Lindner, Ines; Strulik, Holger

INNOVATION AND INEQUALITY IN A SMALL WORLD

International Economic Review

Provided in Cooperation with:
John Wiley & Sons

Suggested Citation: Lindner, Ines; Strulik, Holger (2020) : INNOVATION AND INEQUALITY IN A SMALL WORLD, International Economic Review, ISSN 1468-2354, Wiley, Hoboken, NJ, Vol. 61, Iss. 2, pp. 683-719,
https://doi.org/10.1111/iere.12437

INTERNATIONAL ECONOMIC REVIEW
Vol. 61, No. 2, May 2020 DOI: 10.1111/iere.12437

INNOVATION AND INEQUALITY IN A SMALL WORLD*

BY INES LINDNER AND HOLGER STRULIK[1]

Department of Econometrics and Operations Research, VU University Amsterdam, The Netherlands; Department of Economics, University of Goettingen, Germany

We present a multicountry theory of economic growth and R&D-driven technological progress in which countries are connected by a network of knowledge exchange. Technological progress in any country depends on the state of technology in the countries it exchanges knowledge with. The diffusion of knowledge throughout the world explains a period of increasing world inequality, followed by decreasing relative inequality. Knowledge diffusion through a small world network produces an extraordinary diversity of country growth performances, including the overtaking of individual countries and the replacement of the technologically leading country in the course of world development.

1. INTRODUCTION

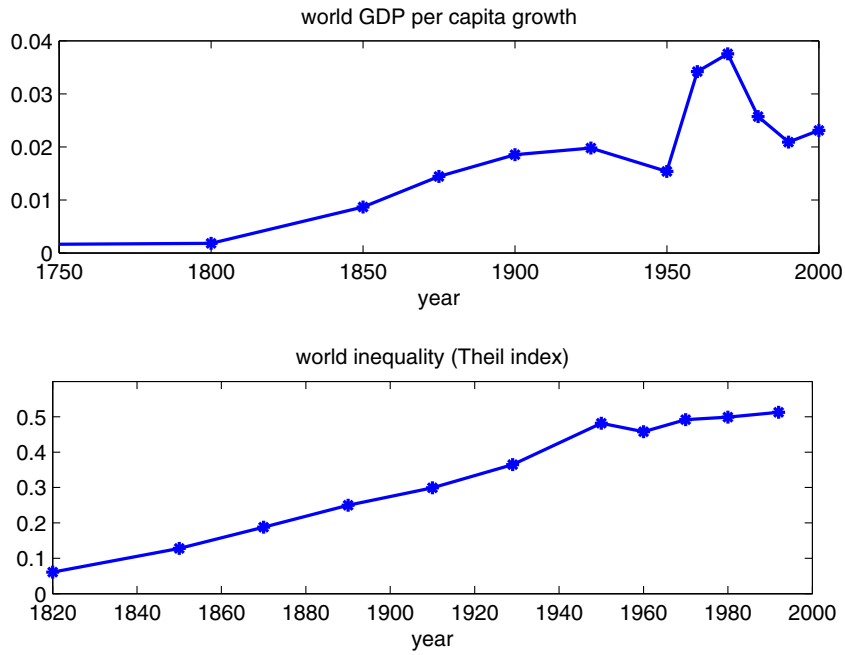
In premodern times, before the take-off to long-run growth of the countries that led the industrial revolution, national income differences were minuscule from today's perspective. Bairoch (1993, Ch. 9) reviews the literature and comes to the conclusion that, even in the mid-18th century, income of the future developed countries exceeded income of the future least developed countries by only factor 1.1 to 1.3. With the beginning of the industrial revolution, the world witnessed the "great divergence." Income inequality between countries, measured by the Theil index, increased from 0.06 in 1820 to 0.25 in 1870 to 0.48 in 1950 to 0.5 in 1980 (according to Bourguignon and Morrisson, 2002). Since then, the increase of inequality has slowed down to a point such that researchers speculate whether it has settled at a steady state or started to decline (e.g., Jones, 1997; Acemoglu and Ventura, 2002; Sutcliffe, 2004). Figure 1 shows the gradual increase of world income growth and the evolution of world inequality since the onset of the industrial revolution.[2]

Country-specific differences in the timing of the gradual take-off from stagnation to long-run growth are a major theme in unified growth theory (Galor, 2005). It is argued that the varying time of the take-off to growth contributed significantly to both increasing world inequality and the emergence of convergence clubs (i.e., clusters of countries that grow similarly with respect to each other but differently to other countries). Unified growth theory, however, largely focuses on countries conceptualized as closed economies, which implies, in particular, the notion that each country independently generated its own impulse for the take-off to growth.[3]

*Manuscript received December 2018; revised August 2019.
[Correction added on 24 October 2020, after first online publication: Copyright line and Projekt Deal funding statement has been added.]
[1] We would like to thank Jean-Louis Arcand, Francois Bourguignon, Clive Bell, Christiane Clemens, Guido Cozzi, Carl-Johan Dalgaard, Oded Galor, Jan Willem Gunning, Sophia Kan, Alessandra Pelloni, Robert Waldmann, David Weil, and three anonymous reviewers for helpful comments. Please address correspondence to: Holger Strulik, Department of Economics, University of Goettingen, Platz der Goettinger Sieben 3, 37073 Goettingen, Germany. Phone: +49(0)551/39-10614. Fax: +49(0)551/39-8173. E-mail: *holger.strulik@wiwi.uni-goettingen.de*.

[2] We refer to the inequality of GDP per capita between countries, which is the relevant measure for the theory developed below. For inequality of income between world citizens, the evidence is stronger that it actually declined since the 1980s due to the take-off of populous China and India (Sala-i-Martin, 2006).

[3] See Galor and Weil (2000), Kögel and Prskawetz (2001), Jones (2001), Galor and Moav (2002, 2004), Boucekkine et al. (2002), Doepke (2004), Strulik and Weisdorf (2008), and many others. The study by Galor and Mountford (2008)



NOTES: Data from De Long (1998) and Bouguignon and Morrisson (2002). Inequality is measured by the Theil index and as inequality between 33 countries or groups of countries whereby each country or country group represents at least 1% of world population or world GDP in 1950.

FIGURE 1

WORLD ECONOMIC GROWTH AND INEQUALITY [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

This article proposes a different approach. It considers a world of many countries connected by a network of knowledge exchange. As knowledge diffuses gradually through the world, more and more countries are "infected," their firms start investing in new technologies and their economy takes off to long-run growth. With more and more countries jumping on the bandwagon of growth, world income per capita increases gradually toward a balanced growth path. The individual timing of the take-off is explained by the countries' closeness to the leaders of the industrial revolution. Knowledge created in the leader countries is adopted earlier by countries connected directly or within only a few links, compared to poorly connected or "remote" countries. Take-off to growth of the forerunners of the industrial revolution is, naturally, accompanied by increasing world inequality, as the income gap with respect to the backward countries gets larger. Eventually, however, knowledge diffuses through the whole world and the remote countries also take off. Because the available knowledge has increased tremendously since the take-off of the original leaders, the latecomers of the industrial revolution have more to learn from and thus they take off faster, at rates that temporarily exceed the balanced growth rate. The feature that the growth rates of latecomers temporarily overshoot the balanced growth rate implies that *relative* world inequality eventually declines.

Most of the related literature focuses exclusively on relative inequality measured, for example, by the conventional Gini index or Theil index. One exception is Atkinson and Brandolini (2010) who discuss alternative measures of absolute inequality based on Kolm (1976) and find that it accelerated since the 1950s, that is, during the period when relative inequality leveled off. We show that our network theory of long-run growth captures this phenomenon as well. We compute the absolute Gini index, defined by the product of the Gini index and average income

is an exception, in that, it considers two interacting countries (or regions) in a unified growth setting. It investigates trade—but not knowledge exchange—and argues that the fact that countries are connected delays the take-off to growth of the initially backward country. Two-country models of imperfect knowledge diffusion and endogenous growth are proposed by Baldwin and Forslid (2000), Baldwin et al. (2001), Strulik (2014), and Davis and Hashimoto (2015).

(Chakravarty, 1988), and show that declining relative income inequality is predicted to be accompanied by increasing income gaps in absolute levels.

In order to focus on the knowledge diffusion process, the underlying economic model is a deliberately simple one. It is built on the multicountry model of knowledge diffusion and endogenous growth developed by Howitt (2000) and simplified by Acemoglu (2009, Ch. 18). The main difference is that in our world countries are not symmetric and do not exchange knowledge with all other countries alike. Instead, we assume that countries exchange knowledge with connected countries (their neighbors). We then investigate knowledge diffusion through a small world network and show how a great diversity of individual growth experiences evolves out of initial similarity between countries.

The network is necessary and sufficient for the complexity of individual growth performances to arise since this behavior does not occur in the original symmetric models of knowledge diffusion. In fact, the complexity of growth performances arises independently from the specific economic model of knowledge creation. In an earlier working paper (Lindner and Strulik, 2014), we considered economic growth based on human capital externalities and knowledge diffusion through networks and found that a small world network generates a similar complexity of growth performances. Our current approach of R&D-based growth and technological progress brought forward by market processes adds more realism and facilitates comparison with standard endogenous growth theory. The decisive element that explains the evolution of world inequality, however, continues to be the knowledge exchange through networks.

There exists a large literature on R&D externalities between countries. This literature usually involves a rather sophisticated modeling of households and firms but the way knowledge is exchanged between countries is straightforward and (most of) the analysis concerns the steady state (e.g., Eaton and Kortum, 1999; Howitt, 2000). In our article, in contrast, the economic model is straightforward but the process of knowledge diffusion is nontrivial and the analysis focuses on transitional dynamics.[4]

Our study is related to the work of Lucas (2000, 2009) on the initial divergence and subsequent convergence of income across countries. An important difference is that in Lucas' studies, countries have either full access or no access to world knowledge. A stochastic mechanism determines when countries gain access to world knowledge. According to our approach, in contrast, the economic take-off and subsequent growth of the leaders, followers, and trailers of the industrial revolution is endogenously explained by R&D effort and the imperfect diffusion of knowledge throughout the world. In Lucas' studies, the driving force of the growth process is assumed (as in neoclassical growth theory), whereas in our approach, it is explained (as in new growth theory). Lucas openly admits that his model is mechanical without much economic content and expresses his confidence that the mechanical predictions could be confirmed in refined theories of growth and development. Here, we propose such a refinement. It is based on a theory of endogenous technological progress and technology diffusion through a world network of knowledge exchange. Aside from addressing aggregate world growth and world inequality, the modeling of endogenous creation and diffusion of technology allow us furthermore to explain a richer set of phenomena than the mechanical model. For example, according to Lucas' approach, the United States would never had outpaced England, the initial industrial leader. In our setup, countries do not only temporarily diverge and converge but they also (occasionally) overtake each other.

Another strand of literature investigates multicountry models in which convergence is driven by capital accumulation and trade (e.g., Acemoglu and Ventura, 2002). Conceptually, the available multicountry growth literature focuses on the question of whether and how countries

[4] Klenow and Rogriguez-Clare (2005) survey the literature on knowledge externalities in economic growth and propose some extensions. In particular, they consider treating knowledge diffusion as being country-pair specific and depending on distance (but they do not pursue this approach very far, cf. pp. 852–3). Comin et al. (2012) propose a microfounded theory of spacial knowledge diffusion based on the random interaction of individuals. Their study is also indirectly supportive of our approach by empirically showing that knowledge diffuses slower to countries farther away from the technological leaders.

at initially different income levels converge while we also investigate how countries that were initially similar diverged. In other words, as with the available multicountry growth literature, we also share an interest in the question of where the steady-state cross-country income distribution lies. Additionally, as with unified growth theory, we share an interest in the question of how the presently observable diversity of growth experiences evolved out of an initial similarity between countries. In a unifying framework, our network theory of knowledge diffusion offers an explanation for both "the great divergence" as well as "the great convergence."[5]

There is a relatively small body of literature on networks in the context of economic growth. Cavalcanti and Giannitsarou (2017) investigate learning externalities between households (or schools) in simple networks and focus on convergence behavior. Fogli and Veldkamp (2012) provide a study on the role of network connectivity for the diffusion of knowledge and diseases. Lindner and Strulik (2015) investigate how economic development is affected by globalization conceptualized as an evolving network, that is, how decreasing local connectivity affects occupational choice and investment behavior through eroding trust and trustworthiness.[6]

The network through which knowledge diffuses is best conceptualized as face-to-face interaction of people. This notion is supported by a series of recent studies documenting the importance of short-term (Andersen and Dalgaard, 2011; Hovhannisyan and Keller, 2015) and long-term (Ortega and Peri, 2014) cross-border flows of people for total factor productivity (TFP) growth and economic growth. These studies find simultaneously little support for openness to trade as a separate channel of knowledge diffusion, therewith corroborating Frankel and Romer's (1999) suspicion that it is the exchange of ideas trough communication and travel instead of the shipment of goods through which openness generates international productivity spillovers. An increasing trend of knowledge exchange through increasing (business) travel is captured by our model by the prediction that the amount of knowledge diffusing through the world network increases over time. In an extension of the basic model, we consider that the network itself evolves over time in the sense that economic development induces the creation of new long-distance links.

The links between countries can be interpreted as geographic proximity (Mexico next to the United States) as well as cultural or genetic similarity (England next to the United States). The latter interpretation captures the notion that a similar language and cultural background facilitates the adoption of knowledge (Spolaore and Waziarg, 2013). These cultural links between countries may well have been established in preindustrial times, before the take-off to growth (Ashraf and Galor, 2013). Given this notion of the links between countries, the model's prediction is that countries that are well connected to similar countries experience an earlier take-off to growth.

The article is organized as follows. The next section sets up the model. Section 3 provides analytical results for comparative statics and comparative dynamics (steady-state, S-shaped transitions, overshooting growth of latecomers, rising and eventually declining world inequality). Proofs of the propositions are delegated to the Appendix where we also investigate the implied growth dynamics for some very simple networks in order to provide a better understanding of the impact of the network architecture on knowledge diffusion. In Section 4, we introduce the small world network (Watts and Strogatz, 1998), and in Section 5 we investigate the distribution and growth of world income when countries are connected by such a network,

[5] The term great divergence was initially coined by Pomeranz (2000) with respect to the divergent evolution of China and the West. It is now more broadly applied to the divergent evolution of income per capita across the world (Galor, 2005).

[6] Recently, Delventhal (2018) proposed a network model to study the evolution of the world income distribution. In contrast to this article and Lindner and Strulik (2014), the economic model is more detailed and the network considered is very specific. It is a weighted complete network, in which the weights are determined by transport costs between countries. The feature that this approach ignores other factors relevant for knowledge exchange (e.g., common language, cultural or genetic proximity) may be the reason why the model fails to explain the fast and early take-off of the United States and the other "neo-European" countries. In the small world model, cultural or genetic proximity of geographically distant countries is captured by the presence of long-distance links (short-cuts).

provide a sensitivity analysis with respect to the network parameters, and discuss the phenomenon of endogenously changing world economic leaders and overtaking in the course of global development. In Section 6, we discuss the robustness of results when the basic model is extended toward country-specific degrees of openness, different assumptions on learning from neighbors, and the endogenous creation of long-distance links in an evolving network. Section 7 concludes.

2. THE MODEL

Consider a world consisting of a number n of countries indexed by i. All Countries are populated by a (nonoverlapping) workforce L. The economic side of the model can be understood as a simplified version of the knowledge diffusion model of Howitt (2000) and Acemoglu (2009, Ch. 18). We follow these studies and assume that there is no international trade in goods or factors in order to focus on technology transfer as the main connection between countries. The novel aspect is the conceptualization of the world as a network of knowledge exchange. The model consists of three sectors: final goods production, intermediate goods production, and R&D. Growth is generated as in Romer (1990) by expanding variety of intermediate goods, which implies increasing productivity of the final goods sector. In deviation from the related literature, we assume that a successful innovation or adaptation of an intermediate good generates a monopoly for that specific good for one period (instead of for an infinite period). This allows for a convenient modeling of a state transition from no R&D toward R&D-based growth. The discrete time period of the model is thus given by the length of monopoly advantage after successful innovation or adaption of a new good (in the numerical part, a period takes 20 years).

2.1. *Goods Production.* The final goods sector produces competitively using labor and a range of intermediate products. At time t there is a continuum of N_{it} intermediate goods available in country i. Let $y_{it}(v)$ denote the quantity of input v. Output of final goods is then given by

$$Y_{it} = \frac{1}{1-\alpha} L^{\alpha} \int_0^{N_{it}} y_{it}(v)^{1-\alpha} \mathrm{d}v. \tag{1}$$

Let $p_{it}(v)$ denote the price of good v. From the first-order condition for optimal factor input, we obtain the demand function

$$y_{it}(v) = p_{it}(v)^{-1/\alpha} L. \tag{2}$$

As Acemoglu (2009), we consider a lab-equipment variant of R&D-based growth, implying that final goods are used for consumption, production of intermediates, and R&D. Implicitly the lab-equipment model assumes that R&D uses labor and other factor inputs in the same proportion as the manufacturing sector (Rivera-Batiz and Romer, 1991). Production of an intermediate good requires the input of $(1-\alpha)$ final goods. The latest vintage of available goods is supplied under monopolistic competition. All other vintages are supplied competitively. This means that the price of all but the latest vintage is given by $p_{it}(v) = (1-\alpha)$ such that factor input is $y_{it}(v) = (1-\alpha)^{-1/\alpha} L$. Profits of the $N_{it} - N_{it-1}$ firms supplying the latest vintage of intermediate goods are given by

$$\pi_{it}(v) = p_{it}(v) y_{it}(v) - (1-\alpha) y_{it}(v) = p_{it}(v)^{-1/\alpha} L[p_{it}(v) - (1-\alpha)]. \tag{3}$$

From the first-order condition for maximum profits, we obtain $p_{it}(v) = 1$ such that $y_{it}(v) = L$ and $\pi_{it}(v) = \alpha L$. We assume that all countries face the same constant interest rate, which is normalized to zero. The study of Caselli and Feyrer (2007) showed that there is indeed little variation of interest rates across countries. A constant world interest rate could be motivated

by perfect international mobility of capital. The assumption of a zero interest rate is made to simplify the algebra and could be relaxed without loss of generality.[7]

2.2. *R&D.* New goods are linearly created by spending z_{it} units of final goods on technology adoption such that the number of newly available products in period $t+1$ is given by

$$N_{it+1} = A_{it} z_{it} + N_{it}. \tag{4}$$

Following Acemoglu (2009), the expenditure on technology adoption z_{it}, may take the form of R&D but may also be conceptualized as other expenditures conducive to product creation and adoption. In short, we call it R&D. Productivity of R&D, A_{it}, is given for the individual firm but endogenously determined through knowledge externalities, $A_{it} = A_{it}(z_{it}, \cdot)$. Since there is free entry into R&D, the output of R&D (blueprints for new goods), given by $N_{it+1} - N_{it}$, is sold to firms in the intermediate goods sector at unit costs. Since there is also free entry to intermediate goods production, this means that the price of a blueprint equals expected profits $\pi_{it}(v) = \alpha L$. Free entry thus implies

$$A_{it}(z_{it}, \cdot) z_{it} \alpha L \leq z_{it}. \tag{5}$$

Whenever there is R&D, the constraint holds with equality.

2.3. *R&D Productivity and the Knowledge Network.* In this section, we describe in detail the features of country-specific productivity in R&D, A_{it}, and explain how it is determined by a network of knowledge exchange. Following Howitt (2000) and Acemoglu (2009), we assume that there is a "standing on shoulders" effect in R&D productivity, which increases in the number of varieties available domestically and in other countries. The related literature assumes that countries have access to the knowledge available worldwide. Here, we assume in contrast that countries do not exchange knowledge with all other countries alike. Specifically, we assume that countries are connected by a network of knowledge exchange. A link between two countries i and j thus means that these countries are open with respect to each other and that they are in mutual knowledge exchange.

Let the network of the world be represented by a matrix $\mathbf{W}$ whose elements indicate whether countries are linked with each other. We assume that links are unweighted and undirected. This means that the entry $w_{ij} = w_{ji}$ is equal to one if countries i and j are linked and zero otherwise. The nodes to which country i is linked are called neighbors of i. Assume country i has d_i links. By definition, each country is not linked to itself such that d_i can assume any value between 0 (isolation) and $n-1$ (connected to all other countries). Let $\tilde{n}_i$ denote the set of countries to which country i is linked to.

Let ϵ denote the share of international knowledge externalities, $\epsilon \in [0, 1]$. For $\epsilon = 0$, the model collapses to a conventional R&D-based growth model, in which countries are treated as if in isolation, and for $\epsilon = 1$ the model collapses to a simplified version of the Howitt (2000)–Acemoglu (2009) model. Knowledge spillovers from abroad are derived from the *externality matrix* $\bar{\mathbf{W}}$, which is obtained by normalizing $\mathbf{W}$ such that for every linked country, $d_i > 0$, the sum of weights to neighbors in $\bar{\mathbf{W}}$ is equal to ϵ, that is $\bar{w}_{ij} = \epsilon / d_i$ for $j \in \tilde{n}_i$. In case of isolation, $d_i = 0$ and we set $\bar{w}_{ij} = 0$ for all $j \neq i$. Finally, we assign $\bar{w}_{ii} = 1 - \epsilon$ for all i. Hence, all rows of $\bar{\mathbf{W}}$ have positive elements and sum up to one if every country has at least one link.

We define the standing-on-shoulders externality for country i as the average number of varieties available in its neighboring countries including the country itself, $\sum_{j=1}^{n} \bar{w}_{ij} N_{jt} = \epsilon \bar{N}_{it} +$

[7] Consideration of imperfect capital mobility and variable country-specific interest rates would require to explicitly introduce a savings motive (in an extension toward an overlapping generations framework) and would blur the focus on adjustment dynamics resulting from knowledge exchange through a network. See Strulik (2014) for a two-country model of imperfect knowledge exchange and imperfect capital mobility.

$(1-\epsilon)N_{it}$, in which the average available varieties of the neighbors of i is denoted by $\bar{N}_{it} = \frac{1}{d_i}\sum_{j\in\bar{n}_i} N_{jt}$. A "stepping on toes" externality captures the fact that R&D success gets harder when there is much spending on R&D (i.e., when z_{it} is large). As in the related literature, it prevents the economy from exploding. Summarizing, productivity in R&D is given by

$$A_{it} = \left(\frac{A\left[\epsilon\bar{N}_{it} + (1-\epsilon)N_{it}\right]}{\beta + z_{it}}\right)^{\phi}. \tag{6}$$

We assume that the productivity parameters A, β, and ϕ are equal across countries such that all differences between countries result either from their position in the network or their initial endowments N_{i0}. The parameter $\beta > 0$ ensures that the marginal productivity of the first unit spent on R&D is less than infinity. As a result of this plausible assumption, there exists for any country an environment in which market R&D and the creation of patented blueprints is not profitable, characterizing the state of the world for most of human history (Mokyr, 2005). The parameter β thus captures barriers to R&D that are important when the level of R&D is small but negligible when it is large. The parameter ϕ controls for potential scale effects in innovation.[8]

The intuition behind the use of country averages is that, at any time increment, any person in country i can exchange knowledge either with a person in country j or country k. If for, example, the English traveler who previously spent all of his travel time in the United States now splits his travel time between the United States and India, he spends only half of the former time on knowledge exchange with the United States, implying that the Indian engagement reduces the knowledge gained from the United States. The fact that aggregate time for knowledge exchange per country is normalized to unity then implies that the total knowledge acquired from abroad is measured by the average varieties developed by one's neighbors. Ceteris paribus, a link to a backward country (with $N_{jt} < N_{it}$) leads to a lower knowledge externality for country i, and a link to a forward country (with $N_{jt} > N_{it}$) implies a higher knowledge externality. This means that initially backward countries that are well connected to initially advanced countries have an advantage in learning from abroad. Notice that the model does *not* predict that a country's productivity worsens when it is connected to a country that has lower productivity than itself. Instead, all international links increase productivity, compared to autarky. The implication of (6) is that productivity decreases when a link to a forward country is replaced by a link to a backward country such that average number of varieties available in the neighboring countries declines.

In Section 6, we consider several alternative assumptions on knowledge diffusion and investigate (numerically) the robustness of results. First, we show that aggregate results are hardly affected if we allow the degree of openness ϵ to be country specific. Second, we assume that countries learn only from their most advanced neighbor. We show that this assumption preserves all qualitative results but generates less variety in individual growth performances, less overtaking, and less inequality than the benchmark model. Finally, we assume that the network is evolving, in the sense that economic advances lead to the creation of more long-distance links, and show that this feature preserves all results and increases the speed of catch-up growth.

[8] For $\bar{N}_{it} = N_t$ and $\epsilon = 1$ (i.e., the special case of learning from all countries in a complete network), the knowledge externality is similar but not equal to Acemoglu (2009). It deviates because the stepping-on-toes effect is measured in terms of research effort (as in Jones, 1995) instead of in terms of existing products and because of $\beta > 0$, that is, the assumption of less than infinite productivity of the first unit spent on R&D, which is necessary in order to allow for the possibility of no R&D and economic stagnation.

3. LONG-RUN DYNAMICS IN CONNECTED NETWORKS

From (5) and (6), we obtain R&D input

$$z_{it} = \max\left\{0,\ A\left[\epsilon\bar{N}_{it} + (1-\epsilon)N_{it}\right](\alpha L)^{1/\phi} - \beta\right\}. \tag{7}$$

Notice that a country is situated at the corner solution (no R&D) if R&D productivity is sufficiently low, which is the case if the country has not (yet) developed or adopted many products (N_{it} is sufficiently low) and—this is the novel result—if the country is badly connected to the rest of the world ($\bar{N}_{it}$ is sufficiently low). The latter happens if the country has no or few links to countries that have reached an advanced state of development characterized by a relatively high number of available products.

Inserting (6) and (7) into (4), we get a description of the world as one vector-valued difference equation,

$$N_{it+1} = N_{it} + \max\left(0,\ (\alpha L)^{-1}\left\{A\left[\epsilon\bar{N}_{it} + (1-\epsilon)N_{it}\right](\alpha L)^{1/\phi} - \beta\right\}\right), \tag{8}$$

$i = 1, \ldots, n$, and a network $\bar{\mathbf{W}}$. Note that in case of isolation, $\bar{N}_{it} = 0$, which is always harmful to growth. Throughout the remainder of the article, we assume that $\bar{\mathbf{W}}$ is connected.[9] If the underlying network were disconnected, its separate components would behave as separate (small) worlds.

Suppose a country invests in R&D. Then, the (gross) growth rate of varieties is given by

$$g^N_{it+1} \equiv \frac{N_{it+1}}{N_{it}} = 1 + (\alpha L)^{-1}\left\{A\left[\epsilon\frac{\bar{N}_{it}}{N_{it}} + (1-\epsilon)\right](\alpha L)^{1/\phi} - \frac{\beta}{N_{it}}\right\}. \tag{9}$$

Output of final goods is computed from (1) using symmetry within vintages of intermediate goods:

$$\begin{aligned} Y_{it} &= \frac{1}{1-\alpha}L^{\alpha}\left[\int_0^{N_{it-1}}\left[(1-\alpha)^{-1/\alpha}L\right]^{1-\alpha}\mathrm{d}j + \int_{N_{it-1}}^{N_{it}} L^{1-\alpha}\mathrm{d}j\right] \\ &= \frac{L}{1-\alpha}\left[N_{it-1}(1-\alpha)^{-(1-\alpha)/\alpha} + N_{it} - N_{it-1}\right]. \end{aligned} \tag{10}$$

Notice that output per worker Y_{it}/L is independent from scale for a given number of varieties. GDP is defined as $Y_{it} - (1-\alpha)X_{it} - Z_{it}$, in which X_{it} and Z_{it} are the aggregate factor demands for intermediate goods production and R&D. GDP can then be computed as $GDP_{it} = Y_{it} - (1-\alpha)L[(1-\alpha)^{-1/\alpha}N_{it-1} + (N_{it} - N_{it-1})] - Z_{it}$.

Let $N^w_t \equiv \sum_{i=1}^n N_{it}$ and $\bar{N}_t \equiv N^w_t/n$ denote the worldwide produced varieties of intermediate goods and the average number of varieties per country, respectively. Furthermore, let $g^N_t \equiv N^w_{t+1}/N^w_t$ denote the (gross) growth rate of the world's varieties and $x_{it} = N_{it}/\bar{N}_t$ country i's relative endowment, that is, the ratio between country i's varieties and average varieties. The dynamics of relative varieties are then determined by $x_{it+1} = g^N_{it}/g^N_t \cdot x_{it}$.

DEFINITION 1. A steady state is defined by $x_{it+1} = x_{it}$ for all $i = 1, \ldots, n$ and all $t \geq 0$. A balanced growth path is defined by each country growing at the same constant growth rate, that is, $g_{it} = g$.

COROLLARY 1. *When all countries start with equal initial endowment N_0, the network is irrelevant and the economy is always at the steady state where $x_{it+1} = x_i = 1$ for all $i \in n$ and $t = 1, 2, \ldots$.*

[9] A network is connected if there exists a path between any two nodes. Formally, this means that for any i and j, $i \neq j$, there exists a $k \geq 1$ such that $(\bar{\mathbf{W}}^k)_{ij} > 0$.

PROPOSITION 1 (LONG-RUN GROWTH). *The world economy converges toward a steady state of growth or stagnation. In case of positive long-run growth, the growth rate of varieties and final goods output is given by* $g^N = (\alpha L)^{\frac{1-\phi}{\phi}} A + 1$. *A sufficient condition for long-run growth is that all countries are endowed with a number of varieties greater than* $\beta/[A(\alpha L)^{1/\phi}]$.

For long-run growth, the proof shows for (9) that $\bar{N}_{it}/N_{it} \to 1$ for $t \to \infty$. This provides g^N. In order to see that Y_{it} grows at the same rate as N_{it}, insert $g^N = N_{it}/N_{it-1}$ into (10) and obtain that along the steady state $Y_{it}/N_{it} = L(g^N)^{-1}[(1-\alpha)^{-(1-\alpha)/\alpha} + g^N - 1]$. The right-hand side of the equation is constant, implying that numerator and denominator on the left-hand side of the equation grow at equal rates.

Since the network is connected, all knowledge is eventually shared by all countries. This feature implies that the world economy converges toward a steady state, that is, a situation in which all countries grow at a common rate. A steady state of positive growth means that all countries of the world irrespective of their backward initial situation are eventually "infected" by knowledge diffusion and will grow eventually at the same rate as the leaders of the industrial revolution.

A stylized fact of long-run development is that countries gradually, with increasing growth rates, take off to modern growth (the new Kaldor fact number 2; Jones and Romer, 2010). Both neoclassical growth theory and conventional endogenous growth theory have difficulties in predicting a gradual take-off. The neoclassical model, for example, predicts that growth is highest at low levels of income, a feature that follows immediately from decreasing returns to factor accumulation. Here, the model generates S-shaped transitions. The growth rate of GDP per capita accelerates gradually during the first phase after take-off. During the second phase growth decelerates, a phenomenon that renders convergence toward the steady state.

In order to see the S-shaped transition of growth rates, inspect the growth equation (9). At the beginning of the take-off to growth, the second term in curly parenthesis, β/N_{it} dominates and growth is barely positive. Productivity in R&D is low because the country has developed (or adopted) relatively few intermediate goods. As the country develops, the second term vanishes to zero. The first term, in contrast, increases initially for followers of the industrial revolution, driven by the externality ratio $\bar{N}_{it}/N_{it}$. At the time of the take-off of economic growth, a country is poorer than the average of its neighborhood, implying that the neighborhood invests more in R&D and has developed more varieties. Altogether, this means that growth is increasing during the early phase after the take-off. As the country gets richer, the second term vanishes and $\bar{N}_{it}/N_{it}$ declines to unity. Along the transition, growth rates of initially backward countries overshoot the balanced growth rate.[10]

PROPOSITION 2 (OVERSHOOTING GROWTH). *Suppose that the network* $\bar{\mathbf{W}}$ *and the initial conditions support a steady state of long-run growth,* $g^N > 1$.

(i) Forerunners of the industrial revolution converge monotonically toward g^N.
(ii) Followers of the industrial revolution converge nonmonotonically at growth rates that are temporarily above g^N *if their initial endowment of varieties is small relative to the neighborhood average.*

Intuitively, for initially backward countries, there is much to learn from other countries or, more precisely, from the countries to which a link of knowledge exchange exists (i.e., the neighbors). The opportunity to tap into a greater pool of knowledge creates an advantage of backwardness (Gerschenkron, 1962). When R&D becomes profitable, these countries reach a phase of above steady-state growth because of the high learning potential from the neighbors. This means that the initially backward countries manage to double their income per capita in a

[10] In an early analytical contribution, Wälde (1996) has shown that in a two-country model of R&D-driven growth and capital market integration, the initially backward country catches up with overshooting growth rates.

much shorter amount of time than the leaders of the industrial revolution (Parente and Prescott, 2005).

In order to develop a comprehensive picture of the evolution of world income inequality, we distinguish between relative and absolute income inequality. In order to see the difference, consider a world of two countries with endowment $(y_1, y_2) = (10, 40)$. Assume the endowment changes to (20,80). This means that the absolute gap increases from 30 to 60, whereas the relative difference 30/50 stays constant. Relative income inequality can be expressed by the Gini index (or the Theil index), whereas absolute income inequality can be measured by the *absolute* Gini index, defined as the product of the Gini index and average income (Chakravarty, 1988). In the present example of two countries, the Gini index is 0.3 for both distributions but the absolute Gini index changes from 7.5 to 15. Lemma A.3 in the Appendix summarizes the main properties of these measures of inequality.

PROPOSITION 3. *[The World Kuznets Curve]*

(i) Relative income inequality between countries eventually vanishes. The Gini index converges to zero such that the world tends to the unique steady state of relative equality $x_i^* = 1$ for all $i \in N$.

(ii) If in case of long-run growth, some countries initially grow and others stagnate at a constant level of income, then relative income inequality increases initially and declines subsequently.

The feature that the network is connected is sufficient for convergence toward a balanced growth path of relative equality. The specific architecture of the network, however, determines whether this will be a path of positive growth and how fast convergence will be. An obstacle to long-run growth could be, for example, when the leader is linked to too many initially poor countries such that there is no take-off to long-run growth and $g^N \leq 1$. The sufficient condition of Proposition 1 represents a lower bound on initial endowments ensuring take-off for all network configurations. Furthermore, the network architecture determines the speed of convergence toward balanced growth (see Proposition 5) and thus also how fast convergence toward relatively equality will be.[11]

The model produces not only a "great divergence" (Pomeranz, 2000), initiated by the take-off of the leaders of industrial revolution, but also a "great convergence" in terms of relative income levels. Convergence occurs after the take-off of the latecomers of the industrial revolution. The latecomers are identified as the countries with inferior initial endowments *and* missing links to the forerunners of the industrial revolution. Since a connected network ensures the existence of a steady state, it implies that eventually all knowledge is shared between all countries, which explains the phenomenon of vanishing relative income inequality. This result is in disagreement with some popular articles on the world income distribution (Jones, 1997; Acemoglu and Ventura, 2002) but it is in line with Lucas' (2000, 2009) vision of the world's future development. However, the decline of relative income inequality does not imply that absolute income levels converge. In fact, as we show later, countries may even overtake each other (several times) and the absolute income gap may increase, whereas income inequality measured by the (relative) Gini or Theil index disappears.

PROPOSITION 4. *[No Convergence in Levels] Suppose that the network $\bar{\mathbf{W}}$ and the initial conditions support a steady state of long-run growth, $g^N > 1$. Despite eventually declining relative world inequality there is not necessarily convergence of income levels.*

[11] The rate of convergence measures the speed at which a convergent sequence approaches its limit, measured by the difference to the stationary distribution in comparison to the previous period. Formally, for multidimensional systems with vector $x^t \in \mathbb{R}^n$ at time t and limit x^*, the rate of convergence is defined as $\lim_{t\to\infty} ||x^{t+1} - x^*||/||x^t - x^*||$.

The result in Proposition 4 means that, along the transition to the steady state, the relative Gini index always tends to zero but not necessarily the absolute Gini index. In order to get an intuition of this insight, consider the left term in the maximum argument of (8),

$$A(\alpha L)^{1/\phi-1}\left[\epsilon \bar{N}_{it} + (1-\epsilon)N_{it}\right], \tag{11}$$

which can be read as a composition of two functions. The "inner" operation in square brackets averages over neighborhoods and therefore contracts the range of different levels of N_{it}. The "outer" operation of multiplication by a constant, $A(\alpha L)^{1/\phi-1}$, magnifies this range again. The classical Gini index is based on relative measures such that the outer operation of multiplying by a constant is always mitigated. The *absolute* Gini index, however, fails to tend to zero if the contraction effect of averaging over neighborhoods is offset by the (repeated) multiplication by the constant $A(\alpha L)^{1/\phi-1}$. The dynamics in (11) gets more complex by subtracting the constant $\beta/(\alpha L)$ and by the fact that the whole right term in (8) is an updating term, added to the current level of N_{it}. Nevertheless, disentangling the effects as in (11) gives a hint as to why the absolute and relative Gini index can behave differently.

Cavalcanti et al. (2016) introduce the notion of *network cohesion* κ for a broad class of dynamic models of endogenous perpetual growth with network externalities. They show that this statistic is relevant for characterizing the stability and the speed of convergence when the analysis is carried out in terms of relative variables like $x_{it} = N_{it}/\bar{N}_t$ in the present article. Network cohesion is defined as one minus the second largest modulus eigenvalue of $\bar{\mathbf{W}}$, in particular,

$$\kappa = 1 - \max_{\lambda_i \in \sigma(\bar{\mathbf{W}}) \setminus \{1\}} |\lambda_i|, \tag{12}$$

where $\sigma(\bar{\mathbf{W}})$ is the spectrum of $\bar{\mathbf{W}}$. Cohesion is generally a measure between 0 and 1. In our case, network cohesion is always positive since the network $\bar{\mathbf{W}}$ is connected. The complete network has the largest possible network cohesion $\kappa = 1$ as all eigenvalues of $\bar{\mathbf{W}}$ besides the largest one are equal to 0. The empty network provides $\kappa = 0$ as in this case the eigenvalue 1 of $\bar{\mathbf{W}}$ has algebraic multiplicity n. It is easy to show that the star network has cohesion $\kappa = 1/2$ (see Corollary 1 of Calvacanti et al., 2016).

PROPOSITION 5. *Higher cohesion implies faster convergence to the balanced growth path. In particular, an upper bound for the rate of convergence is given by*

$$1 - \frac{g^N - 1}{g^N}\kappa.$$

In order to investigate transitional dynamics in more detail, we next turn to a numerical presentation of the model.

4. SETUP OF THE SMALL WORLD MODEL

4.1. *The Initial State of the World.* Suppose that initially there are two distinct groups of countries. A small group of countries with relatively high initial endowments (the rich) and a large group with relatively low endowment (the poor). Initial endowments are such that rich countries are growing, albeit at a very low rate, whereas poor countries stagnate because aggregate productivity is so low that investment in R&D is not worthwhile. This setup is the most interesting case because it allows for evolving country heterogeneity. As time proceeds and knowledge crosses borders, income and productivity of the countries grow differently according to their connections with other countries and countries become more dissimilar with respect to economic growth. Having two different groups of countries is the *minimum setup* to discuss evolving heterogeneity (cf. Corollary 1). We do not ask where the initial differences between countries come from but assume, in line with the historical evidence on economic

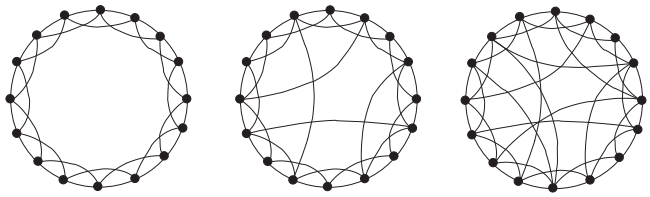


FIGURE 2

SMALL WORLD NETWORK

conditions in premodern times, that the initial differences are small from today's perspective. The challenge is thus to explain how a great variety of growth performances evolves out of small initial differences.[12]

4.2. *The Small World Network.* In Appendix A.2, we explain how knowledge diffuses through selected simple networks (bridge, ring, core-periphery). It is shown that these networks are sufficient to generate a staggered and gradual take-off to growth and overshooting growth patterns of the latecomers in this process. The simple networks are, however, insufficient to generate a variety of distinct growth processes at the country level and they cannot be used to fit predicted world economic growth and income inequality to the actual long-run trends shown in Figure 1. Here, we show that the small world network is capable to produce these features.

The small world network, developed by Watts and Strogatz (1998) is an irregular network that features both local connectivity and long-distance links. Mathematically, it is easily understood but complex enough to allow for an application to a plethora of biological and social phenomena (see Newman, 2003, for an overview). The small world model appears to be particularly suited for our purpose because it retains the importance of local connectivity, capturing the fact that most knowledge diffuses from direct neighbors, but at the same time allows for the establishment of long-distance links between distant countries.

Here, we consider a modification of the Watts and Strogatz model, developed by Newman and Watts (1999), which appears to be more appropriate for our purpose. The idea of the small world model can be illustrated best by considering a network on a one-dimensional lattice. It is constructed from a regular network in which any node (country) is connected with its direct neighbors that are m or fewer lattice spaces away. In the example of Figure 2, $m = 2$. Each country is connected to four neighbors, two at each side. The regular network is then modified by randomly adding long-distance links. The probability for a long-distance link per link of the underlying lattice is denoted by p. The middle panel of Figure 2 shows an example for which p is low and the panel on the right shows an example for larger p.

For international knowledge flows, the feature of local connectivity, created through positioning the countries on a ring, captures the empirical fact that knowledge spillovers, in principle, decline with geographic distance (e.g., Keller, 2002). The presence of long-distance links means that this generality is occasionally broken and that the effective distance is (much) shorter than geographic distance. Figuratively speaking, we could imagine the United States to be geographically only two neighbors away from Guatemala but exchanging much more knowledge with England because both countries are connected with a long-distance link. This may turn out to be crucial for comparative development because the United States benefits directly from knowledge created in England, whereas Guatemala benefits only indirectly via the United States. Moreover, in order for the knowledge to arrive in Guatemala, it has to cross Mexico,

[12] In contrast to conventional unified growth theory, the model is too simple to explain how the leader countries left stagnation. Here, we simply give the leader countries a small head start that supports an initial growth process. We think that this approach is justifiable because of our focus on the differentiated take-offs and growth experiences across countries that happened *after* the take-off of the leader countries.

another initially poor country, such that a part of the knowledge created in England gets "lost in transition."[13]

4.3. *Numerical Specification.* We begin with a benchmark specification of the model. Later on, we discuss the sensitivity of results on parameter choice. Suppose the world consists of 100 countries of which 10% are initially rich. We set $\phi = 1$ in order to eliminate scale effects. We set the labor share α to 0.65 and adjust the value of A such that the implied steady-state growth rate is about 2% annually. The parameter values of ϵ and β are irrelevant for the steady state but shaping adjustment dynamics. Eaton and Kortum (1999) estimate, for a sample of fully developed countries, that between one-half and three-fourths of the knowledge adopted has been generated abroad. We take the benchmark value for our (temporarily) more heterogenous set of countries from the lower bound of their estimates and set $\epsilon = 0.5$. This means that one-half of the knowledge available in a country has been generated by domestic firms and the other half stems from international knowledge diffusion. We set β in order to get the best fit of worldwide economic growth with the historical data.

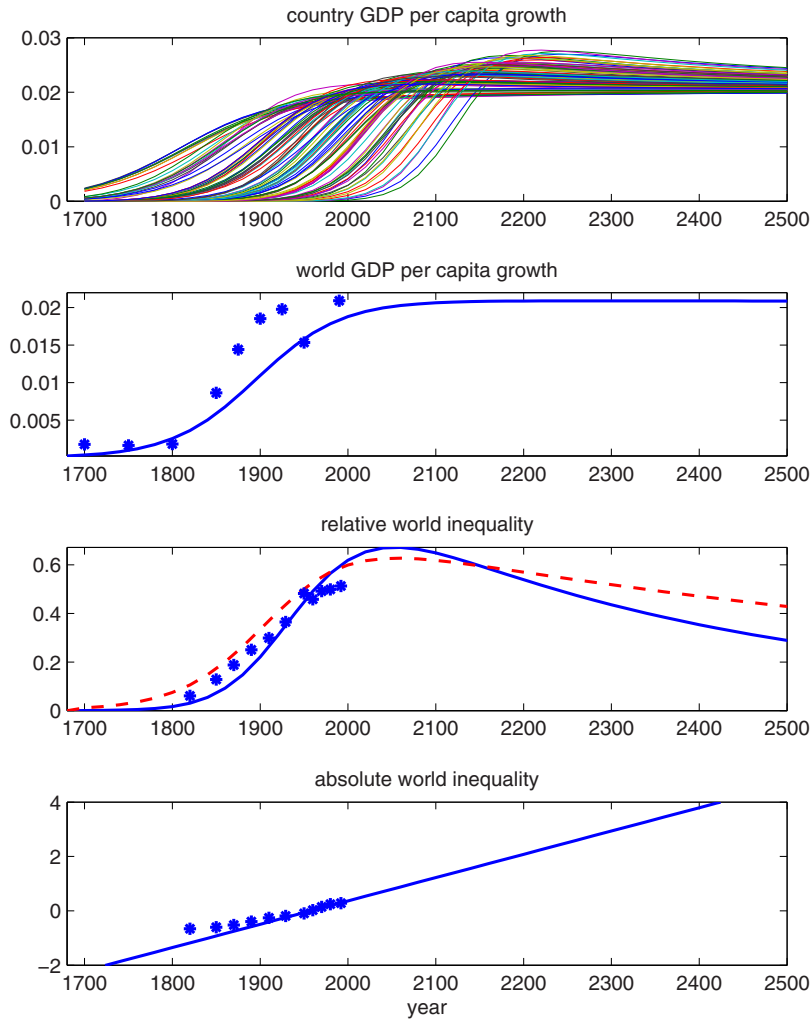
We assume that a period takes 20 years. After running the model, we convert results into annual data for better comparability with real data. We set initial time to the year 1700, that is, shortly before the onset of the first industrial revolution (Encyclopedia Britannica, 2018). We set $N_{it} = 10$ for the poor and $N_{it} = 11$ for the rich countries implying that income in the rich countries is initially about 1.2 times higher than in poor countries. This gap corresponds well with the estimates of the head start of Western European countries vis-à-vis the rest of the world at the dawn of the first industrial revolution (Bairoch, 1998, Ch. 9). Most importantly, this specification means that poor countries initially stagnate, whereas rich countries initially grow at a low rate of around 0.3%. As a benchmark, we consider $p = 0.3$, that is, the case in which 30% of the countries are equipped with a long-distance link. We subsequently provide sensitivity analysis with respect to p and other important parameters. For the basic run, we assume that the initially rich countries are clustered such that no initially rich country is surrounded by poor neighbors.

5. ADJUSTMENT DYNAMICS: THE EVOLUTION OF INNOVATION AND INEQUALITY ACROSS THE WORLD

5.1. *Basic Results.* The predicted adjustment dynamics for the benchmark case are shown in Figure 3. For better comparability with the data, the gross growth rate per 20 years from (9) is converted into net growth per year. In contrast to the simple networks discussed in the Appendix, the small world generates a lot of heterogeneity. Basically, each of the 100 countries follows its own idiosyncratic growth trajectory. Recalling that initially, in the year 1700, there were only two different types of countries and that the initial difference between rich and poor countries was small (about 1.2:1), we conclude that, with industrialization, diversity evolves out of similarity.

Naturally, the initially richer countries take off first. Next, follow the poor countries that are well linked with the rich part of the world, either through geographic proximity on the ring or through long-distance links. The less well-connected countries take off late but experience an "advantage of backwardness" (Gerschenkron, 1962) in the sense that their income growth surpasses the income growth of the forerunners of the take-off to modern growth. Generally, we observe that overshooting growth is higher, the later the take-off time is. Although it took about 200 years for the forerunners of the industrial revolution to reach a growth rate of 2%, the countries taking off in the 1950s needed only about two generations to achieve the same rate of growth. The explanation is that latecomers, once growth is initiated, tap into a greater

[13] For evidence on imperfect knowledge spillovers and their decay with distance see also, among others, Jaffe et al. (1993) and Coe et al. 2009. For the influence of cultural (linguistic, genetic) distance on knowledge diffusion see, among others, Keller (2002), Spolaore and Wacziarg (2009, 2013), and Madsen and Farhadi (2018).



NOTES: Solid Lines: model prediction. Dots in the second panel: world GDP growth according to De Long (1998); Third panel: solid line: Theil-index; dashed line: Gini index; dots: Theil index from Bourguignon and Morrison (2002); Fourth Panel: log of absolute absolute Gini index; dots: absolute Gini from Atkinson and Brandolini (2010). Sample size: 100 Countries. Small world based on ring with 1 neighbor to left- and right-hand side; long-distance links added with probability $p = 0.3$.

FIGURE 3

ECONOMIC GROWTH IN A SMALL WORLD [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

reservoir of world knowledge. This knowledge has been accumulated in the recent past and was not yet available when the forerunners took off. This phenomenon relates the model to the new Kaldor fact no. 1: the increasing flow of ideas via globalization (Jones and Romer, 2010). Globalization here means that an increasing share of countries gets out of stagnation with R&D (or, more broadly, investments in technology adoption) becoming worthwhile and that an increasing stock of knowledge diffuses through the world network.

Comparing the model's predictions with the historical facts (Bairoch, 1993; Galor, 2005), we would imagine the group of initially rich countries as Western Europe, which reaches on average a growth rate of 1% in the mid-19th century, a period in which some of the Latin American countries started to grow. In the 20th century, when the latecomers take off, the initially rich countries grow at an almost constant rate of about 2% annually. It is also interesting to observe that growth of the leaders is already surpassed by growth of some followers in 19th century and

that despite the presence of long-distance links, some countries are predicted to take off only in the 21st century. The differentiated and relatively rapid take-offs of the latecomers of the industrial revolution in the 20th century produce the picture of a great variety of subsequent growth experiences of countries that were almost equally poor just a generation ago.

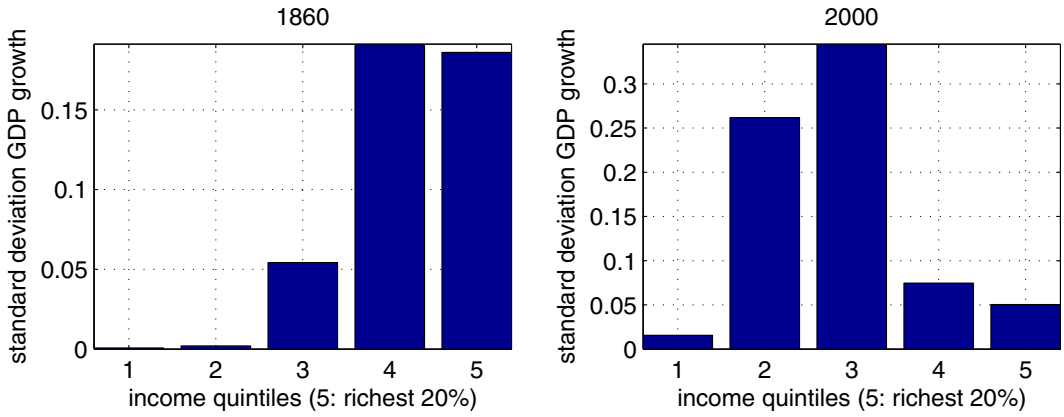
The second panel of Figure 3 shows the implied average economic growth in the world. Dots represent the data points from De Long (1998) shown in Figure 1. The model predicts the take-off of aggregate world growth reasonably well. World growth rises from almost zero to just below 1% in the mid-19th century and to about 1.5% in the mid-20th century. Compared to the data, the take-off is somewhat too slow, an outcome that could be corrected (by assuming a higher p or ϵ) at the expense of predicting a take-off that is "too early" for the latecomers. Altogether, however, the model generates plausible S-shaped transitions. On the individual level, as well as on the global level, the model provides an explanation for the new Kaldor fact no. 2, the gradual increase of the rate of economic growth (Jones and Romer, 2010).

The differentiated take-off of countries produces the Great Divergence: relative world inequality increases strongly from 1800 to 2000. This is shown in the third panel of Figure 3, in which dots represent the data points from Figure 1 (Bourguignon and Morrison, 2002). The solid and dashed line, respectively, show the model's prediction for the evolution of the Gini index and the Theil index, computed from the individual income trajectories of the 100 countries. According to the model, for its benchmark calibration, relative inequality stops growing in the 21st century. From then onward, the model predicts a "great convergence." As more and more latecomers catch up with overshooting growth rates, relative world inequality declines. The inequality curve, however, is skewed. The great convergence is predicted to take several centuries longer than the great divergence. The intuition is straightforward. The fact that the original leaders of the industrial revolution keep growing makes the catch up harder than the quick departure of the leaders from the almost stagnant income of the followers and latecomers two centuries earlier.[14]

The focus on the conventional Gini index, however, conceals that absolute world inequality keeps on rising. The bottom panel of Figure 3 shows the absolute Gini, that is, the relative index from the third panel multiplied by mean income. The log-scaling means that absolute inequality grows exponentially. These findings illustrate Proposition 4 and highlight the importance of distinguishing relative and absolute convergence. The relative income gap between rich and poor tends to zero because the absolute gap grows slower than the total level of income (cf. Lemma A.1). Dots in the bottom panel show the absolute Gini index computed for the Bourguignon and Morrison (2002) data by Atkinson and Brandolini (2010). The network model somewhat underestimates absolute inequality in the early 19th century but gets the exponential increase over the 20th century about right. It predicts this trend to continue in the future.

5.2. *Variation in Growth Rates.* In order to further explore the evolution of growth, we sorted the countries for any time t into income quintiles, with the poorest 20% of countries in the first quintile and the richest 20% in the fifth quintile. The variability of growth rates is shown in Figure 4, exemplarily for the years 1860 and 2000. The figure shows the standard deviation (in percent) of GDP growth for each income quintile. The panel for the year 2000 corresponds with the new Kaldor fact no. 3, stating that the variance of growth rates across countries increases with distance to the technological frontier (Jones and Romer, 2010). The variation in growth rates is low in rich countries compared to the "emerging economies" of the second and third income quintiles. Only in the poorest countries, which are still close to stagnation, variation in growth rates is smaller. The model also highlights that Kaldor fact no. 3 is a phenomenon of the 20th century. In the 19th century, when the frontier countries themselves sequentially experienced their take-offs to growth while the rest of the world was still close to subsistence, the variance of growth rates was highest among the rich countries.

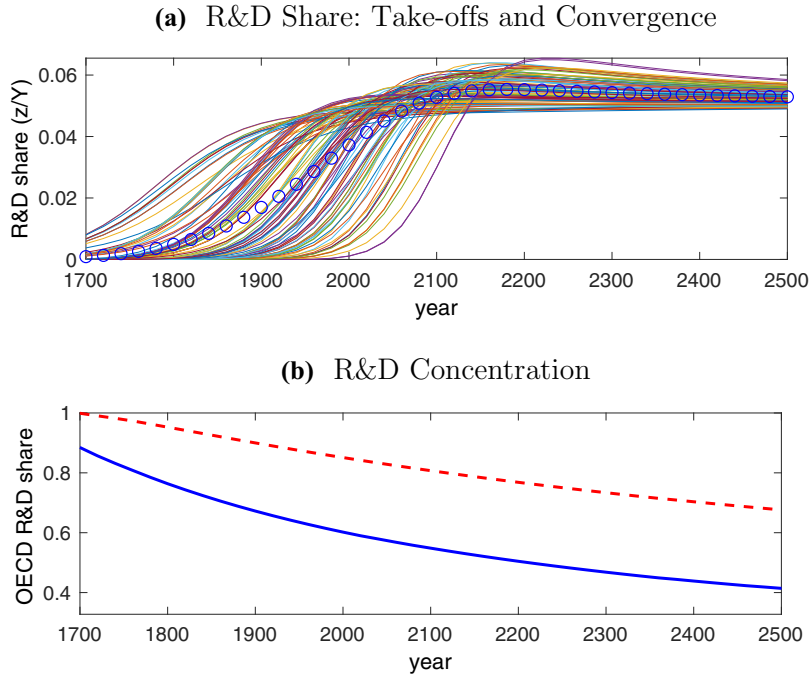
[14] In the Appendix, we provide results on an alternative measure of inequality, σ–convergence. In Subsection 6.2, we show how an evolving network accelerates the speed of convergence.



NOTES: The figure shows the predicted standard deviation (in percent) of GDP growth for income quintiles in the years 1860 and 2000.

FIGURE 4

STANDARD DEVIATION OF GDP GROWTH ACROSS COUNTRIES [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

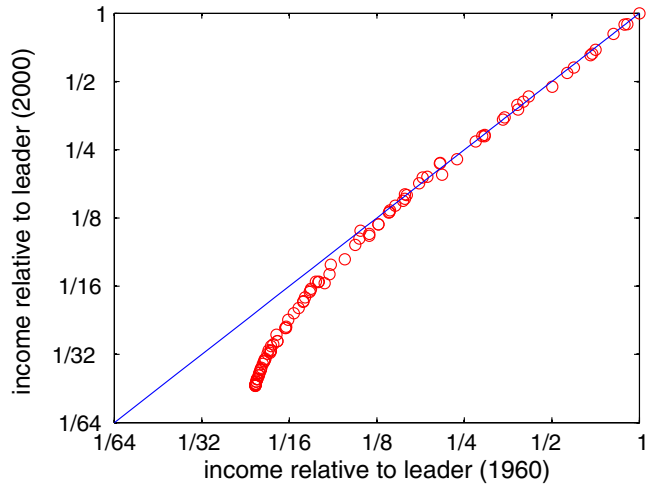


NOTES: (A) Circles indicate the world average R&D share; (B) predicted share in world R&D of the 15 richest countries (blue solid line) and the 30 richest countries (red dashed line).

FIGURE 5

R&D SHARE: TAKE-OFFS AND CONVERGENCE [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

5.3. *R&D Dynamics.* The economic take-off is associated with a take-off of R&D activity, at the country level as well as worldwide. Figure 5 shows the predicted output share of R&D, z_{it}/Y_{it}, for all countries. The world average R&D share is indicated by blue circles. On average, the R&D share rises gradually from about one per mill in 1720 to about 0.5% in 1820 to about 2% in 1950 and about 3% in 2000. Hence, the model gets the order of magnitude of the R&D share about right. The staggered and gradual take-off of R&D is in line with the observation of Lederman and Maloney (2003). The estimates of Lederman and Maloney, however, suggest



NOTE: Country GDP relative to leader country benchmark model; loglog scale.

FIGURE 6

INCOME RELATIVE TO LEADER: 1960 VS. 2000 [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

a convex association of the R&D share and economic development, that is, exploding R&D activity. Our model, in contrast, predicts a convex–concave association with gradual convergence toward an R&D share of about 5%.

The relatively fast catch-up (and overshooting) of the R&D share of latecomer countries is consistent with the observation that, in absolute levels, the bulk of R&D is performed by the richest countries, which are mainly the forerunner countries. According to OECD (2019), in 2015, the OECD countries accounted for about 72% of world R&D spending. In Figure 5(B), the blue (solid) line shows the predicted share in world R&D of the 15 richest countries. The red (dashed) line shows the share of the 30 richest countries. We observe a high concentration of R&D in the sense that most of R&D is performed by the richest countries. Their importance, however, declines over time as more emerging economies not belonging to the OECD club take up R&D. In the year 2015, the 15 richest countries account for about 60% of world R&D and the 30 richest countries account for about 85% of R&D. This means that, in a stylized way, the model gets the actual R&D concentration about right.

5.4. *Is the Present World Income Distribution Close to Its Steady State?* As evidenced in Figure 3, the model predicts that relative income inequality across the world will eventually decline after the take-off of the latecomers of industrialization. In contrast to such an optimistic outlook, some related studies developed theories in order to explain a *constant* world income distribution at a state of high inequality, most notably perhaps the study of Acemoglu and Ventura (2002). Acemoglu and Ventura's work was inspired by the observation of "a relatively stable" world income distribution in the second half of the 20th century.

A "relatively stable" distribution, however, could also be inferred from an actually slowly evolving distribution. This is particularly the case if the window of observation is relatively short and if the observation happens to be taken at a period of time when the trajectory of relative world inequality is flat because it is close to its maximum. In order to verify this claim by way of example, we compute for the outcome from the benchmark economy a relative income plot similar to the one displayed in Jones (1997, Figure 2) and Acemoglu and Ventura (2002, Figure 1).

Specifically, we compute from the time series shown in Figure 3 the relative income with respect to the leader country in the year 1960 and in the year 2000 and plot the result on a loglog scale, as shown in Figure 6. In accordance with the earlier studies, we observe little deviation from the 45 degree line. Aside from the poorest countries (for which the model predicts

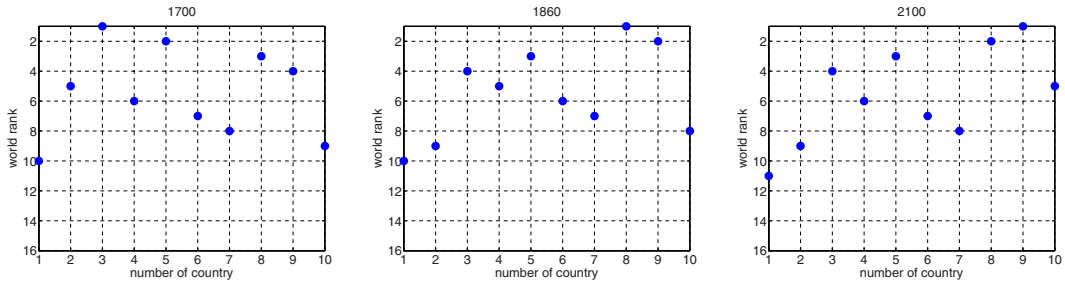


FIGURE 7

WORLD RANKING POSITION FOR COUNTRIES 1–10 [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

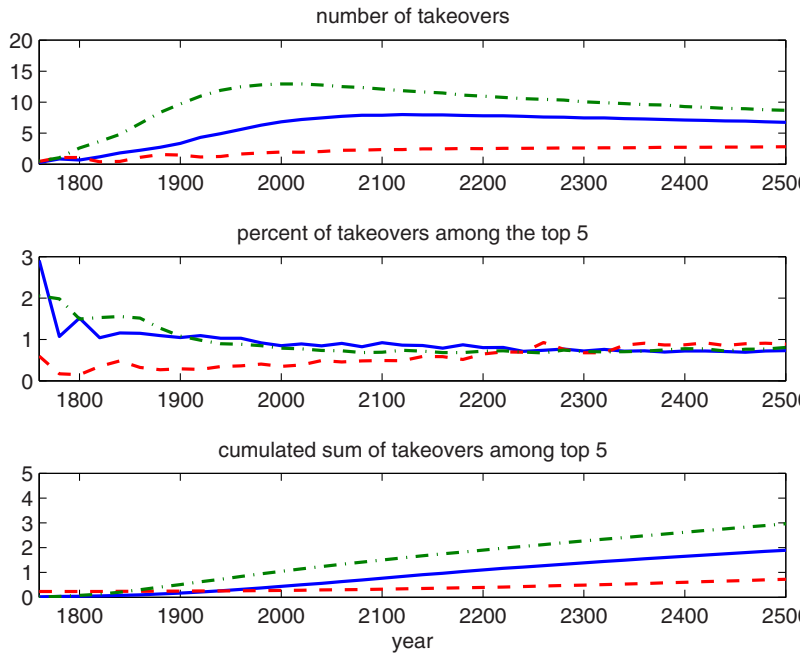
divergence), relative income in 1960 is a good predictor of relative income in 2000. Confronted with this picture alone, one could indeed be tempted to conclude convergence toward a constant unequal world income distribution. In fact, however, we know from Proposition 3 that income relative to the leader country moves to unity for all countries as time goes to infinity. This convergence process, however, is very slow and not discernable within a 40-year time window. The observation of an (almost) stable distribution of high relative inequality is consistent with a moving distribution toward relative equality.

5.5. *Overtaking and Falling Behind.* The phenomenon of income convergence is at the center of modern growth economics. The phenomenon of overtaking, however, is less frequently investigated in the context of endogenous growth. The original leapfrogging literature (Brezis et al., 1993) generated overtaking by the assumption that new technologies are less productive in the leading countries (the leading industry). Some researchers modeled overtaking in a purely stochastic context of Markov chains of income distributions, see, for example, Jones (1997). Others considered overtaking as a one-time event reflecting growth traps for initially leading countries (Acemoglu et al., 2006). Here, in contrast, overtaking is endogenously generated as knowledge flows through the network, that is, it is neither based on technological assumptions or stochastic elements nor does it imply nonconvergence (of relative income levels).

In order to demonstrate this extraordinary behavior, we perform the following numerical experiment with an example economy: We follow the 10 initially richest countries, named $1, 2, \ldots, 10$, along the way toward the steady state and visualize their position in the world income ranking. Figure 7 shows the resulting "income ladders" for three different years. For example, a dot at the (1,10) position in the 1700 diagram means that country 1 was ranked 10th place in the year 1700.

In the numerical example, country 3 leads the world income ranking in the year 1700. Obviously, it was favorably connected with other rich countries. By the year 1860, country 3 gave up the lead to country 8. Interestingly, country 8 is not a direct neighbor of country 3 but obviously it benefitted from favorable connections with quick followers of the industrial revolution. We also observe that country 2 falls behind whereas country 9 advances. In 2100, country 9 is at the top whereas country 1 fell out of the top 10 altogether. These changes in rank are explained by the changing advantage of links as knowledge is accumulated and diffused through the network. For example, an initially rich country connected only to one other initially rich country, which in turn is connected only to latecomers of the industrial revolution, grows initially fast and then slows down. It is overtaken by a country that is connected with initially poor countries, which are, however, well connected and "infected" by the growing knowledge of their neighbors at an early stage of the diffusion process.

In order to develop an intuition for these results, consider a "network" of two countries, one with an initial variety of products N, the other with an initial variety $N + \Delta$. Neglecting the corner solution, and assuming $\phi = 1$, the equation of motion (8) for the first country is



NOTES: Solid lines: benchmark run ($\epsilon = 0.5$). Dashed lines: low international knowledge diffusion ($\epsilon = 0.2$). Dash-dotted lines: high international knowledge diffusion ($\epsilon = 0.8$).

FIGURE 8

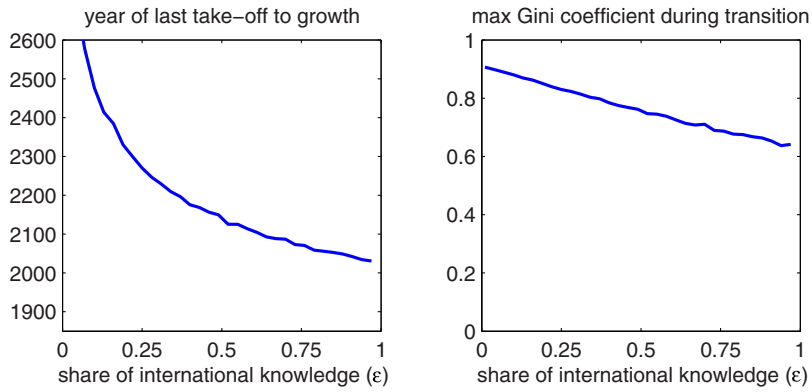
OVERTAKING FREQUENCIES [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

given by $f_1(N, N+\Delta) = N + \{A[\epsilon(N+\Delta) + (1-\epsilon)N] - \bar{c}\}$, with $\bar{c} \equiv \beta(\alpha L)^{-1}$. For the second country, it is given by $f_2(N, N+\Delta) = (N+\Delta) + \{A[\epsilon N + (1-\epsilon)(N+\Delta)] - \bar{c}\}$. Suppose $A > 1$ and consider the implausible yet illuminating case in which *all* knowledge comes from abroad, that is, $\epsilon = 1$. In this case, the two economies have changed their roles in the next period. Now, the first country is the better endowed one but it keeps this status only for one period after which the advantage is again transferred to the second country. There is overtaking in every period.

Generally, overtaking seems to be more likely the greater ϵ. In order to verify this claim for the simple example, it is easy to show that $f_1 > f_2$ for $\epsilon > (1+A)/2A$. For the actual model with a complex network of 100 participating economies, we cannot obtain a simple condition for overtaking. Instead, we investigate overtaking frequencies by way of numerical experiments. For that purpose, we run the model 5,000 times (i.e., for 5,000 alternative specifications of the small world network) and count the average number of overtakings in each period. An overtaking is defined as the advancement by one step in the income ranking of countries. Countries of the same income level are assigned the same rank. If, for example, a country advances from rank 5 to 4 in one period, it is recorded as one overtaking. If it advances from rank 5 to rank 3, we count two overtakings.

The results for the benchmark model are shown by solid lines in Figure 8. The top panel shows the total number of overtakings per period. On average, we observe about eight overtakings. Overtakings are relatively rare during early global development, gradually increasing until they reach a maximum in the late 20th century and then gradually declining to a level of about eight in the long run. This means that overtaking never stops. The world reaches a steady state only in terms of growth rates and relative income levels (see Section 2).[15]

[15] The number of overtakings in actual data (e.g., Maddison, 2003), is much higher. Clearly, there are many reasons that influence the relative performance of countries that are not captured by our model. Our main point is thus of



NOTE: Example with 100 countries; 10% initially rich, $p = 0.3$, varying ϵ.

FIGURE 9

INEQUALITY AND GROWTH—DEGREE OF INTERNATIONAL KNOWLEDGE DIFFUSION [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

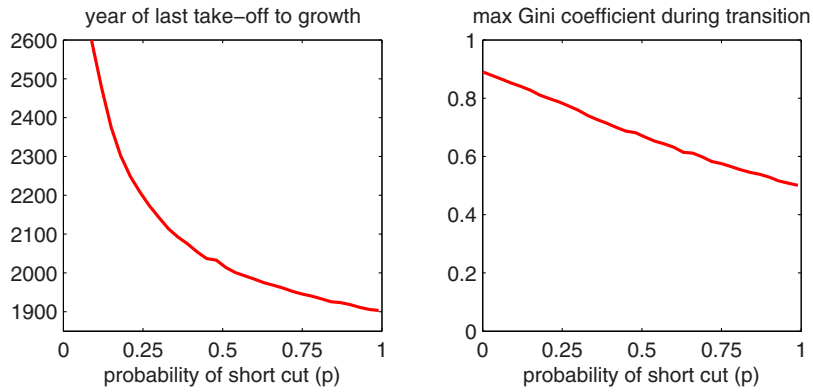
Although overtaking takes place relatively frequently at the world level, it is at the same time quite rare among the world leaders. But even the world leaders cannot expect to maintain their position permanently. This is shown in the middle panel of Figure 8 where we consider the top five countries in terms of GDP per capita. On average, only about 1% of overtaking takes place among the top five. If overtakings were equally distributed among countries, we would have expected about 5% of them taking place in the top five. When 1% of overtaking takes place among the world leaders, and there are on average 10 overtakings, this means that there are on average $10 \times 0.01 = 0.1$ overtakings among the top five. In order to better assess these results quantitatively, the bottom panel shows the cumulated sum of average overtakings among the top five. For the benchmark case (solid lines), there is less than one overtaking happening before the year 2000 and two overtakings before the year 2500.

The incidence of overtaking, naturally, depends on the degree of openness (ϵ). Dashed lines in Figure 8 show that there are fewer overtakings in total and among the top five, when only 20% of productivity advancements are learned from abroad ($\epsilon = 0.2$). More overtakings can be expected when openness is large, as demonstrated by the dashed lines for $\epsilon = 0.8$. We thus find numerical evidence in large networks for the theoretical conclusions about the role of ϵ derived from small (two-country) networks.

5.6. *Network Effects on Global Inequality and Growth.* We next investigate how the specific make up of the network affects the evolution of the world income distribution. For that purpose, we focus on two characteristic numbers, the calendar time when the last country takes off from stagnation and the maximum Gini index reached during the transition. Since long-distance links are set at random in the small world model, we ran each specification of the model 1,000 times and took averages afterward.

Figure 9 shows that a large contribution of international knowledge to productivity (i.e., a large ϵ) increases the pace of world development. Larger international knowledge spillovers are helpful to reduce worldwide inequality faster because a greater share of the initial knowledge advantage of the leaders is passed on through the network. Average network cohesion as defined in (12) starts with $\kappa = 0$ in the isolation case, $\epsilon = 0$, and rises linearly to $\kappa = 0.035$ for $\epsilon = 1$. Note that the modest level of $p = 0.3$ implies that the network is essentially a ring with 30 shortcuts. This explains the relatively low level of cohesion even when the knowledge externality is entirely dependent on the neighborhood for $\epsilon = 1$.

qualitative nature: The network model of knowledge diffusion provides one rational for the observation of overtakings although standard economic growth theory cannot explain the phenomenon.



NOTE: Example with 100 countries; 10% initially rich, $\epsilon = 0.5$, varying p.

FIGURE 10

INEQUALITY AND GROWTH—SHARE OF LONG-DISTANCE LINKS [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

Next, in Figure 10, we investigate how the share of long-distance links affects the evolution of the world income distribution. The year of the last take-off decreases very quickly for low values of p but remains rather insensitive for ps larger than 0.5. The outcome reflects a well-known feature of the small world model, namely, that average path length between nodes decreases sharply at low values of p and not much at high values (Watts and Strogatz, 1998). Maximum inequality also decreases sharply with increasing p, in an almost linear way. If every country had a long-distance link ($p = 1$), the last take-off would have been, according to the model, around the year 1900 with an associated maximum Gini index of 0.5. With this simulation setup average network cohesion starts with $\kappa = 0.0011$ for $p = 0$ (ring with no shortcuts) and rises linearly to $\kappa = 0.0075$ for $p = 1$.[16]

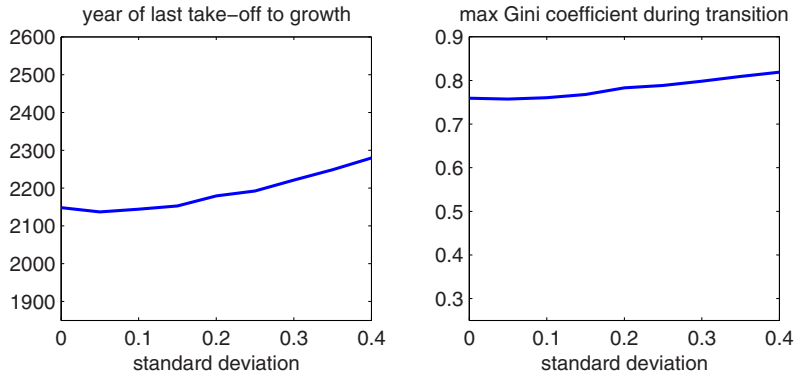
6. EXTENSIONS AND VARIATIONS

6.1. *Country-Specific Degrees of Openness.* It could be argued that the degree of openness to knowledge flows from abroad (ϵ) varies across countries. We thus finally demonstrate that allowing for country-specific openness adds more realism but leaves our main results basically unaffected. The performance of individual countries, of course, depends crucially on their degree of openness. In particular, we expect initially backward countries with high degree of openness to catch up relatively quickly, and relatively closed countries to be latecomers of industrialization. At the world level, however, we expect little change in performance. In order to verify this claim, we assume that the degree of openness is a normally distributed random variable with mean ϵ and standard deviation σ. We then run the small world model 1,000 times for alternative values of σ and record the year of the last take-off and the maximum Gini index during the transition.

Figure 11 shows the outcome for alternative $\sigma \in (0, 0.4)$ and ϵ drawn from a (truncated) normal distribution.[17] There is almost no change in the average maximum Gini along the transition and the year of the last take-off as the standard deviation of the degree of openness increases from 0 (our benchmark case) to 0.4. Allowing for country-specific degree of openness naturally affects country-specific performance but has little impact of the world's aggregate

[16] It can be shown that the year of the last take-off and the maximum Gini also depends quite strongly on the share of initially rich countries. The initial income ratio between rich and poor countries, in contrast, does not much affect the speed of transition and maximum inequality. The reason is that the negative impact on income inequality of an initially higher income gap is almost completely balanced by the fact that more can be learned from initially better endowed economies.

[17] In the rare event when the random draw provided a value above unity or below zero, we assign a value of 0.01 and 0.99, respectively.



NOTES: Example with 100 countries; 10% initially rich. Country-specific degree of openness, ϵ is normally distributed with mean 0.5 and standard deviation $\sigma \in (0, 0.4)$.

FIGURE 11

INEQUALITY AND GROWTH—VARYING DISTRIBUTION OF OPENNESS [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

performance with respect to the diffusion on knowledge and the take-off to growth as well as with respect to inequality.

6.2. *Biggest-Neighbor Learning.* In Section 2, we argued that the assumption that countries benefit from the average knowledge available in its neighbor countries is reasonable from the viewpoint of knowledge exchange through face-to-face communication. However, we acknowledge that alternative assumptions are conceivable. We next consider the assumption that countries learn only from their most advanced neighbor. Formally, this idea is implemented by

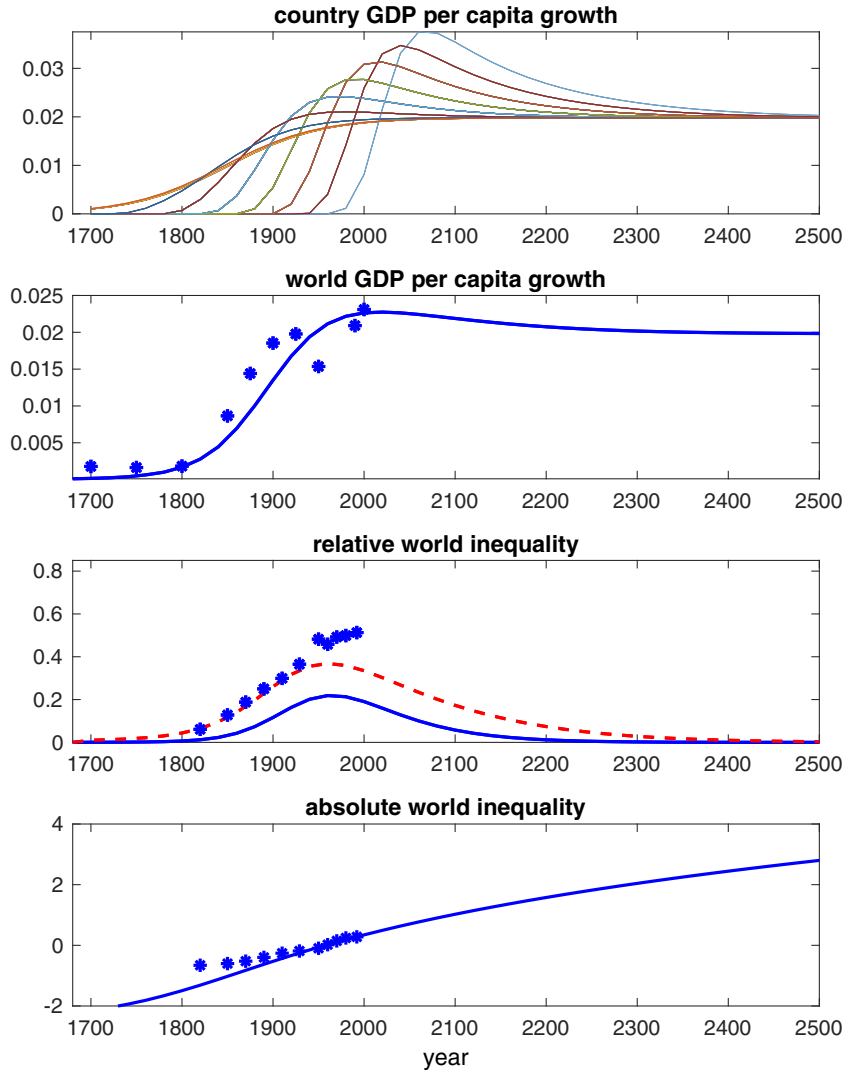
$$A_{it} = \left(\frac{A[\epsilon \max_{j \in \tilde{n}_i} N_{jt} + (1-\epsilon) N_{it}]}{\beta + z_{it}} \right)^{\phi}, \tag{13}$$

which replaces (6).

Figure 12 shows the predicted evolution of the world economy. Obviously, "biggest neighbor" learning generates less diversity in the timing of the country-specific take-offs to growth. The reason is that every country that is a neighbor of a forerunner country (i.e., of an initially rich country) benefits from knowledge from abroad to the same degree, irrespective of whether it is surrounded by other rich countries or surrounded by poor countries and connected via a long-distance link with the rich world (figuratively speaking, irrespective of whether it is France or India). This generates a predictable pattern of take-offs to growth: The initially rich countries take off first, then all countries connected via a long-distance link with the rich world (i.e., India and the United States), then all countries one link away from a link to the rich world, etc. This way of learning from abroad generates in some respect similar results (S-shaped transitions, overshooting of latecomers, etc.) but it generates less variety in individual growth performances. Thus, we also find that overtakings are rare under this scheme of knowledge exchange and that there is less inequality along the transition due to the lower variance of income across countries (third panel in Figure 12). The fit with actual inequality can be improved by reducing the number of long distance links, which would however extend the stagnation period of latecomer countries.

Furthermore, we considered linear combinations between the two cases of "biggest-neighbor" learning and the learning-from-all neighbors such that

$$A_{it} = \lambda \left(\frac{A[\epsilon \max_{j \in \tilde{n}_i} N_{jt} + (1-\epsilon) N_{it}]}{\beta + z_{it}} \right)^{\phi} + (1-\lambda) \left(\frac{A[\epsilon \bar{N}_{it} + (1-\epsilon) N_{it}]}{\beta + z_{it}} \right)^{\phi}. \tag{14}$$



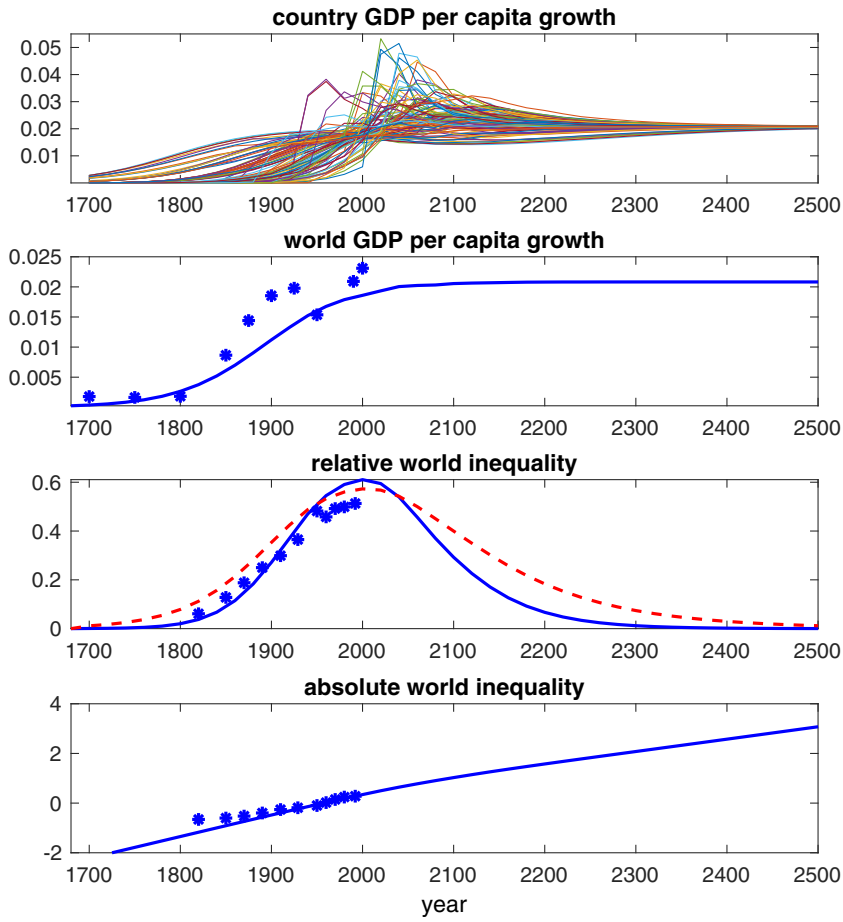
NOTES: Solid Lines: model prediction for learning from biggest neighbor. Dots in the second panel: world GDP growth according to De Long (1998); Third panel: solid line: Theil index; dashed line: Gini index; dots: Theil index from Bourguignon and Morrison (2002); Fourth Panel: log of absolute absolute Gini index; dots: absolute Gini from Atkinson and Brandolini (2010). Parameters as for benchmark model (Figure 3).

FIGURE 12

ECONOMIC GROWTH IN A SMALL WORLD: LEARNING ONLY FROM MOST ADVANCED NEIGHBOR [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

For $0 < \lambda < 1$, this scheme approximates a weighted network in which countries are open to knowledge flows from all neighbors but more so with respect to the most advanced neighbor. Figure A.4 in the Appendix shows results for $\lambda = 0.5$. Although every country has its unique growth trajectory, there are clusters of similar countries that develop "in sync." The mixed learning scheme performs better in the prediction of world inequality but still generates a smaller variety of growth performances and less overtaking than the benchmark model.

6.3. *An Evolving Network.* Finally, we consider the assumption that the network itself is evolving. The idea here is that advances in knowledge are conducive to the creation of new long-distance links between countries. Since countries have different degrees of development,



NOTES: Solid lines: model prediction for learning from biggest neighbor. Dots in the second panel: world GDP growth according to De Long (1998); Third panel: solid line: Theil index; dashed line: Gini index; dots: Theil index from Bourguignon and Morrison (2002); Fourth Panel: log of absolute absolute Gini index; dots: absolute Gini from Atkinson and Brandolini (2010). Parameters as for the benchmark model (Figure 3) and $\psi = 0.0005$.

FIGURE 13

ECONOMIC GROWTH IN A SMALL WORLD: EVOLVING NETWORK [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

this necessarily implies that the probability of long-distance is now country specific. We measure the degree of development by the number of available varieties and assume that in period t the probability of country i having a long-distance link to any arbitrary other country is given by

$$p_{i,t+1} = 1 - \mathrm{e}^{-\psi N_{i,t}}. \tag{15}$$

Consequently, with economic development, a country converges toward complete knowledge exchange with the rest of the world and the world converges toward a complete network. In order to avoid the implausible case of a temporary reduction of long-distance links, which could, in principle be possible since long-distance links are randomly created, we additionally assume that long-distance links can only be created and maintained but not destroyed. This means that, in period $t+1$, countries keep their long-distance links from period t and create new long-distance links with probability $\Delta_{i,t+1} = p_{i,t+1} - p_{i,t}$ if $p_{i,t+1} > p_{i,t}$. After the additional long-distance links have been created for period $t+1$, the new network of knowledge exchange is determined, and $\tilde{N}_{i,t+1}$, $N_{i,t+1}$, and $A_{i,t+1}$ are obtained. Everything else is kept from the basic model.

Figure 13 shows the prediction for an evolving network and $\psi = 0.0005$. Obviously, the endogenous creation of long-distance links amplifies the speed convergence of the latecomer countries of modern growth. Although the extension has a relatively small effect on the evolution of world GDP growth and on the path of increasing inequality in the divergence phase, it greatly amplifies the speed of decline of (relative) inequality in the convergence phase. It also leads to a greater variety in growth performances. In particular, some "lucky" countries experience very high growth after takeover because they become connected to several developed countries during the take-off period. The qualitative behavior of knowledge diffusion and growth, however, is very similar to the benchmark model.

In another extension of the basic model, we considered country differences in size and endogenous population dynamics. In order to implement Malthusian elements we assumed the need of (given) land in production and a positive response of population growth to increases in income when the level of income is low. In this setting, population growth leads to lower labor productivity in goods production and taken for itself (i.e., for constant technology) it leads to lower income and eventually to stagnation. Population growth, however, also induces more research and technological progress (as in Jones, 1995) and leads through this channel to higher growth. By appropriate choice of parameters, both channels balance each other in their impact on economic development such that the extended model fits the actual time paths of world growth and inequality equally well as the basic model. These results are available upon request.

7. CONCLUSION

In this article, we laid out a network-based theory of knowledge diffusion as an explanation for the divergence of countries as well as for their subsequent global convergence. Besides the endogenous evolution of the world income distribution, the theory contributes also to the explanation of the new Kaldor facts (Jones and Romer, 2010). The theory generates S-shaped transition paths with gradual take-off from stagnation as well as overshooting growth rates at later stages of development. In the long run, it thus predicts (slow) convergence of relative income across the globe.

The model could be extended such that it predicts permanent relative income inequality by introducing scale effects or by assuming that some countries use the available knowledge less efficiently than others. From the perspective of the very long run, convergence appears to be more intuitively appealing. However, even with knowledge eventually diffusing through the whole world, inequality vanishes only in relative terms, measured, for example, by the conventional Gini index. Absolute inequality, measured, for example, by the absolute Gini index, is predicted to keep on rising with increasing global development.

Although the underlying economic model has been a deliberately simple one, the theory can explain the long-run evolution of the world income distribution and a great variety of individual growth performances, including the overtaking of countries in the course of global development. In Section 6, we have discussed the robustness of the results with respect to several network-specific extensions. Naturally, further extensions are conceivable. More complex versions of the model could integrate trade in goods or factors or endogenous population dynamics with different phases of endogenous growth, in which learning-by-doing eventually triggers R&D-based growth, as in Strulik et al. (2013). A central result of the basic model, however, is that a complex economic model is not essential in order to generate a great variety of individual growth performances along with a great divergence and convergence of relative income levels across countries. Sufficient is a simple model and imperfect knowledge diffusion in a small world network.

APPENDIX A.

A.1. *Appendix.* For the proof of the propositions from the main text, it is useful to start with proving two lemmas.

LEMMA A.1. *If all countries grow at a positive rate, the system* (8) *is asymptotically given by*

$$x_{it+1} = \frac{g_{it}^N}{g_t^N} x_{it}, \tag{A.1}$$

with

$$g_{it}^N(x_{1t}, x_{2t}, \ldots, x_{nt}) = 1 + A(\alpha L)^{(1-\phi)/\phi}\left[(1/x_{it})\sum_{k\neq i}\bar{w}_{ik}x_{kt} + \bar{w}_{ii}\right] \quad i = 1, \ldots, n, \tag{A.2}$$

$$g_t^N := \frac{1}{n}\sum_{i=1}^{n} x_{it} g_{it} \tag{A.3}$$

$$x_{i0} > 0, \quad \textit{for all } i = 1, \ldots, n, \tag{A.4}$$

$$\frac{1}{n}\sum_{i=1}^{n} x_{i0} = 1. \tag{A.5}$$

Let e_n an $n \times 1$ column vector of ones. Long run equality $x^* = e_n$ is the unique steady state of (A.1)–(A.5) *with corresponding growth rate*

$$g^N = 1 + A(\alpha L)^{(1-\phi)/\phi}. \tag{A.6}$$

PROOF OF LEMMA A.1. In case of long-run growth, the term β/N_{it} becomes negligible such that (9) reads

$$\begin{aligned} g_{it+1}^N &\approx 1 + (\alpha L)^{-1}\left\{A\left[\epsilon\frac{\bar{N}_{it}}{N_{it}} + (1-\epsilon)\right](\alpha L)^{1/\phi}\right\} \\ &= 1 + A(\alpha L)^{(1-\phi)/\phi}\left[\frac{\sum_{k\neq i}\bar{w}_{ik}N_{kt}}{N_{it}} + \bar{w}_{ii}\right] \\ &= 1 + A(\alpha L)^{(1-\phi)/\phi}\left[\frac{\sum_{k\neq i}\bar{w}_{ik}(N_{kt}/N_t)}{(N_{it}/N_t)} + \bar{w}_{ii}\right], \end{aligned} \tag{A.7}$$

which simplifies to (A.2).

From (A.1), we conclude for the steady state that $g_i^N = g^N$ for all $i = 1, \ldots, n$. The growth rate in (A.2) is constant if and only if the term in square brackets is constant. The latter is equivalent to

$$\bar{\mathbf{W}}x = \mu x,$$

with $x = (x_1, \ldots, x_n)$ and $\mu > 0$ is a constant, an eigenvalue of $\bar{\mathbf{W}}$, respectively. Since $\bar{\mathbf{W}}$ is row stochastic, the spectral radius is $\rho(\bar{\mathbf{W}}) = 1$. Recall that we assume throughout the article that the network is strongly connected such that $\bar{\mathbf{W}}$ is irreducible. Note also that $\bar{\mathbf{W}}$ is aperiodic as $\bar{w}_{ii} > 0$ for all $i = 1 \ldots n$. The Perron–Frobenius Theorem (see, e.g., Mayer, 2000) states that

the eigenraum of $\bar{\mathbf{W}}$ to eigenvalue 1 is one dimensional. From this theorem, we also derive that there exists an eigenvector x of $\bar{\mathbf{W}}$ with eigenvalue 1 such that all components of x are positive. Furthermore, the theorem confirms that there are no other positive (moreover nonnegative) eigenvectors except positive multiples of x, that is, all other eigenvectors must have at least one negative or nonreal component. Note that x^* is an eigenvector $\bar{\mathbf{W}}$ with eigenvalue 1. From (A.4) and (A.5), we conclude that long-run equality x^* is the unique solution of (A.1)–(A.5). Finally, inserting x^* into (A.2) provides the steady-state growth rate. ■

LEMMA A.2. *Long-run equality $x^* = e_n$ is a stable solution of* (A.1)–(A.5) *such that there is convergence of initially different countries.*

PROOF OF LEMMA A.2. In order to establish that x^* is stable, we need to evaluate the Jacobian J of (A.1) at x^* and show that all eigenvalues of $J(x^*)$ are inside the unit circle.

Put $J_{ij}(x) := \partial x_{it+1}/\partial x_{jt}(x)$. From (A.1) follows

$$J_{ii}(x) = (1/g_t^N)^2 \left[\left(1 + A(\alpha L)^{\frac{1-\phi}{\phi}} \bar{w}_{ii}\right) g_t^N - \frac{g_{it}^N x_{it}}{n} \left(1 + A(\alpha L)^{\frac{1-\phi}{\phi}} \sum_{k=1}^{n} \bar{w}_{ki}\right)\right], \tag{A.8}$$

$$J_{ij}(x) = (1/g_t^N)^2 \left[A(\alpha L)^{\frac{1-\phi}{\phi}} \bar{w}_{ij}\, g_t^N - \frac{g_{it}^N x_{it}}{n} \left(1 + A(\alpha L)^{\frac{1-\phi}{\phi}} \sum_{k=1}^{n} \bar{w}_{ki}\right)\right] \text{for } i \neq j.$$

Put $C_n := I_n - \frac{1}{n} e_n e_n^T$, where I_n is the $n \times n$ identity matrix and e_n an $n \times 1$ column vector of ones. Furthermore, let F be an $n \times n$ identity matrix with typical element f_{ij}, defined as

$$f_{ij} = \bar{w}_{ij} - \frac{1}{n} \sum_{k=1}^{n} \bar{w}_{kj}.$$

Inserting $x^* = e_n$ into (A.8) provides

$$J(x^*) = \frac{1}{g^N} \left[A(\alpha L)^{\frac{1-\phi}{\phi}} F + C_n \right].$$

Inserting (A.6), this expression simplifies to

$$J(x^*) = \frac{1}{g^N} \left[(g^N - 1)F + C_n \right]. \tag{A.9}$$

From Theorem 1 of Calvacanti et al. (2016) follows that the set of eigenvalues $\sigma(F)$ of F is given by replacing the eigenvalue 1 of $\bar{\mathbf{W}}$ by 0 such that $\sigma(F) = \sigma(\bar{\mathbf{W}}) \setminus \{1\} \cup \{0\}$. Since $\bar{\mathbf{W}}$ is row stochastic, we also know that $0 \leq |\lambda_i| \leq 1$. Without loss of generality put $\lambda_1 = 1$ and assume $|\lambda_2| \geq |\lambda_3| \ldots \geq |\lambda_n|$.

From Lemma A.1 of Calvacanti (2016) follows that set of eigenvalues of $J(x^*)$ are given by

$$\sigma(J(x^*)) = \{b\lambda_2 + d, \ldots, b\lambda_n + d, 0\}, \tag{A.10}$$

with $\lambda_i \in \sigma(F) \setminus \{0\}$ and

$$b = \frac{g^N - 1}{g^N}, \tag{A.11}$$

$$d = \frac{1}{g^N}. \tag{A.12}$$

For $x^* = e_n$ to be stable, we need to show $|b\lambda_i + d| < 1$ for all $i = 1, \dots n-1$. We will first show that the eigenvalues of $\bar{\mathbf{W}}$, of F, respectively, are real. Recall that d_i is the number of links of country i. Put $D := diag\{d_1, \dots, d_n\}$ such that D is a diagonal matrix. Then, the matrix $G := D\bar{\mathbf{W}}$ has elements $(1-\epsilon)$ at the diagonal and $G_{ij} = G_{ji} = \epsilon$ if i and j are linked and zero otherwise. Since G is symmetric, the matrix $\tilde{G} = D^{-1/2}GD^{-1/2}$ is symmetric as well. Finally, for $\bar{\mathbf{W}}$ follows

$$\begin{aligned}\bar{\mathbf{W}} &= D^{-1}G = D^{-1/2}\big(D^{-1/2}GD^{-1/2}\big)D^{1/2} \\ &= D^{-1/2}\tilde{G}D^{1/2}.\end{aligned} \tag{A.13}$$

We conclude from (A.13) that $\bar{\mathbf{W}}$ is similar to the symmetric matrix $\tilde{G}$, which implies that the eigenvalues of $\bar{\mathbf{W}}$ are real. From the triangle equality and since $b, d \geq 0$, a sufficient condition for stability is

$$b|\lambda_i| + d < 1 \text{ for } i = 2 \dots n. \tag{A.14}$$

Since $b + d = 1$, we conclude that (A.14) is equivalent to $|\lambda_2| < 1$. The latter is confirmed by the theorem of Perron Frobenius since $\bar{\mathbf{W}}$ is irreducible and aperiodic. ■

PROOF OF PROPOSITION 1. From (8) follows that N_{it} grows iff

$$A\big[\epsilon\bar{N}_{it} + (1-\epsilon)N_{it}\big](\alpha L)^{1/\phi} - \beta \geq 0,$$

which is equivalent to

$$\big[\epsilon\bar{N}_{it} + (1-\epsilon)N_{it}\big] \geq \frac{\beta}{A(\alpha L)^{1/\phi}}. \tag{A.15}$$

We conclude that there is long-run growth if all countries are endowed with a number of varieties greater than $\beta/[A(\alpha L)^{1/\phi}]$.

In case of long-run growth, Lemma A.2 states that

$$\frac{\bar{N}_{it}}{N_{it}} = \frac{\bar{N}_{it}/N_t}{N_{it}/N_t} = \frac{\sum_{k\neq i}\bar{w}_{ik}x_{kt}}{x_{it}} \to 1. \tag{A.16}$$

Taking into account that $\beta/N_{it} \to 0$, we conclude for (9)

$$\begin{aligned}g^N = \lim_{t\to\infty} g^N_{it} &= 1 + (\alpha L)^{-1}\big\{A[\epsilon 1 + (1-\epsilon)](\alpha L)^{1/\phi}\big\} \\ &= 1 + (\alpha L)^{\frac{1-\phi}{\phi}}A.\end{aligned} \tag{A.17}$$

PROOF OF PROPOSITION 2. We define overshooting as the temporary surpassing of the long-run growth rate g^N (global overshooting).

(i) Assume that country j belongs to the forerunners of the industrial revolution such that $\bar{N}_{jt} \leq N_{jt}$ for all $t \geq 0$. For (9) follows

$$g^N_{it+1} = 1 + (\alpha L)^{-1}\left\{A\underbrace{\left[\epsilon\frac{\bar{N}_{it}}{N_{it}} + (1-\epsilon)\right]}_{\leq 1}(\alpha L)^{1/\phi} - \underbrace{\frac{\beta}{N_{it}}}_{\to 0}\right\}, \tag{A.18}$$

which prevents overshooting.

(ii) From (9) and (A.17), we conclude that overshooting $g_{it}^N > g^N$ simplifies to

$$A\left[\epsilon\frac{\bar{N}_{it}}{N_{it}} + (1-\epsilon)\right](\alpha L)^{1/\phi} - \frac{\beta}{N_{it}} > A(\alpha L)^{1/\phi},$$

which is equivalent to

$$\bar{N}_{it} > N_{it} + \frac{\beta}{\epsilon A(\alpha L)^{1/\phi}}. \tag{A.19}$$

We conclude that there is overshooting if the standing-on-shoulders externality $\bar{N}_{it}$ of country i exceeds its own variety by $\beta/\epsilon A(\alpha L)^{1/\phi}$. ■

Let $D_t = \max_{i,j}(y_{it} - y_{jt})$ denote the absolute income gap between the richest and poorest country and $Y = \sum_{i=1}^n y_i$. Let the relative gap be defined by $d_t = D_t/Y_t$. The following lemma summarizes the main properties of these measures of inequality:

LEMMA A.3.

(i) The Gini index stays constant if income grows at the same rate for all countries. It tends to zero for $t \to \infty$ if and only if $d_t \to 0$.
(ii) The absolute Gini index stays constant if income in every country increases by the same absolute amount. It converges to zero if and only if $D_t \to 0$.
(iii) If the absolute Gini index tends to zero, then the Gini index tends to zero as well, but not vice versa.

Statement (iii) follows immediately from the fact that $d_t \to 0$ occurs if the gap between rich and poor countries D_t increases at a lower rate than growth of Y_t.

PROOF OF LEMMA A.3. For notational convenience, we omit the time index t.

(i) The Gini index is defined by $G = (1 - 2B)$, where B is the area under the Lorenz curve. Without loss of generality, assume the countries are labeled such that $y_1 \le y_2 \le \cdots \le y_n$. Put $Y_k = \sum_{i=1}^k y_i$. The Lorenz curve is a polygonal line defined by the set of points

$$\left\{(0,0), \left(\frac{1}{n}, \frac{Y_1}{Y}\right), \left(\frac{2}{n}, \frac{Y_2}{Y}\right), \ldots, \left(\frac{n-1}{n}, \frac{Y_{n-1}}{Y}\right), (1,1)\right\}.$$

If all countries grow by the same rate, the fractions Y_i/Y stay the same for all $i = 1, \ldots, n$. The Gini index is zero if and only if the Lorenz curve is the identity line, which means, in particular, that the first slope of the polygonal line, ny_1/Y, and the slope of the last polygonal line, ny_n/Y, are identical. We conclude that the relative Gini index converges to zero if and only if

$$n\hat{d} = n\left(\frac{y_n}{Y} - \frac{y_1}{Y}\right) \to 0.$$

(ii) The term $\tilde{B} = nYB$ measures the area under the rescaled Lorenz curve where the horizontal axis ranges from 0 to n and the vertical axis from 0 to Y. The rescaled Lorenz curve is a polygonal line defined by the set of points

$$\{(0,0), (1, Y_1), (2, Y_2), \ldots, (n-1, Y_{n-1}), (n, Y)\}.$$

We get

$$\tilde{B} = \sum_{k=1}^{n} \frac{y_k}{2} \sum_{i=1}^{k-1} y_i = \frac{Y}{2} + \sum_{k=1}^{n} (n-k) y_k.$$

For the absolute Gini index follows

$$\begin{aligned} G \cdot \bar{Y} &= G \cdot (1/n) Y = (1 - 2B)(1/n) Y = (1 - 2\tilde{B}/(nY)) \cdot (1/n) Y \\ &= \frac{2}{n^2} \left[(nY)/2 - \tilde{B} \right]. \end{aligned} \tag{A.20}$$

We are now ready to prove the first claim of (ii) by induction. Note that the first term in (A.20) is just a scaling factor such that it suffices to prove the statement for the term in square brackets

$$T_n = \frac{n}{2} Y_n - \tilde{B}_n,$$

where the index indicates the number of countries. For $n = 2$, follows

$$\begin{aligned} \tilde{B}_2 &= \frac{y_1}{2} + \left(y_1 + \frac{y_2}{2} \right) = y_1 + \frac{y_1 + y_2}{2}, \\ T_2 &= (2 * Y_2)/2 - \tilde{B}_2 = Y_2 - y_1 - \frac{y_1 + y_2}{2} = \frac{y_2 - y_1}{2}. \end{aligned}$$

Hence, T_2 does not change if $y_1$and y_2 change by the same absolute amount. Suppose this holds for T_n with n countries. We get

$$T_{n+1} = \frac{(n+1)}{2} * Y_{n+1} - \tilde{B}_{n+1} = \frac{(n+1)}{2} * (Y_n + y_{n+1}) - \tilde{B}_{n+1},$$

where $\tilde{B}_{n+1}$ is given by $\tilde{B}_{n+1} = \tilde{B}_n + Y_n + y_{n+1}/2$. For T_{n+1} follows

$$\begin{aligned} T_{n+1} &= \frac{(n+1)}{2} * (Y_n + y_{n+1}) - \tilde{B}_{n+1} \\ &= \frac{n}{2} * (Y_n + y_{n+1}) + \frac{1}{2} * (Y_n + y_{n+1}) - \tilde{B}_{n+1} \\ &= \frac{n}{2} * Y_n + \frac{n}{2} * y_{n+1} + \frac{1}{2} * Y_n + \frac{1}{2} * y_{n+1} - \tilde{B}_n - Y_n - \frac{y_{n+1}}{2} \\ &= \frac{n}{2} * Y_n - \tilde{B}_n + \frac{n}{2} * y_{n+1} - \frac{1}{2} * Y_n = T_n + \frac{1}{2}(n * y_{n+1} - Y_n). \end{aligned} \tag{A.21}$$

From (A.21), we conclude that the term in parenthesis does not change if all income levels increase by the same amount.

Finally, note that the relative Gini index is given by the area between the identity line and the Lorenz curve divided by the total area under the identity line from 0 to 1 (which is 1/2). Multiplying the relative Gini index by Y/n is equivalent to studying this index with a rescaling of the vertical axis ranging from 0 to Y/n. Here, the rescaled Lorenz curve is a polygonal line defined by the set of points

$$\left\{ (0,0), \left(\frac{1}{n}, \frac{Y_1}{n} \right), \left(\frac{2}{n}, \frac{Y_2}{n} \right), \ldots, \left(\frac{n-1}{n}, \frac{Y_{n-1}}{n} \right), \left(1, \frac{Y}{n} \right) \right\}.$$

The area between the identity line and the rescaled Lorenz curve is 0 if and only if the first polygonal line of the Lorenz curve has the same slope as the last one. This is equivalent to $y_1 = y_n$.

(iii) From (3), (4), and (7), we conclude $1 \leq \bar{Y}$. Hence, the product $G\bar{Y}$ can only tend to zero if the relative Gini index tends to zero. However, the latter is not a sufficient condition since $G\bar{Y}$ increases if $\bar{Y}$ grows at a higher rate than G declines. ■

PROOF OF PROPOSITION 3.

(i) Follows directly from Lemma A.2.

(ii) If some countries initially grow and others stagnate, the Lorenz curve bends below the identity line such that the Gini index increases. According to Lemma A.2, however, the index tends to zero eventually.

PROOF OF PROPOSITION 4. The proposition is proven by a simple example of a connected network with long-run growth such that the relative Gini index tends to zero according to Proposition 3. We will show, however, that the absolute Gini index keeps growing.

Consider the simple case of $n = 2$. At some time t, put $N_{1t} = x$ and $N_{2t} = \lambda x$. Assume $\lambda > 1$ such that $N_{2t} > N_{1t}$ and

$$D_t = N_{2t} - N_{1t} = x(\lambda - 1) > 0. \tag{A.22}$$

Put $\mu := A(\alpha L)^{1/\phi - 1}$. From (8) follows

$$\begin{aligned} N_{t+1} &= \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \mu \begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix} \right) \begin{pmatrix} x \\ \lambda x \end{pmatrix} - \frac{\beta}{\alpha L} \\ &= \begin{pmatrix} x(\mu(1-\epsilon) + \lambda\mu\epsilon + 1) \\ x(\lambda + \lambda\mu(1-\epsilon) + \mu\epsilon) \end{pmatrix} - \frac{\beta}{\alpha L}. \end{aligned} \tag{A.23}$$

For the distance $D_{t+1} = N_{2t+1} - N_{1t+1}$, we get from (A.23)

$$D_{t+1} = x(\lambda - 1)(\mu - 2\mu\epsilon + 1). \tag{A.24}$$

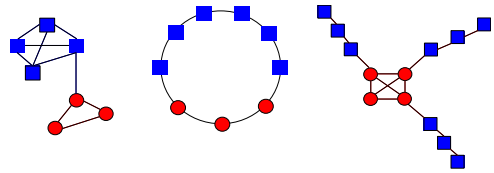
Absolute distance grows iff $D_{t+1} > D_t$. Inserting (A.22) and (A.24) provides

$$x(\lambda - 1)(\mu - 2\mu\epsilon + 1) > x(\lambda - 1),$$

which simplifies to $\epsilon < 1/2$. We conclude that the absolute distance D_t grows in our simple model of two nodes iff the externality weight ϵ is smaller than $1/2$. According to Proposition A.3, this is equivalent to stating that the absolute Gini index tends to infinity. ■

PROOF OF PROPOSITION 5. Since $\bar{\mathbf{W}}$ is irreducible and aperiodic, the Perron Frobenius Theorem confirms that $|\lambda_2| < 1$. In particular, from (A.10)–(A.12), we conclude that

$$\frac{g^N - 1}{g^N}\lambda_2 + \frac{1}{g^N}, \tag{A.25}$$



NOTES: Stylized networks: bridge (left), ring (middle), and core-periphery (right). Rich countries are represented by red circles and poor countries are represented by blue squares.

FIGURE A.1

STYLIZED NETWORKS [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

with $g^N = 1 + (\alpha L)^{\frac{1-\phi}{\phi}} A$, is an upper bound for the rate of convergence. Inserting (12) implies that (A.25) is equivalent to

$$1 - \frac{g^N - 1}{g^N} \kappa.$$

■

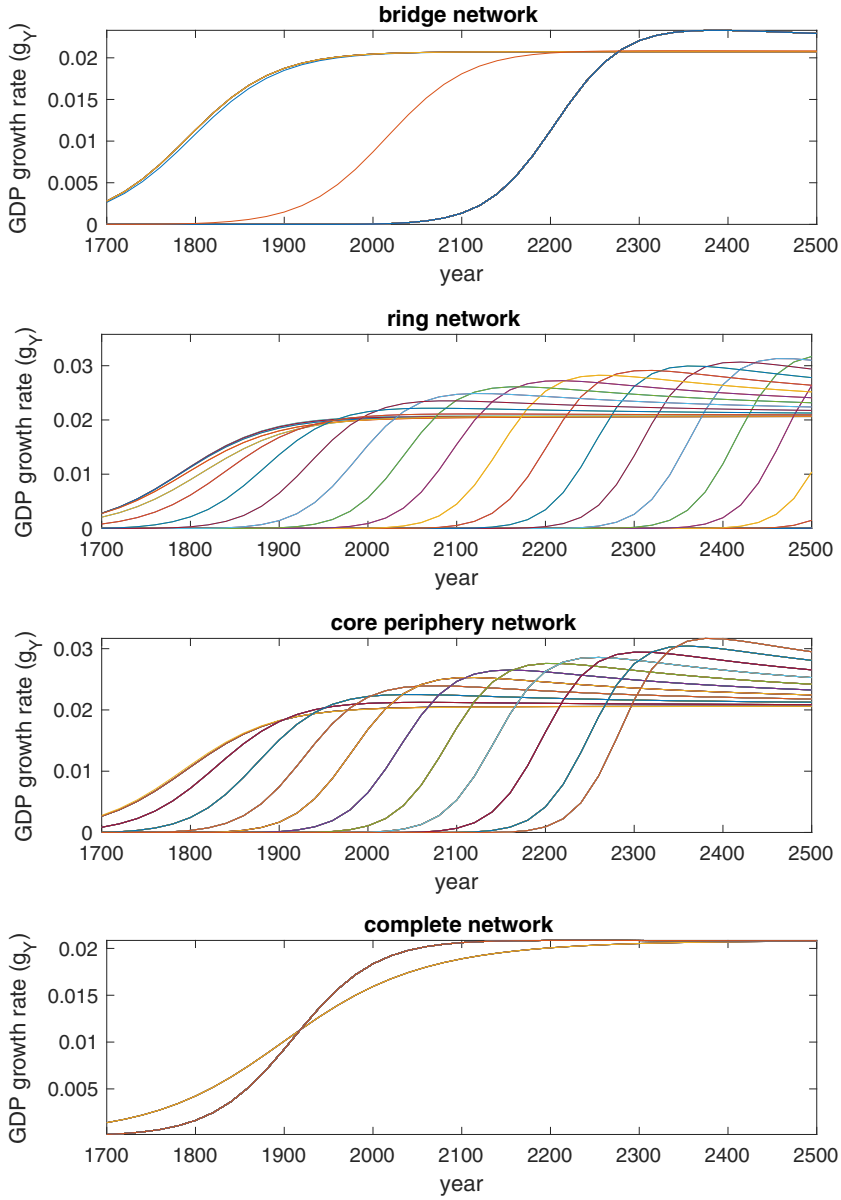
A.2. *Stylized Networks.* In this section, we investigate adjustment dynamics for some particularly simple examples of the network $\bar{\mathbf{W}}$. This allows us to provide an understanding of the main mechanism behind the international flow of knowledge and world income dynamics. Suppose the world network is given alternatively by a stylized network from the set of networks depicted in Figure A.1. Rich countries are represented by red circles and poor countries are represented by blue squares.

A bridge network is partitioned into two components. The rich and the poor are each internally representing a complete network. The two components share exactly one link, the bridge. The bridge network could be understood as a metaphor for a world of different continents connected by a minimum of links.

A ring network is obtained by positioning each country along a line, ordered by country-specific initial endowments. In order to establish a symmetric architecture, the line is closed to form a circle. Each country is connected to its k nearest neighbors (not counting itself as a neighbor). This means that there are $2k$ poor countries connected with rich countries. In the example, we have $k = 1$. The ring network emphasizes the role of geographic proximity for knowledge exchange. The world is "round" and countries are directly connected only with their geographical neighbors.

Finally, we consider the core-periphery network. The core, consisting of initially rich countries, forms a complete network to which a number of peripheries consisting of initially poor countries are connected. The poor countries are connected in series implying that there is one bridge per periphery, linking it with the core. The core-periphery network describes a situation in which a subset of rich countries is fully integrated and another subset of poor countries (the colonies) is less well integrated.

For the simple networks, we keep the economic model as numerically specified in the main text. Figure A.2 shows the evolution of growth predicted by the numerical experiments. The upper panel assumes that the world network is a bridge. Knowledge diffusion through the network generates four visibly distinct adjustment trajectories. Naturally, the rich countries take off first. The rich country linked directly to the poor world takes off a bit later because there is less to learn from the poor neighbor. In contrast, the poor country equipped with a direct link to the rich world experiences a huge advantage vis-à-vis its poor neighbors and takes off about two centuries earlier, fueled by knowledge diffusion from its rich neighbor. The remaining club of less developed countries takes off late but experiences an "advantage of backwardness" (Gerschenkron, 1962) in the sense that their income growth surpasses the income growth of the forerunners of the industrial revolution.

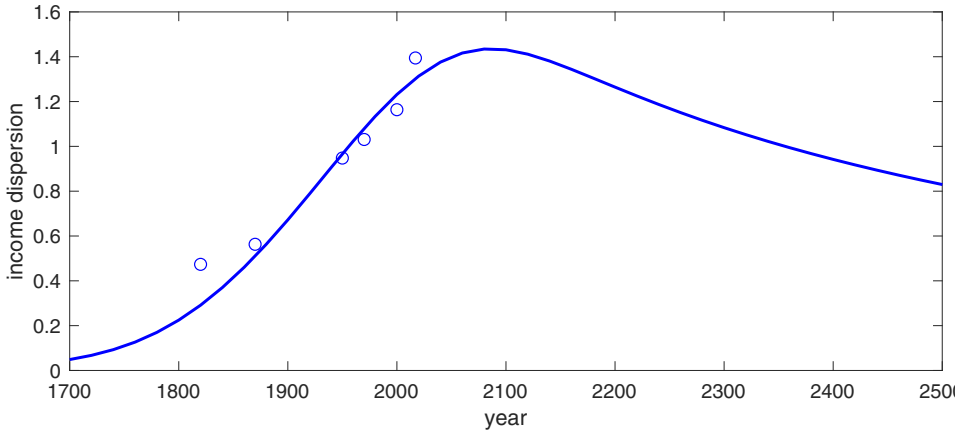


NOTES: For all three "worlds": 10% of countries initially better endowed, $N_0 = 10$ for poor countries, $N_0 = 11$ for rich countries. Parameters: $\alpha = 0.65$, $\phi = L = 1$, $\beta = 3.32$, $A = 0.51$, $\epsilon = 0.5$. Ring: two neighbors per node. Core-periphery: nine peripheries of 10 countries. Among the incomplete networks, cohesion of bridge is highest with $5.05 \cdot 10^{-3}$, followed by core periphery with $1.88 \cdot 10^{-3}$. Ring has lowest cohesion $9.87 \cdot 10^{-4}$.

FIGURE A.2

EVOLUTION OF ECONOMIC GROWTH IN 100 COUNTRIES WHEN THE WORLD IS A . . . [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

The bridge network already displays one important phenomenon of growth in networks, the overshooting growth of latecomers, but it generates insufficient variety of economic performance across countries. This is different for the ring network, as evidenced in the center panel of Figure A.2. The initially rich countries are again experiencing a very similar take-off to growth, in which the countries surrounded by other rich countries perform only slightly better than those at the border to the poor world. The poor countries, on the other hand, experience a very varied take-off. The reason is that new knowledge is "handed over" along the circle.



NOTES: The line shows the cross-sectional standard deviation of the log of per capita output as predicted by the model. Circles: data; 1820 to 2000 from Maddison (2003); 2017 from IMF (2018). See text for details.

FIGURE A.3

σ–CONVERGENCE [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

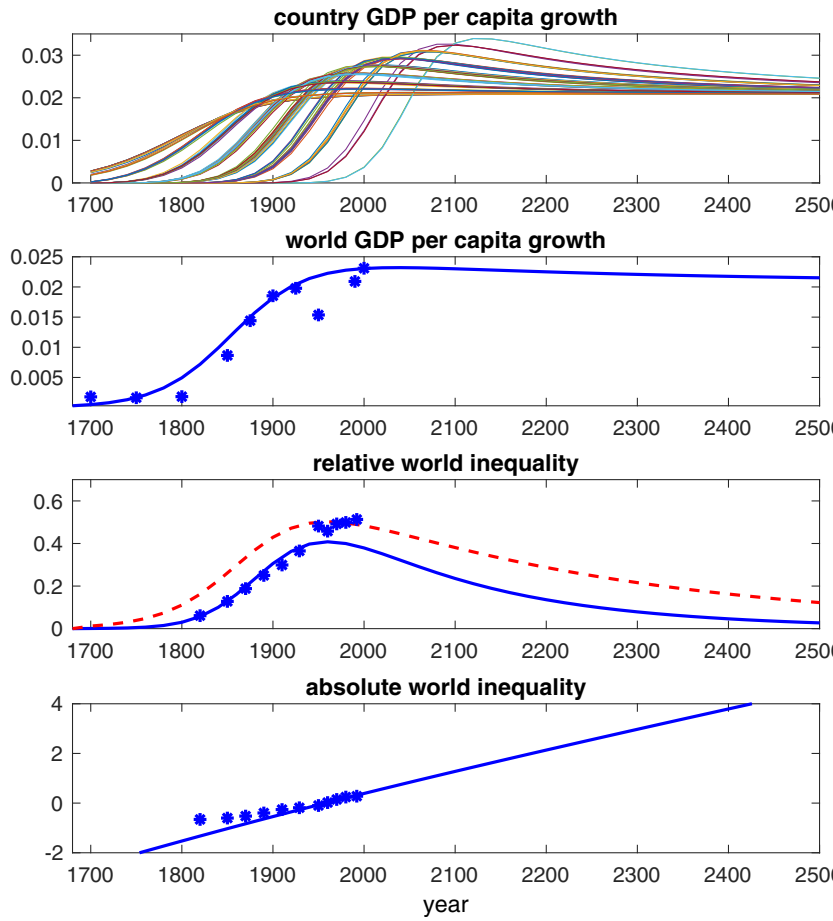
The two countries neighboring the rich take off first among the poor, then the countries next to these countries follow, etc. There is also more variety in growth rates.

Compared with other networks, the ring predicts a very long period of take-offs, implying a very long period of increasing world inequality. This is confirmed by the cohesion values, which is the lowest in the ring case with a value of 9.8710^{-4}. The reason is that it takes time until knowledge is passed on along the circle from neighbor to neighbor toward the most unfortunate country "at the other side of the world." Moreover, the take-offs are "too predictable." Their sequence follows the position of countries on the circle.

The core-periphery network, shown in the third panel of Figure A.2, eliminates some of the flaws of the two previous networks. It produces a variety of growth experiences, largely overshooting growth rates, and a reasonable duration of the "era of take-offs to growth" from 1700 to the mid-21st century. Yet, the growth experience of countries is still too easily predicted. The countries next to the bridges to the core take off just after the initially rich and then we observe departures from stagnation according to the order of countries along the peripheries. Altogether, we observe "only" 10 different growth paths, one for the core countries and one for each position on the periphery. There is still too little heterogeneity in the world. Moreover, the connectivity between the initially rich countries is "too high" in all three simple networks. This is evident from the result that the take-off of the forerunners of the industrial revolution happens too fast in all three panels of Figure A.2. By the year 1800, the forerunners of the industrial revolution are counterfactually predicted to grow already at a rate of 1.5% annually.

Finally, we consider knowledge diffusion in a complete network, in which every country exchanges knowledge with every other country. This means that all countries have full access to world knowledge. The bottom panel in Figure A.2 shows that in a complete network there are only two different paths for the 100 countries. The initially rich countries start growing earlier but the initially poor countries follow suit and catch up. As a consequence of immediate access to knowledge from everywhere, there is very little diversity in growth performance and very little inequality.

σ-convergence. In this section, we consider σ-convergence, an alternative measure of inequality, which is popular in the growth economics literature. σ-convergence is measured by the evolution of the standard deviation of log GDP per capita across countries. Results for the benchmark model from Figure 3 are shown in Figure A.3. The solid line shows the predicted σ



NOTES: Adjustment dynamics for mixed knowledge diffusion as in (14) and $\lambda = 0.5$. Parameters as for benchmark model.

FIGURE A.4

ECONOMIC GROWTH IN A SMALL WORLD: MIXED LEARNING SCHEME [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

and the dots represent data points computed from Maddison (2003). Since the historical data of Maddison (2003) has a lot of missing values, we computed σ-values before 1950 only when there were data for more than 50 countries (i.e., for the years 1820 and 1870). The last data point, for 2017, is obtained from IMF (2018). We see that the model fits the data quite well. Only for the year 1820, the model somewhat underestimates σ. One reason could be that the Maddison sample is censored against poor countries, which all stagnated close to subsistence at that time, thereby driving down the cross-country σ.

ACKNOWLEDGEMENTS

Open access funding enabled and organized by Projekt DEAL.

REFERENCES

ACEMOGLU, D., *Introduction to Modern Economic Growth* (Princeton: Princeton University Press, 2009).
———, AND J. VENTURA, "The World Income Distribution," *Quarterly Journal of Economics* 117 (2002), 659–94.
———, P. AGHION, AND F. ZILIBOTTI, "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association* 4(1) (2006), 37–74.

ANDERSEN, T. B., AND C.-J. DALGAARD, "Flows of People, Flows of Ideas, and the Inequality of Nations," *Journal of Economic Growth* 16 (2011), 1–32.
ASHRAF, Q., AND O. GALOR, "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review* 103(1) (2013), 1–46.
ATKINSON, A. B., AND A. BRANDOLINI, "On Analyzing the World Distribution of Income," *World Bank Economic Review* 24 (2010), 1–37.
BAIROCH, P., *Economics and World History* (Hemel Hempstead: University of Chicago Press, 1993).
BALDWIN, R. E., AND R. FORSLID, "The Core–Periphery Model and Endogenous Growth: Stabilizing and Destabilizing Integration," *Economica* 67(267) (2000), 307–24.
———, P. MARTIN, AND G.I. OTTAVIANO, "Global Income Divergence, Trade, and Industrialization: The Geography of Growth Take-offs," *Journal of Economic Growth* 6(1) (2001), 5–37.
BOUCEKKINE, R., D. DE LA CROIX, AND O. LICANDRO, "Vintage Human Capital, Demographic Trends, and Endogenous Growth," *Journal of Economic Theory* 104 (2002), 340–75.
BOURGUIGNON, F., AND C. MORRISSON, "Inequality Among World Citizens: 1820–1992," *American Economic Review* 92 (2002), 727–44.
BREZIS, E. S., P. R. KRUGMAN, AND D. TSIDDON, "Leapfrogging in International Competition: A Theory of Cycles in National Technological Leadership," *American Economic Review* 83 (1993), 1211–19.
CAVALCANTI, T., AND C. GIANNITSAROU, "Growth and Human Capital: A Network Approach," *Economic Journal* 127 (2017), 1279–1317.
———, ———, AND C. R. JOHNSON, "Network Cohesion," *Economic Theory* 62 (2016), 1–21.
CASELLI, F., AND J. FEYRER, "The Marginal Product of Capital," *Quarterly Journal of Economics* 122(2) (2007), 535–68.
CHAKRAVARTY, S. R., "Extended Gini Indices of Inequality," *International Economic Review* 29 (1988), 147–56.
COMIN, D. A., M. DMITRIEV, AND E. ROSSI-HANSBERG, "The Spatial Diffusion of Technology," NBER Discussion Paper 18534, 2012.
COE, D. T., E. HELPMAN, AND A. W. HOFFMAISTER, "International R&D Spillovers and Institutions," *European Economic Review* 7 (2009), 723–41.
DAVIS, C., AND K. I. HASHIMOTO, "Industry Concentration, Knowledge Diffusion and Economic Growth without Scale Effects," *Economica* 82(328) (2015), 769–89.
DE LONG, B. "Estimates of World GDP, One Million BC – Present," University of California, Berkeley, Department of Economics, 1998, http://delong.typepad.com/print/20061012_LRWGDP.pdf.
DELVENTHAL, M. J. "The Globe as a Network: Geography and the Origins of the World Income Distribution," Mimeo, Claremont McKenna College, 2018.
DOEPKE, M. "Accounting for Fertility Decline During the Transition to Growth," *Journal of Economic Growth* 9 (2004), 347–83.
EATON, J., S. KORTUM, "International Patenting and Technology Diffusion: Theory and Measurement," *International Economic Review* 40 (1999), 537–70.
ENCYCLOPEDIA BRITANNICA, Industrial Revolution, 2018. https://www.britannica.com/event/Industrial-Revolution
FOGLI, A., AND L. VELDKAMP, "Germs, Social Networks and Growth," NBER Working Paper No. w18470, 2012.
FRANKEL, J. A., AND D. ROMER, "Does Trade Cause Growth?" *American Economic Review* 89(3) (1999), 379–99.
GALOR, O. "From Stagnation to Growth: Unified Growth Theory," in P. Aghion and S. Durlauf, eds., *Handbook of Economic Growth*, Volume 1A, (Amsterdam: North-Holland, 2005), 171–293.
———, AND D. N. WEIL, "Population, Technology and Growth: From the Malthusian Regime to the Demographic Transition and Beyond," *American Economic Review* 110 (2000), 806–28.
———, AND O. MOAV, "Natural Selection and the Origin of Economic Growth," *Quarterly Journal of Economics* 117 (2002), 1133–92.
———, AND ———, "From Physical to Human Capital Accumulation: Inequality and the Process of Development," *Review of Economic Studies* 71 (2004), 1001–26.
———, AND A. MOUNTFORD, "Trading Population for Productivity: Theory and Evidence," *Review of Economic Studies* 75 (2008), 1143–79.
GERSCHENKRON, A., *Economic Backwardness in Historical Perspective, A Book of Essays* (Cambridge, Massachusetts: Harvard University Press, 1962).
HOVHANNISYAN, N., AND W. KELLER, "International Business Travel: An Engine of Innovation?" *Journal of Economic Growth* 20(1) (2015), 75–104.
HOWITT, P., "Endogenous Growth and Cross-Country Income Differences," *American Economic Review* 90 (2000), 829–46.
IMF, *World Economic Outlook*, 2018, https://www.imf.org/external/datamapper/NGDPDPC@WEO/OEMDC/ADVEC/WEOWORLD
JAFFE, A., M. TRATJENBERG, AND R. HENDERSON, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics* 63 (1993), 577–98.

JONES, C. I., "R& D-Based Models of Economic Growth," *Journal of Political Economy* 103(4) (1995), 759–84.
———, "On the Evolution of the World Income Distribution," *Journal of Economic Perspectives* 11(3) (1997), 19–36.
———, "Was an Industrial Revolution Inevitable? Economic Growth over the Very Long Run," *Advances in Macroeconomics* 1 (2001), 1–43.
———, AND P. M. ROMER, "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," *American Economic Journal: Macroeconomics* 2 (2010), 224–45.
KELLER, W., "Geographic Localization of International Technology Diffusion," *American Economic Review* 92 (2002), 120–42.
KLENOW, P. J., AND A. RODRIGUEZ-CLARE, "Externalities and Growth," in: P. Aghion and S. N. Durlauf, eds., *Handbook of Economic Growth* Vol. 1 (Amsterdam: Elsevier, 2005), 817–61.
KOLM, S.-C."Unequal Inequalities I," *Journal of Economic Theory* 12 (1976), 416–42.
LEDERMAN, D., AND W. F. MALONEY, "Research and Development (R&D) and Development," Policy Research Working Paper Series 3024, The World Bank, 2003.
LINDNER, I., AND H. STRULIK, "The Great Divergence: A Network Approach," Tinbergen Institute Discussion Paper 14-033/II, 2014.
———, AND ———, "From Tradition to Modernity: Economic Growth in a Small World," *Journal of Development Economics* 109 (2015), 17–29.
LUCAS, R. E., "Some Macroeconomics for the 21st Century," *Journal of Economic Perspectives* 14 (2000), 159–68.
———, "Trade and the Industrial Revolution," *American Economic Journal: Macroeconomics* 1 (2009), 1–25.
MADDISON, A., *The World Economy: A Millennial Perspective* (Paris: Development Center Studies, OECD, 2003).
MADSEN, J. B., AND M. FARHADI, "International Technology Spillovers and Growth over the Past 142 Years: The Role of Genetic Proximity," *Economica* 85(338) (2018), 329–59.
MOKYR, J. "Long-term Economic Growth and the History of Technology," in P. Aghion and S. N. Durlauf, eds., *Handbook of Economic Growth* Vol. 1, Chapter 17 (Amsterdam: Elsevier, 2005), 1113–80.
NEWMAN, M. E. J., "The Structure and Function of Complex Networks," *SIAM Review* 45 (2003), 167–256.
———, AND D. J. WATTS, "Renormalization Group Analysis of the Small-World Model," *Physics Letters A* 263 (1999), 341–46.
OECD, "Gross Domestic Spending on R&D (Indicator)," 10.1787/d8b068b4-en (Accessed on April 07, 2019). https://www.oecd-ilibrary.org/industry-and-services/gross-domestic-spending-on-r-d/indicator/english_d8b068b4-en
ORTEGA, F., AND G. PERI, "Openness and Income: The Roles of Trade and Migration," *Journal of International Economics* 92(2) (2014), 231–51.
PARENTE, S. L. AND E. D. PRESCOTT, "A Unified Theory of the Evolution of International Income Levels, in P. Aghion and S. N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Chapter 21 (Amsterdam: Elsevier, 2005), 1371–1416.
POMERANZ, K., *The Great Divergence: China, Europe, and the Making of the Modern World* (Princeton: Princeton University Press, 2000).
RIVERA-BATIZ, L. A., AND P. M. ROMER, "Economic Integration and Endogenous Growth," *Quarterly Journal of Economics* 106(2) (1991), 531–55.
ROMER, P. M., "Endogenous Technological Change," *Journal of Political Economy* 98 (1990), S71–S102.
SALA-I-MARTIN, X., "The World Distribution of Income: Falling Poverty and Convergence, Period," *Quarterly Journal of Economics* 121 (2006), 351–97.
SPOLAORE, E., AND R. WACZIARG, "The Diffusion of Development," *Quarterly Journal of Economics* 124(2) (2009), 469–529.
———, AND ———, "How Deep Are the Roots of Economic Development?" *Journal of Economic Literature* 51(2) (2013), 325–69.
STRULIK, H., "Knowledge and Growth in the Very Long Run," *International Economic Review* 55(2) (2014), 459–82.
———, AND J. WEISDORF, "Population, Food, and Knowledge: A Simple Unified Growth Model," *Journal of Economic Growth* 13 (2008), 169–94.
———, K. PRETTNER, AND A. PRSKWETZ, "The Past and Future of Knowledge-Based Growth," *Journal of Economic Growth* 18 (2013), 411–37.
SUTCLIFFE, B., "World Inequality and Globalization," *Oxford Review of Economic Policy* 20 (2004), 15–37.
WÄLDE, K., "Proof of Global Stability, Transitional Dynamics, and International Capital Flows in a Two-Country Model of Innovation and Growth," *Journal of Economics* 64(1) (1996), 53–84.
WATTS, D. J., AND S. H. STROGATZ, S. H. "Collective Dynamics of "Small-World" Networks," *Nature* 393 (1998), 440–42.