Center for European, Governance and Economic Development Research

cege

Discussion Papers

Number 313 – June 2017

INNOVATION AND INEQUALITY IN A SMALL WORLD

Ines Lindner and Holger Strulik

Georg-August-Universität Göttingen

ISSN: 1439-2305

Innovation and Inequality in a Small World *

Ines Lindner† and Holger Strulik**

June 2017

Abstract. We present a multi-country 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 after the take-off of the forerunners of the industrial revolution, 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.

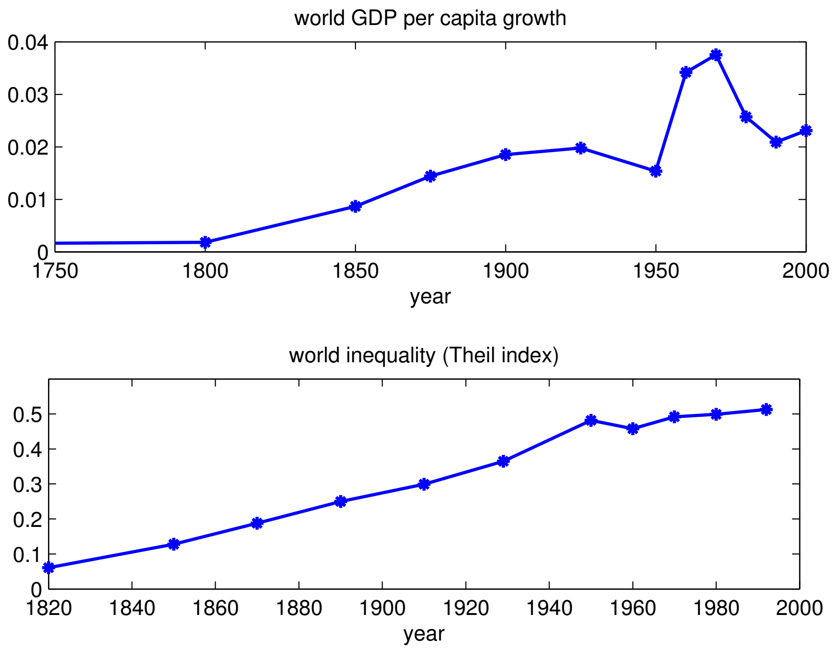
Keywords: networks, knowledge diffusion, economic growth, world income distribution.
JEL: O10; O40; D85; F43.

* 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, and David Weil for helpful comments.
† VU University Amsterdam, Department of Econometrics and Operations Research, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, email: ilindner@feweb.vu.nl
** University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany; email: holger.strulik@wiwi.uni-goettingen.de.

1. Introduction

In pre-modern 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, 2002). Figure 1 shows the gradual increase of world income growth and the evolution of world inequality since the onset of the industrial revolution.[1]

Figure 1: World Economic Growth and Inequality



Data from De Long (1998) and Bouguignon and Morrisson (2002).

The 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, 2011). 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

[1] 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).

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.[2]

This paper 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 towards 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. This 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 e.g. 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, i.e. 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 (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 upon the multi-country 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

[2] 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) is an exception in that it considers two interacting countries (or regions). 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. Alternative two-country models of the differentiated take-off to growth are proposed by Baldwin et al. (2001) and Strulik (2014).

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, 2015b) 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 this paper, in contrast, the economic model is straightforward but the process of knowledge diffusion is non-trivial and the analysis focuses on transitional dynamics.[3]

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 growth of the leaders, followers and trailers of the industrial revolution is endogenously explained and understood by the increasing diffusion of knowledge throughout the world. This allows us to explain a richer set of phenomena. For example, according to Lucas' approach, the US would never had outpaced England, the industrial leader. In our setup, countries do not only temporarily diverge and converge but they also (occasionally) overtake each other.

[3] 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 micro-founded 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.

Another strand of literature investigates multi-country models in which convergence is driven by capital accumulation and trade (e.g. Acemoglu and Ventura, 2002). Conceptually, the available multi-country growth literature focuses on the question of whether and how countries at initially different income levels converge while we also investigate how countries that were initially similar diverged. In other words, as with the available multi-country 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".[4]

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 (2015a) investigate how economic development is affected by globalization conceptualized as an evolving network, i.e. how decreasing local connectivity affects occupational choice and investment behavior through eroding trust and trustworthiness.

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 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 rather than 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.

The links between countries can be interpreted as geographic proximity (Mexico next to the US) as well as cultural or genetic similarity (England next to the US). 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 pre-industrial times, before the take-off to

[4] 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).

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 paper 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. In Section 4 we 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 5, we introduce the Small World network (Watts and Strogatz, 1998) and investigate the distribution and growth of world income when countries are connected by such a network. We argue that the Small World network is already sufficiently complex to generate growth trajectories consistent with the historical development of world inequality and economic growth. We provide a sensitivity analysis with respect to the network parameters and discuss the phenomenon of endogenously changing world economic leaders and overtaking in general in the course of global development. Section 6 concludes.

2. The Model

Consider a world consisting of a number n of countries indexed by i. All Countries are populated by a (non-overlapping) workforce L.[5] 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). 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 towards 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. Final Goods Production.

The final goods sector produces competitively using labor and a range of intermediate products. At time t there are N_{it} intermediate goods

[5] Applying Occam's razor we present the simplest network-R&D model that motivates the evolution of world income inequality. Country differences in size constitute a potentially interesting comparative static for an extended model.

available in country i. Intermediate inputs are continuously measured. Let $x_{it}(j)$ denote the quantity of input j. Output of final goods is then given by

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

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

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

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. 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 $p_{it}(j) = (1-\alpha)$ such that factor input is $x_{it}(j) = (1-\alpha)^{-1/\alpha} L$. Profits of the $N_{it} - N_{it-1}$ firms supplying the latest vintage of intermediate goods are given by

$$\pi_{it}(j) = p_{it}(j) x_{it}(j) - (1-\alpha) x_{it}(j) = p_{it}(j)^{-1/\alpha} L \left[p_{it}(j) - (1-\alpha)\right]. \tag{3}$$

From the first order condition for maximum profits we obtain $p_{it}(j) = 1$ such that $x_{it}(j) = L$ and $\pi_{it}(j) = \alpha L$.

2.2. **R&D.** Suppose 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 such as the purchase of machines embodying new technologies. In shorthand, 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 $\pi_{it}(j) = \alpha L$. 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. **The Knowledge Network.** Following Howitt (2000) and Acemoglu (2009) we assume that there is a "standing on shoulders" effect in R&D productivity which is increasing 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. For simplicity 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 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 $d_i = 0$ 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, $\sum_{j=1}^{n} \bar{w}_{ij} N_{jt} = \epsilon \bar{N}_{it} + (1-\epsilon) N_{it}$, in which the average available varieties in the neighbors of i is denoted by $\bar{N}_{it} = \frac{1}{d_i} \sum_{j \in \tilde{n}_i} N_{jt}$. 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

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

As another application of Occam's razor, we assume that the the productivity parameters A, β, and ϕ are equal across countries. 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 is not profitable (characterizing the state of the world for most of human history). The parameter ϕ controls for potential scale effects in innovation.

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. 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.[6]

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 country specific conditions for R&D are sufficiently bad (A_i is sufficiently low) 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 paper we assume that $\bar{\mathbf{W}}$ is connected.[7] If the underlying network were disconnected, its separate components would behave as separate (small) worlds.

[6] 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.

[7] 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 $\left(\bar{\mathbf{W}}^k\right)_{ij} > 0$.

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

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

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 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 $GDP_{it} = Y_{it} - (1-\alpha)L\left[(1-\alpha)^{-1/\alpha} N_{it-1} + (N_{it} - N_{it-1})\right] - Z_{it}$. Notice that GDP per capita is independent from scale when Z_{it} is independent from scale, i.e. for $\phi = 1$.

Let $N_t^w \equiv \sum_{i=1}^n N_{it}$ and $\bar{N}_t \equiv N_t^w / n$ denote the world-wide produced varieties of intermediate goods and the average number of varieties per country, respectively. Furthermore, let $g_t^N \equiv N_{t+1}^w / N_t^w$ denote the growth rate of the world's varieties and $x_{it} = N_{it}/\bar{N}_t$ country i's relative endowment, i.e. the ratio between country i's varieties and average varieties. The dynamics of relative varieties are then determined by $x_{it+1} = g_{it}^N / g_t^N \cdot x_{it}$.

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

Corollary 1. *When all countries start with equal initial endowment N_0 and there are no isolated countries, 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 $t = 1, 2, ...$*

Proposition 1 (Long-Run Growth). *If the network $\bar{\mathbf{W}}$ is connected, the world economy converges towards 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 $g^N = (\alpha L)^{\frac{1-\phi}{\phi}} A$. A sufficient condition for long-run growth is that all countries are endowed with a number of varieties greater than $\beta / \left[A(\alpha L)^{1/\phi}\right]$.*

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 $Y_{it}/N_{it} = L\left[(g^N)^{-1}(1-\alpha)^{-(1-\alpha)/\alpha} + g^N - 1\right]$. 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.

If the network is connected, all knowledge is eventually shared by all countries. This feature implies that the world economy converges towards a steady state, i.e. 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. 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 which renders convergence towards 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 intermediated 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.

PROPOSITION 2 (Overshooting Growth). *Suppose that the network $\bar{\mathbf{W}}$ is connected and supports a steady state of long-run growth, $g^N > 0$.*
(i) Forerunners of the industrial revolution converge monotonously towards g^N.
(ii) Followers of the industrial revolution converge non-monotonously 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 much shorter amount of time than the leaders of the industrial revolution (Parente and Prescott, 1994, 2005).

In order to develop a comprehensive picture of the evolution of world income inequality, we distinguish between relative and absolute income inequality. 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, while 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. In order to operationalize these ideas, 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 proposition summarizes the main properties of these measures of inequality.

PROPOSITION 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.

PROPOSITION 4. *[The World Kuznets Curve] Suppose the network $\bar{\mathbf{W}}$ is connected.*
(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, then relative income inequality increases initially and declines subsequently.

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 while income inequality measured by the (relative) Gini or Theil index disappears.

PROPOSITION 5. *[No Convergence in Levels] Suppose that the network* $\bar{\mathbf{W}}$ *induces a steady state of long-run growth* $g^N > 0$. *Despite eventually declining relative world inequality there is not necessarily convergence of income levels. In other words, if the network is connected 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}, followed by an "outer" operation of multiplication by a constant, $A(\alpha L)^{1/\phi-1}$, which magnifies this range again. Ignore for a moment that the dynamics in (8) is more complicated than (11). Now, recall that the classical Gini index is based on relative measures such that the outer operation of multiplying by a constant is mitigated. If, however, the contraction effect of averaging over neighborhoods is offset by the (repeated) multiplication by the constant $A(\alpha L)^{1/\phi-1}$, the *absolute* Gini index fails to tend to zero.
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 paper. 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 $\sigma\left(\bar{\mathbf{W}}\right)$ is the spectrum of $\bar{\mathbf{W}}$. Cohesion is a measure between 0 and 1. 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 $\kappa = 1/2$.[8]

Proposition 6. *Network cohesion is positive if the network* $\bar{\mathbf{W}}$ *is connected. Higher cohesion implies faster convergence to the balanced growth path. In particular, an uppper 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. Adjustment Dynamics: The Evolution of Innovation and Inequality across 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, while poor countries are stagnating 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 conditions in pre-modern 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.[9]

---

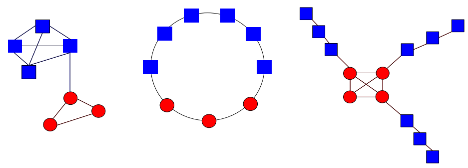
[8] See Corollary 1 of Calvacanti et al. (2016).

[9] 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.

4.1. **Stylized Networks.** We first 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 2. 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.

Figure 2: Stylized Networks



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.

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. Here, 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.

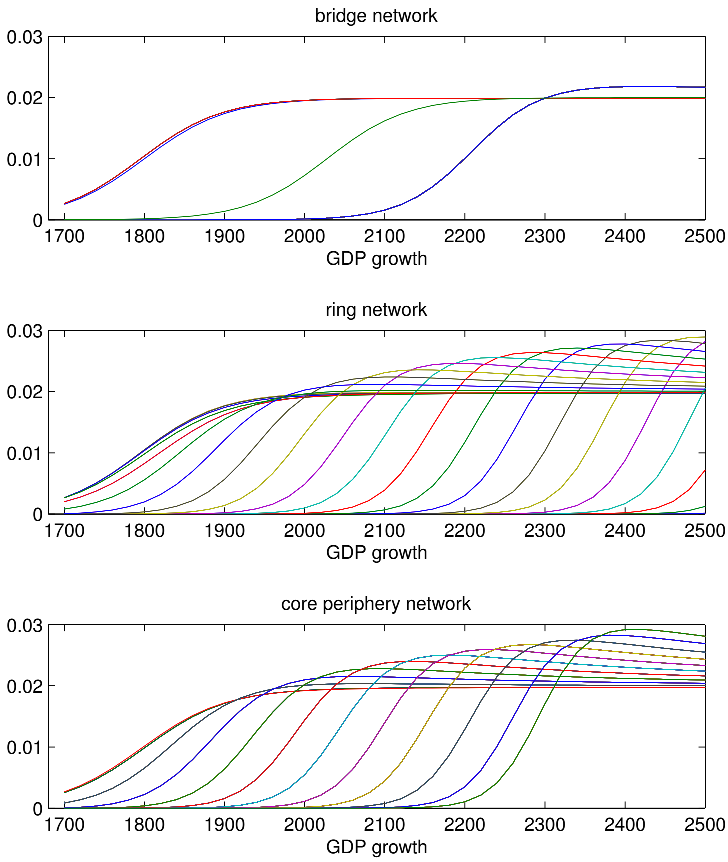
4.2. **Numerical Specification.** We begin with a benchmark specification of the model. Later on, we discuss the sensitivity of results on parameter choice. Suppose the world consists of 100 countries of which 10 percent are initially rich. We set $\phi = 1$ in order to eliminate scale effects. We set the labor share α to 0.65 and adjust the value of A such that the implied steady-state growth rate is about 2 percent 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 world-wide 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, i.e. shortly before the onset of the first industrial revolution. 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-a-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 while rich countries initially grow at a low rate of around 0.3 percent.

Figure 3 shows the evolution of growth predicted by the numerical experiments. For better comparability with the data, the gross growth rate per 20 years period from (9) is converted into net growth per year. 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-a-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.

The fact that growth rates of latecomers overshoot the balanced growth rate means that relative income inequality declines eventually. The explanation is that latecomers, once growth is initiated, tap into a greater 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.

Figure 3: Evolution of Economic Growth in 100 Countries when the World is a...



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: 2 neighbors per node. Core-periphery: 9 peripheries of 10 countries. 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 $9.87 \cdot 10^{-4}$.

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 3. 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. 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. Generally, we observe that overshooting growth is higher, the later the take-off time is. This is the case because there is more to learn from the neighbors once

R&D becomes worthwhile for the latecomers. Consequently, growth during the early take-off phase is much faster for latecomers. While it took about 200 years for the forerunners of the industrial revolution to reach a growth rate of 2 percent, the countries taking off in the 1950s needed only about two generations to achieve the same rate of growth.

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.

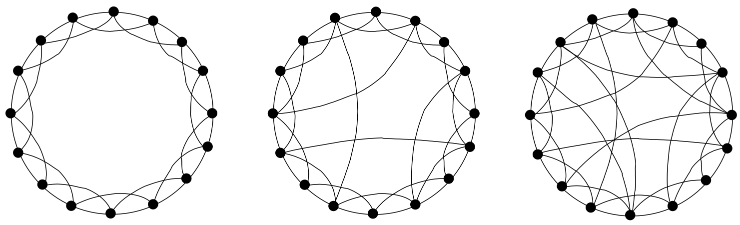
The core-periphery network, shown in the bottom panel of Figure 3, eliminates some of the flaws of the two previous networks. It produces a variety of growth experiences, largely overshooting growth rates, and a reasonable duration of the "era of take-offs to growth" from 1700 to the mid 21st century. 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 3. By the year 1800, the forerunners of the industrial revolution are counterfactually predicted to grow already at a rate of 1.5 percent annually.

5. Knowledge Diffusion and Income Evolution in A Small World

5.1. **Model Setup.** The small world model (Watts and Strogatz, 1998) is a device to investigate 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 4, $m = 2$. Each country is connected to 4 neighbors, 2 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 4 shows an example for which p is low and the panel on the right shows an example for larger p.

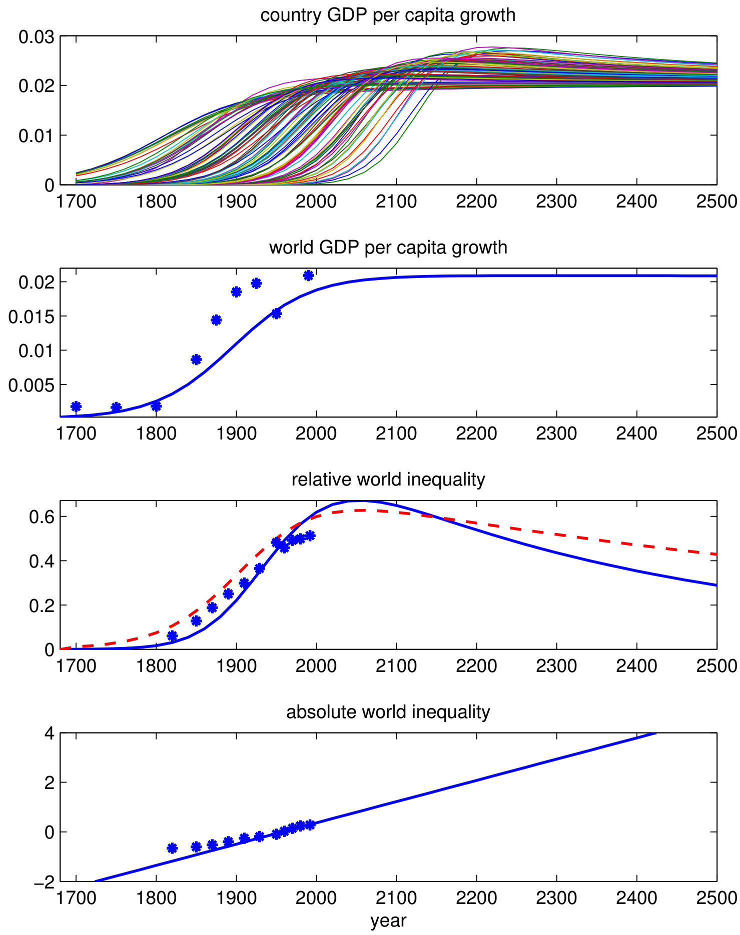
Figure 4: Small World Network



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 US 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 US benefits directly from knowledge created in England while Guatemala benefits only indirectly via the US. Moreover, in order for the knowledge to arrive in Guatemala it has to cross Mexico, another initially backward country, such that a part of the knowledge created in England gets "lost in transition".

5.2. **Results for the Benchmark Model.** As a benchmark we take the specification of the economic model from above and consider $p = 0.3$, i.e. the case in which 30 percent of the countries are equipped with a long-distance link. We subsequently provide sensitivity analysis with respect to p and other important parameters. We assume that the initially rich countries are all direct neighbors. Figure 5 shows the implied adjustment dynamics for the benchmark case. In contrast to the simple networks discussed above, 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.

Figure 5: Economic Growth in a Small World



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. Parameters as for Figure 3; small world based on ring with 1 neighbor to left- and right-hand side; long-distance links added with probability $p = 0.3$.

Comparing the model prediction 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 percent 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 percent 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 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.

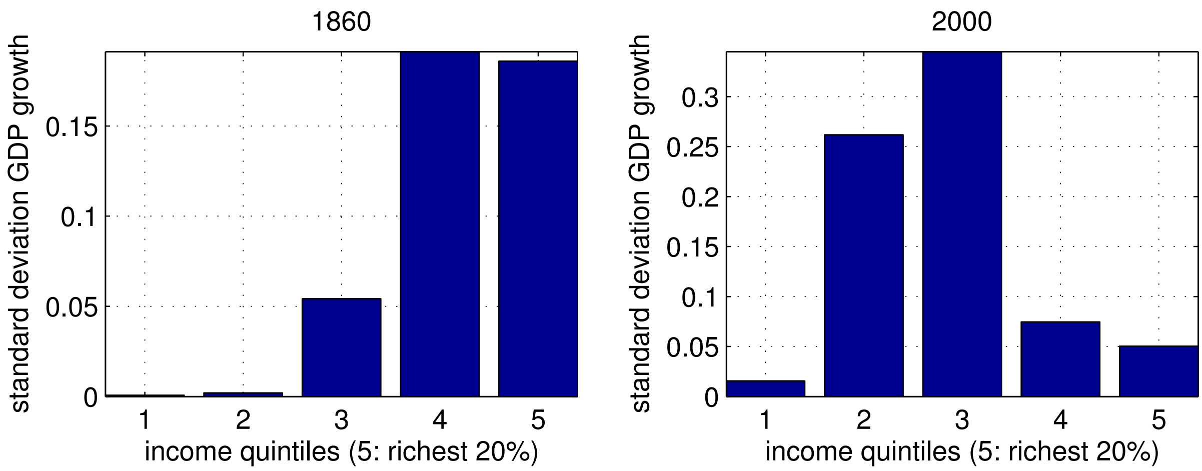
The second panel of Figure 5 shows the implied average economic growth in the world. Dots represent the data points from De Long (1998) shown in Figure 1. The model predicts the take-off of aggregate world growth reasonably well. World growth rises from almost zero to just below 1 percent in the mid-19th century and to about 1.5 percent 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.

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 5, 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 21st century. From then onwards, 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.

The focus on the conventional Gini index, however, conceals that absolute world inequality keeps on rising. The bottom panel of Figure 5 shows the absolute Gini, i.e. the relative index from the third panel multiplied by mean income. The log-scaling means that absolute inequality grows exponentially. These findings illustrate Proposition 5 and highlights 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 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.

5.3. **Variation in Growth Rates.** In our simple model of constant factor endowments, TFP growth coincides with GDP growth. In order to further explore the evolution of growth, we sorted the countries for any time t into income quintiles, with the poorest 20 percent of countries in the first quintile and the richest 20 percent in the 5th quintile. The variability of growth rates is shown in Figure 6, 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 2nd and 3rd income quintile. 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.

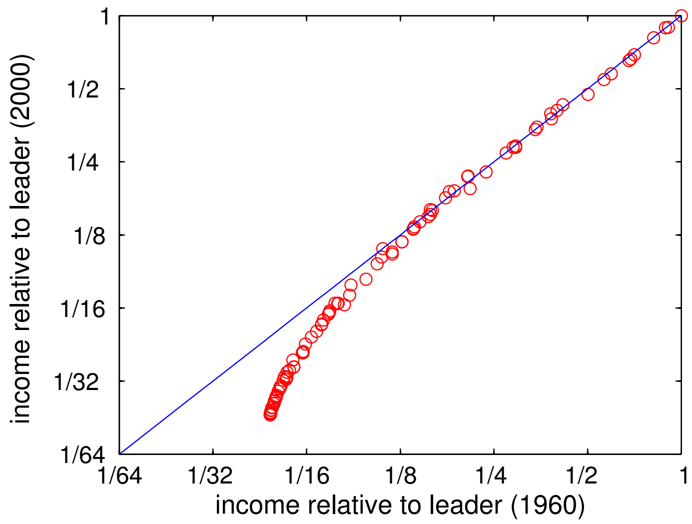
Figure 6: Standard Deviation of GDP Growth Across Countries



5.4. **Is the Present World Income Distribution Close to Its Steady State?** As evidenced in Figure 5, 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

Figure 7: Income Relative to Leader: 1960 vs. 2000



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

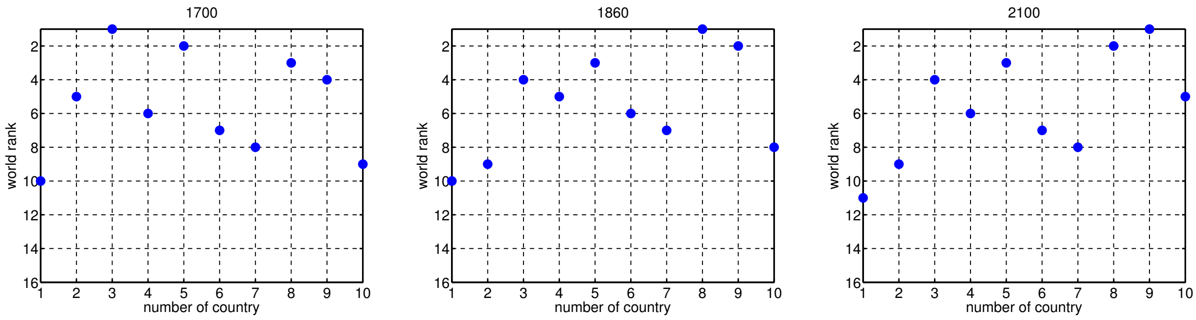
relatively short and if the observation happens to be taken at a period of time when the trajectory of relative world inequality is flat because it is close to its maximum. In order to verify this claim by way of example, we compute for the outcome from the benchmark economy a relative income plot similar 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 5 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 7. 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 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 towards a constant unequal world income distribution. In fact, however, we know from Proposition 7 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 e.g. 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, i.e. it is neither based on technological assumptions or stochastic elements nor does it imply non-convergence (of relative income levels).

To demonstrate this extraordinary behavior, we perform the following numerical experiment. We follow the 10 initially richest countries, named $1, 2, \ldots, 10$, along the way towards the steady state and visualize their relative position in the world income ranking. Figure 8 shows the resulting "income ladders" for four 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.

Figure 8: World Ranking Position for Country 1 - 10



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 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 and the original number one dropped to the sixth place while 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.

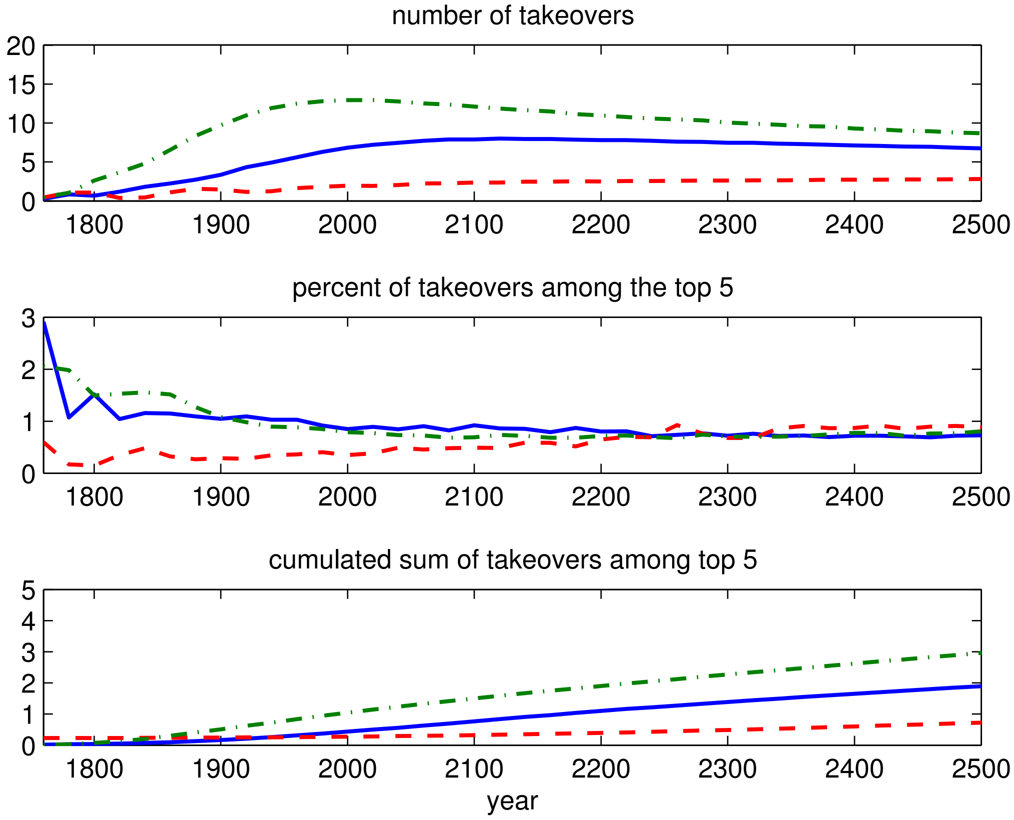
In order to develop an intuition for these results, consider a “network” of two countries, one with an initial variety of products N, the other with an initial variety $N+\Delta$. Neglecting the corner solution, and assuming $\phi = 1$, the equation of motion (8) for the first country is given by $f_1(N, N+\Delta) = N+\{A\left[\epsilon(N+\Delta)+(1-\epsilon)N\right]-\bar{c}\}$, with $\bar{c} \equiv \beta(\alpha L)^{-1}$. For the second country, it is given by $f_2(N, N+\Delta) = (N+\Delta)+\{A\left[\epsilon N+(1-\epsilon)(N+\Delta)\right]-\bar{c}\}$. Consider the implausible yet illuminating case in which $A > 1$ and *all* knowledge comes from abroad, i.e. $\epsilon = 1$. 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. 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 one hundred 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 5000 times (i.e. for 5000 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 1 overtaking. However, if it advances from rank 5 to rank 3 we count 2 overtakings.

The results for the benchmark model are shown by solid lines in Figure 9. The top panel shows the total number of overtakings per period. On average, we observe about 8 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 8 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).

Although overtaking takes place surprisingly 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 9 where we consider the top 5 countries in terms of GDP per capita. On average, only about 1 percent of overtaking takes place among the top 5. If overtakings were equally distributed among countries, we would have expected about 5 percent of them taking place in the top 5. When 1 percent 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 5. In order to better assess these results quantitatively, the bottom panel shows the cumulated sum of average overtakings among the top 5. For the benchmark

Figure 9: Overtaking Frequencies



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

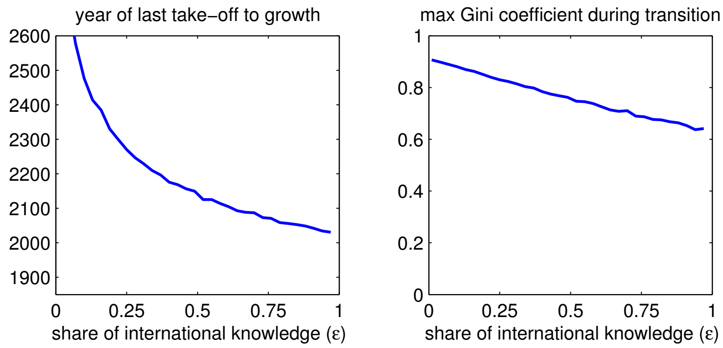
case (solid lines), there is less than 1 overtaking happening before the year 2000 and 2 overtakings before the year 2500.

The incidence of overtaking, naturally, depends on the degree of openness (ϵ). Dashed lines in Figure 10 show that there are fewer overtakings in total and among the top 5, when only 20 percent 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 1000 times and took averages afterwards.

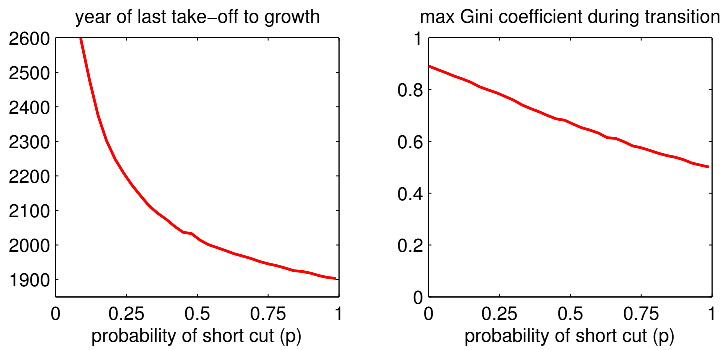
Figure 11 shows that a large contribution of international knowledge flows to productivity (that is large ϵ) increases the pace of world development. Larger international knowledge spillovers are helpful to reduce world-wide inequality faster because a greater share of the initial knowledge advantage of the leaders is passed on through the network. Average network cohesion as defined in (6) 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 knowledge externality (7) is entirely dependent on the neighborhood for $\epsilon = 1$.

Figure 10: Inequality and Growth – Degree of International Knowledge Diffusion



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

Figure 11: Inequality and Growth – Share of Long-Distance Links



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

Next, in Figure 11, 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 p's 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 $p = 1$.[10]

5.7. **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 1000 times for alternative values of σ and record the year of the last take-off and the maximum Gini index during the transition.

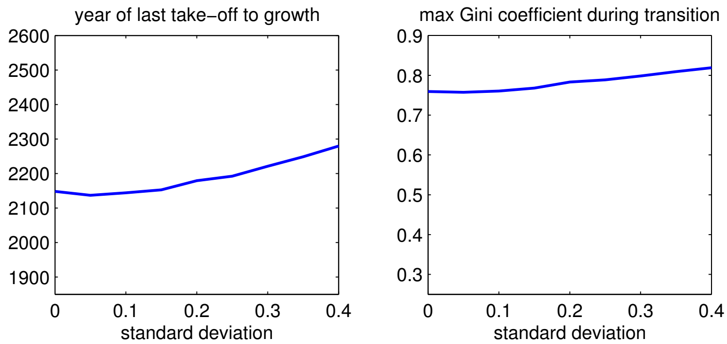
Figure 12 shows the outcome for alternative $\sigma \in (0, 0.4)$ and ϵ drawn from a (truncated) normal distribution.[11] For better comparison, we kept the scaling of Figure 11. There is almost no change in the average maximum Gini along the transition and the year of the last take-off as the standard deviation of the degree of openness increases from 0 (our benchmark case) to 0.4.

Summarizing the results from the last two subsections, we observe that the speed of industrialization and the evolution of world inequality crucially depends on the aggregate structure of the network, as determined by the (average) degree of openness (ϵ) and the share of long-distance links (p). The distribution of comparative advantage of countries as measured, for example, by the standard deviation of country-specific ϵ, has comparatively minor effects on aggregate behavior.

[10] 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.

[11] 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.

Figure 12: Inequality and Growth – Varying Distribution of Openness



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

6. Conclusion

In this paper, 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, however, 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. Naturally, further extensions are conceivable. In the present paper, we conceptualized globalization as the increasing flow of knowledge through the world. The network itself, however, may be subject to globalization as well. Integrating an increasing share of long-distance links over time, as in Lindner and Strulik (2015a), could be an interesting extension of our network-based theory of global knowledge diffusion and growth.

References

Acemoglu, D. (2009). *Introduction to Modern Economic Growth.* Princeton University Press.

Acemoglu, D. and Ventura, J., 2002, The world income distribution, *Quarterly Journal of Economics* 117, 659-694.

Acemoglu, D., Aghion, P., and Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association* 4(1), 37-74.

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

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

Ashraf, Q., and Galor, O., 2013, The "Out of Africa" hypothesis, human genetic diversity, and comparative economic development. *American Economic Review* 103(1), 1-46.

Atkinson, A.B. and Brandolini, A., 2010, On analyzing the world distribution of income, *The World Bank Economic Review* 24, 1-37.

Bairoch, P., 1993, *Economics and World History*, University of Chicago Press, Hemel Hempstead.

Baldwin, R.E., Martin, P., and Ottaviano, G.I. (2001). Global income divergence, trade, and industrialization: The geography of growth take-offs. *Journal of Economic Growth* 6(1), 5-37.

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

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

Bourguignon, F. and Morrisson, C., 2002, Inequality among world citizens: 1820-1992, *American Economic Review* 92, 727-744.

Brezis, E.S., Krugman, P.R., and Tsiddon, D. (1993). Leapfrogging in international competition: A theory of cycles in national technological leadership. *American Economic Review* 83, 1211-1219.

Cavalcanti, T. and Giannitsarou, D., 2017, Growth and Human Capital: A Network Approach, *Economic Journal*, forthcoming.

Cavalcanti, T., Giannitsarou C. and Johnson, C.R., 2016, Network cohesion, *Economic Theory* 62, 1-21.

Chakravarty, S. R., 1988, Extended Gini indices of inequality, *International Economic Review* 29, 147-156.

Comin, D. A., Dmitriev, M., and Rossi-Hansberg, E., 2012. The spatial diffusion of technology, NBER Discussion Paper 18534.

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

De Long, B., 1998, Estimates of World GDP, One Million BC – Present. University of California, Berkeley, Department of Economics, `http://delong.typepad.com/print/20061012_LRWGDP.pdf`.

Diamond, J., 1997, *Guns, Germs, and Steel: The Fates of Human Societies* WW Norton, New York.

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

Easterly, W. and Levine, R., 2001, What have we learned from a decade of empirical research on growth? It's not Factor accumulation: stylized facts and growth models, *The World Bank Economic Review* 15, 177-219.

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

Eurostat (2012), Educational expenditure statistics (epp.eurostat.ec.europa.eu/statistics_explained/index.php/Educational_expenditure_statistics)

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

Frankel, J.A., and Romer, D., 1999, Does trade cause growth? *American Economic Review* 89(3), 379-399.

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

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

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

Galor, O. and Moav, O., 2004, From physical to human capital accumulation: Inequality and the process of development, *Review of Economic Studies* 71, 1001-1026.

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

Gerschenkron, A., 1962, Economic backwardness in historical perspective, in: Hoselitz, B. (ed.) *The Progress of Underdeveloped Areas*, Chicago University Press, Chicago.

Hovhannisyan, N., and Keller, W., 2015, International business travel: an engine of innovation? *Journal of Economic Growth* 20(1), 75-104.

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

Jackson, M.O. and Yariv, L., 2007, Diffusion of Behavior and Equilibrium Properties in Network Games, *American Economic Review* (Papers and Proceedings), 97(2), 92-98.

Jones, C. I., 1997, On the evolution of the world income distribution, *Journal of Economic Perspectives* 11(3), 19-36.

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

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

Kaldor, N., 1961, Capital accumulation and economic growth, in: *The Theory of Capital*, ed. F.A. Lutz and D.C. Hague, St. Martins Press, New York.

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

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

Kolm, S.-C., 1976, Unequal inequalities I, *Journal of Economic Theory* 12, 416-442.

Kuznets, S., 1966, *Modern Economic Growth: Rate, Structure and Spread*, Yale University Press, New Haven.

Lindner, I. and Strulik, H., 2015a, From Tradition to Modernity: Economic Growth in a Small World, *Journal of Development Economics* 109, 17-29.

Lindner, I. and Strulik, H., 2015b, The Great Divergence: A Network Approach (with Ines Lindner), CESifo WP No. 5638.

Lucas, R.E., 2000, Some Macroeconomics for the 21st century, *Journal of Economic Perspectives* 14, 159-168.

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

Meyer, C.D., 2000, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelhpia.

Newman, M.E.J. and Watts, D.J., 1999, Renormalization group analysis of the small-world model, *Physics Letters A* 263, 341-346.

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

Ortega, F., and Peri, G., 2014, Openness and income: The roles of trade and migration. *Journal of International Economics* 92(2), 231-251.

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

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

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

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

Sala-i-Martin, X., 2006, The world distribution of income: falling poverty and convergence, period, *Quarterly Journal of Economics* 121, 351-397.

Spolaore, E., and Wacziarg, R., 2013, How deep are the roots of economic development? *Journal of Economic Literature* 51(2), 325-369.

Strulik, H. (2014). Knowledge and growth in the very long run. *International Economic Review* 55(2), 459-482.

Strulik, H. and Weisdorf, J., 2008, Population, food, and knowledge: A simple unified growth model, *Journal of Economic Growth* 13, 169-194.

Strulik, H., Prettner, K., and Prskwetz, A., 2013, The past and future of knowledge-based growth, *Journal of Economic Growth*, 18, 411-437.

Sutcliffe, B., 2004, World inequality and globalization, *Oxford Review of Economic Policy* 20, 15-37.

Ventura, J., 2005, A global view of economic growth, in: P. Aghion and S. Durlauf, *Handbook of Economic Growth* Vol 1B, Amsterdam, North-Holland.

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

Appendix

For the proof of the propositions from the main text it is useful to start with proving two lemmas.

Lemma 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

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

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

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

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

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 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 paper 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 non-negative) eigenvectors except positive multiples of x, i.e., all other eigenvectors must have at least one negative or non-real 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 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 2. 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

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

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 of Calvacanti (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 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 $i = 1, \ldots n-1$. We will first show that the eigenvalues of $\bar{\mathbf{W}}$, of F respectively, are real. Put $D := diag\{d_1, \ldots, d_n\}$ such that D is a diagonal matrix with and d_i is the number of links of country i. 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

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

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

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

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

Proof of Proposition 1. From (8) follows that N_{it} grows iff

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

which is equivalent to

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

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

In case of long-run growth, Lemma 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)

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

□

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

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

which prevents overshooting.

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

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

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

Proof of Proposition 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 $y_1 \leq y_2 \leq ... \leq y_n$. Put $Y_k = \sum_{i=1}^{k} y_i$. The Lorenz curve is a polygonal line defined by the set of points

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

If all countries grow by the same rate, the fractions Y_i/Y stay the same for all $i = 1, ..., n$. 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 re-scaled Lorenz curve where the horizontal axis ranges from 0 to n and the vertical axis from 0 to Y. The re-scaled Lorenz curve is a polygonal line defined by the set of points

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

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

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

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

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

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

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

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 4. (i) Follows directly from Lemma 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 2, however, the index tends to zero eventually. □

Proof of Proposition 5. 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 4. 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

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

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

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

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

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

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 3 this is equivalent to stating that the absolute Gini index tends to infinity. □

Proof of Proposition 6. 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

$$
\frac{g^N - 1}{g^N}\lambda_2 + \frac{1}{g^N}, \qquad \text{(A.25)}
$$

with $g^N = (\alpha L)^{\frac{1-\phi}{\phi}} A$, is an upper bound for the rate of convergence. Inserting (6) implies that (A.25) is equivalent to

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

□