2.4; Jones and Romer, 2010). Here, I propose a different conclusion and suggest a simple modification of conventional growth theory by a feedback mechanism between capital accumulation and knowledge diffusion such that it gets the historical adjustment dynamics right. The feedback mechanism assumes that there exists a knowledge externality in production which is not set in stone but which is itself determined by the state of the economy, represented by the accumulated capital stock. In other words, it is assumed that knowledge travels more easily in an economy in which capital is abundant (including “vehicles of knowledge exchange” like horses, railways, or cars).

The view that capital accumulation and embodied technological change promote productivity and facilitate the diffusion of knowledge is a central theme in historians’ narratives of economic

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).

development and in particular in the work of Mokyr (1990, 2002, 2005). One crucial element of Mokyr's view on historical technological evolution is to distinguish between existing knowledge and accessible knowledge: "[I]f one person possesses a certain knowledge, how costly is it for others to acquire it? This question, indeed is at the heart of the idea of a 'technological society'. Knowledge is shared and distributed, and its transmission through learning is essential for such a society to make effective use of it. Between the two extremes of a society in which all knowledge acquired by one member is 'episodic' and not communicated to any other member, and the other extreme in which all knowledge is shared instantaneously to all members through some monstrous network, there was a reality of partial and costly sharing and access. But these costs were not historically invariant, and the changes in them are one of the keys to technological change." (Mokyr, 2005, p. 1120).

Following Arrow (1962) and taking the capital stock of an economy as a measure of existing knowledge, we could say – with reference to Mokyr – that the neoclassical model without knowledge externality implicitly assumes "episodic knowledge" and that the learning-by-doing variant of the Ak growth model, implicitly assumes a "monstrous network" by allowing any firm to have instant access to all knowledge. One interpretation of the here proposed theory is that it fills the void in between these extremes. Allowing the knowledge externality σ_t to vary with time t makes it possible to distinguish between existing knowledge (still measured by aggregate capital k_t) and the degree to which this knowledge is accessible an individual firm, $0 < \sigma_t < 1$. A localized, backward economy would be characterized by a low degree of knowledge diffusion and thus not far from one described by the neoclassical model. A fully integrated, advanced economy in turn is better described by a high degree of knowledge diffusion. The associated production function of any firm can then be thought of as providing output net of access costs to knowledge.

The model is closed by acknowledging that the knowledge externality, though time-varying, is not a direct function of time but state-dependent. The state of an economy in simple models of economic growth is given by the aggregate capital stock, implying $\sigma_t = \sigma(k_t)$. A more advanced economy is characterized by a larger capital stock and thus, *ceteris paribus*, by an easier flow of knowledge in the economy. This link is perhaps best visible for capital goods like cars, railways and other means of transportation as well as for IT and communication technologies but Mokyr (2002) describes also other more subtle links between capital accumulation and access costs to knowledge like the increasing use of interchangeable parts in mass production.

As shown below, the mutually enforcing power of capital accumulation and knowledge diffusion

can explain a long period during which an economy grows at glacier speed because both forces are low. Over time, however, increasing social returns on investment manifest themselves in gradually increasing rates of factor accumulation, knowledge diffusion, and productivity growth. The phase of gradually increasing growth comes to an end when the degree of knowledge diffusion reaches its upper bound and the economy reaches a balanced growth path.

There exist corroborating evidence that knowledge externalities are state-dependent and increasing with economic development. For contemporaneous poor countries the literature in development economics shows that farmers are learning mostly from their neighbors and that even at that localized level learning is incomplete (e.g. Foster and Rosenzweig, 1995; Conley and Udry, 2001). In fully developed countries and modern industrial production, with contrast, knowledge spillovers are more extensive within and across countries (Irwin and Klenow, 1995). Although knowledge spillovers in developed countries are still localized and decaying with geographic distance (Bottazzi and Peri, 2003) the degree of localization has shrunk substantially over time (Keller, 2002). The observation that, across countries, aggregate TFP and GDP per worker are strongly associated with the intensity of international travel (Andersen and Dalgaard, 2011) suggests that personal contact is an important determinant of knowledge exchange. The long-run history of means of transportation and human travel and the associated reduction of effective distance is surveyed by Asubel et al. (1998). 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”. For the time around the Industrial Revolution Clark (2007) reports diffusion lags for the steam engine, the cotton mill and the railway of on average 13 years for Western European countries, 22 years for southern and eastern Europe, 35 years for India, and 52 years for Latin America. Dittmar (2011) documents the slow diffusion of the printing press – another capital good alleviating knowledge exchange – and the associated local knowledge externalities expressed in city growth across medieval Europe.

The ease of knowledge diffusion is, of course, not fully determined by technology and capital accumulation but it is also affected, at any given capital stock, by institutions (as, for example, barriers to travel and trade). Comparing both channels, Keller and Shiue (2008) conclude 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 perfor-

mance in that better institutions improved the rate of adoption of steam trains. For the modern world Coe et al. (2009) provide evidence that institutional differences are an important determinant for the national appropriation of international R&D spillovers. The present paper includes a parameter (\bar{A}) 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.

The paper is organized as follows. Section 2 briefly discusses related literature. Section 3 sets up the basic model. Section 4 provides the steady-state analysis and a qualitative discussion of adjustment dynamics. It will be shown how the model explains the gradual take-off to growth and the potential existence of a poverty trap. Section 4 also offers an explanation for why capital does not flow to poor countries and why there may exist a productivity slowdown during the transition towards balanced growth. Section 5 calibrates the model for England and compares the long-run adjustment dynamics predicted by the theory with real data. It will be shown that the time series for income per capita and TFP predicted by the model approximate the historical time series for England 1700-2000 reasonably well. In particular, the theory gets the gradual historical take-off to growth about right. Section 6 considers an extension towards a two-region world economy in which knowledge and capital is allowed to move across borders. It will be shown that the presence of international knowledge flows implies a unique steady-state growth rate of income and TFP. At the same time the presence of country-specific access to knowledge explains interesting dynamics of temporary divergence and catch-up growth for the leaders, followers, and trailers of the Industrial Revolution.

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 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 present approach less appropriate. In an earlier version of this paper (Strulik, 2009) I thus investigated the accumulation-diffusion mechanism also within a Romer (1990)-style, R&D-driven growth model. I briefly discuss results from there in the Conclusion.

2. RELATED LITERATURE

The issue of long-run economic development relates the present paper to unified growth theory.² Unified growth theory emphasizes the interaction between demographic and economic forces and their impact on education as essential for our understanding of long-run economic growth. The present paper shows that the gradual-take off of income and TFP growth can as well be explained by the interaction of capital accumulation and knowledge diffusion, i.e. by another, completely different, and so far unexplored mechanism. In order to make this obvious the present theory shuts off all mechanisms for economic development proposed by unified growth theory. This does of course not mean that I am denying the importance of demographic change and education for long-run economic development. But disregarding them 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 other driving forces of economic development.

Another stand-alone mechanism that generates gradually increasing income growth is a positive feedback between income and savings rates.³ The easiest way to generate increasing savings rates is to impose a non-homothetic utility function. But other, more sophisticated mechanisms, for example working through consumption habits (Carroll et al., 2000) or wealth effects on patience (Strulik, 2012), have been proposed as well. The main difference to the present paper is that the savings channel cannot motivate the gradual rise of TFP growth observable in association with the take-off to modern growth. In a multi-country context, it cannot motivate cross-country variations in TFP growth and variation in income growth is completely explained by varying savings rates. In Strulik (2009) I extend the present model by non-homothetic utility. While all the main features remain unchanged, there appears additionally a phase during which income growth and knowledge diffusion are further amplified through increasing investment rates such that overshooting behavior occurs. Income dependent savings rates could thus be helpful to explain periods of extraordinary high growth with subsequent decline (the Roaring Twenties).

The feature that capital accumulation impacts on TFP growth relates the present paper to the

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

³ This mechanism has been used by, among others, Steger (2000), Carroll et al. (2000), Voigtlaender and Voth (2006), and Strulik (2010). Voigtlaender and Voth (2006) also consider knowledge externalities. The degree of knowledge spillovers, however, is treated as constant, as usual in the so far available literature. This feature shuts off the positive feedback between capital accumulation and knowledge diffusion proposed in the present paper.

literature on embodied technological change.⁴ This literature emphasizes the (firm-specific) knowledge that can be bought with the installation of new vintages of capital goods. The present paper, in contrast, emphasizes the effect of investment on access to knowledge at the aggregate level.

The two-region version of the model is related to a large literature on international knowledge diffusion and catch-up growth.⁵ 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.

3. MODEL SETUP

3.1. Households. Consider an economy populated by two overlapping generations. The concept of overlapping generations is useful to simplify the analysis but it is not driving the results (see Strulik, 2009). 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).

$$s_t = \frac{\beta}{1 + \beta} \cdot w_t. \quad (1)$$

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

3.2. 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 $y_t(i) = A_t(i)k_t(i)^\alpha \ell_t(i)^{1-\alpha}$ 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

⁴ See, among others, De Long and Summers (1991), Hulten (1992), Greenwood et al. (1997).

⁵ See, among 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.

implies that production factors are demanded such that factor prices equal the (private) marginal product, i.e. $w_t = (1 - \alpha)A_t(i)k_t(i)^\alpha$, $r_t = \alpha A_t(i)k_t(i)^{\alpha-1}$ where w_t denotes wages and r_t denotes the interest rate. Aggregate employment is denoted by $k_t = \int_0^1 k_t(i)di$ and $\ell_t = \int_0^1 \ell_t(i)di = 1$.

3.3. Knowledge Diffusion. As proposed by Arrow (1962) we think of knowledge embodied in capital goods such that the aggregate capital stock approximates the existing knowledge. In the so far available literature the degree of knowledge diffusion was assumed to be constant; the association between $\log A_t(i)$ and $\log k_t$ was assumed to be linear (see e.g. Arrow, 1962, Frankel, 1962, Romer, 1986). Whereas Arrow focussed on the case of incomplete knowledge diffusion by assuming an exponent smaller than $1 - \alpha$, the standard model of endogenous growth, i.e. the so called “ Ak growth model”, is based on the assumption that the exponent equals $1 - \alpha$ such that the knowledge of each firm changes in proportion to the aggregate capital stock, $dA_t(i)/A_t(i) = dk/k$. In any case the degree of knowledge diffusion was assumed to be independent from the state of the economy.⁶

Here, we allow the knowledge externality to vary over time. Equation (2) is the simplest way to formalize this feature.

$$A_t(i) = \bar{A}k_t^{\sigma_t(1-\alpha)}, \quad \sigma_t \in [0, 1]. \quad (2)$$

One interpretation of this assumption, inspired by Mokyr (2005), is that we distinguish between *existing* knowledge, which is, 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 degree of knowledge diffusion, denoted by σ_t . In other words, $A_t(i)$ measures the productivity of a firm net of access costs to knowledge.

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 the Ak growth model 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.

⁶ Romer (1986) actually allowed the knowledge externality to be larger than $1 - \alpha$ and avoided the otherwise arising explosion of the economic system by introducing an exogenous upper bound for the growth of k . For obvious reasons this case got less attention in the literature and is also subsequently ignored.

3.4. 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).

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

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 , $y_t \equiv \bar{A}k_t^{\alpha+\sigma_t(1-\alpha)}$.

3.5. 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).

$$k_{t+1} = ak_t^{\alpha+\sigma_t(1-\alpha)}, \quad a \equiv (1 - \alpha)\bar{A}\beta/(1 + \beta). \quad (4)$$

Although time-variant, the degree of knowledge diffusion is certainly not a *function* of time. Instead the state of the economy determines how easily knowledge diffuses. The only state variable in this simple model is the capital stock. The final element is thus to introduce a positive feedback effect from the size of the capital stock to the degree of knowledge diffusion, $\sigma_t = \sigma(k_t)$. One interpretation of this assumption is that capital goods alleviate travel and information exchange. 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 and firms have better access to the available knowledge. 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 $\lim_{k \rightarrow \infty} \sigma(k) = 1$.*

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

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

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, $\lim_{k \rightarrow \infty} k_{t+1} = ak_t$. 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.

The case of $\sigma > 1$ is ignored because 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. But, of course, the upper bound of σ could be less than one. This case is plausible but less interesting since it prevents long-run growth and the inclusion of the popular Ak -model as the limiting case. I discuss this possibility in the Conclusion.

The time-invariant efficiency parameter \bar{A} controls the diffusion of knowledge for any *given* state of the economy. This means that the size of \bar{A} is determined by “something” outside the model that changes the ease of knowledge diffusion at any level capital stock and income. Thus \bar{A} can be thought of capturing the effect of persistent determinants of access to knowledge as, for example, institutional barriers to travel and trade.

The condition that σ approaches one in the limit is necessary but not sufficient for economic growth. For long-run growth to be feasible we need $a > 1$.

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

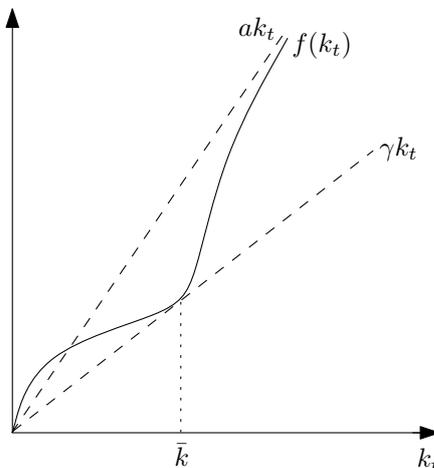
If a were smaller than one, we would be back in the neoclassical world without sustainable long-run growth. The interesting question here is not whether a balanced growth path exists but if and how an economy arrives there in the long-run. Finally we impose a mild assumption on the curvature of $\sigma(k)$.

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.

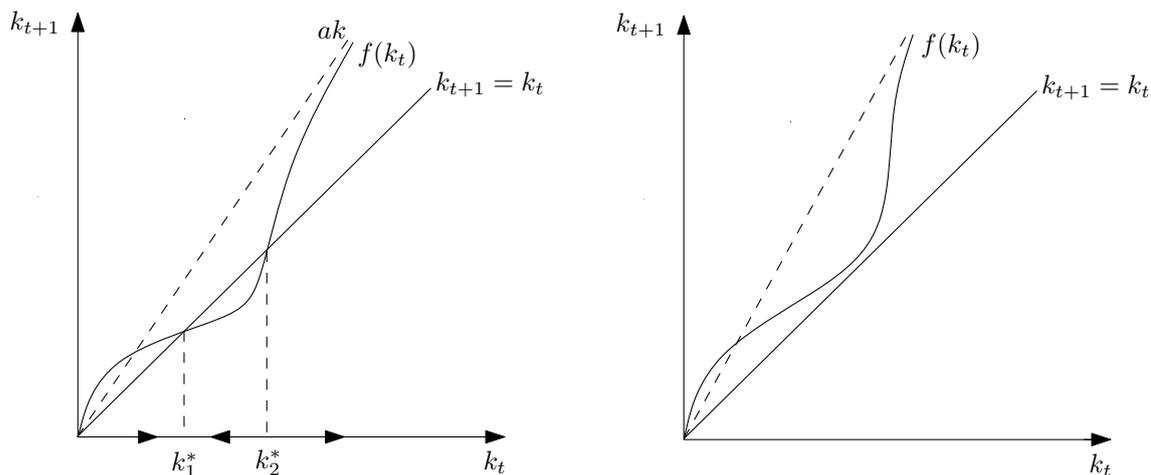
Figure 1 visualizes the result. When capital stock is low and there is little knowledge diffusion, the neoclassical channel of decreasing individual returns to investment dominates and $f(k_t)$ is concave, as suggested by the neoclassical approach. At capital stock \bar{k} the modern channel capturing 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.

4. STEADY-STATES AND ADJUSTMENT DYNAMICS

4.1. Steady-States. The hyperbolic shape of $f(k)$ provides some plausible steady-state features and adjustment dynamics which can neither be captured by the neoclassical model nor 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 $\gamma = 1$ for which 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. For sufficiently large k , $k > \bar{k}$, the growth rates of capital and income are increasing during the adjustment towards the balanced growth path and are eventually converging towards the constant $a - 1$.*

Figure 2: Phase Diagram



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 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.

4.2. Comparative Statics. Observe that a change of the compound constant $a \equiv (1-\alpha)\bar{A}\beta/(1+\beta)$ operates like a re-scaling of $f(k_t)$ along the k_{t+1} axis: with rising a the supporting tangent γk and the 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 knowledge flow-promoting institutions or survival prospects can initiate an escape from poverty and convergence towards balanced growth.

4.3. The Poverty Trap. The equilibrium of stagnation differs qualitatively from the standard Malthusian poverty trap (as, for example, investigated in Ashraf and Galor, 2011). In the standard Malthusian model, income at the steady-state is pinned down by the weight of children in the utility function of parent and child costs. The crucial feature of this kind of steady-state is that an (exogenous) increase in productivity would not change the position of the steady-state, i.e. it would not improve steady-state income. In the present model an exogenous increase of productivity (\bar{A}) would draw the $f(k)$ curve in direction of the k_{t+1} axis (cf. Figure 2) and thus raise income per capita at the steady-state. This means that the model can explain stagnating income levels far above those that can be motivated by the cost of children. Another implication is that the present model can easily explain a great variance of income levels across stagnating economies, a fact which is hard to square with the standard Malthusian model. In particular, stagnation may occur in midst of a

process of increasing economic integration at an intermediate degree of knowledge diffusion, a result that may help to explain the poor growth performance of many contemporaneous countries which are less appropriately characterized as traditional societies trapped in a Malthusian equilibrium.⁷

4.4. Why Doesn't Capital Flow to Poor Countries? 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. An economy in the neighborhood of the equilibrium of stagnation k_1^* would not attract capital flows from a rich country in which $k > k_2^*$ because productivity is too low. In turn, productivity in the poor country 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 is not sustainable in the long-run. We take up issues of catch-up growth and international flows of knowledge and capital in Section 6.

4.5. 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 toward balanced growth the rate of economic growth is gradually increasing until it levels off at a high level and approaches a balanced growth path. In the figure this can be seen by the convex-concave shape of the $f(k_t)$ curve. The gross growth rate of k_t is given by $f(k_t)/k_t$, i.e. the ratio between $f(k)$ and the identity line. For $k_t > \bar{k}$ this expression first increases gradually and then levels off as $f(k)$ approaches ak . This means that along the path towards balanced growth the growth rates of capital and income increase perpetually although the savings rate remains constant.

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

⁷ See Kraay and Raddatz, 2007, for the difficulties incurred by calibrating actual economies to get support of stagnation at subsistence level. There exist a few recent studies, which have extended the Malthusian model such that it can explain a greater variance of income levels across countries while still relying on the Malthusian mechanism to explain stability of the poverty trap; see Dalgaard and Strulik (2010), Sharp et al. (2012), and Voigtlaender and Voth (2013).

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.

4.6. The 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.

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

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

$$g_{At} \approx a^{\sigma_t(1-\alpha)} \cdot k_t^{\sigma_t(\sigma_t-1)(1-\alpha)^2} - 1.$$

The first term $a^{\sigma_t(1-\alpha)}$ is monotonously increasing in the degree of knowledge diffusion σ_t . The second term, however, exerts a non-monotonous, hump-shaped, influence on productivity growth. The exponent vanishes for $\sigma_t = 0$ as well as for $\sigma_t = 1$ and assumes a maximum for $\sigma_t = 1/2$. Accordingly, increasing capital accumulation has no influence on TFP growth for $k_t = 0$ (when there is no knowledge diffusion) and for $k_t \rightarrow \infty$ (when knowledge diffuses completely and TFP growth assumes a constant). In between, for intermediate values of k_t the second term reaches a maximum. This effect is dominating TFP growth if k_t is sufficiently large.

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.

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.

Specifically I try to match the actual evolution of income and productivity in England 1700-2000. I begin with setting the capital share $\alpha = 1/3$ as in many related studies. Then I set \bar{A} and β such that along the balanced growth path the savings rate is 0.175 and income per capita grows at 2.0 percent annually. The value of 0.175 corresponds with the average gross savings rate in the United Kingdom 1970-2000 (World Bank, 2012). The value of 2.0 percent corresponds with the average growth rate of income per capita in England 1970-2000 according to Clark (2009).⁸ These values implies $\bar{A} = 12.6$ and $\beta = 0.21$. In the Appendix I provide robustness checks for a higher and lower value of the capital share. Compared to the capital share the model responds relatively insensitively to changes in the other parameter values.

In an earlier version of this paper I derived a network-based foundation of the function for knowledge diffusion $\sigma(k_t)$.⁹ Here I follow a different approach and try to infer $\sigma(k_t)$ from the data. For that purpose I use data for England 1700-2000 provided at decennial intervals by Clark (2009). From Clark’s time series for income per capita, the return on capital, and the capital share I compute the implied time series of capital stock \hat{k}_t . Clark also provides an decennial estimate TFP \hat{A}_t . Since these values are normalized such that $\hat{A}_t(1860) = 100$ there exists one degree of freedom concerning

⁸ Maddison (2001) estimates a similar average growth rate of 2.1 percent for the UK during the same period.

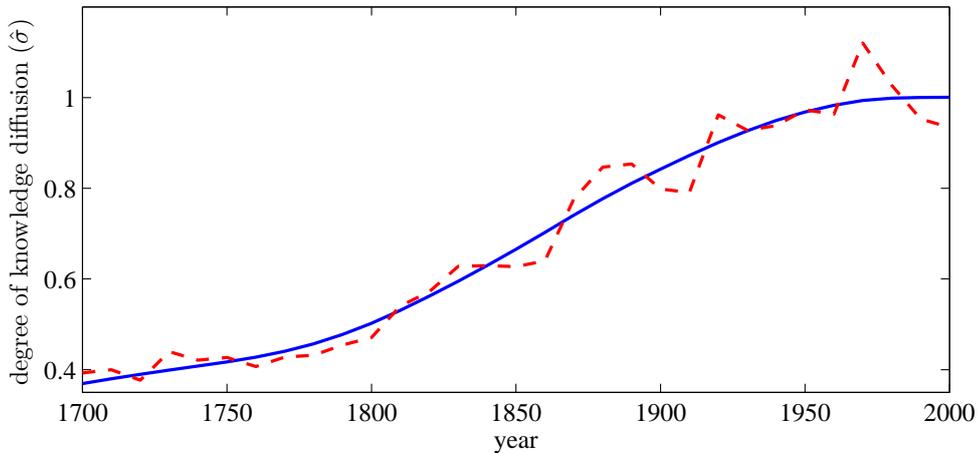
⁹ This part is still available as an online Appendix at <http://holger-strulik.uni.me>

the scale at which TFP is measured implying that “model TFP” should be a constant multiple of “data TFP”, that is $A_t = \mu \hat{A}_t$. Taking logs, equation (2) implies $\sigma_t = (\log A_t - \log \bar{A}) / [(1 - \alpha) \log k_t]$. Using Clark’s data I thus obtain an estimate of the degree of knowledge spillovers as

$$\hat{\sigma}_t = \frac{\log \hat{A}_t - \log \bar{A} + \log \mu}{(1 - \alpha) \log(k_t)}. \quad (6)$$

I plug into (6) the calibrated values of α and \bar{A} and adjust the scale parameter μ such that the upper limit of $\hat{\sigma}$ equals one. As explained in Section 2, an upper limit of one for σ is needed in order to imbed the Ak growth model as a limiting case and in order to allow for perpetual growth. Because there is a lot of fluctuation in the data, in particular because TFP is measured as a residual, I apply the Hodrick–Prescott filter and a smoothing parameter of 50 to uncover the long-term trend. Figure 3 shows the estimated time series of σ for England 1700-2000. The red dashed line shows the raw estimate and the blue solid line the smoothed time series. Figure 3 shows that over the last three hundred years the estimated degree of knowledge diffusion in England was far from constant. It increased by about factor 2.5 with the steepest rise between the mid 19th and mid 20th century.

Figure 3: The Degree of Knowledge Diffusion: England 1700-2000



Dashed line: raw estimate. Solid line: smoothed estimate using the Hodrick–Prescott filter ($\lambda = 50$). Parameters: $\alpha = 1/3$, $\bar{A} = 12.6$, and $\mu = 0.28$. Data from Clark (2009).

The next step is to find and calibrate a function $\sigma(k_t)$ that approximates $\hat{\sigma}_t$ reasonably well. In the following I shall use the function

$$\sigma(k_t) = 1 - \left(\frac{\omega}{\omega + \exp(k_t)} \right)^3. \quad (7)$$

This function has three desirable properties. Firstly, it can be derived from first principles under the

assumption that the economy is a small world network in which the number of long-distance links depends positively on the capital stock (see the Online-Appendix). Secondly, it fulfills Assumption 3. Thirdly, it is very parsimonious. It leaves only one parameter, ω , to calibrate. It turns out that the best fit of the estimated time series is obtained for $\omega = 76.8$. Table 1 in the Appendix summarizes the calibration of the model.

For the numerical solution I assume that the economy starts in year 1 CE with a capital stock of 1 unit. After solving the model numerically I 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. The predicted time series for knowledge diffusion, income, income growth, and TFP growth are shown by solid lines in Figure 4. Dashed lines show the corresponding data for England, based on Clark (2009).

The calibration implies that there exists no steady-state of stagnation. The take-off to growth is endogenously initiated. The funnel between the implied $f(k)$ curve and the identity line, however, is quite narrow, implying that the economy grows only at glacier speed for an extended period of time. For better visibility the period from year 1 to year 1700 is not shown in the figure. Basically, the medieval period looks very much like the early 18th century. Growth is so exceedingly slow such that it looks like stagnation for the uninformed spectator.

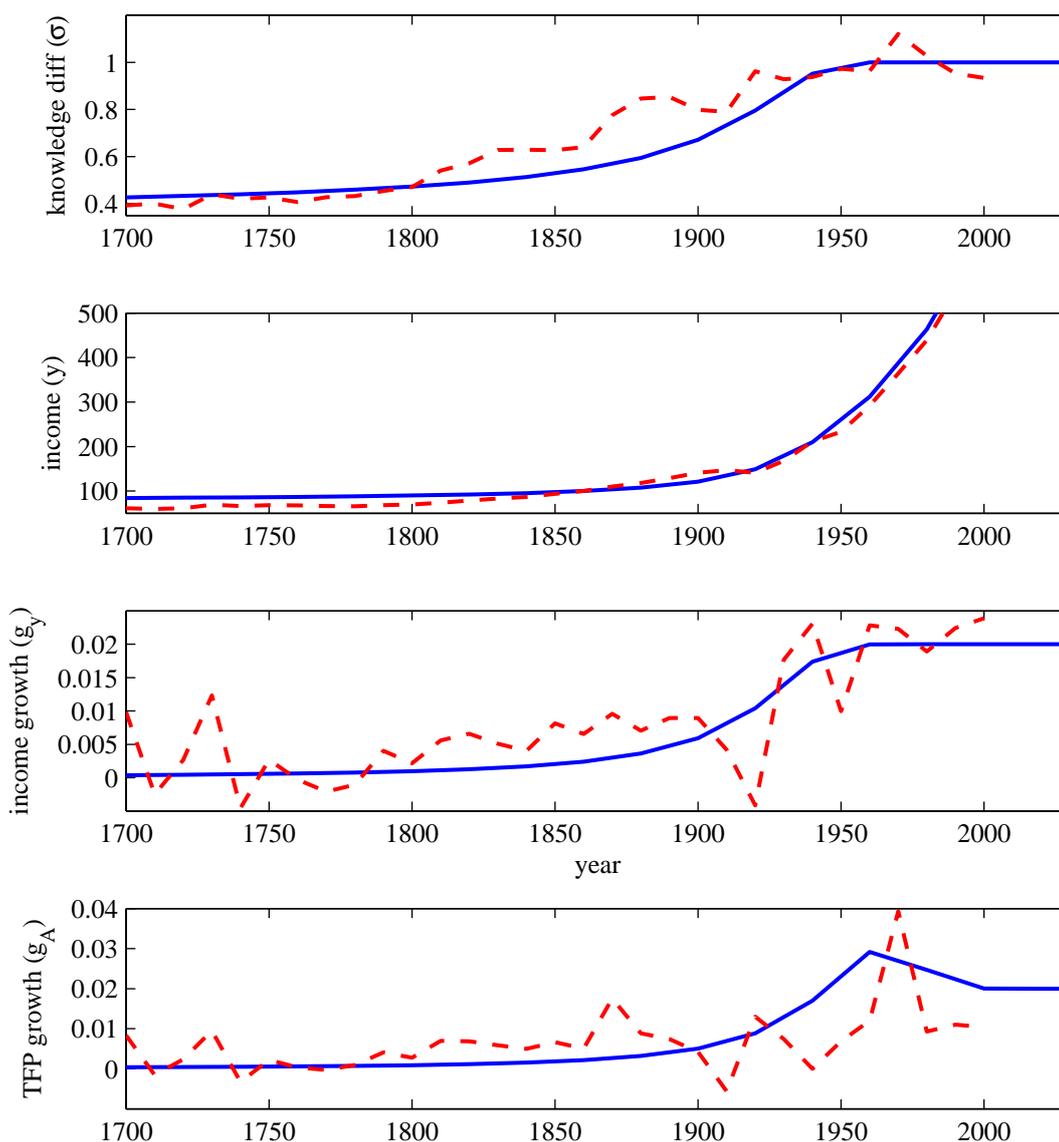
The main take away from Figure 4 is that the model predicts a gradual take-off towards balanced growth, in line with the actual time series. Because growth is so low in pre-modern times an uninformed spectator could also have difficulties in distinguishing the growing economy from another economy resting at the equilibrium of stagnation. 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).¹⁰

As shown in the first panel of Figure 4 the predicted time series for knowledge diffusion underestimates the time series inferred from the data during the 19th century. Given that there was only one parameter to calibrate the $\sigma(k)$ -curve, however, the predicted time series for knowledge diffusion approximates the estimated time series reasonably well.

The second panel of Figure 4 confronts the income series predicted by the model with the income per capita estimates of Clark (2009). For comparison, I followed Clark and normalized the generated time series such that $y_{1860} = 100$. The time series predicted by the model accords quite well with

¹⁰ In Strulik (2009) I perform a numerical exercise of the Great Divergence.

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



Solid lines: model prediction. Dashed lines: data for England based on Clark (2009). Income has been normalized such that income $y_{1860} = 100$.

the historical times series. In the third panel we look at the same data, now in terms of growth rates. The focus on rates emphasizes the fluctuations in the historical data and makes the fit of the two series visually less impressive. The third panel, however, is useful in documenting the gradual take-off of economic growth, which is less well visible in the panel showing level data. Again, the model underestimates somewhat economic performance during the 19th century. But model and data agree in that growth during the 19th century was comparatively low from the perspective of the 21st century.

The lower panel of Figure 4 compares the model's prediction of TFP growth with the historical

estimates from Clark (2009).¹¹ Naturally, there is again much fluctuation in the historical series but the rising trend of TFP growth during the first half of the 20th century is discernable. According to Clark's estimate, TFP growth peaks in 1970 and falls sharply afterwards. The model predicts TFP to reach the peak a decade earlier and to decline less sharply afterwards.

6. A TWO-REGION MODEL OF KNOWLEDGE DIFFUSION

6.1. Model Setup. The closed economy model of knowledge diffusion and growth displayed multiple equilibria and suggested that low productivity associated with 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 it 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.¹² 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).

$$A_t^A(i) = \bar{A} \cdot (k_t^A + k_t^B)^{\sigma(k_t^A)(1-\alpha)}, \quad A_t^B(i) = \bar{B} \cdot (k_t^A + k_t^B)^{\sigma(k_t^B)(1-\alpha)} \quad (8)$$

where \bar{A} and \bar{B} capture the influence of knowledge-flow promoting institutions in region A and B .

¹¹ The series is called Efficiency (P_{DE}) in Clark (2009). Clark computes also an alternative TFP series, which is, however, disbanded after 1860. This fact settles the problem which one to use for the comparison in Figure 4.

¹² 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.

World capital stock, $k_t^A + k_t^B$, approximates the existing knowledge whereas region-specific capital 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 $y_t^A = A_t^A (k_t^A)^\alpha$ and $y_t^B = A_t^B (k_t^B)^\alpha$. 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.

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

In order to avoid case differentiation about international capital mobility we assume that capital is flowing internationally such that $r_t^A = \eta r_t^B$. Here, $\eta > 0$ 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, $\eta = 1$. Despite this relatively crude treatment of capital flows, the model is capable to display very rich and plausible dynamics of regional capital accumulation and productivity growth.¹³ Inserting the regional rates of return, $r_t^A = \alpha A_t^A (k_t^A)^{\alpha-1}$ and $r_t^B = \alpha A_t^B (k_t^B)^{\alpha-1}$ into the interest-parity, $r_t^A = \eta r_t^B$ we get the correlation between regional capital stocks.

$$k_t^B = \left(\eta \frac{A_t^B}{A_t^A} \right)^{\frac{1}{1-\alpha}} \cdot k_t^A. \quad (10)$$

Naturally, there is relatively more capital allocated to region B if TFP is relatively high in B . 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:

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

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

¹³ 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.

$$k_{t+1} = s_t^A + s_t^B. \quad (12)$$

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 region, say A , grows perpetually while the other region stagnates at the poverty trap identified in Section 4 (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.*

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 (asymptotically) 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, $\sigma(k_t^A) \rightarrow 1$ and $\sigma(k_t^B) \rightarrow 1$ for $t \rightarrow \infty$ and thus from (8) along the steady-state $\lim_{t \rightarrow \infty} (A_{t+1}^B/A_{t+1}^A) = (\bar{B}k_t)/(\bar{A}k_t) = \bar{B}/\bar{A}$. Inserting this into (11) and applying $\lim_{t \rightarrow \infty} \sigma(k_t^A) = 1$, we see that, along the steady-state, capital and productivity in country A are growing in proportion to the world capital stock.

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

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

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

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

Having obtained the region-specific savings rates as a function of 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.

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

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. metaphysical 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 (that is higher β^A or higher β^B) increases world-wide growth. An improvement of regional institutions (higher \bar{A} or \bar{B}) does not generally increase world-wide growth.*

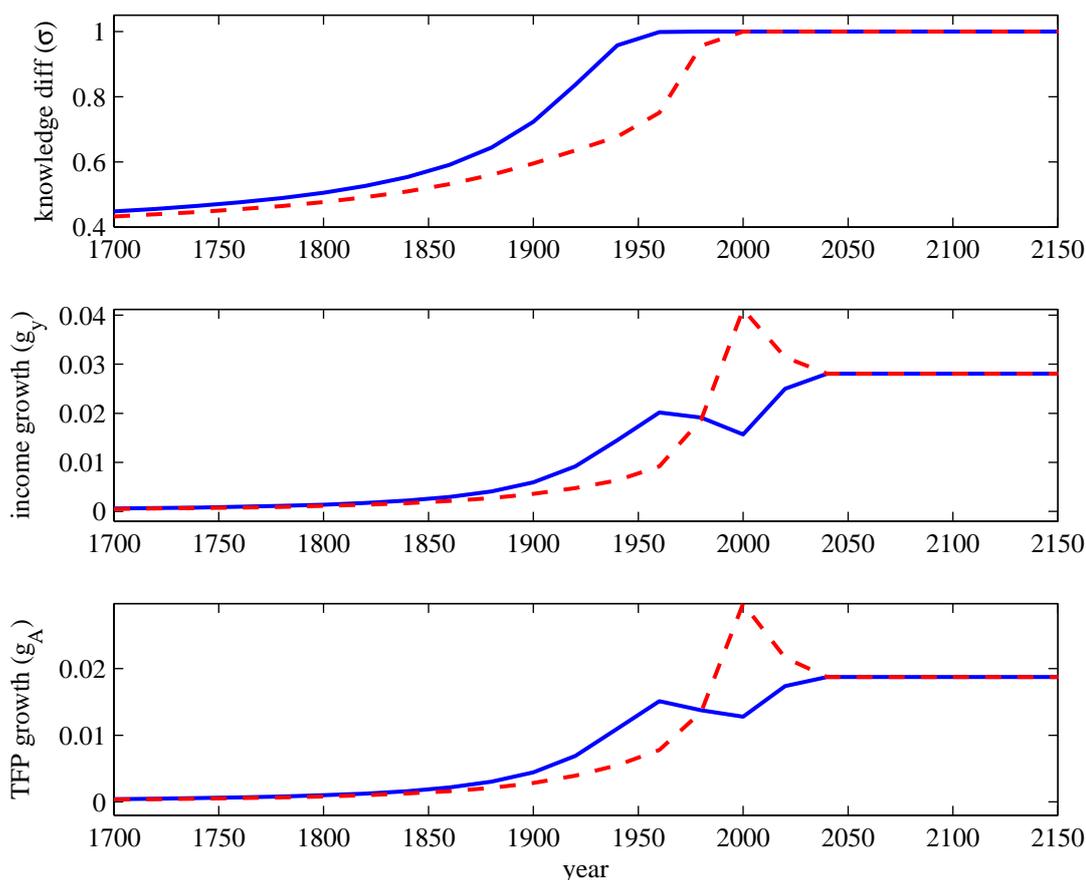
To see why better knowledge-flow promoting institutions can deteriorate growth, assume an extreme case of $\beta^A = 0$, take the derivative of growth with respect to \bar{A} , and obtain $\partial g / \partial \bar{A} < 0$. In this extreme case only the citizens of B are saving. If knowledge-flow promoting institutions in region A improve, savings are redirected 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.

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 diffused 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.¹⁴ 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 values of α and β from the one country experiment and fix the remaining parameter values, \bar{A} , \bar{B} , and ω such that (i) the West takes off in the late 18th century, (ii) the West grows at a rate of about 2 percent during the second half of the 20th century, (iii), income growth in the East catches up with the West around 1980.

Figure 5: West–East Adjustment Dynamics



Parameters: $\alpha = 0.33$, $\beta = 0.21$, $\omega = 38$, $k_0 = 1$ for both countries. $\bar{A} = 9.33$, $\bar{B} = 9.30$.

¹⁴ 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”.

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.

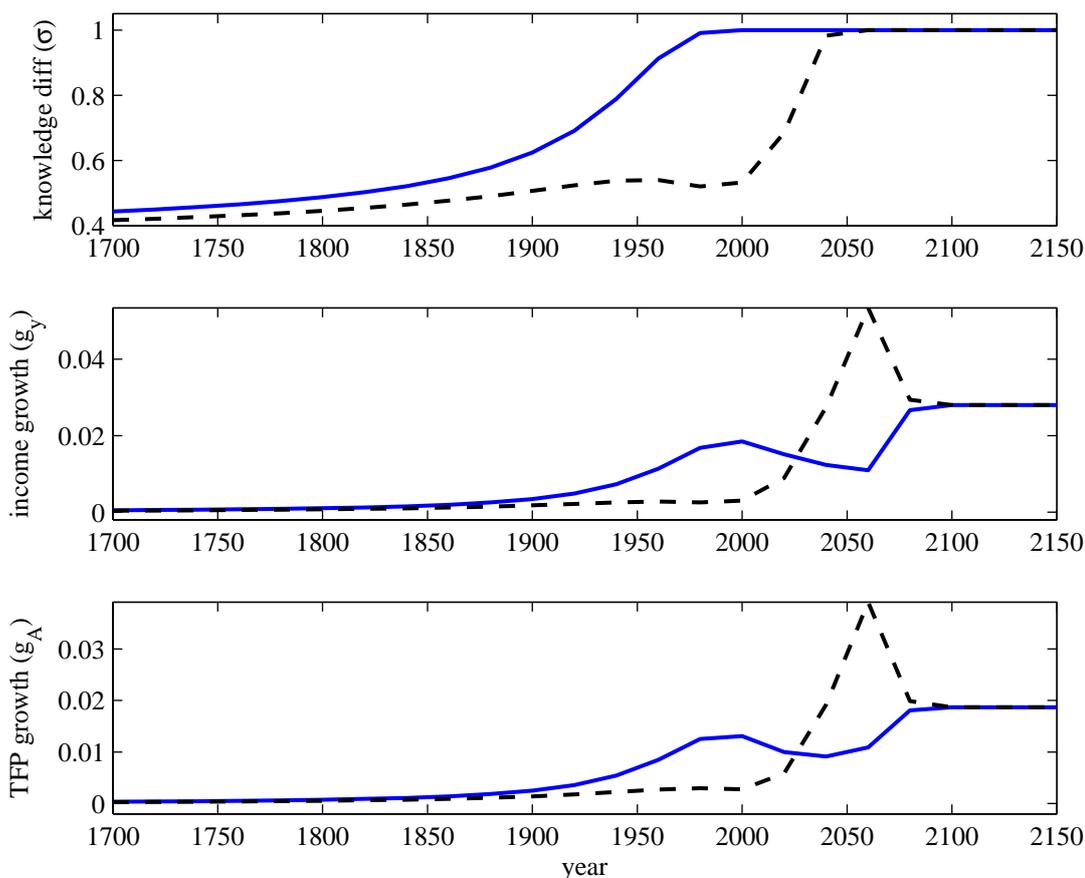
At the time when growth loses momentum in the West because firms are almost completely integrated and knowledge diffusion is close to its maximum degree, the East takes off. Now the East experiences a phase of temporarily increasing returns to scale (convex shape of $f(k)$) and TFP growth surpasses that 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 and the West experiences also a mild decrease of economic growth. 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. In other words the model supports an “advantage of backwardness” (Gerschenkron, 1962). The follower of the Industrial Revolution experiences temporarily higher growth than the leader.¹⁵ In the end, however, the West benefits from higher integration: growth experienced during the 20th century is surpassed by growth along the steady-state of a fully integrated world.

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, I kept almost all parameters from the previous case but reduced \bar{B} further. As a consequence the 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

¹⁵ 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: $\alpha = 0.33$, $\beta = 0.21$, $\omega = 38$, $k_0 = 1$ for both countries. $\bar{A} = 9.32$, $\bar{B} = 9.26$.

half of the 20th century when growth in the North reaches unprecedented levels, income in the South declines.

The reasons for divergence are capital flight and deteriorating capital stock in the South. It is worthwhile for the 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). This can be seen in the σ -Panel of Figure 6 where the diffusion of knowledge in the South declines during the second half of the 20th century. 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 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 emerge 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 be 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 engine” growth model it is possible to abandon the assumption of complete knowledge diffusion through learning-by-doing ($\lim_{k \rightarrow \infty} \sigma(k) = 1$ in the current setup) and yet generate perpetual growth. Intuitively, spillovers in learning-by-doing need only to be strong enough to generate a certain *level* of TFP from which onwards market-based R&D

becomes profitable.

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

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

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 $k \neq \bar{k}$. Compute

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

And thus at \bar{k}

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

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 , $k \rightarrow \infty$, $f(k)$ 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 (\bar{k}, ∞) , 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.

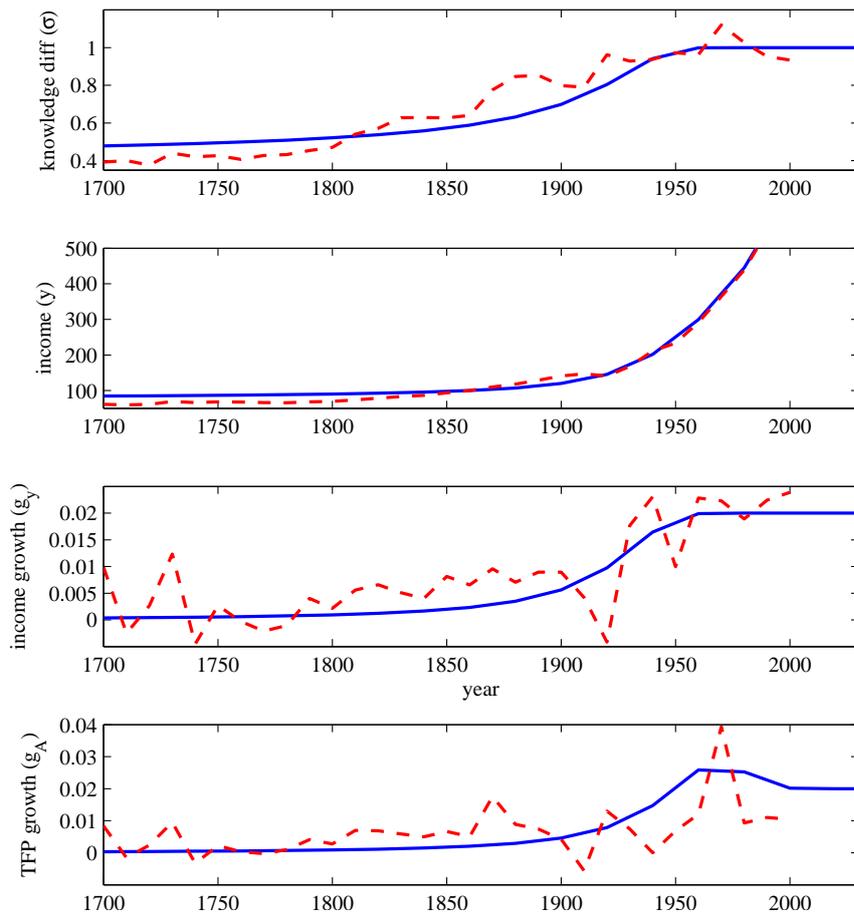
If $\gamma > 1$, $f(k)/k > 1$ everywhere and there exists no fixed point. In that case, and for $k > k_2^*$

if $\gamma < 1$, the capital stock is perpetually growing. For $k \rightarrow \infty$, $\sigma \rightarrow 1$ and $f(k) \rightarrow ak$. 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 $k > \bar{k}$.

Table 1: Numerical Specification of the Economy

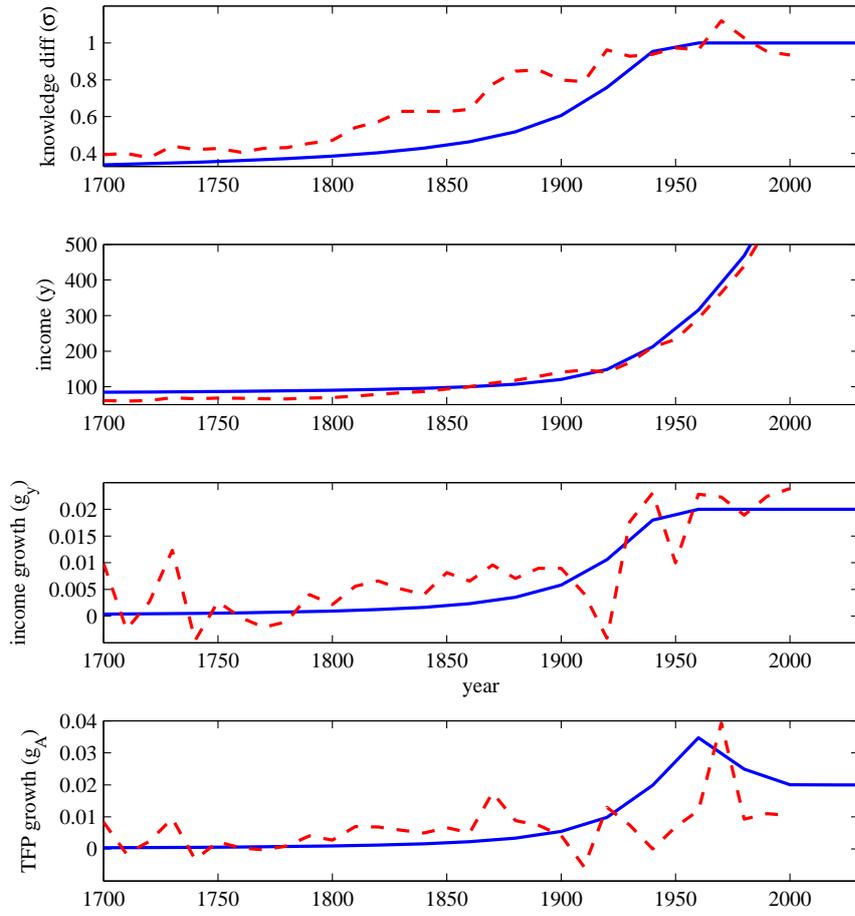
one country model		East West model		North South model	
α	0.33	α	0.33	α	0.33
β	0.21	$\beta^A = \beta^B$	0.21	$\beta^A = \beta^B$	0.21
ω	76.8	ω	38.0	ω	37.8
\bar{A}	12.6	\bar{A}	9.33	\bar{A}	9.33
		\bar{B}	9.30	\bar{B}	9.26
		η	1.00	η	1.00

Figure A.1: Sensitivity Analysis: Low Capital Share



Parameters as for Figure 4 except $\alpha = 0.2$ and $\omega = 52.3$. Solid lines: model prediction. Dashed lines: data from Clark (2009). See main text for details.

Figure A.2: Sensitivity Analysis: High Capital Share



Parameters as for Figure 4 except $\alpha = 0.5$ and $\omega = 172.5$. Solid lines: model prediction. Dashed lines: data from Clark (2009). See main text for details.

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.
- Aghion, P. and P. Howitt, 2009, *The Economics of Growth*, MIT Press, Cambridge, MA.
- 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.
- Bottazzi, L. and G. Peri, 2003, Innovation, demand, and knowledge spillovers: Evidence from 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.
- Clark, G., 2009, The macroeconomic aggregates for England, 1209 - 2008, Working Papers, University of California, Department of Economics, No. 09,19, <http://hdl.handle.net/10419/58383>.
- Conley, T. and Udry, C., 2001, Social learning through networks: The adoption of new agricultural technologies in Ghana, *American Journal of Agricultural Economics* 83, 668-673.
- 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, 2009, International R&D spillovers and institutions, *European Economic Review* 7, 723-741.
- 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.
- Dalgaard, C.-J. and Strulik, H., 2010, The physiological foundations of the wealth of nations, University of Copenhagen DP 10-05.

- De Long, B. and Summers, L.H., 1991, Equipment investment and economic growth, *Quarterly Journal of Economics* 106, 445–502.
- Dittmar, J., 2011, Information technology and economic change: The impact of the printing press, *Quarterly Journal of Economics* 126, 1133-1172.
- 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.
- 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.
- Frankel, M., 1962, The production function in allocation of growth: A synthesis, *American Economic Review* 52, 995–1022.
- 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.
- Gerschenkron, A., 1962, Economic backwardness in historical perspective, in: Hoselitz, B. (ed.) *The Progress of Underdeveloped Areas*, Chicago University Press, Chicago.
- Greenwood, J., Hercowitz, Z. and Krusell, P., Long-run implications of investment-specific technological change, *American Economic Review* 87, 342-362.
- Grossman, G.M. and E. Helpman, 1991, *Innovation and Growth in the Global Economy*, MIT Press, Cambridge, MA.
- Ha, J. and P. Howitt, 2006, Accounting for trends in productivity and R&D, *Journal of Money, Credit, and Banking* 39, 733-774.
- Howitt, P., 2000, Endogenous growth and cross-country income differences, *American Economic Review* 90, 829-846.
- Hulten, C.R., 1992, Growth accounting when technical change is embodied in capital, *American Economic Review* 84-964-980.
- Irwin, D.A. and Klenow, P.J., 1994, Learning-by-doing spillovers in the semiconductor industry, *Journal of Political Economy* 102, 1200-1227.
- 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 diffusion of the industrial revolution, *American Economic Journal: Macroeconomics* 1, 1-25.
- Maddison, A., 2001, *The World Economy – A Millennial Perspective*, Development Center Studies, OECD, Paris.
- Mankiw, G., D. Romer and D. Weil, 1992, A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107, 407-37.
- Mokyr, L., 1990, *The Lever of Riches: Technological Creativity and Economic Progress*, Oxford University Press, Oxford.
- 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.
- 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.
- Sharp, P., Strulik, H., and Weisdorf, J., 2012, The determinants of income in a Malthusian equilibrium, *Journal of Development Economics* 97, 112-117.
- 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 Dynamics* 14, 763-777.
- Strulik, H., 2012, Patience and prosperity, *Journal of Economic Theory* 147, 336-352.
- Strulik, H. and Weisdorf, J., 2008, Population, food, and knowledge: A simple unified growth model, *Journal of Economic Growth* 13, 169-194.
- Strulik, H. and Weisdorf, J., 2012, The determinants of income in a Malthusian equilibrium, *Journal of Development Economics* 97, 112-117.
- Temin, P., 1997, Two views of the British Industrial Revolution, *Journal of Economic History* 57, 63-82.
- Voigtlaender, N. and Voth, H.-J., 2006, Why England? Demographic factors, structural change and physical capital accumulation during the industrial revolution, *Journal of Economic Growth* 11, 319-361.
- Voigtlaender, N. and Voth, H.-J., 2013, The Three Horsemen of Riches: Plague, War and Urbanization in Early Modern Europe, *Review of Economic Studies*, forthcoming.
- World Bank, 2012, World Bank Development Indicators, <http://data.worldbank.org/country/united-king>