If Technology Has Arrived Everywhere, Why Has Income Diverged?

Diego Comin
Harvard University and NBER

Martí Mestieri
TSE

November 15, 2012
PRELIMINARY

Abstract

We study the dynamics of the lags with which new technologies are adopted over the last two centuries, and how intensively new technologies are demanded once they are adopted. We document two new facts: there has been convergence in adoption lags between rich and poor countries, while there has been divergence in the intensive margin of adoption. Using a model of adoption and growth, we show that these changes in the pattern of technology diffusion account for at least two thirds of the Great Income Divergence between rich and poor countries since 1800.

Keywords: Technology Diffusion, Transitional Dynamics, Great Divergence.
JEL Classification: E13, O14, O33, O41.

1 Introduction

There is increasing evidence that cross-country differences in technology are key to account for existing cross-country differences in productivity. Rich countries tend to adopt new technologies faster than poor countries. Faster adoption allows rich countries to enjoy the productivity gains that new technologies bring, leading to higher TFP and labor productivity. Comin and Hobijn (2010) have quantified these effects and concluded that they account for at least 25% of observed cross-country differences in productivity.

The rate of arrival of technologies to countries has evolved over the last two centuries. Adoption lags have declined dramatically.[1] Technologies invented in the nineteenth century such as telegrams or railways often took many decades to first arrive to countries. In contrast, new technologies such as computers, cellphones or the internet have arrived on average within a few decades (in some cases less than one) after their invention. The decline in adoption lags has surely not been homogeneous across countries. Anecdotal evidence suggests that the reduction in adoption lags has been particularly significant in developing countries, where technologies have traditionally arrived with longer lags.[2]

If this anecdotal evidence is representative and indeed adoption lags have converged, then we should wonder how to square the cross-country dynamics of technology with the cross-country dynamics of income. Cross-country differences in per-capita income have increased dramatically over the last 200 years. A phenomenon known as the Great Divergence (Pritchett, 1997, and Pomeranz, 2000). Maddison (2004) shows that the income gap between countries at the technology frontier, which he labeled as Western, and the rest of the world was seven-fold in 2000, and that around 75% of this gap emerged in the last 180 years.[3] If we focus just on the last 50 years for which we have more precise estimates of income, it is well known that cross-country income differences have not declined and probably they have increased. How is that possible given the evolution of the distribution of adoption lags across countries?

In this paper, we explore the cross-country dynamics of technology and income over the last two centuries. In particular, we investigate two questions. First, how have the relevant dimensions of technology diffusion evolved across countries over the last 200 years. Second, how have the cross-country evolution of technology dynamics affected the evolution of income growth in different countries over the last 200 years. And, do they help us explain the Great Divergence.

The contribution of the technology to a country's productivity growth can be decomposed in two parts. One part is related to the range of technologies used in the country. New

[1]See Comin and Hobijn (2010).

[2]Khalba (2007), Dholakia, Dholakia and Kshetri (2003).

[3]Maddison (2004) defines Western countries as the following 17 countries: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, Untied Kingdom, Australia, New Zealand, Canada and the United States of America. To be more precise, the relative median income per-capita Western to non-Western countries was 1.85 in 1820 and 7 in 2000.

technologies embody higher productivity. Therefore, an acceleration in the rate at which new technologies arrive in the country raises productivity growth. Productivity is also affected by the penetration rate of new technologies. The more units of any new technology (relative to income) a country uses, the higher the number of workers or units of capital that can benefit from the productivity gains brought by the new technology. It follows then that increases in the penetration of technology (or as we call it below, the intensive margin of adoption) raise the growth rate of productivity.

To identify adoption lags (extensive margin) and penetration rates (intensive margin) of technology, we follow Comin and Hobijn (2010) and Comin and Mestieri (2011). To illustrate their strategy, consider Figures **??** and **??** which plot respectively the (log) of the tonnage of steam and motor ships over real GDP in the UK and Indonesia and the (log) number of computers over real GDP for the U.S. and Vietnam. One feature of these plots is that the diffusion curves for different countries are similar, but displaced vertically and horizontally. Comin and Hobijn (2010) show that, to a first approximation, this is a general property for the diffusion curves of a given technology across countries. Given the common curvature of diffusion curves, the relative position of a curve can be characterized by only two parameters. The horizontal shifter informs us about when the technology was introduced in the country. The vertical shifter captures the penetration rate the technology will attain when it has fully diffused.

These intuitions are formalized with the help of a simple model of technology adoption and growth. The model features both adoption margins, and has predictions about how variation in these margins affect the curvature and level of the diffusion curve of specific technologies. Using the CHAT data set,[4] we identify the extensive and intensive adoption margins for twenty significant technologies invented over the last 200 years in an (unbalanced) sample that covers over 150 countries.

Figures **??** and **??** reveal our two main findings on the evolution of adoption patterns. First, the horizontal shifter between the diffusion curves for steam and motor ships in the UK and Indonesia is much larger than the horizontal shifter between the U.S. and Vietnam for computers (131 years vs. 11). More generally, we show that cross-country differences in adoption lags have narrowed. That is, adoption lags have declined more in poor/slow adopter countries than in rich/fast adopter countries. This trend is not a recent phenomenon, but a secular phenomenon that started 200 years ago. Second, the vertical gap between the curves for ships in the UK and Indonesia are smaller than the vertical gap between the diffusion curves of computers in the U.S. and Vietnam (0.9 vs. 1.6). In section 3, we show the generality of this observation by documenting the divergence in the degree of penetration of technologies across countries over the last 200 years.

After characterizing the dynamics of technology, we explore its consequences for the cross-

[4] Comin and Hobijn (2004) and Comin, Hobijn and Rovito (2008)

country dynamics of income. To this end, we take advantage of the fact that our model of adoption and growth has an aggregate representation that we can use to study its transitional dynamics. Unlike the neoclassical growth model, ours does not have capital. The evolution of the growth rate of the economy is instead associated with the evolution in the stock of unadopted technologies, and with changes in adoption lags and in the intensive margin, which we take as given. These forces help us better understand the dynamics of income growth during the last two centuries.

Our model of adoption generates extremely protracted transitional income dynamics despite not having physical capital, habit formation or other mechanisms used in macroeconomics to generate slower transitions. For example, after a one-time permanent shift in the growth rate of invention of new technologies (which captures the industrial revolution) the half-life for income is approximately 80 years while the half-life for the growth rate in income is 100 years. By way of comparison, the equivalent numbers in the neoclassical growth model are 14 and 1 year.

After feeding in the estimated dynamics of technology adoption, the model generates cross-country patterns of income growth that resemble very much those observed in the data over the last two centuries. In particular, in developed economies, it took approximately one century to reach the long-run growth rate of productivity (2%) while in developing economies it takes twice as much, if not more. As a result, the model generates an increase in the income gap between rich and poor countries by a factor of 5 which is approximately three-fourths of the actual increase observed over the last two centuries. Finally, our preliminary results suggest that enriching the model to have capital accumulation does not affect significantly these findings.

Our findings are related to several lines of research. Our exploration is motivated by Klenow and Rodríguez-Clare (1997) and Clark and Feenstra (2003) who find that neoclassical dynamics are not very relevant to account for the cross-country distribution of productivity growth over long intervals (the period 1960-85 for Klenow and Rodríguez-Clare (1997) and the last 150 years for Clark and Feenstra (2003)). The model we use is a version of the models used in Comin and Hobijn (2010) and Comin and Mestieri (2011). However, there are stark differences with these previous attempts to explore empirically diffusion dynamics. First, these authors do not study the cross-country dynamics of the adoption margins. Second, they do not explore the transitional dynamics of the growth and adoption model. Third, they do not study the implications of technology dynamics for the dynamics of income, they just study the relationship between technology and the **level** of income in the **steady state**.

This paper is also related to the literature that has explored the drivers of the Great Divergence. Part of this literature has emphasized the role of the expansion of international trade during the second half of the nineteenth century. Galor and Mountford (2006) argue that trade affected asymmetrically the fertility decisions in developed and developing economies, due to their different initial endowments of human capital, leading to different evolutions in productivity growth. O'Rourke et al. (2012) elaborate on this perspective and argue that the direction of technical change, in particular the fact that after 1850 it became skilled bias (Mokyr, 2002), also contributed to the increase in income differences across countries. Trade-based theories of the great divergence, however, need to confront two facts. Prior to 1850, the technologies brought by the industrial revolution were unskilled-bias rather than skilled bias (Mokyr, 2002). Why did incomes diverge also during this period. Furthermore, the trade boom ended abruptly in 1913 with WWI and world trade reverted back and did not expand until the 1970s. Yet, the great divergence continued at similar rates until now. Probably motivated by these questions, another strand of the literature has emphasized studied the cross-country evolution of Solow residuals and has found that they account for the majority of the divergence (Easterly and Levine, 2000, Clark and Feenstra, 2003). These authors have interpreted this finding as evidence on the importance of technology differences for the great divergence. Our paper takes on that claim and measures the dynamics of technology across countries to directly assess its importance.[5]

The rest of the paper is organized as follows. Section 2 presents the expository model. Section 3 presents and implements the identification of the extensive and intensive margins of adoption, and describes the trends we observe in the cross-country evolution of both adoption margins. Section 4 characterizes the transitional dynamics of the model. Section 5 quantifies the effect of the technology dynamics on the cross-country growth dynamcis and discusses the results. Section 6 conducts some robustness checks, and section 7 concludes.

2 Model

Next we present a simple model of technology adoption and growth. Our model serves four purposes.[6] First, it precisely defines the intensive and extensive margins of adoption. Second, it illustrates how variation in these margins affect the evolution of the diffusion curves for individual technologies. Third, it helps us develop the strategies used to identify the extensive and intensive margins of adoption in the data. Fourth, because ours is a general equilibrium model with a simple aggregate representation, the expository model can be used to study the dynamics of productivity growth. In particular, we use it to understand what factors account for the different evolutions of productivity growth we have observed between rich and poor countries over the last 200 years. For the sake of clarity, we assume that the adoption of

[5]Our analysis is also related to a strand of the literature that has studied the productivity dynamics after the industrial revolution. Galor and Weil (2000), Hansen and Prescott (2002), Tamura (2003), Crafts (1997), among others, provide different reasons why there was a slow growth acceleration in productivity after the industrial revolution. These arguments are different from our mechanism but perfectly consistent with it.

[6]This model is inspired by Comin and Hobijn (2010) and Comin and Mestieri (2012).

technologies occurs exogenously.[7]

2.1 Preferences and Endowments

There is a unit measure of identical households in the economy. Each household supplies inelastically one unit of labor, for which they earn a wage w. Households can save in domestic bonds which are in zero net supply. The utility of the representative household is given by

$$U = \int_{t_0}^{\infty} e^{-\rho t} \ \ln(C_t) dt \tag{1}$$

where ρ denotes the discount rate and c, consumption. The representative household, maximizes its utility subject to the budget constraint (2) and a no-Ponzi game condition (3)

$$\dot{b}_t + c_t = w_t + r_t b_t, \tag{2}$$

$$\lim_{t\to\infty} b_t e^{\int_{t_0}^{t} -r_s ds} \geq 0, \tag{3}$$

where b denotes the bond holdings of the representative consumer, $\dot{b}$ is the increase in bond holdings over an instant of time, and r_t its return on bonds.

2.2 Technology

World technology frontier .– At a given instant of time t, the world technology frontier is characterized by a set of technologies and a set of vintages specific to each technology. Each instant, a new technology, τ, exogenously appears. To simplify notation, we omit time subscripts, t, whenever possible. We denote a technology by the time it was invented. Therefore, the range of invented technologies is $(-\infty, t]$.

For each existing technology, a new, more productive, vintage appears in the world frontier every instant. We denote vintages generically by v. The productivity of a technology-vintage pair has two components. The first component, $Z(\tau, v)$, is common across countries and it is purely determined by technological attributes. In particular,

$$Z(\tau, v) = e^{\chi\tau + \gamma v}, \tag{4}$$

where $(\chi + \gamma)\tau$ is the productivity level associated with the first vintage of technology τ and $\gamma(v - \tau)$ captures the productivity gains associated with the introduction of new vintages.

In addition, we allow for a technology-country specific productivity, a_τ. a_τ may differ across countries and, as we shall see below, it determines the relative penetration of a technology across countries in the long term. Consequently, we call it the intensive margin of adoption.

[7] See Comin and Hobijn (2010) and Comin and Mestieri (2011) for straightforward ways to enodegnize these adoption margins.

Adoption lags .– Economies typically are below the world technology frontier. Let D_τ denote the age of the best vintage available for production in a country for technology τ. D_τ reflects the time lag between when the best vintage in use was invented and when it was adopted for production in the country; that is, the *adoption lag.* The set of available vintages for technology τ in this economy is $V_\tau = [\tau, t - D_\tau]$.[8]

Intensive margin .– Production consists of several layers. Each technology-vintage (τ, v) – vintage, for brevity – is embodied in an intermediate good. Intermediate goods are produced competitively using final output. We assume that it takes one unit of final output to produce one unit of intermediate good.

Intermediate goods are combined with labor to produce services associated with a given vintage. In particular, let $x_{\tau,v}$ be the number of units of intermediate good (τ, v) used in production, and $L_{\tau,v}$ be the number of workers that use them to produce services. Then, the amount of services associated with technology (τ, v), $Y_{\tau,v}$, is given by

$$Y_{\tau,v} = a_\tau Z(\tau, v) X_{\tau,v}^{\alpha} L_{\tau,v}^{1-\alpha} \tag{5}$$

The term a_τ in (5) represents factors that reduce the effectiveness of a technology in a country. A key difference between a_τ and $Z(\tau, v)$ is that while $Z(\tau, v)$ is specific to a vintage, a_τ affects symmetrically the productivity of all the vintages associated with a technology. As we show below, this asymmetry has implications on how different adoption margins affect diffusion curves.

Taken literally, a_τ introduces cross-country differences in the productivity of the technology in (5). However, we regard this formulation as a shortcut for a variety of factors that may *affect the penetration of the technology in the long term.* Hence, we refer to a_τ as the *intensive margin* of adoption of a technology.

More generally, a_τ can be driven by differences in the costs of producing the intermediate goods associated witha technology, taxes, relative abundance of complementary inputs or technologies, frictions in capital, labor and goods markets, barriers to entry for producers that want to develop new uses for the technology, etc.[9] Regardless of the specific nature of a_τ, what is relevant for our characterization of technology is that a_τ affects the level of both the output produced with the technology and the inputs associated with it. As we shall see below, the identification strategy we implement allow us to compute the effect that a_τ has on productivity *though its impact in the penetration of technology.*

*Production .–*The services associates with different vintages of the same technology are imperfect substitutes and can be combined to produce sectoral output, $Y_{\tau,}$ as follows

[8] Here, we are assuming that vintage adoption is sequential. Comin and Hobijn (2010) provide a micro-founded model in which this is an equilibrium result rather than an assumption.

[9] Comin and Mestieri (2012) discuss how a wide variety of distortions result in wedges in technology adoption that imply a reduced form as in (5).

$$Y_\tau = \left(\int_\tau^{t-D_\tau} Y_{\tau,v}^{\frac{1}{\mu}} \, dv \right)^\mu, \qquad \text{with} \quad \mu > 1. \tag{6}$$

Final good production is an aggregate of the output produced using different technologies

$$Y = \left(\int_{-\infty}^{\bar{\tau}} Y_\tau^{\frac{1}{\theta}} d\tau \right)^\theta, \qquad \text{with} \quad \theta > 1. \tag{7}$$

where $\bar{\tau}$ denotes the most advanced technology adopted in the economy, that is the technology τ for which $\tau = t - D_\tau$.[10]

2.3 Factor Demands and Final Output

We take the price of final output as numeraire. The demand for output produced with a particular technology is

$$Y_\tau = Y p_\tau^{-\frac{\theta}{\theta-1}} \tag{8}$$

where p_τ is the price of sector τ output. *Both* the income level of a country and the price of a technology affect the demand of output produced with a given technology. Because of the homotheticity of the production function, the income elasticity of technology τ output is one. Similarly, the demand for output produced with a particular technology vintage is

$$Y_{\tau,v} = Y_\tau \left(\frac{p_\tau}{p_{\tau,v}} \right)^{-\frac{\mu}{\mu-1}}, \tag{9}$$

where $p_{\tau,v}$ denotes the price of the (τ, v) intermediate good. The demands for labor and intermediate goods at the vintage level are

$$(1-\alpha) \frac{p_{\tau,v} Y_{\tau,v}}{L_{\tau,v}} = w \tag{10}$$

$$\alpha \frac{p_{\tau,v} Y_{\tau,v}}{X_{\tau,v}} = 1 \tag{11}$$

Perfect competition in the production of intermediate goods implies that the price of intermediate goods equals their marginal cost,

$$p_{\tau,v} = \frac{w^{1-\alpha}}{Z(\tau,v) a_\tau} (1-\alpha)^{-(1-\alpha)} \alpha^{-\alpha} \tag{12}$$

Combining (9), (10) and (11), the total output produced with technology τ can be ex-

[10] Again, we assume, for notational simplicity, that older vintages are adopted earlier than newer ones. Our simulations do not impose this constraint.

pressed as

$$Y_\tau = Z_\tau L_\tau^{1-\alpha} X_\tau^\alpha, \tag{13}$$

where L_τ denotes the total labor used in sector τ,

$$L_\tau = \int_\tau^{t-D_\tau} L_{\tau,v} dv, \tag{14}$$

X_τ is the total amount of intermediate goods in sector τ,

$$X_\tau = \int_\tau^{t-D_\tau} X_{\tau,v} dv, \tag{15}$$

and the productivity associated to a technology is

$$\begin{aligned} Z_\tau &= \left(\int_\tau^{\max\{t-D_\tau,\tau\}} Z(\tau,v)^{\frac{1}{\mu-1}} dv \right)^{\mu-1} \\ &= \left(\frac{\mu-1}{\gamma} \right)^{\mu-1} \underbrace{a_\tau}_{\text{Intensive Mg}} \underbrace{e^{(\chi\tau+\gamma\max\{t-D_\tau,\tau\})}}_{\text{Embodiment Effect}} \underbrace{\left(1 - e^{\frac{-\gamma}{\mu-1}(\max\{t-D_\tau,\tau\}-\tau)}\right)^{\mu-1}}_{\text{Variety Effect}} \end{aligned} \tag{16}$$

This expression is quite intuitive. The productivity of a technology, Z_τ, is determined by the intensive margin, the productivity level of the best vintage used (i.e., embodiment effect), and the productivity gains from using more vintages (i.e., variety effect). Adoption lags have two effects on Z_τ. The shorter the adoption lags, D_τ, the more productive are, on average, the vintages used. In addition, because there are productivity gains from using different vintages, the shorter the lags, the larger variety gains are.

The price deflator of technology-τ output is

$$\begin{aligned} p_\tau &= \left(\int_\tau^{t-D_\tau} p_{\tau,v}^{-\frac{1}{\mu-1}} dv \right)^{-(\mu-1)} \\ &= \frac{w^{1-\alpha}}{Z_\tau} (1-\alpha)^{-(1-\alpha)} \alpha^{-\alpha} \end{aligned} \tag{17}$$

There exists an aggregate production function representation in terms of the aggregate labor (which is normalized to one),

$$Y = AX^\alpha L^{1-\alpha} = AX^\alpha = A^{1/(1-\alpha)} (\alpha)^{\alpha/(1-\alpha)}, \tag{18}$$

with

$$A = \left(\int_{-\infty}^{\bar\tau} Z_\tau^{\frac{1}{\theta-1}} d\tau \right)^{\theta-1} \tag{19}$$

where $\bar\tau$ denotes the most advanced technology adopted in the economy.

2.4 Equilibrium

Given a sequence of adoption lags and intensive margins $\{D_\tau, a(\tau)\}_{\tau=0}^{\infty}$, a competitive equilibrium in this economy is defined by consumption, output, and labor allocations paths $\{c_t, L_{\tau,v}(t), Y_{\tau,v}(t)\}_{t=0}^{\infty}$ and prices $\{p_\tau(t), p_{\tau,v}(t), w_t, r_t\}_{t=0}^{\infty}$, such that

1. Households maximize utility by consuming according to the following Euler equation:

$$\frac{\dot{C}}{C} = r - \rho \tag{20}$$

2. Firms maximize profits taking prices (equation 12) as given. This optimality condition gives the demand for labor and intermediate goods for each technology and vintage, equations (10) and (11) , for the output produced with a vintage (equation 9) and for the output produced with a technology (equation 8).

3. Labor market clears

$$L = \int_{-\infty}^{\bar{v}} \int_{\tau}^{\bar{v}} L_{\tau,v} dv d\tau = 1 \tag{21}$$

4. The resource constraint holds:

$$\begin{aligned} Y &= C + X & (22) \\ C &= (1-\alpha)Y & (23) \end{aligned}$$

Combining (21) and (10), it follows that the wage rate is given by

$$w = (1-\alpha)Y/L \tag{24}$$

Combining the Euler equation (20) and the resource constraint (23) we obtain that the interest rate depends on output growth and the discount rate

$$r = \frac{\dot{Y}}{Y} + \rho.$$

Equation (18) implies that output dynamics are completely determined by the dynamics of aggregate productivity, A. Below, we explore in depth how productivity has evolved in response to changes in χ, γ, adoption lags, and the intensive margin. For the time being, it is informative to understand the growth rate of the economy along the balanced growth path. To this end, suppose that D_τ and a_τ are constant across technologies. Further, let's make the simplifying (and empirically relevant)[11] assumption that $\theta = \mu$. Then, omitting technology subscripts,

[11] As we show below, this is what we observe in our estimation.

$$A = \left(\frac{(\theta - 1)^2}{(\gamma + \chi)\chi} \right)^{\theta - 1} a \; e^{(\chi + \gamma)(t - D)}. \tag{25}$$

Naturally, a higher intensity of adoption, a, and shorter adoption lags (D) lead to higher aggregate productivity. Along this balanced growth path, productivity and output grow at rate $(\chi + \gamma)/(1 - \alpha)$.

3 Technology Dynamics

To assess the effect of changes in technology adoption on income dynamics first it is necessary to uncover the evolution of the extensive and the intensive margin. In this section, we describe the estimation procedure we use to measure the intensive and extensive margins of adoption for each technology-country pair. Then, we explore whether there are any significant trends in the evolution of these adoption margins.

3.1 Estimation strategy

As in Comin and Hobijn (2010), we derive our estimating equation by combining the demand for sector τ output, (8), the sectoral price deflator (17), the expression for the equilibrium wage rate (24), and the expression for Z_τ, (16). Taking logs we obtain

$$y_\tau = y + \frac{\theta}{\theta - 1} \left[z_\tau - (1 - \alpha)(y - l) \right] \tag{26}$$

where lowercase letters denote logs.

It is easy to see from expression (16) that, to a first order approximation γ only affects y_τ through the linear trend. As we show in the appendix, this allows us to approximate the log of Z_τ, to a second order around the starting adoption date, as follows:

$$z_\tau \approx \ln a_\tau + (\chi + \gamma)\tau + (\mu - 1) \ln (t - \tau - D_\tau) + \frac{\gamma}{2} (t - \tau - D_\tau). \tag{27}$$

Substituting (27) in (26) gives us the following estimating equation

$$y^c_{\tau t} = \beta^c_{\tau 1} + y^c_t + \beta_{\tau 2} t + \beta_{\tau 3} \left((\mu - 1) \ln(t - D^c_\tau - \tau) - (1 - \alpha)(y^c_t - l^c_t) \right) + \varepsilon^c_{\tau t}, \tag{28}$$

where $y^c_{\tau t}$ denotes the log of the output produced with technology τ, y^c_t is the log of output, $y^c_t - l^c_t$ is the log of output per capita, $\varepsilon^c_{\tau t}$ is an error term, and the country-technology specific intercept, β^c_1, is equal to

$$\beta^c_{\tau 1} = \beta_{\tau 3} \left(\ln a^c_\tau + \left(\chi + \frac{\gamma}{2} \right) \tau - \frac{\gamma}{2} D^c_\tau \right). \tag{29}$$

It is clear from (28) that the adoption lag is the only determinant of the curvature of the diffusion curve. In particular, longer lags imply that fewer vintages available for production and, because of the diminishing gains from variety, the steepness of the diffusion curve declines faster than if more vintages had been already adopted. It is also clear that, for a given adoption lag, the only driver of cross-country differences in the intercept $\beta^c_{\tau 1}$ is the intensive margin, a^c_τ. Intuitively, a lower level a^c_τ generates a downward shift of the diffusion curve which, ceteris paribus, leads to lower output associated with technology τ throughout its diffusion and, in particular in the long-run.

Formally, we can identify differences in the intensive margin relative to a benchmark, which we take to be the U.S., as

$$\ln a^c_\tau = \frac{\beta^c_{1,\tau} - \beta^{US}_{1,\tau}}{\beta_{3,\tau}} + \frac{\gamma}{2}(D^c_\tau - D^{US}_\tau). \tag{30}$$

When bringing the model to the data, we shall see that some of the technology measures we have in our data set correspond to the output produced with a specific technology, and therefore equation (28) is the appropriate model counterpart. Other technology measures, instead, capture the number of units of the input that embody the technology (e.g. number of computers). The model counterpart to those measures is X_τ. Towards deriving an estimating equation for these measures, we can integrate (11) across vintages, take logs and obtain

$$x^c_\tau = y^c_\tau + p^c_\tau + \ln(\alpha).$$

Substituting in for equation (28), we obtain the following expression which we use to estimate the diffusion of the inputs that embody technology.[12]

$$x^c_{\tau t} = \beta^c_{\tau 1} + y^c_t + \beta_{\tau 2} t + \beta_{\tau 3}\left((\mu - 1)\ln(t - D^c_\tau - \tau) - (1-\alpha)(y^c_t - l^c_t)\right) + \varepsilon^c_{\tau t}, \tag{31}$$

The procedure we use to estimate (28) and (31) consists in two parts. For each technology, we first estimate the equation for the U.S. (our baseline country). Then, we re-estimate the equation for each technology-country pair, imposing the technology specific estimates of β_2^{US} and β_3^{US} we have obtained for the U.S.[13,14]

[12] Note that there are two minor differences between (28) and (31). The first difference is that in the first $\beta_{\tau 3}$ is $\theta/(\theta-1)$, while in the second it is $1/(\theta-1)$. The second difference is that in the second the intercept $\beta^c_{\tau 1}$ has an extra term equal to $\beta_{\tau 3}\ln(\alpha)$.

[13] Note that the coefficients β_2 and β_3 in (28) are functions of parameters that common across countries, and therefore their estimates should be independent of the sample used to estimate them. Our procedure is less computationally intensive than estimating simultaneously the system of diffusion equations for all countries imposing the restriction that β_2 and β_3 are common across countries.

[14] Comin and Hobijn (2010) show that for a large majority of technology-country pairs, it is not possible to reject the null that β_3 is common across countries. Furthermore, the estimated adoption lags are virtually unaffected by this restriction.

Before presenting the estimates we obtain from implementing this procedure, it is important discussing one extension that result from relaxing the homotheticity in production implied by equation (7). In particular, after relaxing this assumption, we obtain the following equation

$$y_{\tau t}^{c}=\beta_{\tau 1}^{c}+\beta_{y}y+\beta_{\tau 2}t+\beta_{\tau 3}\left((\mu-1)\ln(t-D_{\tau}^{c}-\tau)-(1-\alpha)(y_{t}^{c}-l_{t}^{c})\right)+\varepsilon_{\tau t}^{c} \tag{32}$$

where the key difference is that now the income elasticity of technology is β_y rather than imposing it to be equal to 1.[15] Our two part estimation procedure allows to estimate β_y (along with β_2 and β_3) from the diffusion curve in the baseline country and then to impose these estimates when re-estimating the equation for all the technology-country pairs. Effectively what this means is that we β_y from the time series variation in technology and output for the baseline country and then assume that the slope of the Engel curve is constant across countries. Given that the baseline country has long time series that for many technologies cover much of its development experience, we consider this to be a reasonable approximation.[16]

3.2 Data and estimation results

We implement our estimation procedure using data on the diffusion of technologies from the CHAT data set (Comin and Hobijn, 2009), and data on income and population from Maddison (2004). The CHAT data set covers the diffusion of many technologies for 171 countries over the last 200 years. Because of the unbalanced nature of the data set we focus on a sub-sample of technologies that have a wider coverage over rich and poor countries and for which the data captures the initial phases of diffusion. The 25 technologies that meet these criteria are listed in the Appendix and cover a wide range of sectors in the economy. Their invention dates also span quite evenly over the last 200 years. It is worthwhile remarking that the specific measures of technology diffusion in CHAT match the dependent variables in specification (28) or in the equivalent specification in the extended model developed below. In particular these measures capture either the amount of output produced with the technology (e.g., tons of steel produced with electric arc furnaces) or the number of units of capital that embody the technology (e.g. number of computers per capita).

The value added of this paper is not the estimation of the two adoption margins, as we have done this elsewhere (Comin and Hobijn, 2010; Comin and Mestieri, 2012)–albeit with fewer technologies. Accordingly, we describe the fit and summary statistics of the margins briefly. As in Comin and Hobijn (2010), we use the plausibility and precision of the estimates of the adoption lags from equation (28) as a pre-requisite to utilize the technology-country

[15] An elasticity of 1 is required for a balanced growth path to exist.
[16] See Comin and Mestieri (2011).

pair in our analysis.[17] We find that these two conditions are met for the majority of the technology country-pairs (62%). For these technology country-pairs, we find that equation (28) provides a very good fit for the data with a detrended R^2 of 0.80 across countries and technologies (Table 5).[18]

Tables 1 and 2 report summary statistics for the estimates of the adoption lags and the intensive margin for each technology. The average adoption lag across all technologies (and countries) is 50 years. We find significant variation in average adoption lags across technologies. The range goes from 8 years for the Internet to 128 years for spinning spindles. There is also considerable cross-country variation in adoption lags for any given technology. The range for the cross-country standard deviations goes from 2 years for the Internet to 51 years for steam and motor ships.

We also find significant cross-country variation in the intensive margin. On average, differences in the log intensive margin (relative to the U.S.) are on the order of -1, which implies 40% level of adoption relative to the U.S.. More generally, there is significant cross-country dispersion in the intensive margin. The range goes from 0.4 for mail to 1.79 to cellphones. These summary statistics for the estimates of adoption lags and the intensive margin of adoption are very consistent with those in Comin and Hobijn (2010) and Comin and Mestieri (2012) which use smaller technology samples and estimate other versions of the diffusion equation (28).

3.3 Evolution of adoption lags and intensive margin

One key goal of our analysis consists in studying the *evolution* of the cross-country dispersion of the adoption lags and the intensive margins. To do that, we divide the countries in our sample in two groups. We follow Maddison, and define a group of 12 Western, frontier, countries[19] and lump the rest on the category "Rest of the World."For brevity we may refer to the first group as "rich" and the second as "poor." Then, we estimate the trends in the adoption margins for both samples. These allows us to see whether there has been convergence in adoption patterns between rich and poor countries.

Figure 1 plots, for each technology in our sample, the median adoption lag among the western countries and among the rest of the countries in the world. The Figure suggests that cross-country differences in adoption lags have narrowed. Table 3 formalizes this intuition by regressing (log) adoption lags on their year of invention (and a constant). Column (1) reports

[17] Plausible adoption lags are those with an estimated adoption date of no less than ten years before the invention date that is significant at 5% level. Precise are those with an standard error smaller than $\sqrt{2003 - \text{invention date}}$. This allows for older technologies to be more imprecisely estimated.

[18] To compute the detrended R^2, we partial out the linear trend $\gamma\tau$ and compute the R^2 of the de-trended data.

[19] These are the following: Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, Untied Kingdom, Australia, New Zealand, Canada and the United States of America.

Table 1: Estimated Adoption Lags

	Invention Year	Obs.	Mean	SD	P10	P50	P90	IQR
Spindles	1779	28	111	55	38	102	171	105
Ships	1788	41	122	54	47	144	179	104
Railways Freight	1825	39	77	33	31	75	123	56
Railways Passengers	1825	29	74	41	15	80	123	70
Telegraph	1835	38	50	31	18	42	96	44
Mail	1840	43	44	38	8	33	108	59
Steel (Bessemer, Open Hearth)	1855	39	67	33	14	78	108	45
Telephone	1876	49	51	31	8	54	91	51
Electricity	1882	74	51	22	18	56	72	33
Cars	1885	61	41	21	15	37	65	33
Trucks	1885	54	38	21	14	35	64	32
Tractor	1892	114	66	11	57	68	69	5
Aviation Freight	1903	36	43	14	28	44	63	19
Aviation Passengers	1903	40	30	15	16	25	53	18
Electric Arc Furnace	1907	39	49	19	22	56	78	35
Fertilizer	1910	85	47	9	39	48	54	7
Harvester	1912	56	40	15	20	42	52	16
Synthetic Fiber	1924	48	38	4	34	39	41	2
Blast Oxygen Furnace	1950	37	15	8	8	13	26	11
Kidney Transplant	1954	24	13	7	3	13	25	5
Liver Transplant	1963	20	19	4	15	18	25	3
Heart Surgery	1968	17	13	4	9	13	20	4
Cellphones	1973	79	13	4	9	14	18	5
PCs	1973	66	17	3	13	16	20	3
Internet	1983	57	7	3	2	7	11	3
Total		1213	45	35	10	39	85	47

the unconditional trend, that is, for the whole sample of countries. We confirm the finding in Comin and Hobijn (2010) that it is downward sloping. That is, newer technologies have diffused faster. Then, we run the same regression separately for the two groups of countries. We find that the rate of decline in adoption lags is almost a 50% higher in poor than in rich countries. In particular the annual rate of decline is around .9% for rich countries (see column (3)) versus a 1.3% for poor countries (column (2)). Hence, there has been *convergence* in adoption lags between rich and poor countries.

Figure 2 shifts the attention to the cross-country evolution of the penetration rates. In particular, it plots for each technology the median intensive margin (relative to the U.S.) among the western and non-western countries. The Figure suggests that the gap between rich countries and the rest of the world in the intensive margin of adoption was smaller

Table 2: Estimated Intensive Margin

	Invention Year	Obs.	Mean	SD	P10	P50	P90	IQR
Spindles	1779	28	0.0	0.7	-1.0	0.0	1.1	0.8
Ships	1788	41	-0.1	0.7	-0.8	-0.1	0.8	0.8
Railways Freight	1825	39	-0.1	0.4	-0.5	-0.1	0.5	0.5
Railways Passengers	1825	29	0.0	0.3	-0.5	-0.1	0.4	0.3
Telegraph	1835	38	-0.2	0.5	-1.0	-0.2	0.4	0.5
Mail	1840	43	-0.1	0.3	-0.7	-0.1	0.2	0.4
Steel (Bessemer, Open Hearth)	1855	39	-0.2	0.5	-0.7	-0.1	0.3	0.6
Telephone	1876	49	-0.8	0.9	-1.9	-0.6	0.2	0.9
Electricity	1882	74	-0.5	0.6	-1.2	-0.4	0.1	0.8
Cars	1885	61	-1.1	1.1	-2.2	-1.0	0.1	1.6
Trucks	1885	54	-0.8	1.0	-1.7	-0.8	0.2	1.2
Tractor	1892	114	-0.9	0.8	-2.0	-0.9	0.1	1.3
Aviation Freight	1903	36	-0.3	0.6	-1.3	-0.1	0.3	0.6
Aviation Passengers	1903	40	-0.4	0.7	-1.3	-0.3	0.3	0.7
Electric Arc Furnace	1907	39	-0.2	0.5	-1.1	-0.1	0.4	0.7
Fertilizer	1910	85	-0.9	0.8	-2.0	-0.8	0.1	1.3
Harvester	1912	56	-1.2	1.1	-3.0	-1.1	0.1	1.7
Synthetic Fiber	1924	48	-0.6	0.8	-1.8	-0.5	0.3	1.0
Blast Oxygen Furnace	1950	37	-0.8	1.0	-2.3	-0.4	0.1	1.3
Kidney Transplant	1954	24	-0.2	0.4	-0.9	-0.1	0.1	0.4
Liver Transplant	1963	20	-0.4	0.7	-1.7	-0.1	0.1	0.5
Heart Surgery	1968	17	-0.4	0.8	-1.8	-0.1	0.2	0.4
Cellphones	1973	79	-0.8	0.7	-1.9	-0.6	0.1	1.2
PCs	1973	66	-0.6	0.6	-1.4	-0.6	0.1	0.9
Internet	1983	57	-1.0	1.2	-2.2	-0.9	0.1	1.6
Total		1213	-0.6	0.8	-1.7	-0.4	0.2	1.0

for technologies invented at the beginning of the nineteenth century than for technologies invented at the end of the twentieth century. Table **??** establishes this finding. We first look at all countries of our sample. Column (1) shows that, on average, differences in the intensive margin relative to the U.S. have widened over time. Then, we look at the evolution of the intensive margin separately for rich and poor countries. Column (2) shows that, for rich countries, the intensive margin has barely diverged from the U.S. declining at an annual rate of .15%. This is in stark contrast to what we find in column (3) for poor countries, for which the intensive margin (relative to the U.S.) has declined at a much faster speed, a 1.26% annual rate. Hence, there has been a *divergence* in the intensive margin of adoption between rich and poor countries.

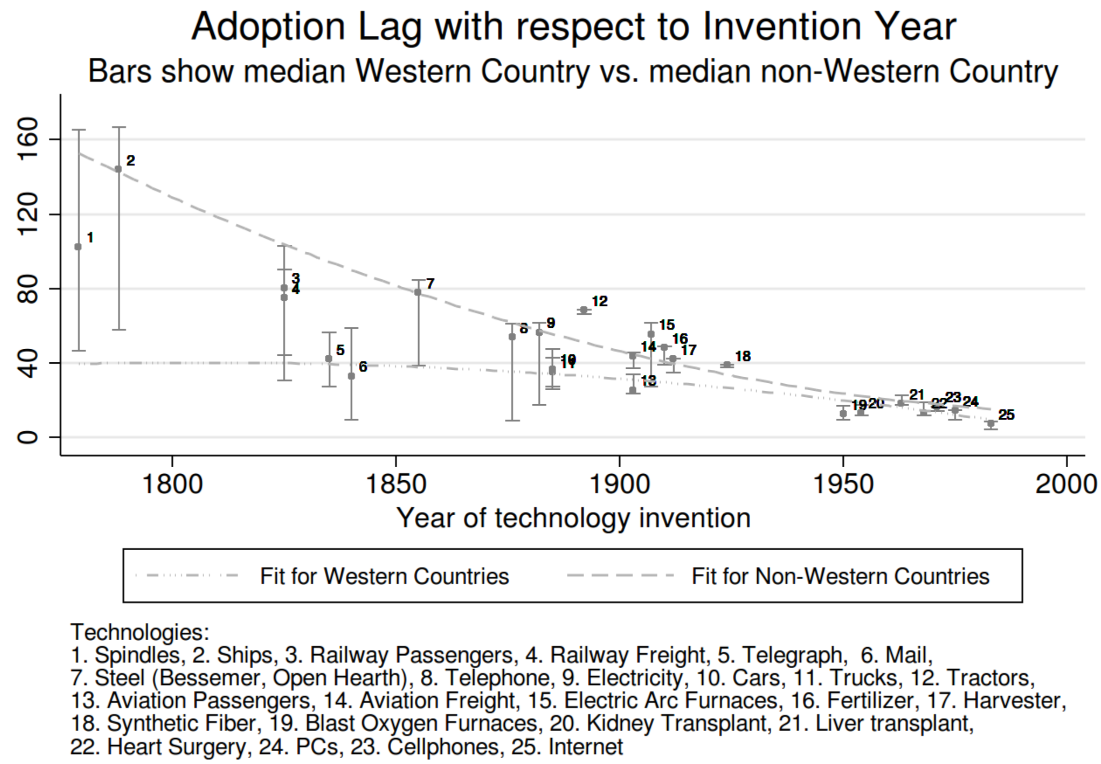


Figure 1: Convergence of Adoption Lags.

4 Income dynamics: Analytic Results

The final goal of this paper is to explore how the technology dynamics we have uncovered affect the evolution of productivity growth across countries. Given the novelty of the model, we first provide some analytic intuitions about the growth dynamics in the model. Then, in the next section, we evaluate quantitatively its ability to generate the observed cross-country income growth dynamics over the last 200 years with the help of simulations. In this section, we analyze the special case $\alpha = 0$. Recall from equation (18) that $Y = A^{1/(1-\alpha)}(\alpha)^{\alpha/(1-\alpha)}$. This simplifying assumption means final output is produced one-for-one with technology, as opposed of having some curvature ($\alpha > 0$). Indeed, the results presented in this section extend to the case $\alpha > 0$, but the same insights are obtained with this simpler model.

In our model, the relevant transitional dynamics are driven by three variables: adoption lags, intensive margins and the growth of the technology frontier. Before studying the transitional dynamics, it is helpful to discuss the sources of growth in our model when these three variables are held constant.

As described in Section 2, at each instant of time, it appears the first vintage of a new technology and a new vintage for all past technologies. Thus, the set of technologies available

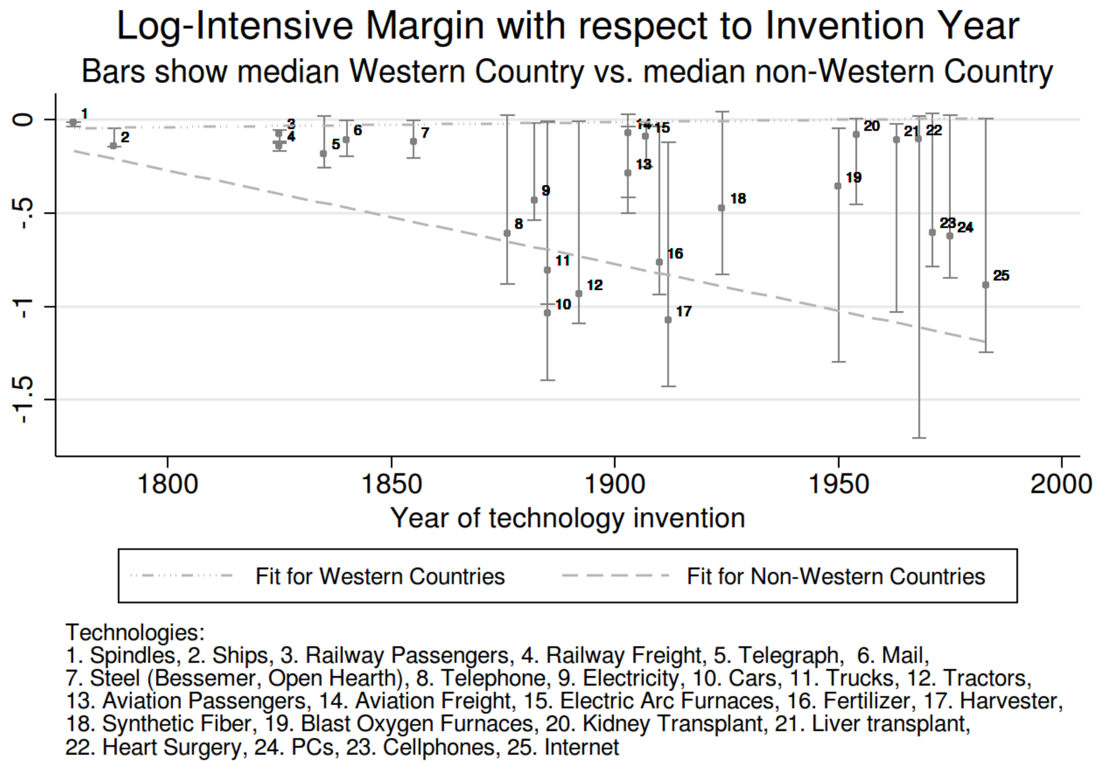


Figure 2: Divergence of the Intensive Margin

to an economy at time t is given by $[-\infty, t - D_t)$, and that the set of vintages of a given technology is $[\tau, t - D_\tau)$ where τ is time of invention of the technology and D_τ the adoption lag. Taking the time derivative of (19) and using (**??**), denoting a time derivative by a dot and growth rates by g, we find that

$$g_Y = \underbrace{(\theta - 1)\left(\frac{Z_{t-D_t}}{Y}\right)^{\frac{1}{\theta-1}}(1 - \dot{D}_t)}_{\text{New Technology}} + \underbrace{\int_{-\infty}^{t-D_t}\left(\frac{Z_\tau}{Y}\right)^{\frac{1}{\theta-1}} g_{Z_\tau} d\tau}_{\text{Old Technologies}}, \tag{33}$$

where

$$g_{Z_\tau} = \gamma\left(1 + \frac{e^{\frac{-\gamma}{\mu-1}(t-\tau-D_\tau)}}{1 - e^{\frac{-\gamma}{\mu-1}(t-\tau-D_\tau)}}\right). \tag{34}$$

The first term in (33) captures the growth imputable to a new technology being introduced in the economy. This term has three parts. $(1 - \dot{D}_t)$ captures the number of new technologies introduced at instant t. If the adoption lag D_t does not change (i.e., $\dot{D}_t = 0$), only one new technology arrives in the economy at instant t. But if adoption lags decline (i.e., $\dot{D}_t < 0$), the flow of new technologies in the economy is greater than one. The effect on growth of the arrival

Table 3: Evolution of the Adoption Lag

Dependent Variable is:	(1) Log(Lag) World	(2) Log(Lag) Western Countries	(3) Log(Lag) Rest of the World
Year-1820	-0.011***	-0.0076***	-0.0113***
	(0.004)	(0.000441)	(0.0004)
Constant	4.37***	3.64***	4.55***
	(0.07)	(0.07)	(0.06)
Observations	1198	330	868
R-squared	0.44	0.33	0.59

Note: robust standard errors in parentheses,*** p<0.01. Each observation is re-weighted so that each technology carries equal weight.

of new technologies depends on two factors. First, the (inverse) of elasticity of substitution between technologies ($\theta - 1$). The more substitutable are different technologies, the smaller the gains from having a new technology available for production. Second, the share of the new technologies in output (i.e. $(Z_{t-D_t}/Y)^{1/(\theta-1)}$).[20] The higher the productivity embodied in a technology, the larger the impact of its arrival on GDP growth. Note from (??) that the share of a new technology in GDP depends both on its intensive margin and its vintage $(t - D_t)$.

The second term in (33) captures the increases of productivity due to the introduction of new vintages in already present technologies. The contribution to overall growth is an average of different sectoral growths g_{Z_τ} weighted by the sector's share in total output. Note from (34) that the productivity of new technologies grows faster than for older ones because of the larger gains from variety when few vintages have been adopted (i.e., for small $t - \tau - D_\tau$). Eventually, g_{Z_τ} converges to γ, the long-run growth rate of productivity embodied in new vintages.

4.1 Transitional dynamics after an acceleration in frontier growth

To make the mechanics of the model more transparent, we introduce the dynamics generated by each mechanism sequentially. The first parameter change we consider is a permanent, instantaneous increase in the growth rate of the technology frontier, from g_{Old} to $\gamma + \chi$, that takes place at time T. In our view, the acceleration in the growth rate of the technology frontier is a key property of the Industrial Revolution. Thus, we study how an economy

[20]Recall from (19) that $Y_t = \left(\int_{-\infty}^{t-D_t} Z_\tau^{\frac{1}{\theta-1}} d\tau \right)^{\theta-1}$

Table 4: Evolution of the Intensive Margin

Dependent Variable is:	(1) Intensive World	(2) Intensive Western Countries	(3) Intensive Rest of the World
Year-1820	-0.0034*** (0.0005)	0.0000*** (0.000229)	-0.0062*** (0.0006)
Constant	-0.24*** (0.06)	-0.00*** (0.0598)	-0.30*** (0.08)
Observations	1213	341	872
R-squared	0.06	0	0.15

Note: robust standard errors in parentheses,*** p<0.01. Each observation is re-weighted so that each technology carries equal weight.

transitions from an original balanced growth path with growth g_{Old} coming from the usage of pre-Modern technologies to a new balanced growth path with growth $\chi + \gamma$. We keep the intensive and extensive margins constant at their pre-Industrial levels in this initial exercise.

To explore the dynamics after an acceleration in frontier growth, it is convenient to decompose output as follows

$$Y(t) = \left(\int_{-\infty}^{T} Z_\tau^{\frac{1}{\theta-1}} + \int_{T}^{t-D} Z_\tau^{\frac{1}{\theta-1}} \right)^{\theta-1} \equiv \left(X_{\text{Old}}^{\frac{1}{\theta-1}} + X_{\text{Modern}}^{\frac{1}{\theta-1}} \right)^{\theta-1}, \tag{35}$$

$$g_Y = (1-s)\, g_{\text{Old}} + s\, g_{\text{Modern}}, \tag{36}$$

where X_{Old} denotes the output produced with "old", pre-Industrial Revolution, technologies, X_{Modern} the output produced with Modern technologies. T denotes the advent of the Industrial Revolution, s is the output share of modern technologies $\left(\frac{X_{\text{Modern}}}{Y}\right)^{\frac{1}{\theta-1}}$, and g_i denotes the growth rate of i.

It is clear from (36) that the dynamics can come from the evolution of the sectoral growth rates, g_{Old} and g_{Modern}, or from changes in the output share of the modern sector, s. The next proposition characterizes the evolution of output produced with modern and pre-industrial technologies.

Proposition 1 *After the economy starts adopting Modern technologies, Modern and pre-*

Industrial output are

$$X_{Old}(t) = aA_{Old}e^{g_{Old}(t-D)}, \tag{37}$$

$$X_{Modern}(t) = aA_{Modern}e^{(\chi+\gamma)(t-D)}h(t)^{\theta-1}, \tag{38}$$

where a is the intensive margin, A_{Old}, A_{Modern} are positive constants and $h(t)$ is an S-shaped function. It is increasing, convex for $t < \frac{\theta-1}{\gamma}\ln\left(\frac{\chi+\gamma}{\chi}\right) + T + D$ and concave thereafter, its initial value is 0 and it reaches a plateau, $\lim_{t\to\infty} h(t) = 1$. Finally, it approaches smoothly to its minimum and maximum values, $h'(T+D) = \lim_{t\to\infty} h'(t) = \lim_{t\to\infty} h''(t) = 0$.[21]

Note that the output produced using Old technologies grows at rate g_{Old}. Modern output, instead, has two components that change over time.[22] First, there is a log-linear trend, $(\chi+\gamma)t$, and second, a transient source of growth, $h(t)$. The log-linear trend captures the higher productivity embodied in new technologies and vintages (embodiment effect). This term drives long-run growth. The transient term $h(t)$ is S-shaped and eventually reaches a ceiling, so it does not contribute to long-run output growth. This term originates from the gains from variety of having more vintages and more technologies in production. In an initial phase, the increment in productivity from the arrival of vintages is larger the more modern technologies have arrived in the economy. Hence, the initial convexity of $h(t)$. At some point, though, the decreasing gains from variety to the number of modern technologies and to the number of vintages within existing modern technologies kick in and $h(t)$ becomes concave and eventually plateaus.

Next we describe the shape of the transition to the new balanced growth path.

Proposition 2 *The transition of the growth rate from the pre-Industrial balanced growth path to Modern growth is S-shaped.*

From (36), we know that the growth rate in the economy is a weighted average of the growth of the modern and old sectors. The weights correspond to the output share of Modern and pre-Modern technologies. We show in the Appendix that the share of the Modern sector inherits the properties of the transient component $h(t)$, so that the weight on modern output is increasing and has an S-shape.

[21] The expression for $h(t)$ is

$$h(t) = \frac{\chi(\chi+\gamma)}{\gamma}\left(\frac{1}{\chi}\left(1 - e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right), \tag{39}$$

where $\Delta t \equiv t - D - T$.

[22] Here we are assuming that output produced with pre-Modern technologies keeps increasing independently from the advent of the Industrial revolution. In Section B of the Appendix, we show how to embed this in the framework of Section 2. The differences that we obtain are qualitatively minor, and quantitatively insignificant for the relevant parameter range.

If growth in the modern sector was only given by the embodiment effect (the log-linear trend, $\chi + \gamma$), modern output would grow at a constant rate. In this case, output growth would be given by $g_Y = (1-s)g_{\text{Old}} + s(\chi+\gamma)$. It follows from this expression that aggregate output growth would be increasing over time reaching $(\chi+\gamma)$ asymptotically. Furthermore, g_Y would mimic the S-shape of the modern sector output share, s.

However, modern output grows faster than the log-linear trend. Thus, aggregate output growth will overshoot its long-run level $(\chi+\gamma)$, as illustrated in Figure 4. Whether this over-shooting is quantitatively important depends on whether when the weight on modern growth becomes close to one, the growth rate of the modern sector is substantially higher than $(\chi+\gamma)$.

Next, we assess the protractedness of the transition to the new balance growth path.

Proposition 3 *The half-life in terms of levels and growth rates are approximately*

$$t_{1/2}^{level} \simeq D + \frac{1}{\chi+\gamma-g_{Old}} \ln\left(\frac{1}{2}\frac{A_{Old}}{A_{Modern}}\right), \tag{40}$$

$$t_{1/2}^{growth} \simeq D + \frac{1}{\chi+\gamma-g_{Old}} \ln\left(\frac{A_{Old}}{A_{Modern}}\right). \tag{41}$$

The first terms in both equations capture the fact that there is a lag between the advent of the Industrial Revolution and when a country starts adopting Modern technologies. The second terms capture the evolution of the transition conditional on having started to adopt modern technologies. In particular, the term inside the brackets reflects the ratio of the productivity of pre-modern output at the time of the Industrial Revolution to the Modern sector (and, hence, long-term level of output). Intuitively, if the output produced with pre-Modern technologies is "high", it takes longer for modern output to become the major driver of output per capita. This slows down the transition to the new balanced growth path.

4.2 Changes in the Adoption Lags and Intensive Margin

Next we study how changes in adoption lags and intensive margin affect the transitional dynamics. Perhaps surprisingly, we show that the qualitative results we derived in the previous section remain.

We start by considering a one period permanent change of adoption lags and the intensive margin from its pre-Modern levels to their average Modern levels. Formally,

$$D_\tau = \begin{cases} D_{\text{Old}} & \text{for } \tau < T \\ D_{\text{Modern}} & \text{for } \tau \geq T \end{cases} \qquad a_\tau = \begin{cases} a_{\text{Old}} & \text{for } \tau < T \\ a_{\text{Modern}} & \text{for } \tau \geq T \end{cases} \tag{42}$$

where T denotes the first Modern technology.

Proposition 4 *Let the evolution of the adoption lag and the intensive margin be given by (42), then pre-Industrial and Modern output are*

$$X_{Old}(t) = a_{Old}A_{Old}e^{g_{Old}(t-D_{Old})}, \tag{43}$$
$$X_{Modern} = a_{Modern}A_{Modern}e^{(\chi+\gamma)(t-D_{Modern})}h(t)^{\theta-1}, \tag{44}$$

where A_{Old}, A_{Modern} are the same positive constants as in Proposition 3 and $h(t)$ is the same S-shaped function as in Proposition 3. $h(t)$ is increasing, initially convex and concave thereafter, reaching a plateau.[23] *Proposition 2 hold in this case. The half-lives of the system in levels and growth rates are*

$$t_{1/2}^{level} \simeq D_{Modern} - \frac{g_{Old}D_{Old}}{\chi+\gamma-g_{Old}} + \frac{1}{\chi+\gamma-g_{Old}}\ln\left(\frac{1}{2}\frac{a_{Old}A_{Old}}{a_{Modern}A_{Modern}}\right), \tag{45}$$
$$t_{1/2}^{growth} \simeq D_{Modern} - \frac{g_{Old}D_{Old}}{\chi+\gamma-g_{Old}} + \frac{1}{\chi+\gamma-g_{Old}}\ln\left(\frac{a_{Old}A_{Old}}{a_{Modern}A_{Modern}}\right). \tag{46}$$

The changes in the adoption margins do not affect the pre-Modern sector. Growth in the Modern sector does not depend on whether the adoption margins are the same before and after the industrial revolution. Hence, the shape of the transition to the Modern growth era is not affected by the changes introduced in (42). However, the changes in the adoption margins have a quantitative impact on the transitional dynamics.

Because of the quantitative effect of changes in the adoption margins, we have delay using our propositions to estimate the protractedness of the transitional dynamics of the model. Now we are in a position to make such calculations. To this end, we calibrate D and a using information on the averages of both margins over the Modern period (50 years and 40% of the US intensive margin, respectively).[24]Using Propositions 4, we estimate the speed of convergence to the new balanced growth path.As discussed above, the first component of the half lives is the adoption lag, D, which is 50 years. The second components reflects the dynamics once adoption has started. Our calibrations imply a value for this second term of 60 years for the half life in levels and 100 years for the half life of the growth rate. Hence, the resulting halflives are on the order of a hundred years.[25]

As shown in section 3, the evolution of the intensive margin and adoption lags has been

[23]More precisely, $h(t)$ is increasing, convex for $t < \frac{\theta-1}{\gamma}\ln\left(\frac{\chi+\gamma}{\chi}\right)$ and concave thereafter, $h(T+D_{\text{Modern}}) = 0$, $\lim_{t\to\infty} h(t) = \gamma/\chi(\chi+\gamma)$, $h'(T+D_{\text{Modern}}) = \lim_{t\to\infty} h'(t) = \lim_{t\to\infty} h''(t) = 0$.

[24]The rest of the parameters we take from our baseline calibration, which is explained in Section 5. These parameters are $\chi = \gamma = 1\%$, $\theta = 1.45$ and initial output (normalized by a productivity term) of 75.

[25]To derive these simple analytic expressions we have neglected the effect of the evolution of the transient component $h(t)$, so the reader may wonder how much does it matter. In our simulations we find that the effect of the transient component is important and it almost halves the contribution of the second term. However, even so, the contribution of the second term remains important (30 and 50 years). The half-lives that we find for the average country in our simulation are around 80 and 100 years for the levels and the growth, respectively –which makes the dynamics still very protracted.

smoother than in (42). A more realistic characterization of the evolutions of the extensive and intensive margins is given by:[26]

$$D_\tau = \begin{cases} d_o & \text{for } \tau < T, \\ d_o - d_1\tau & \text{for } \tau \in [T, \bar{T}], \\ d_m & \text{for } \tau > \bar{T}, \end{cases} \qquad \ln a_\tau = \begin{cases} a_o & \text{for } \tau < T, \\ a_o - a_1\tau & \text{for } \tau \in [T, \bar{T}], \\ a_m & \text{for } \tau > \bar{T}, \end{cases} \tag{47}$$

where $d_1 = \frac{d_o - d_m}{\bar{T} - T}$ and $a_1 = \frac{a_o - a_m}{\bar{T} - T}$ are the trends in the adoption lags and intensive margin, respectively. We conclude our analytic exploration of the transitional dynamics of the model by characterizing the evolution of output after adoption margins change as in (47).

Proposition 5 *Pre-Industrial output is described by $X_{Old}(t) = A_{Old}e^{g_{Old}t}$. Modern Output is a continuous, increasing function,*

$$X_{Modern}(t) = \begin{cases} A_0 \; e^{(\chi+\gamma+g_a)t} \; h_0(t)^{\theta-1} & \text{for } t \in [T + d_0, d_0 + \bar{T}/(1+d_1)], \\ A_1 \; e^{(\chi+\gamma)t} \; h_1(t)^{\theta-1} & \text{for } t > d_0 + \bar{T}/(1+d_1), \end{cases} \tag{48}$$

where $g_a = d_1(\chi + (1+d_1)\gamma) - (1+d_1)a_1$, $h_0(t)$ is S-shaped in the sense that it is continuous, increasing, convex for any $t < t^c$ and concave thereafter, reaching a ceiling value as time approaches infinity. $h_1(t)$ is a continuous function defined as the CES aggregator (with elasticity $\frac{1}{2-\theta}$) of $e^{-\chi t}h_0(\bar{T})$ and $h(t - \bar{T})$. In the case that $\chi, \gamma \gg a_1, d_1$ and $d_0 < T$ it is S-shaped (increasing, initially convex and eventually concave reaching a ceiling). The transition from the old growth rate to modern growth has two S-shaped transitions.[27,28]

The most noticeable property of the evolution of modern output is that, the evolution of adoption margins affects trend growth in Modern sector output during the transition. Specifically, the decline in the adoption lags accelerates the embodiment effect at the rate g_a

[26] The specification we have estimated in section 3 differs slightly from (42) in that in section 3 we fit a linear trend to the log adoption lag while in (42) the trend is fit to the level. Both approaches seem sensible to us and, quantitatively, there are no significant differences between them.

[27] The expression for $h_0(t)$ is very similar to $h(t)$,

$$h_0(t) = \frac{1}{\chi + \gamma d_1 - a_1}\left[1 - e^{-\frac{\chi+\gamma d_1 - a_1}{\theta-1}(1+d_1)(t-d_0)}\right] - \frac{e^{-\frac{d_1^2\gamma}{\theta-1}(t-d_0)}}{\chi+\gamma-a_1}\left[1 - e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]. \tag{49}$$

See the Appendix for the expression of h_1. The reason for having two S-shaped transitions is that we effectively have two regimes and the transition is S-shaped for both. Hence, it can be the case that if g_a is not very close to zero (which is what happens in our calibration for the non-Western country), we observe a transition to balanced growth $\chi + \gamma$ in two steps. First, while we are in the regime $\tau \in [T, \bar{T}]$ the growth rate converges to $\chi + \gamma + g_a$ (in an S-shaped way), and once we enter the regime $\tau > \bar{T}$, the economy grows from $\chi + \gamma + g_0$ to g_m an that transition looks again as an S-shape. In the case that $\chi + \gamma + g_0 > \chi + \gamma$, we would observe an inverse S-shape.

[28] The Appendix characterizes the half-lives of the model with trends in the adoption margins. (See Proposition 7).

because more technologies and vintages are brought into production. This raises trend growth by $d_1(\chi + (1 + d_1)\gamma)$. Similarly, an acceleration in the intensive margin of new technologies increases the productivity embodied in new technologies increasing trend growth by $-(1 + d_1)a_1$.[29]

Proposition 5 points to the sources of cross-country differences in growth patterns. In particular, it highlights, at least, three relevant dimensions. Differences in the initial adoption lag, d_0, generate differences in the growth acceleration brought by the arrival of modern production technologies. Differences in the trends in adoption lags, d_1, and in the intensive margin, a_1, affect the magnitude of the growth acceleration,g_a, along the transition. As shown in next section, these three factors are important to understand the cross-country patterns in growth over the last two centuries.

Finally, one further implication of Proposition 5 is that the growth effects of a gradual reduction in adoption lags depend separately on χ and γ beyond its sum. In other words, productivity gains embodied in new technologies and in new vintages are not isomorphic. This is the case because productivity gains embodied in new vintages, γ, lead to higher productivity for both new and already adopted (modern) technologies, while increases in the productivity embodied in new technologies only affects output growth through the productivity of newly adopted technologies. This observation motivates the robustness checks we perform in next section to the calibration of χ and γ.

5 Income Dynamics: Simulation Results

We use the expository model outlined in section 2 to evaluate quantitatively the effects of dynamics in technology diffusion on the cross-country evolution of economic growth. In particular, we will focus on the two groups of countries defined by Maddison (2004) as "Western" countries and the rest of the world .[30]

To simulate the model we need to calibrate a few parameters. First, we need to specify the path for the world technology frontier. Prior to year $T = 1765$ (year in which James Watt developed his steam engine),[31] the technology frontier grows at 0.2% which is the growth rate of western Europe according to Maddison (2004) from 1500 to 1800. After 1765, the frontier grows at 2% per year. As shown in equation (25), the growth rate along the balanced growth is equal to $(\gamma + \chi)/(1 - \alpha)$. A priori, it is difficult to divide up the growth of the

[29] Where a_1 is teh rate of decline of the intensive margin.

[30] Western countries consists of 12 (northern) European countries, Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, United Kingdom, the Western off-shots, Australia, New Zealand, Canada, United States, and Japan. In Section **??** also explore the implications of the model for the group of countries with income in the bottom third of the world distribution according to Maddison (2004).

[31] Alternatively, we can set, without any significant change to our findings, the beginning of the industrial revolution at 1779, year of invention of the first technology in our sample, the mule spindle.

frontier between these two parameters. Therefore, in our baseline simulation, we split evenly the sources of growth in the frontier between γ and χ and conduct robustness checks to show the robustness of our findings.

As shown in Proposition 3, aggregate productivity at time T, A_T, affects the speed of the acceleration induced by the industrial revolution. We set A_T to match the relative income in the U.S. in 1820 and in 2000. It is important to note that A_T is the same across countries.

Finally, we need to calibrate the elasticities of substitution between vintages or between technologies, which we assume are the same and equal to $1/(\theta - 1)$. We back out the value of θ from the estimates of $\beta_{\tau 3}$. The implied values for θ are around 1.3, which are similar to the values implied by the estimates of price markups from Basu and Fernald (1997) and Norbin (1993).

Initial income differences .– Before quantifying the transitional dynamics in labor productivity, it is worthwhile noting that our model has precise implications for productivity gap between rich and poor countries at the time of the Industrial Revolution. In particular, expressions (18) and (25) indicate how initial differences in adoption lags and in the intensive margin affect the relative income of these countries in the pre-industrial revolution balanced growth path. To this end, we make the reasonable assumption that the pre-industrial lags and intensive margins are similar to those we estimate for the first technologies in our sample. . Our estimates from Tables 3 and 4 imply that the difference between the average adoption lag in the sample of Western countries and in other countries is 56 years in 1820. Further, the average gap in the (log) intensive margin is 0.3. Assuming a pre-industrial growth rate of the world technology frontier of .2% and using these estimates to proxy for the pre-industrial productivity gap between the sample of rich countries and the rest, we can use equation (25) to obtain that the income differences predicted by the model are of 51% $exp(.2\% \cdot 56 + .3) = 1.51$. In Maddison's data, the productivity gap between these countries in 1820 is 80%. Hence, our model does generate sizeable pre-industrial income differences that account for more than 60% of those observed in the data.

Protracted dynamics .– Before exploring the model predictions for the cross-country dynamics of income, it is worthwhile evaluating the protactedness of the model transitional dynamics. To this end, we consider the average country in our sample, and suppose that there is an increase in the growth of the world technology frontier $(\gamma + \chi)$ like the one we observed in the industrial revolution (from 0.2% to 2%). The average country is parametrized so that its adoption lag and its degree of penetration (a_τ) are constant and equal to the average adoption lag and intensive margins across countries and over our sample of technologies. In particular, the resulting D is 50 years and the intensive margin is 40% of the U.S. level. Figure 5 plots the transition of the output in this representative economy to the new balanced growth path (normalized by the long term trend). In the figure, we can see that the model generates a very slow convergence to the new balanced growth path. The half-life of the output gap relative to the Modern balanced growth path is 77 years while for output growth it is 104 years. These half-lives are almost an order of magnitude higher than the typical half-life in neoclassical growth models (e.g., Barro and Sala-i-Martin 2003).

There are three reasons why our model generates such protracted dynamics. First, the long adoption lags (50 years) imply that it takes this amount of time for the new technologies (which embody the higher productivity gains) to arrive in the economy. Until then, there is no effect whatsoever in output growth. Second, for a given growth in the modern sector output, its impact in GDP depends on the share of the modern sector. Since the modern sector's share increases slowly, so does aggregate output. Third, as shown in Proposition 1, the growth of the modern sector increases progressively since it is initially convex.

Cross-country evolution of income growth .– To evaluate the model's power to account for the Great Divergence, we simulate the evolution of output for Western countries and the rest of the world after feeding in a (common) one time permanent increase in the frontier growth, as well as the estimated evolutions for adoption lags and the intensive margin for each group of countries (see Tables 3 and **??**). The results from this exercise are reported in Figure 6 and Table 6.

The model generates sustained differences in the growth rates of Western and non-Western countries for long periods of time. Output growth starts to accelerate at the beginning of the nineteenth century in the Western economy reaching its peak around 1900 at a rate slightly above 2%. At this point, it slowly converges to the steady state growth of 2%. For the non-Western country, instead, growth does not increase from the pre-industrial rate until the end of the nineteenth century. Growth in the poor country slowly accelerates, but it is still around 1.5%, by year 2000. The gap in growth between the rich and poor countries is considerable. Annual growth rates differ by more than 0.6% for 80 years. The peak gap is reached at around 1900 at 1.3%. From then, the gap declines monotonically until reaching 0.5% by 2000. Table 6 reports the average growth and growth gaps of our simulation and Maddison (2004). The patterns and levels in our data trace quite well Madison's.

The sustained cross-country gap in growth produced by the model leads to a substantial gap in income per capita. In particular, our model generates a 2.8 income gap between the Western countries and the rest of the world. Maddison (2004) reports an actual income widening by a factor of 3.8 between Western countries and the rest of the world since the Industrial Revolution.[32] Hence, most of the variation in the income gap between Western and

[32]We have computed the increase in relative output per capita from Maddison (2004). We restrict our attention to countries for which we have data on technology adoption. In our sample, the median non-Western countries had approximately 68% of the income per capita of the median Western country in 1700. In 2000, this figure was 14%. If instead we looked at the ratio of the 90th to the 10th percentile in our data, we would find that in 1700, it was 2.6 and in 2000, it was 17.9. (so the relative increase is 6.9. Lucas (2004) numbers for English speaking countries, Japan and northern Europe relative to the rest of the world are in the same ballpark. The Western countries increase their income from 1750 to 2000 by 20, while the rest of the word multiplied it by around 4, which again gives a increase of the relative gap of 5.

non-Western countries in the last two centuries is accounted for.

The role of initial conditions and changes in adoption lags .– After showing that the model does a remarkable job in reproducing the cross-country dynamics of income growth over the last two centuries, it is worthwhile dissecting the mechanisms at work. We start this task by simulating the dynamics of our model after a common acceleration of the technology frontier for both countries. To this end, we keep constant at their initial level the adoption lags and intensive margin in each country. Figure 7 shows that these initial conditions are an important source of cross-country income divergence. In particular, longer adoption lags in the non-Western country imply a delay of 80 years to start benefiting from the productivity gains of the Industrial Revolution. As a result, the income gap increases by a factor of 2.7 by year 2000.

Of course, this estimate does not provide an accurate assessment of the contribution of adoption lags to the great divergence because adoption lags did not remain constant over the last 200 years. As we have shown in Section 3.2, cross-country differences in adoption lags have declined. To assess more precisely the role of adoption lags in cross-country growth dynamics, we simulate the evolution of our two model economies after a common acceleration in frontier growth allowing for the specific evolutions of adoption lags that we observe in the data. The intensive margins are kept constant at the pre-industrial levels. Figure 8a presents the results from this simulation. It is clear that cross-country differences in adoption lags are a key driver of income divergence during the nineteenth century. In particular, prior to the non-Western country start adopting Modern technologies and quickly overtaking the Western country, the income gap reaches a level of 1.5. However, after that, the faster reduction in adoption lags in the poor country induces higher growth rates in the non-Western country. As a result, during the twentieth century income converges, and the relative income between the two countries is .95 by 2000.

The role of the intensive margin .– The income dynamics induced by adoption lags suggest that the evolution of the intensive margin may be necessary to explain why the great divergence continued during the twentieth century. To explore this hypothesis more rigorously, we simulate the evolution of the two economies following the acceleration of the common technology frontier, and feeding in the estimated dynamics of the intensive margin. In this simulation, we keep adoption lags constant at the pre-industrial levels.

Figure 8b presents the dynamics of income growth in each country. The first observation is that the divergence in the intensive margin of technology generates a very significant divergence in income growth between the rich and the poor country. In this simulation, the growth acceleration in the poor country starts much later than in the baseline (Figure 6). This is a consequence of omitting the productivity gains from a reduction in adoption lags in the poor country. Another perspective on this same issue is that the decline in the intensive margin reduces productivity growth by a magnitude that, initially, is equivalent to the gains brought by the industrial revolution to the poor countries.

It is also evident from Figure 8b that the rich country grows less, especially during the nineteenth century, than in the baseline. This is a reflection of the productivity gains brought by the reduction in adoption lags for rich countries. Furthermore, as shown in the bottom panel of Figure 8b, the growth gap between rich and poor countries during the nineteenth century is smaller when we omit the evolution of the adoption lags. Despite that, the poor country's growth rate falls behind the rich, and this gap does not begin to close until the second half of the twentieth century. By 2000, the income gap between the rich and the poor country would have increased by a factor of 4.7.[33]

To sum up, the findings from our simulations are as follows:

1. The model is capable of generating a Great Divergence where income per capita between rich and poor countries increases by a factor of five over the last 200 years, which represents most of the actual increase in the income gap observed in the data.

2. The presence of long adoption lags generates very protracted transitional dynamics.

3. Large cross-country differences in adoption lags explain much of income divergence during the nineteenth century between rich and poor countries.

4. The Great Divergence continued during the twentieth century because of the divergence in the penetration rates (i.e., intensive margin of adoption) between rich and poor countries.

6 Robustness

Next we show that these findings are robust to alternative (i) calibrations, (ii) definitions of the samples of rich and poor countries, and (iii) estimates of the intensive and extensive margin.

6.1 Calibration of γ and χ

The results discussed above assumed that the productivity growth after the Industrial Revolution was equally shared between the productivity growth of new technologies (χ) and of new vintages (γ). Given the difficulty of calibrating the contribution of these two sources of growth, it is necessary to study the robustness of our findings to the relative contributions

[33]One side comment we find interesting is that, shutting down the dynamics of adoption lags leads to more protracted dynamics transitional dynamics, as illustrated by Figure 8b. As anticipated above, this shows that the protractedness of the transitional dynamics is related to the length of the adoption lags. One prediction of the model is that the reduction in adoption lags we have observed in the last 200 years should lead to faster transitions at the (low) frequencies we are focusing on.

of new technologies and new varieties to balanced growth. To this end, we redo our baseline simulation under two polar assumptions. We first assume that all steady state growth comes from the development of better vintages, and then assume that all growth comes from the arrival of new, more productive, technologies.

Figure 9a depicts the dynamics of productivity growth in the first case, i.e., when all growth comes from new varieties being introduced in the economy. The most remarkable observation is that the poor country growth rate catches-up with the rich faster than in the baseline calibration. The reason is intuitive. Once the differences in adoption lags becomes small, both countries adopt new technologies at similar rates. Note that in this economy, growth comes from (i) expanding the range of varieties available, (ii) new vintages of a technology being more productive than older ones. There are no additional productivity gains from introducing new technologies. In other words, new technologies are not more productive than older technologies. Vintages of new technologies are as productive as vintages from old technologies (conditional on a vintage of the same age). Hence, in this scenario, differences in the intensive margin have a less damaging effect than in our baseline exercise. The reason is that the only gains from adopting new technologies come from increasing the varieties available in the economy. As the adoption lags of the poor countries converge to the rich countries, the growth rate of the two economies tends to converge because, conditional on adopting a technology, both countries adopt new vintages at the same rate. Hence, both economies expand the range of varieties of new technologies at the same rate and enjoy the same productivity gains, because the only source of growth in this economy comes from expanding the number of varieties and newer varieties being more productive. The income gap generated over two hundred years in this case is 2.5, which implies that the simulation accounts for 60% of the 4 fold difference that was actually generated over this time period.

Figure 9b shows the polar case, in which all productivity growth comes from the adoption of new technologies. In this case, the poor country lags behind more than in the baseline case. The reason is that in this case the effect of differences in the intensive margin of adoption are magnified. Note that in this case, all the productivity gains from adopting a new technology come from (i) new technologies being more productive than older technologies, (ii) gains from variety from adopting new vintages. Crucially, new vintages of a technology are not more productive than older ones. Hence, the marginal gains from expanding the range of varieties for a given technology are decreasing over time. This implies that the gains from convergence in adoption lags (i.e., vintages of new technologies being adopted at the same rate between rich and poor countries) has very little bite in this set up.

The modest gains from variety in the model imply that the key engine for long-run growth is the gains from adoption of new technologies. This makes the divergence in the intensive margin of adoption across countries much more salient. Recall that differences in the intensive margin of adoption are isomorphic to differences in the productivity of a technology. Hence, the divergence in the intensive margin makes the new technologies being adopted in poor countries relatively less productive than in rich countries. Given that this is the primary source of growth in this economy, this makes poor countries grow at a very slow rate. In other words, new technologies being less adopted in the intensive margin prevent poor countries from reaping all the benefits of new technologies –even when the differences in adoption lags are negligible between the two countries. The simulation generates a 3.1-fold income gap differential between the two countries, which coincides with the magnitude found by Maddison (2004).

6.2 Alternative definitions of non-western

We analyze how sensitive our results are to aggregating all non-Western. We compare re-do our exercise comparing now Western countries and countries in the bottom third of the income distribution in 2000.

6.3 Non-homotheticities in production

In this section, we explore how robust our predictions are once we allow for non-homotheticities in the production function. We depart from our baseline estimating equation (28) by allowing the coefficient on income to be different than one. We follow the same strategy as Comin and Mestieri (2012). We use the time series variation in GDP in the United States to compute the elasticity of income, and then use this estimate for all countries. When estimating the income elasticity for the United States, we want to distinguish between the short and long run income elasticities, since the former is likely to capture cyclical variation in the demand for investment goods. This presumption is confirmed by our estimates, where we find that the long-run income elasticity is around 2, while the short-run is three times higher.

As in Comin and Mestieri (2012), the estimates obtained allowing for non-homotheticities are very similar to our baseline estimates. In fact, the correlation between them is almost .9. The additional flexibility allowed in the model comes at the cost of a lower precision in the estimates of the adoption lag for three U.S. technologies: ships, mail and electricity. This creates the minor problem of having a less precise estimate for the United States in the intensive margin. Since we do not want to have as baseline intensity for the technology an imprecise estimate, we use the average intensive margin of Western countries as a reference point rather than the U.S.. Which country is taken as baseline is irrelevant for computing the cross-country dispersion measures. However, the mean intensive margin of adoption may be affected; therefore, the average intensive margin is not directly comparable with the homothetic case.

7 Conclusions

TO BE DONE

References

Barro, R. J. and Sala-i-Martin, X. (1997). Technological diffusion, convergence, and growth. *Journal of Economic Growth*, 2(1):1–26.

Barro, R. J. and Sala-i-Martin, X. (2003). *Economic Growth, 2nd Edition*, volume 1. The MIT Press.

Basu, S. and Fernald, J. G. (1997). Returns to scale in u.s. production:estimates and implications. *Journal of Political Economy*, 105(2):249–283.

Clark, G. and Feenstra, R. C. (2003). Technology in the great divergence. In *Globalization in Historical Perspective*, NBER Chapters, pages 277–322. National Bureau of Economic Research, Inc.

Comin, D. and Hobijn, B. (2009). The CHAT dataset. Working Paper 15319, National Bureau of Economic Research.

Comin, D. and Hobijn, B. (2010). An exploration of technology diffusion. *American Economic Review*, 100(5):2031–59.

Comin, D. and Mestieri, M. (2012). The intensive margin of technology adoption. *Forthcoming, Handbook of Economic Growth*.

Klenow, P. and Rodríguez-Clare, A. (1997). The neoclassical revival in growth economics: Has it gone too far? In *NBER Macroeconomics Annual 1997, Volume 12*, NBER Chapters, pages 73–114. National Bureau of Economic Research, Inc.

Lucas, R. E. (2000). Some macroeconomics for the 21st century. *Journal of Economic Perspectives*, 14(1):159–168.

Lucas, R. E. (2004). Industrial revolution: Past and future. *The Region*, 18(1):4.

Maddison, A. (2004). Contours of the world economy and the art of macro-measurement 1500-2001. *Ruggles Lecture*.

Norbin, S. (1993). The relation between price and marginal cost in u.s. industry: A contradiction. *Journal of Political Economy*, 101(6):1149–64.

Pritchett, L. (1997). Divergence, big time. *Journal of Economic Perspectives*, 11(3):3–17.

Dholakia, N, R. Dholakia and N. Kshetri (2003), "Internet Diffusion," *The Internet Encyclopedia*, edited by Hossein Bidgoli, New York: Wiley.

Khalba, K., (2007), "The Adoption of Mobile Phones in Emerging Markets: Global Diffusion and the Rural Challenge," 6th Annual Global Mobility Roundtable 2007 Center for Telecom Management Marshall School of Business, University of Southern Califonia.

A Detailed derivation of the Equilibrium Conditions

Derivation of equation (16): This follows from

$$
\begin{aligned}
Z_\tau &= \left(\int_\tau^{\max\{t-D_\tau,\tau\}} Z(\tau,v)^{\frac{1}{\mu-1}} dv \right)^{\mu-1} && (50) \\
&= a_\tau e^{(\chi+\gamma)\tau} \left(\int_\tau^{t-D_\tau} e^{\frac{\gamma}{\mu-1}(v-\tau)} dv \right)^{\mu-1} \\
&= \left(\frac{\mu-1}{\gamma} \right)^{\mu-1} a_\tau e^{(\chi+\gamma)\tau} \left(e^{\frac{\gamma}{\mu-1}(t-D_\tau-\tau)} - 1 \right)^{\mu-1} && (51)
\end{aligned}
$$

Derivation of equation (25): Using the definition of the production function and integrating, we have that

$$
\begin{aligned}
A &= \left(\int_{-\infty}^{\bar{\tau}} Z_\tau^{\frac{1}{\theta-1}} d\tau \right)^{\theta-1} \\
&= \left(\frac{\theta-1}{\gamma} \right)^{\theta-1} \left(\int_{-\infty}^{\bar{\tau}} \left[a_\tau e^{(\chi+\gamma)\tau} e^{\gamma(t-D_\tau-\tau)} \right]^{\frac{1}{\theta-1}} \left(1 - e^{-\frac{\gamma}{\theta-1}(t-D_\tau-\tau)} \right) d\tau \right)^{\theta-1}.
\end{aligned}
$$

With a constant D and a, we find

$$
A = a \left(\frac{\theta-1}{\gamma} \right)^{\theta-1} \left(\frac{\theta-1}{\chi} - \frac{\theta-1}{\chi+\gamma} \right)^{\theta-1} e^{(\chi+\gamma)(t-D)}, \qquad (52)
$$

after rearranging, we obtain (25).

Derivation of equation (27): Start considering a second order approximation of Z_τ around $t - D_\tau - \tau = 0$,

$$
Z_\tau \simeq a_\tau e^{(\chi+\gamma)\tau} \left[\Delta t \left(1 + \frac{1}{2} \frac{\gamma}{\mu-1} \Delta t \right) \right]^{\mu-1} \qquad (53)
$$

We can further simplify the expression of $\ln Z_\tau$ by using the first order Taylor approximation $\ln(1+x) \simeq x$ for small x, yielding

$$
\ln Z_\tau \simeq \ln a_\tau + (\chi+\gamma)\tau + (\mu-1)\ln \Delta t + \frac{\gamma}{2}\Delta t. \qquad (54)
$$

Equation (27) is obtained then by direct substitution.

B Description of the evolution of the output produced with Old technologies over time

Suppose that the change in the technology frontier is an instantaneous increase in the growth rate of the technology frontier, from $\chi_o + \gamma_o$ to $\gamma + \chi$, that takes place at time T. We keep the intensive and extensive margins constant at their pre-Industrial levels in this initial exercise.

Proposition 6 *Before adoption of Modern technologies, the economy is in a balanced growth path with growth rate $\chi_o + \gamma_o$,*

$$Y(t) = Ae^{(\chi_o+\gamma_o)t}. \tag{55}$$

After an economy starts adopting Modern technologies, output produced with Old technologies is

$$X_{Old}(t) \quad = \quad A_o e^{\gamma_o t} d(t)^{\theta-1}, \tag{56}$$

where A_o is a positive constant, $d(t)$ is an increasing, concave function, with initial value $\frac{\gamma_o}{\chi_o+\gamma_o}$ and $\lim_{t\to\infty} d(t) = 1$.

Proof For output produced before a country starts adopting modern technologies, $\tau < T$, we have that equation (52) holds, and hence

$$Y = X_{\text{Old}} = a_o \left(\frac{(\theta-1)^2}{\chi_o(\chi_o+\gamma_o)} \right)^{\theta-1} e^{(\chi_o+\gamma_o)(t-D_o)}. \tag{57}$$

Once the adopted technologies are $\tau > T$, the output produced with technologies with $\tau < T$ grows only due to new vintages appearing and being more productive

$$\begin{aligned}
X_{\text{Old}} &= a \left(\int_{-\infty}^{T} d\tau \int_{\tau}^{t-D_o} dv e^{\frac{\chi_o \tau + \gamma_o v}{\theta-1}} \right)^{\theta-1} && (58)\\
&= a \left(\frac{\theta-1}{\gamma_o} \int_{-\infty}^{T} d\tau \left(e^{\frac{\chi_o\tau+\gamma_o(t-D_o)}{\theta-1}} - e^{\frac{(\chi_o+\gamma_o)\tau}{\theta-1}} \right) \right)^{\theta-1} && (59)\\
&= a \left(\frac{(\theta-1)^2}{\gamma_o \chi_o} \right)^{\theta-1} e^{(\chi_o+\gamma_o)T} e^{\gamma_o(t-D_o-T)} \left(1 - \frac{\chi_o}{\chi_o+\gamma_o} e^{\frac{-\gamma_o(t-D_o-T)}{\theta-1}} \right)^{\theta-1}. && (60)
\end{aligned}$$

Equation (56) follows from arranging the terms appropriately. It is immediate to verify that $d(D_o + T) = \frac{\gamma_o}{\chi_o+\gamma_o}$ and $\lim_{t\to\infty} d(t) = 1$. Taking the derivative of $d(t)$, we have that it is positive and the second derivative negative, which completes the proof. ∎

Note that if we assume that pre-Modern technologies were equally productive $\chi_o = 0$, we obtain exactly equation (37). If $\chi_o > 0$ then there is an adjustment after the industrial revolution coming from the fact that only better vintages contribute to growth in the pre-

Modern output. In any case, as $\chi_o + \gamma_o \ll \chi + \gamma$, this transition is of no significance for the transition to Modern growth and can be disregarded.

C Proofs and Derivations of Section 4

Proof of Proposition 1: For the output produced with pre-Modern technologies, we have that by assumption,it grows at rate g_{Old}, so the solution to $dY/dt = g_{\text{Old}}$ with the boundary condition that at time D output is $Y(D)$ is given (37) with $A_o = Y(D)$. For the output produced with modern technologies, applying (19) for only modern technologies, we have that

$$
\begin{aligned}
X_{\text{Modern}} &= a\left(\int_T^{t-D} d\tau \int_\tau^{t-D} dv e^{\frac{\chi\tau+\gamma v}{\theta-1}}\right)^{\theta-1} && (61)\\
&= a\left(\frac{\theta-1}{\gamma}\int_T^{t-D} d\tau \left(e^{\frac{\chi\tau+\gamma(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)\tau}{\theta-1}}\right)\right)^{\theta-1} && (62)\\
&= a\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(e^{\frac{\chi(t-D)+\gamma(t-D)}{\theta-1}} - e^{\frac{\chi T+\gamma(t-D)}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(e^{\frac{(\chi+\gamma)(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)T}{\theta-1}}\right)\right\}\right]^{\theta-1} && (63)\\
&= ae^{(\chi+\gamma)(t-D)}\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(1 - e^{\frac{\chi(T-(t-D))}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(1 - e^{\frac{(\chi+\gamma)(T-(t-D))}{\theta-1}}\right)\right\}\right]^{\theta-1} && (64)\\
&= ae^{(\chi+\gamma)(t-D)}\left[\frac{(\theta-1)^2}{\gamma}\left\{\frac{1}{\chi}\left(1 - e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right\}\right]^{\theta-1} && (65)
\end{aligned}
$$

where $\Delta t \equiv t - D - T$. This last expression can be identified with (38), where

$$
h(t) = \frac{\chi(\chi+\gamma)}{\gamma}\left(\frac{1}{\chi}\left(1 - e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right). \quad (66)
$$

It is readily verified that $h(D+T) = 0$ and $\lim_{t\to\infty} h(t) = 1$. The derivative of $h(t)$ can be expressed as

$$
\frac{\gamma(\theta-1)}{\chi(\chi+\gamma)}h'^{-\frac{\chi\Delta t}{\theta-1}} - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}, \quad (67)
$$

from where it is apparent that $h'(D+T) = 0$ and $\lim_{t\to\infty} h'(t) = 0$. The second time derivative verifies

$$
\frac{\gamma(\theta-1)^2}{\chi(\chi+\gamma)}h'^{-\frac{\chi\Delta t}{\theta-1}} + (\chi+\gamma)e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}. \quad (68)
$$

It is readily verified that $\lim_{t\to\infty} h''(t) = 0$. Algebraic manipulation shows that $h(t)$ is convex for $\Delta t < \frac{\theta-1}{\gamma}\ln\left(\frac{\chi+\gamma}{\chi}\right)$. ∎

Proof of Proposition 2: First we show that the weights on the modern growth sectors have an S-shape. Note that

$$\left(\frac{X_{\text{Old}}}{Y}\right)^{\frac{1}{\theta-1}} = \frac{1}{1+\left(\frac{X_1}{X_0}\right)^{\frac{1}{\theta-1}}}, \tag{69}$$

$$\frac{X_{\text{Modern}}}{X_{\text{Old}}} \propto e^{(\chi+\gamma-g_o)t}\left[\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right)-\frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right]^{\theta-1}. \tag{70}$$

Taking the time derivative of (69) it is readily verified that this share declines over time. Moreover, the sign of the second derivative of (69) coincides with the sign of

$$\left(\frac{d}{dt}\left(\frac{X_{\text{Modern}}}{X_{\text{Old}}}\right)^{\frac{1}{\theta-1}}\right)^2-\left(1+\left(\frac{X_{\text{Modern}}}{X_{\text{Old}}}\right)^{\frac{1}{\theta-1}}\right)\frac{d^2}{dt^2}\left(\frac{X_{\text{Modern}}}{X_{\text{Old}}}\right)^{\frac{1}{\theta-1}}. \tag{71}$$

Note that in the case that $\left(\frac{X_{\text{Modern}}}{X_{\text{Old}}}\right)^{\frac{1}{\theta-1}}$ is concave, the share is unambiguously convex. As we previously discussed for the damp factor, this occurs for sufficiently large t. To see that, denoting by $g \equiv \frac{\chi+\gamma-g_o}{\theta-1}$, abusing notation substituting $t \equiv \Delta t$ and taking the explicit derivatives of the share, it can be verified that the sign of (71) coincides with the sign of

$$e^{gt}\left(h'^2-h''(t)h(t)\right)-\left(g^2h(t)+2gh'(t)+h''(t)\right). \tag{72}$$

Using the properties derived in Proposition 1 for $h(t)$ that the first and second derivative vanish for large t, it is immediate to verify that the limit as t approaches infinity of (72) is positive. Similarly, when a country starts to adopt technologies of the industrial revolution (for $t=0$ after the change of variables), equation (72) simplifies to $-h''(0)<0$. So we he have that the share on pre-Modern output is initially concave and eventually becomes convex. Hence, the share on Modern output is initially convex and eventually concave.[34]

From (36) we can see how the S-shape of the weights translates into an S-shape for the growth rate of output during the transition to the new balanced growth path. If $g_{\text{Modern}} = \chi+\gamma$, then we would have an exact S-shape for the growth rate of aggregate output. However, from our discussion of the damp-factor, we know that g_{Modern} can grow at a faster than $\chi+\gamma$ for some transient period, while the dynamics of the damp factor are relevant. Hence, this can give rise to some over-shooting of the long-run growth rate if when the weight on modern growth becomes close to one, the growth rate of the modern sector is substantially higher than $\chi+\gamma$. ∎

[34]We have not shown that there exists a t^* below which a share is convex and concave thereafter, even though out simulations suggest so.

Proof of Proposition 3: Start with the half-life of the growth rate. The definition of the half-life is

$$\frac{\chi+\gamma}{2} = \left(\frac{X_M(t_{1/2})}{Y(t_{1/2})}\right)^{\frac{1}{\theta-1}} g_M(t_{1/2}) + \left(\frac{X_O(t_{1/2})}{Y(t_{1/2})}\right)^{\frac{1}{\theta-1}} g_O, \tag{73}$$

where we are shortening the subindices, M for Modern, O for Old and

$$g_M = \chi + \gamma + (\theta-1)\frac{h'(t)}{h(t)}. \tag{74}$$

Rearranging, equation (73) becomes

$$(\chi+\gamma+2(\theta-1)g_h(t_{1/2}))X_M(t_{1/2})^{\frac{1}{\theta-1}} = (\chi+\gamma-2g_O)X_O(t_{1/2})^{\frac{1}{\theta-1}}. \tag{75}$$

This is a transcendental equation, which cannot be solved analytically. Before proceeding, we state the following result. The average value of the function $e^{-\beta t}$ for $t \in [0, T]$ is

$$< e^{-\beta t} >= \frac{1}{T}\int_0^T e^{-\beta t} dt = \frac{1-e^{-\beta T}}{\beta T}. \tag{76}$$

We proceed by averaging $h(t)$ and $h'(t)$ to make (75) analytically solvable,

$$(\chi+\gamma+2(\theta-1) < g_h >)X_M(t_{1/2})^{\frac{1}{\theta-1}} = (\chi+\gamma-2g_O)X_O(t_{1/2})^{\frac{1}{\theta-1}}. \tag{77}$$

Denoting by

$$\alpha \equiv \left(\frac{(\chi+\gamma+2(\theta-1) < g_h >}{\chi+\gamma-2g_O}\right)^{\theta-1}, \tag{78}$$

equation (77) is

$$\alpha A_M e^{(\chi+\gamma)(t-D)} h(t)^{\theta-1} = A_O e^{g_O(t-D)}, \tag{79}$$

where we are taking the normalization $T = 0$. As stated before, we proceed by averaging $h(t)$ to make the problem analytically solvable, which yields,

$$t = D + \frac{1}{\chi+\gamma-g_O}\ln\left(\frac{A_O}{\alpha A_M < h(t) >^{\theta-1}}\right). \tag{80}$$

Finally, note that if in the approximation of the averages we would have taken a large T, we would have obtained that $< h(t) >\simeq 1$ and that $g_h \simeq 0$. In this case, $\alpha \simeq 1$ (as $\chi+\gamma \gg g_O$) and we would obtain the result reported in the paper. This shows the result for the half-life of the growth.

Calibration of the half-life of growth We discuss the making of the quantitative exercise next. We take $\chi = \gamma = 1\%$, $\theta = 1.4$. To approximate $h(t)$, we take $T = 80$ years (which is in line from what we obtain in the simulations for the time it takes $h(t)$ to reach 1). If $\beta = \frac{\chi}{\theta-1} \sim 1/40$, we have that $\beta T = 2$, $< e^{-\beta t} >= \frac{1-e^{-2}}{2} \simeq .43$, while if $\beta = \frac{\chi+\gamma}{\theta-1} \sim 1/20$ which gives $\beta T = 4$ and $< e^{-\beta t} >= \frac{1-e^{-4}}{4} \simeq .25$. Thus we have that

$$\frac{1}{\chi}\left(1 - e^{-\frac{\chi \Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right) \simeq 100(1-.43) - 50(1-.25) = 19.5, \quad (81)$$

$$\frac{1}{\theta-1}\left(e^{-\frac{\chi \Delta t}{\theta-1}} - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right) \simeq \frac{10}{4}(.43 - .25) = .45. \quad (82)$$

This implies that $g_h = .45/19.5 \simeq 2.3\%$, thus

$$\alpha = \left(\frac{2\% + 2 \cdot .4 \cdot 2.3\%}{2\% - .4\%}\right)^{.4} = \left(\frac{3.84\%}{1.6\%}\right)^{.4} = 1.42.$$

Next, we have that

$$A_M < h(t) >^{\theta-1} = \alpha a \xi \left(\frac{(\theta-1)^2}{\gamma} < \frac{1}{\chi}\left(1 - e^{-\frac{\chi \Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right) >\right)^{\theta-1} \quad (83)$$

where ξ is a parameter in the production function that ensures that output in 1820 is consistent with the data, its value is our baseline calibration is around 5. Quantitatively this takes the value of

$$A_M < h(t) >^{\theta-1} = 1.42 \cdot \frac{4}{10} \cdot 5 \left(\frac{.4^2}{.01} 19.5\right)^{.4} \simeq 30 \quad (84)$$

Finally, taking the initial average income per capita before the industrial revolution at subsistence levels ($400) we can compute expression (80)

$$t_{1/2}^{\text{growth}} - D = \frac{1}{2\%} \ln\left(\frac{400}{30}\right) \simeq 130 \text{ years} \quad (85)$$

If instead of using the averages, we would have used the approximation of a large T, we would have obtained

$$A_M < h(t) >^{\theta-1} = a\xi \left(\frac{(\theta-1)^2}{\chi(\chi+\gamma)}\right)^{\theta-1} \simeq 30 \text{ years}. \quad (86)$$

Thus, this assumption seems quite innocuous as yields very similar results.

Derivation of the half-life of output Next, we derive the half-life of output. Define

$$\tilde{Y}(t) = \frac{Y(t)}{a_M A_M e^{(\chi+\gamma)(t-D)}}. \quad (87)$$

By construction, $\lim_{t\to\infty}\tilde{Y}(t)=1$. Hence, the definition of the half-life is

$$\tilde{Y}(0)+\frac{1}{2}(1-\tilde{Y}(0))=\tilde{Y}(t_{1/2}), \tag{88}$$

where we are taking the normalization $T=0$. This is a transcendental equation, so to make further progress we substitute the transient part of modern output

$$\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right)-\frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right) \tag{89}$$

for its average value (more on this on the calibration below), which we denote by κ. Denoting the left hand side of (88) by C, we can re-write as

$$\frac{a_O A_O}{a_M A_M\left(C^{\frac{1}{\theta-1}}-\kappa^{\frac{1}{\theta-1}}\right)^{\theta-1}}=e^{(\chi+\gamma-g_O)(t-D)}. \tag{90}$$

Solving for t we have that

$$t=D+\frac{1}{\chi+\gamma-g_O}\ln\left(\frac{a_O A_O}{a_M A_M\left(C^{\frac{1}{\theta-1}}-\kappa^{\frac{1}{\theta-1}}\right)^{\theta-1}}\right). \tag{91}$$

Note that the expression reported in the paper has a 2 instead of the constant $\left(C^{\frac{1}{\theta-1}}-\kappa^{\frac{1}{\theta-1}}\right)^{\theta-1}$. To obtain the exact expression reported in the paper, see below, in which we use an alternative normalization.

Calibration of the half-life of levels We take $Y(0)$ to be the subsistence level in Maddison (2004) of \$400, taking $\chi=\gamma=1\%$ and $\theta=1.3$, $\left(\frac{(\theta-1)^2}{\chi(\chi+\gamma)}\right)^{\theta-1}=14.5$ which implies that

$$\tilde{Y}(O)=\frac{400}{.4\cdot 14.5}=69. \tag{92}$$

This implies that

$$C=69-.5\cdot 68=35 \tag{93}$$

The average value of the transient term (89) we already calculated in (81) to be 19.5. Thus, we have that

$$\left(C^{\frac{1}{\theta-1}}-\kappa^{\frac{1}{\theta-1}}\right)^{\theta-1}=(35^{\frac{1}{.4}}-19.5^{\frac{1}{.4}})^{.4}=31.5 \tag{94}$$

Substituting into (164) we have that

$$t-D=\frac{1}{1.8\%}\ln\left(\frac{400}{.4\cdot 14.5\cdot 31.5}\right)=40\text{ years}. \tag{95}$$

Alternative definition of half-life. If we study the half life of the output gap relative to BGP,

$$\bar{Y}(t) = \frac{a_M A_M e^{(\chi+\gamma)(t-D)}}{Y(t)}, \tag{96}$$

we obtain similar results, but we can obtain a sharper characterization. As before, by construction, $\lim_{t\to\infty} \tilde{Y}(t) = 1$. However, now $\bar{Y}(t)$ is an increasing function. The definition of the half-life is

$$\bar{Y}(0) + \frac{1}{2}(1 - \bar{Y}(0)) = \bar{Y}(t_{1/2}), \tag{97}$$

but now, using (92), we know that $\bar{Y}(0) \simeq 0$, which allows to further simplify the half-life definition to

$$\frac{1}{2} \simeq \bar{Y}(t_{1/2}). \tag{98}$$

Following the previous steps and approximating the average of $h(t)$ by its long-run level, one obtains that

$$\frac{a_O Y_O}{a_M A_M \left(2^{\frac{1}{\theta-1}} - 1\right)^{\theta-1}} = e^{(\chi+\gamma-g_O)(t-D)}. \tag{99}$$

As $2^{\frac{1}{\theta-1}} \gg 1$, we can approximate the last equation by (40). ∎

Proof of Proposition 4 With the definition of the evolution of the intensive and extensive margins (42), we have that Old and Modern output is calculated as in Proposition 1. Applying the definition of evolution of margins we just have to substitute D for D_{Old} in the computation for Old output and D for D_{Modern} in the computation of Modern output. The rest of the claims in the proposition, can be derived analogously to Propositions (42), (35) and (**??**) replacing D for D_{Modern}. An additional correction appears linearly, $t_{1/2} = D_{Modern} - \frac{g_O}{\chi+\gamma} D_{Old} + \dots$ when re-doing the algebra. Note, however that $D_{Modern} \gg \frac{g_O}{\chi+\gamma} D_{Old}$, so this could be in principle neglected.

Proof of Proposition 5: The derivation for output of the Old sector is as in Proposition 1. Next, we characterize Modern output. First, we analyze the case in which $\tau < \bar{T}$. Note that the range of integration for a given technology that is being used goes from $[\tau, t - D_\tau]$, where t denotes current time and D_τ is the lag of technology τ. Without loss of generality, normalize the advent of the Industrial revolution $T = 0$. Recall the parametrization on the evolution of the margins of adoptions, which in this range we simply denote by $D_\tau = d_0 - d_1\tau$ and $\ln a_\tau = a_0 - a_1\tau$. To map D_τ into the time space, note that the first technology will be adopted at time $t = d_0$ and that the range of available technologies at time t can be written as $[0, t - (d_0 - d_1(t - d_0))] = [0, (1 + d_1)(t - d_0)]$. The range of vintages of technology τ at time t is given by the difference between the time the last adopted vintage and the time of adoption of the first one, $t - (\tau + D_\tau)$, $v_\tau \in [\tau, t - D_\tau]$. The output produced using modern technologies can be written as

$$
\begin{aligned}
X_m &= \left(\int_0^{t-D_t} d\tau \int_\tau^{t-D_\tau} dv \left[a_\tau Z(\tau,v)\right]^{\frac{1}{\theta-1}}\right)^{\theta-1} && (100)\\
&= e^{a_0}\left(\int_0^{(1+d_1)(t-d_0)} d\tau \left[e^{\frac{(\chi-a_1)\tau}{\theta-1}} \int_\tau^{t-(d_0-d_1\tau)} dv e^{\frac{\gamma v}{\theta-1}}\right]\right)^{\theta-1} && (101)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\int_0^{(1+d_1)(t-d_0)} d\tau e^{\frac{(\chi-a_1)\tau}{\theta-1}}\left[e^{\frac{\gamma(t-(d_0-d_1\tau))}{\theta-1}} - e^{\frac{\gamma\tau}{\theta-1}}\right]\right)^{\theta-1} && (102)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\int_0^{(1+d_1)(t-d_0)} d\tau\left[e^{\frac{(\chi+d_1\gamma-a_1)\tau+\gamma(t-d_0)}{\theta-1}} - e^{\frac{(\gamma+\chi-a_1)\tau}{\theta-1}}\right]\right)^{\theta-1} && (103)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(^{\dots\theta-1}\ e^{\frac{\gamma(t-d_0)}{\theta-1}}\left(\frac{\theta-1}{\chi+\gamma d_1-a_1}\right)\left[e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}((1+d_1)(t-d_0))}-1\right]\right)-\dots\\
&\quad \left(^{\dots\theta-1}\ \dots - \frac{\theta-1}{\chi+\gamma-a_1}\left[e^{\frac{\chi+\gamma-a_1}{\theta-1}((1+d_1)(t-d_0))}-1\right]\right)^{\theta-1} && (104)\\
&= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1}\exp\left[\left((\gamma+\chi-a_1)(1+d_1)+d_1^2\gamma\right)(t-d_0)\right] && (105)\\
&\quad \left(\frac{1}{\chi+\gamma d_1-a_1}\left[1-e^{-\frac{\chi+\gamma d_1-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]-\frac{e^{-\frac{d_1^2\gamma}{\theta-1}(t-d_0)}}{\chi+\gamma-a_1}\left[1-e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]\right)^{\theta-1}.
\end{aligned}
$$

This last expression can be rewritten as

$$
X_m(t) = A_m e^{g_m t} f(t) \qquad (106)
$$

where

$$
\begin{aligned}
A_m &= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1} e^{-d_0\left((\gamma+\chi-a_1)(1+d_1)+d_1^2\gamma\right)} && (107)\\
g_m &= (\gamma+\chi-a_1)(1+d_1)+d_1^2\gamma && (108)\\
f(t) &= \left(\frac{1}{\chi+\gamma d_1-a_1}\left[1-e^{-\frac{\chi+\gamma d_1-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]-\frac{e^{-\frac{d_1^2\gamma}{\theta-1}(t-d_0)}}{\chi+\gamma-a_1}\left[1-e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]\right)^{\theta-1}
\end{aligned}
$$

Next we analyze the properties of $f(t)$. First, note that the first instant of time in which technology is adopted, $t = d_0$, $f(d_0) = 0$ and $\lim_{t\to\infty} f(t) = \left(\frac{1}{\chi+\gamma d_1-a_1}\right)^{\theta-1}$. To further analyze the behaviour of the damp factor $f(t)$ it is useful to rewrite it as $f(t) = h(t)^{\theta-1}$, note that

$$f'^{\theta-2}h'(t), \tag{109}$$

$$f''^{\theta-2}\left[(\theta-2)h(t)^{-1}h'^2+h''(t)\right], \tag{110}$$

$$g_f \equiv (\ln f(t))' = (\theta-1)\frac{h'(t)}{h(t)}, \tag{111}$$

$$g_f' = (\theta-1)\frac{h''(t)h(t)-h'^2}{h(t)^2}. \tag{112}$$

The time derivative of $h(t)$ is

$$\begin{aligned}(\theta-1)h'(t) &= (1+d_1)e^{-\frac{\chi+\gamma d_1-a_1}{\theta-1}(1+d_1)(t-d_0)}\dots \\ &\dots - e^{-\frac{\gamma d_1^2(t-d_0)}{\theta-1}}\left[(1+d_1)e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}-\frac{\gamma d_1^2}{\chi+\gamma-a_1}\left(1-e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right)\right]\end{aligned}$$

Result 1: $h'(t)>0$ for $t>d_0$, $h'(d_0)=0$ and $\lim_{t\to\infty}h'(t)=0$. Proof: By direct substitution it is verified that $h'(d_0)=0$. To show that $h'(t)>0$. Suppose that it is true, and rearrange,

$$\begin{aligned}(1+d_1)e^{-\frac{\chi+\gamma d_1-a_1}{\theta-1}(1+d_1)(t-d_0)} &> e^{-\frac{\gamma d_1^2(t-d_0)}{\theta-1}}(1+d_1)e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\dots \\ &\dots - e^{-\frac{\gamma d_1^2(t-d_0)}{\theta-1}}\frac{\gamma d_1^2}{\chi+\gamma-a_1}\left(1-e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right) \\ (1+d_1)e^{\frac{-(\chi+\gamma d_1-a_1)(1+d_1)+d_1^2\gamma}{\theta-1}(t-d_0)} &> (1+d_1)e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}-\frac{\gamma d_1^2}{\chi+\gamma-a_1}\left(1-e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right) \\ (1+d_1)e^{\frac{\gamma}{\theta-1}(t-d_0)} &> (1+d_1)-\frac{\gamma d_1^2}{\chi+\gamma-a_1}\left(e^{\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}-1\right)\end{aligned}$$

Note that the left hand side is an increasing function of t while the right hand side is decreasing. Moreover the left hand side equals the right hand side at $t=d_0$, establishing the result claimed for $t>d_0$. Finally, the result that $\lim_{t\to\infty}h'(t)=0$ follows directly from taking the limit of $h'(t)$. QED

Result 2: h(t) is convex for $t_0\le t<t^*$ and concave thereafter. Moreover, $\lim_{t\to\infty}h''(t)=0$. Proof: The expression for $(\theta-1)^2h''(t)$ is

$$-(1+d_1)^2(\chi+d_1\gamma-a_1)e^{-\frac{\chi+d_1\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}+e^{-\frac{(1+d_1)(\chi+\gamma-a_1)+d_1^2\gamma}{\theta-1}(t-d_0)}\cdot$$
$$\left((1+d_1)(2\gamma d_1^2+(1+d_1)(\chi+\gamma-a_1))+\frac{\gamma d_1^2}{\chi+\gamma-a_1}-\frac{\gamma d_1^2}{\chi+\gamma-a_1}e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right).$$

Evaluating this expression at $t=d_0$ yields $(1+d_1)(1+d_1^2)\gamma>0$. Next, conjecture that

$\theta-1)^2h''(t)>0$. This implies that

$$(1+d_1)^2(\chi+d_1\gamma-a_1)e^{-\frac{\chi+d_1\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)} < e^{-\frac{(1+d_1)(\chi+\gamma-a_1)+d_1^2\gamma}{\theta-1}(t-d_0)}.$$
$$\left((1+d_1)(2\gamma d_1^2+(1+d_1)(\chi+\gamma-a_1))+\frac{\gamma d_1^2}{\chi+\gamma-a_1}-\frac{\gamma d_1^2}{\chi+\gamma-a_1}e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right)$$
$$\Longleftrightarrow$$
$$(1+d_1)^2(\chi+d_1\gamma-a_1)e^{\gamma(t-d_0)} <$$
$$\left((1+d_1)(2\gamma d_1^2+(1+d_1)(\chi+\gamma-a_1))+\frac{\gamma d_1^2}{\chi+\gamma-a_1}-\frac{\gamma d_1^2}{\chi+\gamma-a_1}e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right)$$

This last expression is indeed satisfied for $t=d_0$ (as is the same expression we evaluated before). Note that the left hand side is an increasing function that tends to infinity, while the right hand side is a decreasing function that tends to minus infinity. Thus, at some $t^*\geq t_0$ this inequality will cease to be true, and $(\theta-1)^2h''(t)<0$ in that range. Finally, the result that $\lim_{t\to\infty}h''(t)=0$ follows directly from taking the limit of $h''(t)$. QED

We briefly discuss how the behaviour of $h(t)$ can inform our analysis on $f(t)$ and its derivatives. Using equation (109) it is immediate to verify that $f'(t)$ inherits the properties of $h'(t)$, and hence, $f'(t)$ is increasing and $f'(d_0)=\lim_{t\to\infty}f'(t)=0$. Similarly, using (112), we conclude that g_f is increasing and $g_f'(d_0)=\lim_{t\to\infty}g_f'(t)=0$. Moreover $\lim_{t\to\infty}f''(t)=0$. It can be verified too that $\lim_{t\to d_0}f'(t)=\frac{(1+d_1+d_1^2+d_1^3)\gamma}{(\theta-1)^2}>0$.

Next we analyze the case in which $\tau>T$. (the time corresponding to the transition is $t=d_0+T/(1+d_1)$. Note that the output produced with Modern technologies can be divided in the output produced using technologies $\tau\in[0,T]$ and the subsequent technologies, $\tau>T$. The output produced using the first range of technologies can be computed as we have done before,

$$
\begin{aligned}
X_{m0}(t) &= \left(\int_0^T d\tau \int_\tau^{t-D_\tau} dv\,[a_\tau Z(\tau,v)]^{\frac{1}{\theta-1}}\right)^{\theta-1} && (113)\\
&= e^{a_0}\left(\int_0^T d\tau \left[e^{\frac{(\chi-a_1)\tau}{\theta-1}} \int_\tau^{t-(d_0-d_1\tau)} dv e^{\frac{\gamma v}{\theta-1}}\right]\right)^{\theta-1} && (114)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\int_0^T d\tau e^{\frac{(\chi-a_1)\tau}{\theta-1}}\left[e^{\frac{\gamma(t-(d_0-d_1\tau))}{\theta-1}} - e^{\frac{\gamma\tau}{\theta-1}}\right]\right)^{\theta-1} && (115)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\int_0^T d\tau \left[e^{\frac{(\chi+d_1\gamma-a_1)\tau+\gamma(t-d_0)}{\theta-1}} - e^{\frac{(\gamma+\chi-a_1)\tau}{\theta-1}}\right]\right)^{\theta-1} && (116)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(e^{\frac{\gamma(t-d_0)}{\theta-1}}\left(\frac{\theta-1}{\chi+\gamma d_1-a_1}\right)\left[e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}T}-1\right] - \frac{\theta-1}{\chi+\gamma-a_1}\left[e^{\frac{\chi+\gamma-a_1}{\theta-1}T}-1\right]\right)^{\theta-1} && (117)\\
&= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1}\exp\left[\gamma(t-d_0)\right] && (118)\\
&\quad \left(\left(\frac{\theta-1}{\chi+\gamma d_1-a_1}\right)\left[e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}T}-1\right] - e^{-\frac{\gamma}{\theta-1}(t-d0)}\frac{\theta-1}{\chi+\gamma-a_1}\left[e^{\frac{\chi+\gamma-a_1}{\theta-1}T}-1\right]\right)^{\theta-1}
\end{aligned}
$$

Note that this is an increasing function. Write output produced $X_{m0}(t) = C(Ae^{gt}-B)^{\theta-1}$

$$X_{m0}'^{gt}(\theta-1)(Ae^{gt}-B)^{\theta-2} > 0. \tag{119}$$

The second derivative is

$$X_{m0}''^{gt} + B)^{\theta-2} g^2 A\left(A(\theta-2)(Ae^{gt}-B)^{-1}+1\right). \tag{120}$$

It is clear that (120) is asymptotically positive. Whether or not it is always positive, depends on whether

$$Ae^{gT} - B > A(\theta-2),$$

which depends on parametric assumptions.

Next we derive the output produced with $\tau > T$ using in equation (**??**) in which case the adoption margins are constants and we denote $T \equiv t_i$

$$
\begin{aligned}
X_{m1} &= a\left(\int_{t_i}^{t-D} d\tau \int_{\tau}^{t-D} dv e^{\frac{\chi\tau+\gamma v}{\theta-1}}\right)^{\theta-1} && (121)\\
&= a\left(\frac{\theta-1}{\gamma}\int_{t_i}^{t-D} d\tau\left(e^{\frac{\chi\tau+\gamma(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)\tau}{\theta-1}}\right)\right)^{\theta-1} && (122)\\
&= a\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(e^{\frac{\chi(t-D)+\gamma(t-D)}{\theta-1}} - e^{\frac{\chi t_i+\gamma(t-D)}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(e^{\frac{(\chi+\gamma)(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)t_i}{\theta-1}}\right)\right\}\right]^{\theta-1} && (123)\\
&= ae^{(\chi+\gamma)(t-D)}\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(1-e^{\frac{\chi(t_i-(t-D))}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(1-e^{\frac{(\chi+\gamma)(t_i-(t-D))}{\theta-1}}\right)\right\}\right]^{\theta-1} && (124)\\
&= ae^{(\chi+\gamma)(t-D)}\left[\frac{(\theta-1)^2}{\gamma}\left\{\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right\}\right]^{\theta-1} && (125)
\end{aligned}
$$

where $\Delta t \equiv t - D - t_i$ and $a = e^{a_0}$

Note that the results we have derived for $t < T$ apply directly to X_{m1} because the transient part of (125) is a particular case of the case analyzed previously for $d_1 = a_1 = 0$. In this case, taking the second derivative of $h(t)$ one can find a closed form expression for the threshold t^* above which $h(t)$ becomes convex. It is $t^* = \frac{\theta-1}{\gamma}\ln\left(\frac{\chi+\gamma}{\chi}\right)$.

Next, note that the total modern output produced when technologies $\tau > T$ have been adopted is

$$
X_m = \left(X_{m0}^{\frac{1}{\theta-1}} + X_{m1}^{\frac{1}{\theta-1}}\right)^{\theta-1}. \qquad (126)
$$

As both X_{m0} and X_{m1} are increasing functions of time, it is immediate to verify that X_m is increasing over time. To gain further insight on its behaviour, note that (118) and (125) can be written as

$$
\begin{aligned}
X_{m0} &= Ae^{(\chi+\gamma)t}\left(Be^{-\frac{\chi}{\theta-1}t} - Ce^{-\frac{\chi+\gamma}{\theta-1}t}\right)^{\theta-1}, && (127)\\
X_{m1} &= Ae^{(\chi+\gamma)t}\left(De^{-\frac{\chi+\gamma}{\theta-1}t} - Ee^{-\frac{\chi}{\theta-1}t} + F\right)^{\theta-1}, && (128)\\
A &= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1}, && (129)\\
B &= \frac{e^{-\frac{\gamma}{\theta-1}d0}}{\chi+\gamma d_1 - a_1}\left(e^{\frac{\chi+\gamma d_1 - a_1}{\theta-1}T} - 1\right), && (130)\\
C &= \frac{1}{\chi+\gamma-a_1}\left(e^{\frac{\chi+\gamma-a_1}{\theta-1}T} - 1\right), && (131)\\
D &= \frac{e^{\frac{\chi+\gamma}{\theta-1}T}}{\chi+\gamma}, && (132)\\
E &= \frac{e^{\frac{\chi T - \gamma d_m}{\theta-1}}}{\chi}, && (133)\\
F &= \frac{\gamma e^{-\frac{\chi+\gamma}{\theta-1}d_m}}{\chi(\chi+\gamma)}. && (134)
\end{aligned}
$$

Using (126), we have that

$$
X_m = Ae^{(\chi+\gamma)t}\left(\underbrace{(D-C)e^{-\frac{\chi+\gamma}{\theta-1}t} - (E-B)e^{-\frac{\chi}{\theta-1}t} + F}_{h(t)}\right)^{\theta-1} \tag{135}
$$

Denoting by $h(t)$ the terms inside the parenthesis, we have that

$$
\begin{aligned}
(\theta-1)h'^{-\frac{\chi+\gamma}{\theta-1}t} + \chi(E-B)e^{-\frac{\chi}{\theta-1}t}, && (136)\\
(\theta-1)^2 h''^2(D-C)e^{-\frac{\chi+\gamma}{\theta-1}t} - \chi^2(E-B)e^{-\frac{\chi}{\theta-1}t}, && (137)
\end{aligned}
$$

where

$$
\begin{aligned}
D-C &= \frac{e^{\frac{\chi+\gamma}{\theta-1}T}}{\chi+\gamma} - \frac{1}{\chi+\gamma-a_1}\left(e^{\frac{\chi+\gamma-a_1}{\theta-1}T} - 1\right), && (138)\\
&= \frac{1}{\chi+\gamma-a_1}\left[1 + e^{\frac{\chi+\gamma}{\theta-1}T}\left(1 - e^{-\frac{a_1}{\theta-1}T}\right) - \frac{a_1 e^{\frac{\chi+\gamma}{\theta-1}T}}{\chi+\gamma}\right], && (139)\\
E-B &= \frac{e^{\frac{\chi T-\gamma d_m}{\theta-1}}}{\chi} - \frac{e^{-\frac{\gamma}{\theta-1}d0}}{\chi+\gamma d_1 - a_1}\left(e^{\frac{\chi+\gamma d_1 - a_1}{\theta-1}T} - 1\right), && (140)
\end{aligned}
$$

In general, the properties of $h'(t)$ and $h''(t)$ depends on the combination of several parameters.

To gain some insight, consider the case in which $\chi, \gamma \gg a_1, d_1$,

$$D - C = \frac{1}{\chi + \gamma}, \tag{141}$$

$$E - B = \frac{1}{\chi}\left(e^{-\frac{\gamma d_0}{\theta-1}} + e^{\frac{\chi T}{\theta-1}}\left(e^{-\frac{\gamma d_m}{\theta-1}} - e^{-\frac{\gamma d_0}{\theta-1}}\right)\right). \tag{142}$$

In this case, $(\theta-1)h'(T)$,

$$-e^{-\frac{\chi+\gamma}{\theta-1}T} + e^{-\frac{\chi T + \gamma d_0}{\theta-1}} + e^{-\frac{\gamma d_m}{\theta-1}}\left(1 - e^{-\frac{\gamma}{\theta-1}(d_o - d_m)}\right). \tag{143}$$

A sufficient condition for $h'(t)$ to be increasing for all $t \geq T$ is that $d_0 < T$. That is the initial lag has to be relatively small compared to the transition period. For the second derivative, we have that

$$(\chi+\gamma)e^{-\frac{\chi+\gamma}{\theta-1}t} - \chi e^{-\frac{\chi}{\theta-1}t}\left(e^{-\frac{\gamma d_0}{\theta-1}} + e^{\frac{\chi T - \gamma d_m}{\theta-1}}\left(1 - e^{-\frac{\gamma}{\theta-1}(d_o-d_m)}\right)\right). \tag{144}$$

This shows already that asymptotically, (i.e., for large t) $h''(t) < 0$. Similar to the analysis of the first derivative, we have that evaluated at $t = T$, a sufficient condition for equation (144) to be positive is $d_0 < T$. In this case we would have an S-shape.

Next, we study the behavior of the share

$$s = \frac{1}{1 + \left(\frac{X_m}{X_o}\right)^{\frac{1}{\theta-1}}}.$$

The quotient in the previous expression can be written as,

$$y \equiv \left(\frac{X_m}{X_o}\right)^{\frac{1}{\theta-1}} = \frac{X_{m0}^{\frac{1}{\theta-1}} + X_{m1}^{\frac{1}{\theta-1}}}{X_o^{\frac{1}{\theta-1}}} = C_0 e^{g_0 t} + C_1 e^{g_1 t} h(t) \tag{145}$$

where C_0, C_1 are two constants, $g_0 < g_1$, and $h(t)$ is given by equation (125),

$$\frac{1}{\chi}\left(1 - e^{-\frac{\chi \Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right) \tag{146}$$

with $\Delta t = t - T$. It is immediate to verify that y is an increasing function (as it is the sum of two increasing functions). Thus, s is decreasing over time. Taking the second derivative over time of s, one finds that the sign of the second derivative is the same as the sign of $\left[\dot{y}^2 - (1+y)\ddot{y}^2\right]$, using that

$$\dot{y} = g_0 C_0 e^{g_0 t} + g_1 C_1 e^{g_1 t} h + C_1 e^{g_1 t}\dot{h} \tag{147}$$

$$\ddot{y} = g_0^2 C_0 e^{g_0 t} + g_1^2 C_1 e^{g_1 t} h + 2 g_1 C_1 e^{g_1 t}\dot{h} + C_1 e^{g_1 t}\ddot{h} \tag{148}$$

Using the fact that $h(T) = \dot{h}(T) = 0$, $\ddot{h}(T) = \gamma$, we have that $\left[\dot{y}(T)^2 - (1 + y(T))\ddot{y}(T)^2\right]$ is

$$g_0^2 C_0^2 e^{2g_0 T} - (1 + C_0 e^{g_0 t})(g_0^2 C_0 e^{g_0 T} + C_1\gamma) < 0. \quad (149)$$

Thus, s is initially concave. That is, there exist a $\varepsilon > 0$ such that if $t \in [T, T + \varepsilon]$ then $\dot{y}(t)^2 - (1 + y(t))\ddot{y}(t)^2 < 0$.

Next, using that $\lim_{t\to\infty} h(t) = \chi/\gamma(\chi+\gamma)$, $\lim_{t\to\infty} \dot{h}(t) = 0$ and $\lim_{t\to\infty} \ddot{h}(t) = 0$, we find that

$$\lim_{t\to\infty} \dot{y}(t)^2 - (1 + y(t))\ddot{y}(t)^2 \sim \lim_{t\to\infty} \left(\frac{\chi}{\gamma(\chi+\gamma)} - 1\right) e^{2g_1 t}. \quad (150)$$

Hence, the asymptotic behavior depends on whether $\chi \lesseqgtr \chi(\chi+\gamma)$. Note that given that both χ and γ are on the order of $1/100$, we have that $\chi > \chi(\chi+\gamma)$, and hence s is asymptotically convex.[35] ∎

Proposition 7 *Suppose that the half-life of the system for levels and growth is reached for $\tau \in [T, \bar{T}]$. Then, the half-life of the system and the half-life of the growth rate for the Western countries can be approximated by*

$$t_{1/2}^{level} = d_o - \frac{\chi+\gamma}{g_a}(d_o - d_m) + \frac{1}{g_a} \ln\left(\frac{C_1\left(\frac{(\theta-1)^2}{\chi(\gamma+\chi)}\right)^{\theta-1}}{e^{a_o - a_m}}\right) \quad (151)$$

$$t_{1/2}^{growth} = d_0 + \frac{1}{\chi+\gamma+\gamma_a - g_0} \ln\left(\frac{\left(\frac{\chi+\gamma-2g_O}{\chi+\gamma+2g_a}\right)^{\theta-1} Y(d_0)}{\left(\frac{(\theta-1)^2}{\gamma(\chi+\gamma d_1 - a_1)}\right)^{\theta-1} e^{a_0}}\right) \quad (152)$$

(153)

where $Y(d_0)$ denotes the income level when Industrial Revolution technologies starts to be adopted and C is some positive constant.

Proof of Proposition 7: In this proof we will not provide the level of detail of the Proof of Proposition 3, as the derivations are analogous. We assume that the half life is achieved in the regime where $\tau < T$. We assume directly that the transient part $h(t) \simeq 1$ (which we only assumed in the end of 3. Under these assumptions, the definition of the half-life for the

[35] For example, in the baseline case, we have that $\chi = \gamma = 1\%$, so that the asymptotic behavior is convex $\frac{1}{100} > \frac{1}{100}\frac{2}{100}$. In fact, under the assumption that $\chi = \gamma$, the condition for an asymptotic convex behavior is that $\chi < 50\%$.

growth rate is

$$\left(\frac{X_M(t_{1/2})}{Y(t_{1/2})}\right)^{\frac{1}{\theta-1}} g_M(t_{1/2}) + \left(\frac{X_O(t_{1/2})}{Y(t_{1/2})}\right)^{\frac{1}{\theta-1}} g_O = \frac{\chi+\gamma}{2}, \tag{154}$$

$$2X_M(t_{1/2})^{\frac{1}{\theta-1}} g_M + 2X_O(t_{1/2})^{\frac{1}{\theta-1}} g_O \simeq (\chi+\gamma)Y(t_{1/2})^{\frac{1}{\theta-1}}, \tag{155}$$

$$(\chi+\gamma+2g_a)X_M(t_{1/2})^{\frac{1}{\theta-1}} = (\chi+\gamma-2g_O)X_O(t_{1/2})^{\frac{1}{\theta-1}}, \tag{156}$$

$$e^{a_0}\left(\frac{(\theta-1)^2}{\gamma(\chi+\gamma d_1 - a_1)}\right)^{\theta-1} e^{g_m(t-d_0)} = \left(\frac{\chi+\gamma-2g_O}{\chi+\gamma+2g_a}\right)^{\theta-1} Y(d_0)e^{g_O(t-d_0)} \tag{157}$$

Thus,

$$t_{1/2}^{\text{growth}} = d_0 + \frac{1}{g_m - g_0}\ln\left(\frac{\left(\frac{\chi+\gamma-2g_O}{\chi+\gamma+2g_a}\right)^{\theta-1} Y(d_0)}{\left(\frac{(\theta-1)^2}{\gamma(\chi+\gamma d_1-a_1)}\right)^{\theta-1} e^{a_0}}\right) \tag{158}$$

Next, we derive the half-life in levels. Define

$$\tilde{Y}(t) = \frac{Y(t)}{e^{a_m}\left(\frac{(\theta-1)^2}{\chi(\chi+\gamma)}\right)^{\theta-1} e^{(\chi+\gamma)(t-d_m)}}. \tag{159}$$

By construction, $\lim_{t\to\infty}\tilde{Y}(t) = 1$. Hence, the definition of the half-life is

$$\tilde{Y}(0) + \frac{1}{2}(1-\tilde{Y}(0)) = \tilde{Y}(t_{1/2}), \tag{160}$$

where we are taking the normalization $T = 0$. This is a transcendental equation, so to make further progress we assume that the transient part of growth is $h(t) \simeq 1$,[36] Denoting the left hand side of (159) by C, we can re-write as

$$C^{\frac{1}{\theta-1}} = \left(\frac{e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1} e^{(\chi+\gamma+g_a)(t-d_0)}}{e^{a_m}\left(\frac{(\theta-1)^2}{\chi(\chi+\gamma)}\right)^{\theta-1} e^{(\chi+\gamma)(t-d_m)}}\right)^{\frac{1}{\theta-1}} + \left(\frac{Y(d_0)e^{g_O(t-d_O)}}{e^{a_m}\left(\frac{(\theta-1)^2}{\chi(\chi+\gamma)}\right)^{\theta-1} e^{(\chi+\gamma)(t-d_m)}}\right)^{\frac{1}{\theta-1}} \tag{161}$$

$$C^{\frac{1}{\theta-1}} = \frac{e^{\frac{g_a(t-d_o)-(\chi+\gamma)(d_o-d_m)}{\theta-1}}}{\frac{(\theta-1)^2}{\chi(\gamma+\chi)}e^{\frac{(a_m-a_o)}{\theta-1}}}\left(\frac{(\theta-1)^2}{\gamma} + \left(e^{-a_0}Y(d_0)e^{(g_O-\chi-\gamma-g_a)(t-d_O)}\right)^{\frac{1}{\theta-1}}\right). \tag{162}$$

To solve for this equation analytically, we approximate the second term in parenthesis for its average value. Using the following notation

$$\kappa = \frac{(\theta-1)^2}{\gamma} + < \left(e^{-a_0}Y(d_0)e^{(g_O-\chi-\gamma-g_a)(t-d_O)}\right)^{\frac{1}{\theta-1}} > \tag{163}$$

[36] One could do as in the proof of Proposition 3 and take the average values.

t can be expressed as

$$t_{1/2}^{\text{level}} = d_o - \frac{\chi + \gamma}{g_a}(d_o - d_m) + \frac{1}{g_a} \ln \left(\frac{C \left(\frac{(\theta-1)^2}{\kappa\chi(\gamma+\chi)} \right)^{\theta-1}}{e^{a_o - a_m}} \right) \tag{164}$$

D Data

The twenty-five particular technology measures, organized by broad category, that we consider are:

1. **Steam and motor ships:** Gross tonnage (above a minimum weight) of steam and motor ships in use at midyear. *Invention year:* 1788; the year the first (U.S.) patent was issued for a steam boat design.

2. **Railways - Passengers:** Passenger journeys by railway in passenger-KM.
   Invention year: 1825; the year of the first regularly schedule railroad service to carry both goods and passengers.

3. **Railways - Freight:** Metric tons of freight carried on railways (excluding livestock and passenger baggage).
   Invention year: 1825; same as passenger railways.

4. **Cars:** Number of passenger cars (excluding tractors and similar vehicles) in use. *Invention year:* 1885; the year Gottlieb Daimler built the first vehicle powered by an internal combustion engine.

5. **Trucks:** Number of commercial vehicles, typically including buses and taxis (excluding tractors and similar vehicles), in use. *Invention year:* 1885; same as cars.

6. **Tractor:** Number of wheel and crawler tractors (excluding garden tractors) used in agriculture. *Invention year:* 1892; John Froelich invented and built the first gasoline/petrol-powered tractor.

7. **Aviation - Passengers:** Civil aviation passenger-KM traveled on scheduled services by companies registered in the country concerned. *Invention year:* 1903; The year the Wright brothers managed the first successful flight.

8. **Aviation - Freight:** Civil aviation ton-KM of cargo carried on scheduled services by companies registered in the country concerned. *Invention year:* 1903; same as aviation - passengers.

9. **Telegraph:** Number of telegrams sent. *Invention year:* 1835; year of invention of telegraph by Samuel Morse at New York University.

10. **Mail:** Number of items mailed/received, with internal items counted once and crossâ€border items counted once for each country. *Invention year:* 1840; the first modern postage stamp, Penny Black, was released in Great Britain.

11. **Telephone:** Number of mainline telephone lines connecting a customer's equipment to the public switched telephone network. *Invention year:* 1876; year of invention of telephone by Alexander Graham Bell.

12. **Cellphone:** Number of users of portable cell phones. *Invention year:* 1973; first call from a portable cellphone.

13. **Personal computers:** Number of self-contained computers designed for use by one person. *Invention year:* 1973; first computer based on a microprocessor.

14. **Internet users:** Number of people with access to the worldwide network. *Invention year:* 1983; introduction of TCP/IP protocol.

15. **Spindles:** Number of mule and ring spindles in place at year end. *Invention year:* 1779; Spinning Mule invented by Samuel Crompton.

16. **Synthetic Fiber:** Weight of synthetic (noncellulosic) fibers used in spindles *Invention year:* 1924; Invention of rayon.

17. **Steel:** Total tons of crude steel production (in metric tons). This measure includes steel produced using Bessemer and Open Earth furnaces. *Invention year:* 1855; William Kelly receives the first patent for a steel making process (pneumatic steel making).

18. **Electric Arc Furnaces:** Crude steel production (in metric tons) using electric arc furnaces. *Invention year:* 1907; invention of the Electric Arc Furnace.

19. **Blast Oxygen Furnaces:** Crude steel production (in metric tons) in blast oxygen furnaces (a process that replaced Bessemer and OHF processes). *Invention year:* 1950; invention of Blast Oxygen Furnace.

20. **Electricity:** Gross output of electric energy (inclusive of electricity consumed in power stations) in Kw-Hr. *Invention year:* 1882; first commercial power station on Pearl Street in New York City.

21. **Fertilizer:** Metric tons of fertilizer consumed. Aggregate of 25 individual types, corresponding to broadly Ammonia and Phosphates. *Invention year:* 1910; Haber-Bosch process to produce ammonia is patented in 1910.

22. **Harvester:** Number of selfpropelled machines that reap and thresh in one operation. *Invention year:* 1912; The Holt Manufacturing Company of California produces a self-propelled harvester. Subsequently, a selfpropelled machine that reaps and threshes in one operation appears.

23. **Kidney Transplant:** Number of kidney transplants performed. *Invention year:* 1954; Joseph E. Murray and his colleagues at Peter Bent Brigham Hospital in Boston performed the first successful kidney transplant.

24. **Liver Transplant:** Number of liver transplants performed. *Invention year:* 1963; Dr. Thomas Starzl performs the first successful liver transplant in the United States.

25. **Heart Transplant:** Number of heart transplants performed *Invention year:* 1968; Adrian Kantrowitz performed the first pediatric heart transplant in the world on December 6, 1967 at Maimonides Hospital.

E Tables

Technology	Invention Year	Total	Implausible	Imprecise	Precise	% Precise	Detrended R^2		
							$R^2 > 0$	Mean	SD
Spindles	1779	34	5	1	28	82	22	0.66	0.23
Ships	1788	61	17	3	41	67	41	0.81	0.17
Railways Freight	1825	85	40	6	39	46	39	0.80	0.13
Railways Passengers	1825	80	46	5	29	36	29	0.88	0.08
Telegraph	1835	62	19	5	38	61	29	0.60	0.24
Mail	1840	67	20	4	43	64	43	0.89	0.08
Steel (Bessemer, Open Hearth)	1855	52	9	4	39	75	39	0.75	0.17
Telephone	1876	139	84	6	49	35	48	0.89	0.15
Electricity	1882	134	52	8	74	55	74	0.91	0.12
Cars	1885	124	54	9	61	49	61	0.77	0.21
Trucks	1885	108	46	8	54	50	54	0.80	0.20
Tractor	1892	135	16	5	114	84	109	0.71	0.18
Aviation Freight	1903	93	50	7	36	39	36	0.90	0.09
Aviation Passengers	1903	96	52	4	40	42	40	0.90	0.07
Electric Arc Furnace	1907	75	27	9	39	52	37	0.65	0.25
Fertilizer	1910	132	39	8	85	64	74	0.63	0.24
Harvester	1912	104	41	7	56	54	48	0.69	0.22
Synthetic Fiber	1924	49	1	0	48	98	46	0.65	0.27
Blast Oxygen Furnace	1950	49	10	2	37	76	30	0.62	0.29
Kidney Transplant	1954	27	3	0	24	89	24	0.82	0.17
Liver Transplant	1963	21	0	1	20	95	19	0.83	0.13
Heart Surgery	1968	18	0	1	17	94	16	0.65	0.19
Cellphones	1973	84	2	3	79	94	79	0.91	0.07
PCs	1973	70	4	0	66	94	66	0.94	0.06
Internet	1983	59	1	1	57	97	57	0.96	0.04
		1958	638	107	1213	62	1160	0.80	0.20

Table 5: Quality of the Estimates

		Time Period	
		1820-1900	1900-2000
Simulation	Western Countries	.69%	2.16%
	Non-Western Countries	.27%	1.45%
	Difference	.42%	.71%
Maddison	Western Countries	1.1%	2.0%
	Rest	.4%	1.2%
	Difference West-Rest	.7%	.8%

Table 6: Growth rates of GDP per capita. Simulation results and growth rates from Maddison (2004)

F Simulation and Calibration of the Production Function in Section 5

The simulations reported in the paper are a discrete time model, in which at each period of time a new technology and a new vintage of all technologies appear. We run first the model forward to reach a BGP at a growth rate of .2% holding the adoption margins constant. We normalize the output at the initial point of the initial revolution so that it is equal for all countries that we simulate. Then we start reducing the adoption lag and the intensive margin for the new technologies that appear, holding the evolution of the productivity of the "Old" ones at pre-Industrial levels (as explained in the model).

The simulations reported in the paper normalize the output produced using the pre-Modern technologies as follows. The production function we are working with could be written in general as

$$Y(t) = C\left[(\xi_{Old}A_{Old}(t))^{\frac{1}{\theta-1}} + (\xi_{Modern}A_{Modern}(t))^{\frac{1}{\theta-1}}\right]^{\theta-1}. \tag{165}$$

We normalize $\xi_{Modern} = 1$, as the presumption is that we have a good description of the Modern growth process,

$$Y(t) = C\left[(\xi_{Old}A_{Old}(t))^{\frac{1}{\theta-1}} + A_{Modern}(t)^{\frac{1}{\theta-1}}\right]^{\theta-1}. \tag{166}$$

Next, we use the fact that for the Western countries we know much income grew since the our "start" of the industrial revolution (1765) –we interpolate output growth from Maddison (2004) to obtain this number– up to an arbitrary date $\bar{T}$, which we take to be 1820. Denote by α the ratio $Y(\bar{T})/Y(1765)$. Then, to match the increase in relative output, we have that

$$\xi_{Old} = \frac{A_{Modern}(\bar{T})}{\left[(\alpha A_{Old}(1765))^{\frac{1}{\theta-1}} - A_{Old}(\bar{T})^{\frac{1}{\theta-1}}\right]^{\theta-1}}. \tag{167}$$

Finally, to match the level of output at the "start" of the industrial revolution, we have that

$$C = \frac{Y(1765)}{\xi_{Old} A_{Old}(1765)}. \tag{168}$$

Note that the dynamics are independent of the term C. So for our purposes, we can re-scale the output at the start of the Industrial Revolution to 1.

Given that we simulate the evolution of a system to its initial pre-Modern growth rate, the level of output at a given point of time is different depending on the levels of adoption. We can rewrite the production function as $\left[(C\xi_{Old} A_{Old}(t))^{\frac{1}{\theta-1}} + (CA_{Modern}(t))^{\frac{1}{\theta-1}}\right]^{\theta-1}$. Hence, the normalizing initial term $C\xi_{Old}$ is different for each of our countries. Indeed, the term ξ_{Old} is, on the contrary technology specific, as it contains information on the relative productivity of pre-Industrial to Modern technologies. In other words, the ratio of $C\xi_{Old}$ to C Hence, once we pin it down for the Western countries from equation (167), we use it in the rest of the simulations.

G Figures

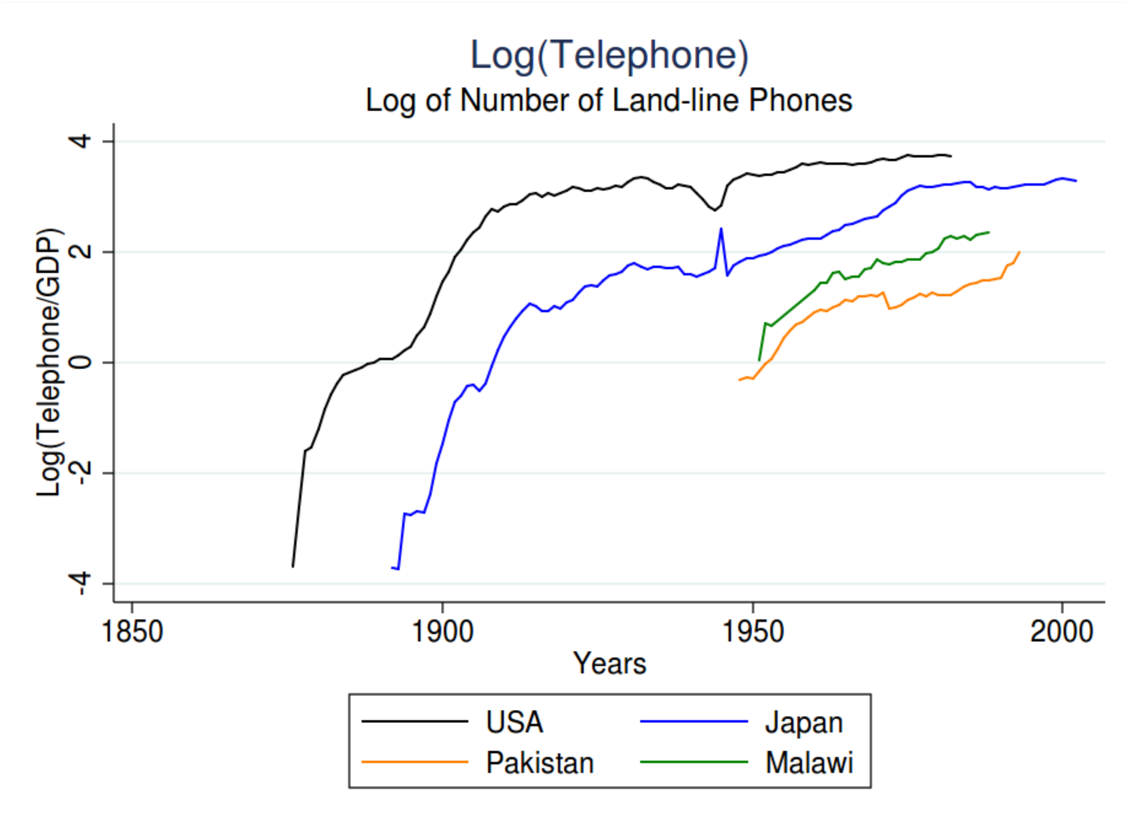


Figure 3: Adoption of Telephone for four countries.

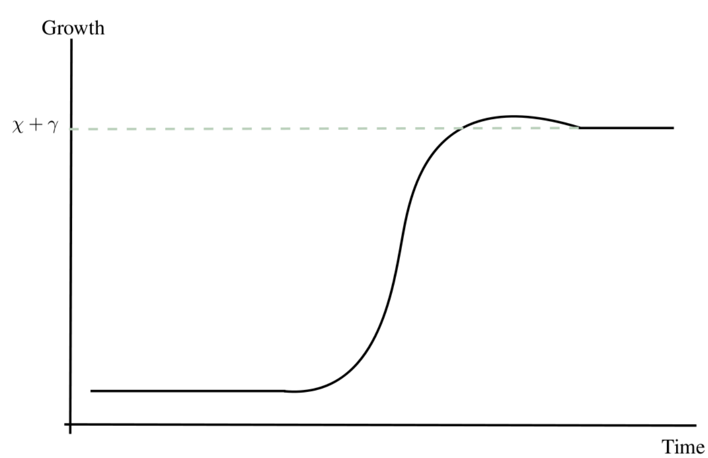
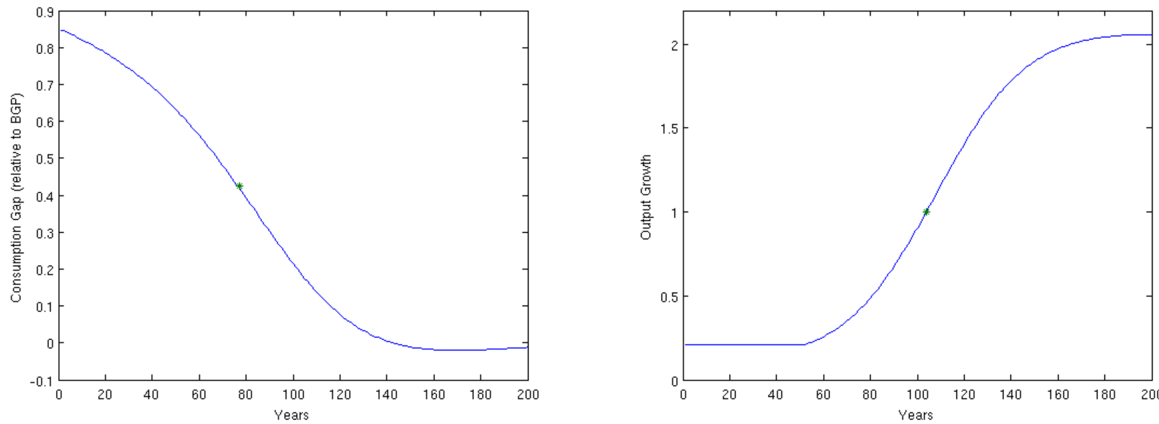


Figure 4: Generic Evolution of the growth rate in the transition to Modern growth.

Figure 5: Slow transitional dynamics.

(a) Consumption gap relative to the Modern BGP (b) Growth path to Modern BGP



This simulation corresponds to the transition to the new balanced growth path after an acceleration of the technological frontier from .2% to 2% for a country with a constant lag as the average lag in our sample (50 years) and average intensive margin (40% of the U.S. productivity level). The star * denotes the half-life.

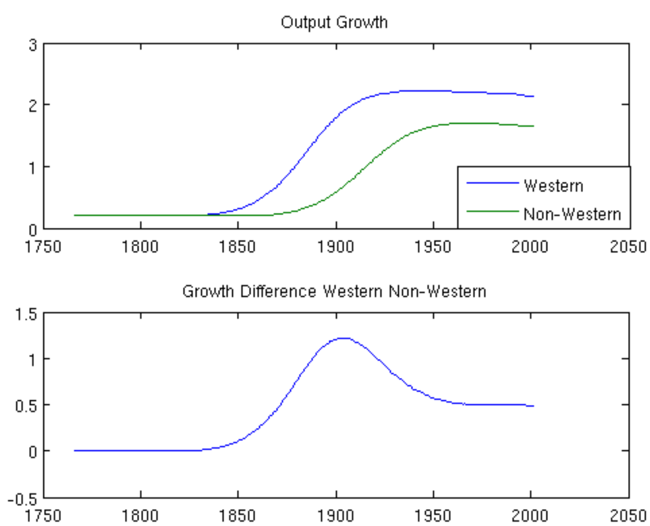


Figure 6: Growth of Western and non-Western countries imputing the estimated evolution of the intensive and extensive margins.

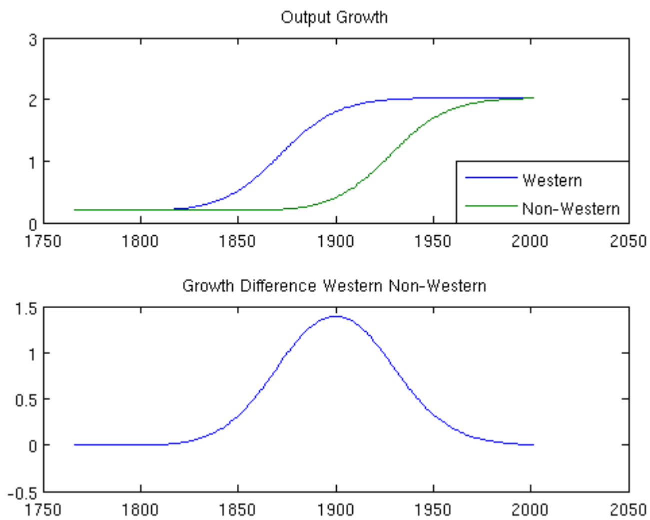
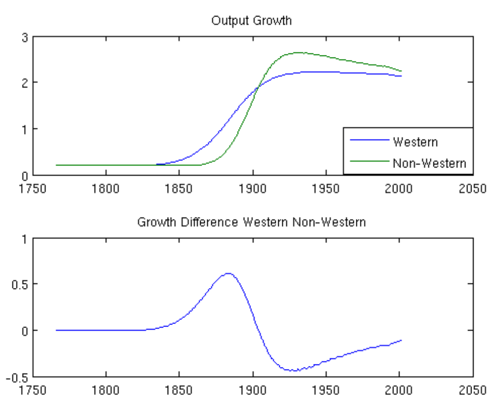
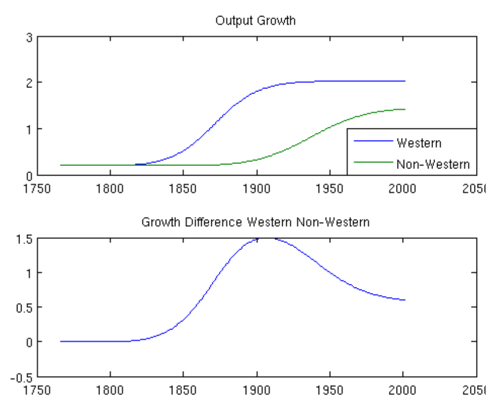


Figure 7: Growth of Western and non-Western countries with *only* an acceleration of the technology frontier. Both margins of adoption are held constant.

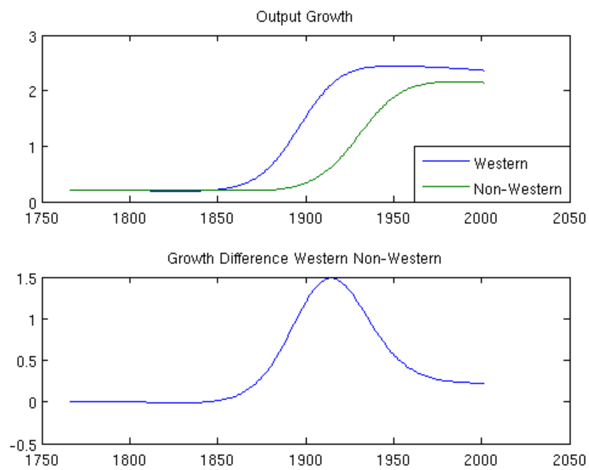


(a) Dynamics due only to a decline in lags.

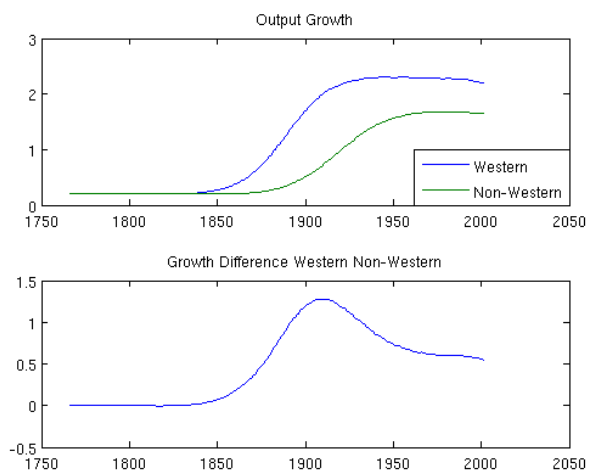


(b) Dynamics due to the divergence in the intensive margin.

Figure 8: Role played by the different margins of adoption.



(a) Dynamics with productivity gains from new varieties only.



(b) Dynamics with productivity gains from new technologies only.

Figure 9: Role played by the different margins of productivity gains