If Technology Has Arrived Everywhere, Why Has Income Diverged?

Diego Comin
Harvard University and NBER

Martí Mestieri
Toulouse School of Economics

May 3, 2013*

Abstract

We study the lags with which new technologies are adopted across countries, and their long-run penetration rates once they are adopted. Using data from the last two centuries, we document two new facts: there has been convergence in adoption lags between rich and poor countries, while there has been divergence in penetration rates. Using a model of adoption and growth, we show that these changes in the pattern of technology diffusion account for 80% of the Great Income Divergence between rich and poor countries since 1820.

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

*We are grateful to Thomas Chaney, Christian Hellwig, Chad Jones, Pete Klenow, Franck Portier, Mar Reguant and seminar participants at Boston University, Brown University, CERGE-EI, Edinburgh, HBS, LBS, Minnesota Fed, SITE, Stanford GSB, UAB, Univeristy of Toronto and TSE for useful comments and suggestions. Comin acknowledges the generous support of INET. All remaining errors are our own.

Comin: dcomin@hbs.edu, Mestieri: marti.mestieri@tse-fr.eu

1 Introduction

We have very limited knowledge about the drivers of growth over long periods of time. Klenow and Rodríguez-Clare (1997) show that factor accumulation (physical and human capital) accounts only for 10% of cross-country differences in productivity growth between 1960 and 1985. Clark and Feenstra (2003) find similar results for the period 1850-2000. What accounts for the bulk of growth dynamics over the long term, and why do these drivers differ across countries?

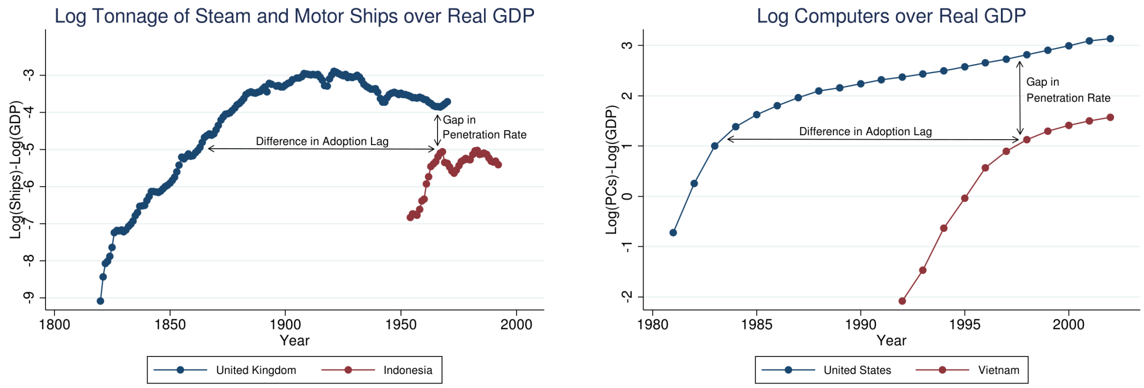
This paper explores whether the dynamics of technology can help us account for the cross-country evolution of productivity and income growth over the last 200 years. We are particularly interested in understanding if the technology channel can account for the dramatic increase in cross-country differences in per-capita income over the last 200 years, a phenomenon known as the Great Divergence (e.g., Pritchett, 1997 and Pomeranz, 2000).

The strategy we follow to study the role of technology on income dynamics has two parts. The first part consists in exploring the evolution of technology adoption patterns. Using a comprehensive dataset with direct measures of the diffusion of the large number of technologies, we study how various dimensions of technology adoption have evolved over time in different groups of countries. With this analysis, we uncover some general cross-country trends in technology adoption. Second, we study the implications of these adoption trends for the evolution of income growth across-countries over the last 200 years.

Technology has probably arrived everywhere. Comin and Hobijn (2010) find that the lags with which new technologies arrive to countries have dropped dramatically over the last 200 years. 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.[1] This evidence would imply that adoption lags have converged across countries. But, if technology has arrived everywhere, why has income diverged over the last two centuries?

To explore this puzzle, we recognize that the contribution of technology to a country's productivity growth can be decomposed in two parts. One part is related to the range of technologies used, or equivalently to the lag with which they are adopted. New technologies embody higher productivity. Therefore, an acceleration in the rate at which new technologies arrive in the country raises aggregate productivity growth. Second, productivity is also affected by the penetration rate of new technologies. The more units of any new technology

[1] See Khalba (2007) and Dholakia and Kshetri (2003).



(a) Diffusion of Steam and Motor ships for the UK and Indonesia.

(b) Diffusion of PCs for the US and Vietnam.

Figure 1: Examples of diffusion curves.

(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.[2] Thus, increases in the penetration of technology (or as we call it below, the intensive margin of adoption) also raise the growth rate of productivity.

To identify adoption lags (extensive margin) and penetration rates (intensive margin) of technology, we follow a strategy similar to Comin and Hobijn (2010) and Comin and Mestieri (2010). To illustrate our strategy, consider Figures 1a and 1b 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, for a given technology, the diffusion curves for different countries have similar shapes, but displaced vertically and horizontally. Comin and Hobijn (2010) show that this property holds generally for a large majority of the technology-country pairs. 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 a model of technology adoption and growth. Crucial for our purposes, the model provides a unified framework for measuring the diffusion of specific technologies and assessing their impact on income 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. This allows us to take these predictions to the data and estimate adoption lags and penetration rates fitting the diffusion curves derived

[2]In our context, this is isomorphic to differences in the efficiency with which producers use technology.

from our model.

Using the CHAT data set,[3] we identify the extensive and intensive adoption margins for 25 significant technologies invented over the last 200 years in an (unbalanced) sample that covers 132 countries. Then, we use our estimates to study the cross-country evolution of the two adoption margins. We uncover two new empirical regularities. First, cross-country differences in adoption lags have narrowed over the last 200 years. That is, adoption lags have declined more in poor/slow adopter countries than in rich/fast adopter countries. Second, the gap in penetration rates between rich and poor countries has widened over the last 200 years, inducing a divergence in the intensive margin of technology adoption. Figure 1 illustrates these patterns. The horizontal gap between the diffusion curves for steam and motor ships in the UK and Indonesia is much larger than the horizontal gap between the U.S. and Vietnam for computers (131 years vs. 11 years). In contrast, 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).

After characterizing the dynamics of technology, we explore their consequences for the cross-country dynamics of income both analytically and with simulations. Taking advantage of the simple aggregate representation of our model, we derive two analytical results. *(i)* The dynamics of technology adoption in our model generate S-shaped transitions for the growth rate of productivity. *(ii)* Simple approximate expressions for the half-life of the system are derived. Despite not having physical capital, habit formation or other mechanisms to generate slow transitions, half-lifes are an order of magnitude larger than in the neoclassical model when evaluated at our estimated parameter values.[4]

We also use simulations to evaluate quantitatively the model's predictions for the cross-country income dynamics. In particular, we simulate the dynamics of income in two representative economies (one "advanced" and one "developing"). After feeding in the dynamics of technology adoption we have uncovered in the data, 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 modern 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 a 3.2-fold increase in the income gap between rich and developing countries, which represents 80% of the actual fourth-fold increase observed over the last two centuries. The model also does well in reproducing the income gap between rich and developing countries circa 1820, and the observed growth dynamics for the countries in the bottom quarter and tenth of the world income distribution,

[3]See Comin and Hobijn (2009) for a description of the data set. Comin and Hobijn (2004) and Comin *et al.* (2008) have used it in alternative set-ups.

[4]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 120 years. A standard measure of the half-life for the neoclassical growth model is 14 years.

and for the different continents.

It is important to emphasize that, when evaluating the role of technology for cross-country differences in income, we take into account that income affects demand for goods and services that embody new technologies. In our baseline model, the restriction that our model has a balanced growth path implies that the income elasticity of technology demand is equal to one. Because this implication of the model may be restrictive, we study the robustness of our findings to allowing for non-homotheticities in the demand for technology. That is, we allow the income elasticity of technology demand to differ from one. Taking advantage of the three-dimensional nature of our data set, we identify this elasticity from the time variation in income and technology adoption in three countries (U.S., UK and France) for which we have the longest time series for technology.[5]

Our findings are robust to allowing for non-homotheticities. In particular, *(i)* we obtain very similar trends for the cross-country evolution of adoption lags and the intensive margin, and *(ii)* we observe that the productivity gap between rich and developing countries increases by a factor of 2.6 over the last 200 years, which accounts for 67% of the Great Divergence.

This paper is related to the literature that has explored the drivers of the Great Divergence. One stream of the 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 of 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 skill-biased (Mokyr, 2002), contributed to the increase in income differences across countries, as Western countries benefited relatively more from them. 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). Yet, incomes diverged also during this period. Second, trade globalization ended abruptly in 1913. With WWI, world trade dropped and did not reach the pre-1913 levels until the 1970s. In contrast, the Great Divergence continued throughout the twentieth century.

Probably motivated by these observations, another strand of the literature has studied the cross-country evolution of Solow residuals and has found that they account for the majority of the divergence (Easterly and Levine, 2002, and Clark and Feenstra, 2003). Our paper takes a strong stand on the nature of the Solow residuals over protracted periods of time –technology–, measures it directly, and shows the direct importance of technology dynamics for cross-country income dynamics.[6]

[5] We allow the income elasticity of technology demand to vary across technologies according to their invention date. Specifically, we estimate common income elasticities for technologies invented in fifty year intervals (i.e., pre-1850, 1850-1900, 1900-1950 and post-1950).

[6] Our analysis is also related to a strand of the literature that has studied the productivity dynamics after

Finally, our paper makes three contributions to the analysis of technology dynamics. First, building on the class of models and estimation developed by Comin and Hobijn (2010) and Comin and Mestieri (2010), we are the first to document the evolution of adoption margins across countries. In particular, we uncover the convergence in adoption lags and the divergence in the intensive margin of adoption in the last 200 years. Second, this is the first paper that studies analytically transitional dynamics within this class of models. Third, this paper evaluates quantitatively the role of technology dynamics in shaping cross-country income dynamics. It shows that the transitional dynamics generated by the model are very protracted and that they play a key role in generating the Great Divergence. In contrast, previous quantitative analyses only explored how technology levels affected cross-country steady-state income levels.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 estimates the extensive and intensive margins of adoption and documents the cross-country evolution of both adoption margins. Section 4 characterizes key features of the model transitional dynamics. Section 5 simulates the model to quantify the effect of the technology dynamics on the cross-country growth dynamics. Section 6 conducts robustness checks, and Section 7 concludes.

2 Model

We present a simple model of technology adoption and growth. Our model serves four purposes. First, it precisely defines the intensive and extensive margins of adoption. Second, it illustrates how variation in these margins affects the evolution of the diffusion curves for individual technologies. Third, it helps develop the identification strategy of the extensive and intensive margins of adoption in the data. Fourth, because ours is a general equilibrium model with a simple aggregate representation, it can be used to study the dynamics of productivity growth.

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

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

the Industrial Revolution. Crafts (1997), Galor and Weil (2000), Hansen and Prescott (2002), Tamura (2002), among others, provide different reasons why there was a slow growth acceleration in productivity after the Industrial Revolution. The mechanisms in these papers are complementary to ours.

where ρ denotes the discount rate and C_t, consumption at time t. The representative household, maximizes its utility subject to the budget constraint (2) and a no-Ponzi scheme 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_t denotes the bond holdings of the representative consumer, $\dot{B}_t$ is the increase in bond holdings over an instant of time, and r_t the 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. To simplify notation, we omit time subscripts, t, whenever possible. Each instant, a new technology, τ, exogenously appears. We denote a technology by the time it was invented. Therefore, the range of invented technologies at time t is $(-\infty, t]$.

For each existing technology, a new, more productive, vintage appears in the world frontier every instant. We denote vintages of technology-τ generically by v_τ. Vintages are indexed by the time in which they appear. Thus, the set of existing vintages of technology-τ available at time $t(> \tau)$ is $[\tau, t]$. The productivity of a technology-vintage pair has two components. The first component, $Z(\tau, v_\tau)$, is common across countries and it is purely determined by technological attributes. In particular,

$$\begin{aligned} Z(\tau, v) &= e^{(\chi+\gamma)\tau+\gamma(v_\tau-\tau)} && (4) \\ &= e^{\chi\tau+\gamma v_\tau}, && (5) \end{aligned}$$

where $(\chi + \gamma)\tau$ is the productivity level associated with the first vintage of technology τ and $\gamma(v_\tau - \tau)$ captures the productivity gains associated with the introduction of new vintages $v_\tau \geq \tau$.[7]

The second component is a technology-country specific productivity term, a_τ, which we further discuss below.

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*.[8] The set of technology-τ

[7] In what follows, whenever there is no confusion, we omit the subscript τ from the vintage notation and simply write v.

[8] Adoption lags may result from a cost of adopting the technology in the country that is decreasing in the proportion of not-yet-adopted technologies as in Barro and Sala-i-Martin (1997), or in the gap between aggregate productivity and the productivity of the technology, as in Comin and Hobijn (2010).

vintages available in this economy is $V_\tau = [\tau, t - D_\tau]$.[9] Note that D_τ is both the time it takes for an economy to start using technology τ and its distance to the technology frontier in technology τ.

Intensive margin.– New vintages (τ, v) are incorporated into production through new intermediate goods that embody them. Intermediate goods are produced competitively using one unit of final output to produce one unit of intermediate good.

Intermediate goods are combined with labor to produce the output associated with a given vintage, $Y_{\tau,v}$. 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, $Y_{\tau,v}$ is given by

$$Y_{\tau,v} = a_\tau Z(\tau, v) X_{\tau,v}^{\alpha} L_{\tau,v}^{1-\alpha}. \tag{6}$$

The term a_τ in (6) represents factors that reduce the effectiveness of a technology in a country. These may include differences in the costs of producing the intermediate goods associated with a 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.[10] As we shall see below, a_τ determines the long-run penetration rate of the technology in the country. Hence, we refer to a_τ as the *intensive margin* of adoption of a technology.

The goal of the paper is to measure these two adoption margins in the data and then study how they affect productivity growth. The nature of the drivers of adoption of the equilibrium adoption margins is irrelevant for this goal. Therefore, we can simplify the analysis by treating these margins of adoption as exogenous parameters.[11]

Production.– The output associated with different vintages of the same technology can be combined to produce competitively sectoral output, $Y_{\tau,}$ as follows

$$Y_\tau = \left(\int_\tau^{t-D_\tau} Y_{\tau,v}^{\frac{1}{\mu}} \, dv \right)^{\mu}, \qquad \text{with} \quad \mu > 1. \tag{7}$$

Similarly, final output, Y, results from aggregating competitively sectoral outputs Y_τ as follows

$$Y = \left(\int_{-\infty}^{\bar{\tau}} Y_\tau^{\frac{1}{\theta}} d\tau \right)^{\theta}, \qquad \text{with} \quad \theta > 1. \tag{8}$$

[9] 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. We do not impose this condition when we simulate the model in Section 5.

[10] Comin and Mestieri (2010) discuss how a wide variety of distortions result in wedges in technology adoption that imply a reduced form as in (6).

[11] See Comin and Mestieri (2010) and Comin and Mestieri (2010) for ways to endogenize these adoption margins as equilibrium outcomes.

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

2.3 Factor Demands and Final Output

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

$$Y_\tau = Y p_\tau^{-\frac{\theta}{\theta-1}}, \tag{9}$$

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

$$Y_{\tau,v} = Y_\tau \left(\frac{p_\tau}{p_{\tau,v}}\right)^{-\frac{\mu}{\mu-1}}, \tag{10}$$

where $p_{\tau,v}$ denotes the price of the (τ, v) intermediate good.[12] The demands for labor and intermediate goods at the vintage level are

$$(1-\alpha)\frac{p_{\tau,v}Y_{\tau,v}}{L_{\tau,v}} = w, \tag{11}$$

$$\alpha\frac{p_{\tau,v}Y_{\tau,v}}{X_{\tau,v}} = 1. \tag{12}$$

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

$$p_{\tau,v} = \frac{w^{1-\alpha}}{Z(\tau,v)a_\tau}(1-\alpha)^{-(1-\alpha)}\alpha^{-\alpha}. \tag{13}$$

Combining (10), (11) and (12), the total output produced with technology τ can be expressed as

$$Y_\tau = Z_\tau L_\tau^{1-\alpha} X_\tau^\alpha, \tag{14}$$

where L_τ denotes the total labor used in sector τ, $L_\tau = \int_\tau^{t-D_\tau} L_{\tau,v}dv$, and X_τ is the total amount of intermediate goods in sector τ, $X_\tau = \int_\tau^{t-D_\tau} X_{\tau,v}dv$. The productivity associated

[12] Even though older technology-vintage pairs are always produced in equilibrium, the value of its production relative to total output is declining over time.

to a technology is

$$
\begin{aligned}
Z_\tau &= \left(\int_\tau^{\max\{t-D_\tau,\tau\}} Z(\tau,v)^{\frac{1}{\mu-1}} dv\right)^{\mu-1} \\
&= \left(\frac{\mu-1}{\gamma}\right)^{\mu-1} \underbrace{a_\tau}_{\text{Intensive Mg}} \underbrace{e^{(\chi\tau+\gamma\max\{t-D_\tau,\tau\})}}_{\text{Embodiment Effect}} \underbrace{\left(1-e^{\frac{-\gamma}{\mu-1}(\max\{t-D_\tau,\tau\}-\tau)}\right)^{\mu-1}}_{\text{Variety Effect}}. \qquad (15)
\end{aligned}
$$

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 more varieties are used in production and the higher Z_τ is.

The price index of technology-τ output is

$$
\begin{aligned}
p_\tau &= \left(\int_\tau^{t-D_\tau} p_{\tau,v}^{-\frac{1}{\mu-1}} dv\right)^{-(\mu-1)} \\
&= \frac{w^{1-\alpha}}{Z_\tau}(1-\alpha)^{-(1-\alpha)}\alpha^{-\alpha}. \qquad (16)
\end{aligned}
$$

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

$$
Y = AX^\alpha L^{1-\alpha} = AX^\alpha = A^{1/(1-\alpha)}(\alpha)^{\alpha/(1-\alpha)}, \qquad (17)
$$

with

$$
A = \left(\int_{-\infty}^{\bar{\tau}} Z_\tau^{\frac{1}{\theta-1}} d\tau\right)^{\theta-1}, \qquad (18)
$$

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

2.4 Equilibrium

Given a sequence of adoption lags and intensive margins $\{D_\tau, a(\tau)\}_{\tau=-\infty}^{\infty}$, a competitive equilibrium in this economy is defined by consumption, output, and labor allocations paths $\{C_t, L_{\tau,v}(t), Y_{\tau,v}(t)\}_{t=t_0}^{\infty}$ and prices $\{p_\tau(t), p_{\tau,v}(t), w_t, r_t\}_{t=t_0}^{\infty}$, such that

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

$$
\frac{\dot{C}}{C} = r - \rho, \qquad (19)
$$

satisfying the budget constraint (2) and (3).

2. Firms maximize profits taking prices as given (equation 13). This optimality condition gives the demand for labor and intermediate goods for each technology and vintage, equations (11) and (12), for the output produced with a vintage (equation 10) and for the output produced with a technology (equation 9).

3. Labor market clears

$$L = \int_{-\infty}^{\bar{\tau}} \int_{\tau}^{\bar{v}_\tau} L_{\tau,v} dv d\tau = 1, \tag{20}$$

where $\bar{v}_\tau$ denotes the last adopted vintage of technology τ.

4. The resource constraint holds:

$$Y = C + X, \tag{21}$$

$$C = (1-\alpha)Y. \tag{22}$$

Combining (20) and (11), it follows that the wage rate is given by

$$w = (1-\alpha)Y/L. \tag{23}$$

Combining the Euler equation (19) and the resource constraint (22) we obtain that the interest rate depends on output growth and the discount rate $r = \frac{\dot{Y}}{Y} + \rho$.

Equation (17) 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 study the growth rate of the economy along the balanced growth path. A sufficient condition to guarantee its existence, which we take as a benchmark, is when D_τ and a_τ are constant across technologies.[13] In the case that we make the simplifying (and empirically relevant) assumption that $\theta = \mu$, aggregate productivity can be computed in closed form.[14] Omitting technology subscripts, we find that

$$A = \left(\frac{(\theta-1)^2}{(\gamma+\chi)\chi} \right)^{\theta-1} a \, e^{(\chi+\gamma)(t-D)}. \tag{24}$$

Naturally, a higher intensity of adoption, a, and shorter adoption lags, D, lead to higher aggregate productivity. Along this balanced growth path, productivity grows at rate $\chi + \gamma$ and output grows at rate $(\chi+\gamma)/(1-\alpha)$.[15]

[13] Comin and Mestieri (2010) show in their microfounded models of adoption that this is a necessary and sufficient condition.

[14] As we discuss below, this is what we observe in our estimation.

[15] For utility to be bounded, this requires the parametric assumption that $(\chi+\gamma)/(1-\alpha) < \rho$.

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. This is not the main goal of the paper, as a similar exercise has already been done (albeit with less technologies) in Comin and Hobijn (2010) and Comin and Mestieri (2010). Then, we explore the novel question of whether there are any significant trends in the evolution of these adoption margins across countries.

3.1 Estimation strategy

We derive our estimating equation by combining the demand for sector τ output, (9), the sectoral price deflator (16), the expression for the equilibrium wage rate (23), and the expression for Z_τ, (15). Taking logs we obtain

$$y_\tau = y + \frac{\theta}{\theta - 1}\left[z_\tau - (1-\alpha)\,(y - l)\right], \tag{25}$$

where lowercase letters denote logs.

From expression (15) we see that, to a first order approximation, γ only affects y_τ through the linear trend. As we show in the Appendix C, this allows us to do a second-order approximation of $\log Z_\tau$ around the starting adoption date $\tau + D_\tau$ as

$$z_\tau \simeq \ln a_\tau + (\chi + \gamma)\tau + (\mu - 1)\ln(t - \tau - D_\tau) + \frac{\gamma}{2}(t - \tau - D_\tau)\,. \tag{26}$$

Substituting (26) in (25) gives us the following estimating equation

$$y^c_{\tau t} = \beta^c_{\tau 1} + y^c_t + \beta_{\tau 2} t + \beta_{\tau 3}\left((\mu - 1)\ln(t - D^c_\tau - \tau) - (1-\alpha)(y^c_t - l^c_t)\right) + \varepsilon^c_{\tau t}, \tag{27}$$

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

$$\beta^c_{\tau 1} = \beta_{\tau 3}\left(\ln a^c_\tau + \left(\chi + \frac{\gamma}{2}\right)\tau - \frac{\gamma}{2} D^c_\tau\right). \tag{28}$$

Equation (27) shows that the adoption lag D^c_τ is the only determinant of horizontal shifts in the curvature of the diffusion curve. Intuitively, longer lags imply 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. Equation (28) shows 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_τ^c. A lower level of a_τ^c 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.

We can identify differences in the intensive margin relative to a benchmark. For consistency with our simulations below, we take as benchmark the average value for the 17 developed countries that Maddison (2004) defines as Western countries.[16] Formally, we idnetify the intensive margin as:

$$\ln a_\tau^c = \frac{\beta_{1,\tau}^c - \beta_{1,\tau}^{Western}}{\beta_{3,\tau}} + \frac{\gamma}{2}(D_\tau^c - D_\tau^{Western}). \tag{29}$$

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 (27) 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. To derive an estimating equation for these measures, we integrate (12) across vintages to obtain (in logs) $x_\tau^c = y_\tau^c + p_\tau^c + \ln\alpha$. Substituting in for equation (27), we obtain the following expression which we use to estimate the diffusion of the inputs that embody technology[17]

$$x_{\tau t}^c = \beta_{\tau 1}^c + y_t^c + \beta_{\tau 2}t + \beta_{\tau 3}\left((\mu - 1)\ln(t - D_\tau^c - \tau) - (1-\alpha)(y_t^c - l_t^c)\right) + \varepsilon_{\tau t}^c. \tag{30}$$

The procedure we use to estimate (27) and (30) consists in two parts. For each technology, we first estimate the equation jointly for the U.S., the U.K. and France, which are the countries for which we have the longest time series.[18] From this estimation, we take the technology-specific parameters $\hat{\beta}_{2\tau}$ and $\hat{\beta}_{3\tau}$. Then, for each technology-country pair, we re-estimate $\beta_{\tau 1}^c$ and D_τ^c imposing the technology specific estimates of $\hat{\beta}_{2\tau}$ and $\hat{\beta}_{3\tau}$ we have obtained in the first stage.[19]

[16]These are Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, Untied Kingdom, Japan, Australia, New Zealand, Canada and the United States.

[17]Note that there are two minor differences between (27) and (30). The first difference is that in the first equation $\beta_{\tau 3}$ is $\theta/(\theta-1)$, while in the second it is $1/(\theta-1)$. The second difference is that, in the second equation, the intercept $\beta_{\tau 1}^c$ has an extra term equal to $\beta_{\tau 3}\ln\alpha$.

[18]In the case of railways, we substitute data of the UK with German data because we lack the initial phase of diffusion of railways in the UK. In the case of tractors, we substitute U.S. data with German data for the same reason. We proceed as Comin and Hobijn (2010) and calibrate $\mu = 1.3$ to match price markups from Basu and Fernald (1997) and Norbin (1993). We take $\alpha = .3$ to match the labor income share.

[19]Note that the coefficients $\beta_{2\tau}$ and $\beta_{3\tau}$ in (27) are functions of parameters that are common across countries (θ and γ). Therefore their estimates should be independent of the sample used to estimate them. An advantage of using this two-step procedure is to avoid the problem that in the estimation of a system of equations, data problems from one country can contaminate the estimation of the common parameters across equations, and thus, the estimates for all countries. Using a small set of countries for which data are most reliable to identify the common technological parameters circumvents this problem. Reassuringly, Comin and Hobijn (2010) show that for a large majority of technology-country pairs, it is not possible to reject the null that $\beta_{3\tau}$ is common

In section 6.2, we relax the homotheticity in production implied by equation (8) and allow the elasticity of $y^c_{\tau t}$ with respect to income to differ from one. Our two-step estimation procedure allows to estimate the income elasticity, $\beta_{\tau y}$, (along with β_2 and β_3) from the diffusion curve in the baseline countries and then to impose these estimates when estimating the equation for all the technology-country pairs. Effectively, what this means is that we estimate $\beta_{\tau y}$ from the time series variation in technology and output for the baseline countries and then assume that the slope of the Engel curve is constant across countries. Given that the baseline countries have long time series that for many technologies cover much of its development experience, we consider this to be a reasonable approximation.

3.2 Data, estimation and summary statistics

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 104 technologies for 161 countries over the last 200 years. Due to 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 Appendix A and cover a wide range of sectors in the economy (transportation, communication and IT, industrial, agricultural and medical sectors). 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 (27) or (30). 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).

We only use in our analysis the estimates of adoption lags that satisfy plausibility and precision conditions.[20] These two conditions are met for the majority of the technology country-pairs (67%). For these technology country-pairs, we find that equation (27) provides a good fit for the data with an average detrended R^2 of 0.79 across countries and technologies (Table 9).[21]

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 44 years. We find significant variation in average adoption lags across technologies.

across countries when estimated separately country by country.

[20] As in Comin and Hobijn (2010), plausible adoption lags are those with an estimated adoption date of no less than ten years before the invention date (this ten year window is to allow for some inference error). Precise are those with a significant estimate of adoption lags D^c_τ at a 5% level. Most of the implausible estimates correspond to technology-country cases when our data does not have the initial phases of diffusion. This makes it hard to separately identify the log-linear trend from the logarithmic component of the diffusion curve.

[21] To compute the detrended R^2, we partial out the linear trend component, γt, of the estimation equation and compute the R^2 of the de-trended data.

Table 1: Estimated Adoption Lags

	Invention Year	Obs.	Mean	SD	P10	P50	P90	IQR
Spindles	1779	31	119	48	51	111	171	89
Steam and Motor Ships	1788	45	121	53	50	128	180	104
Railways Freight	1825	46	74	34	31	74	123	50
Railways Passengers	1825	39	72	39	16	70	123	63
Telegraph	1835	43	45	32	10	40	93	43
Mail	1840	47	46	37	8	38	108	62
Steel (Bessemer, Open Hearth)	1855	41	64	34	14	67	105	51
Telephone	1876	55	50	31	8	51	88	51
Electricity	1882	82	48	23	15	53	71	38
Cars	1885	70	39	22	11	34	65	36
Trucks	1885	62	36	22	9	34	62	32
Tractor	1892	88	59	20	18	67	69	12
Aviation Freight	1903	43	40	15	26	42	60	19
Aviation Passengers	1903	44	28	16	9	25	52	18
Electric Arc Furnace	1907	53	50	19	27	55	71	34
Fertilizer	1910	89	46	10	35	48	54	7
Harvester	1912	70	38	18	10	41	54	17
Synthetic Fiber	1924	48	38	5	33	39	41	2
Blast Oxygen Furnace	1950	39	14	8	7	13	26	11
Kidney Transplant	1954	24	13	7	3	13	25	5
Liver Transplant	1963	21	18	6	14	18	24	3
Heart Surgery	1968	18	12	6	8	13	20	4
Cellphones	1973	82	13	5	9	14	17	6
PCs	1973	68	16	3	12	15	19	3
Internet	1983	58	7	4	1	7	11	3
All Technologies		1306	44	35	9	38	86	46

The range goes from 7 years for the internet to 121 years for steam and motor ships. There is also considerable cross-country variation in adoption lags for any given technology. The range for the cross-country standard deviations goes from 3 years for PCs to 53 years for steam and motor ships.

We also find significant cross-country variation in the intensive margin. The intensive margin is reported as log differences relative to the average adoption of Western countries. To compute the intensive margin we follow Comin and Mestieri (2010) and calibrate $\gamma = (1-\alpha)\cdot 1\%$, $\alpha = 0.3$, and use a value of $\beta_{3,\tau}$ that results from setting the elasticity across technologies, θ, to be the mean across our estimates, which is $\theta = 1.28$. The average intensive margin is -.62, which implies that the level of adoption of the average country is 54% of the Western countries. More generally, there is significant cross-country dispersion in the

Table 2: Estimated Intensive Margin

	Invention Year	Obs.	Mean	SD	P10	P50	P90	IQR
Spindles	1779	31	-0.02	0.6	-0.8	-0.1	0.8	0.7
Steam and Motor Ships	1788	45	-0.01	0.6	-0.6	0.0	0.7	0.6
Railways Freight	1825	46	-0.17	0.4	-0.6	-0.2	0.4	0.6
Railways Passengers	1825	39	-0.24	0.5	-0.9	-0.2	0.2	0.5
Telegraph	1835	43	-0.26	0.5	-1.0	-0.2	0.3	0.7
Mail	1840	47	-0.19	0.3	-0.6	-0.1	0.1	0.4
Steel (Bessemer, Open Hearth)	1855	41	-0.22	0.4	-0.7	-0.1	0.2	0.6
Telephone	1876	55	-0.91	0.9	-2.2	-0.8	0.1	1.2
Electricity	1882	82	-0.58	0.6	-1.2	-0.5	0.1	0.9
Cars	1885	70	-1.13	1.1	-2.1	-1.1	0.1	1.6
Trucks	1885	62	-0.86	1.0	-1.7	-0.8	0.1	1.1
Tractor	1892	88	-1.02	0.9	-2.3	-0.9	0.1	1.5
Aviation Freight	1903	43	-0.39	0.6	-1.3	-0.2	0.2	0.9
Aviation Passengers	1903	44	-0.45	0.7	-1.3	-0.4	0.2	0.9
Electric Arc Furnace	1907	53	-0.29	0.5	-0.9	-0.2	0.3	0.8
Fertilizer	1910	89	-0.83	0.8	-1.9	-0.7	0.1	1.3
Harvester	1912	70	-1.10	1.0	-2.7	-1.0	0.2	1.5
Synthetic Fiber	1924	48	-0.52	0.7	-1.6	-0.4	0.2	0.9
Blast Oxygen Furnace	1950	39	-0.81	0.9	-2.3	-0.4	0.1	1.8
Kidney Transplant	1954	24	-0.19	0.4	-0.8	-0.1	0.1	0.3
Liver Transplant	1963	21	-0.33	0.7	-1.6	-0.1	0.1	0.5
Heart Surgery	1968	18	-0.44	0.8	-1.7	-0.1	0.2	0.6
Cellphones	1973	82	-0.75	0.7	-1.8	-0.6	0.1	1.2
PCs	1973	68	-0.60	0.6	-1.4	-0.6	0.1	0.9
Internet	1983	58	-0.96	1.1	-2.1	-0.8	0.1	1.5
All Technologies		1306	-0.62	0.8	-1.7	-0.4	0.2	1.0

intensive margin. The range goes from 0.3 for mail to 1.1 for cars and the internet. 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 (2010) which use smaller technology samples and estimate other versions of diffusion equations (27) and (30).

3.3 Evolution of adoption lags and the intensive margin

To explore the cross-country *evolution* of the adoption margins, we follow Maddison (2004) and divide the countries into two groups: "Western countries", and the rest, labeled "Rest of the World"or, simply, non-Western. Figure 2 plots, for each technology and country groups, the median adoption lag among Western countries and the rest of the world. This figure suggests that adoption lags have declined over time, and that cross-country differences in adoption lags have narrowed. Table 3 formalizes these intuitions by regressing (log) adoption lags on their year of invention (and a constant). Column (1) reports this regression for the whole sample of countries. We confirm the finding in Comin and Hobijn (2010) that adoption lags have declined with the invention date, on average. Then, we run the same regression separately for the two groups of countries. (See columns 2 and 3.) We find that the rate of decline in adoption lags is almost a 40% higher in non-Western than in Western countries (i.e., 1.12% vs. .81%). Hence, there has been *convergence* in adoption lags between Western and non-Western countries.

Do we observe a similar pattern for the intensive margin? To explore this question, Figure 3 plots, for each technology, the median intensive margin among Western and non-Western countries. This figure suggests that the gap between Western countries and the rest of the world in the intensive margin of adoption was smaller for technologies invented at the beginning of the nineteenth century than for technologies invented at the end of the twentieth century. Table 4 provides econometric evidence for this finding. It reports the regression of the intensive margin on the invention year and a constant. Column (3) shows that, for non-Western countries, the intensive margin has declined at a .54% annual rate. Recall that we define the intensive margin in equation (29) relative to the Western countries. As one would expect, column (2) shows that, for Western countries there is no trend in the intensive margin. Hence, Table 4 documents the *divergence* in the intensive margin of adoption between Western and non-Western countries over the last 200 years.

4 Income Dynamics: Analytic Results

After uncovering new cross-country patterns of technology diffusion, in the remaining of the paper, we study their implications for the evolution of income growth. Given the novelty of the model, we start by providing some analytic intuitions about the growth dynamics in the model. Then, in the next section, we use simulations to quantify the consequences of technology dynamics for income.

Our previous analysis of balanced growth (equation 24) showed that changes in the growth rate of the technology frontier, $\chi + \gamma$, generate changes in long-run growth. Moreover, any change in adoption margins is a source of additional transient growth. In this section, we analyze the transitional dynamics that follow from changes in these parameters. For the sake of clarity, we proceed sequentially. First, we study the sources of growth when the growth rate of the technology frontier is constant. Then, we study the transitional dynamics generated after an acceleration in the growth rate of the technological frontier. We conclude our analysis by exploring the effects of a one-time change in the adoption lag and the intensive margin.

Figure 2: Convergence of Adoption Lags

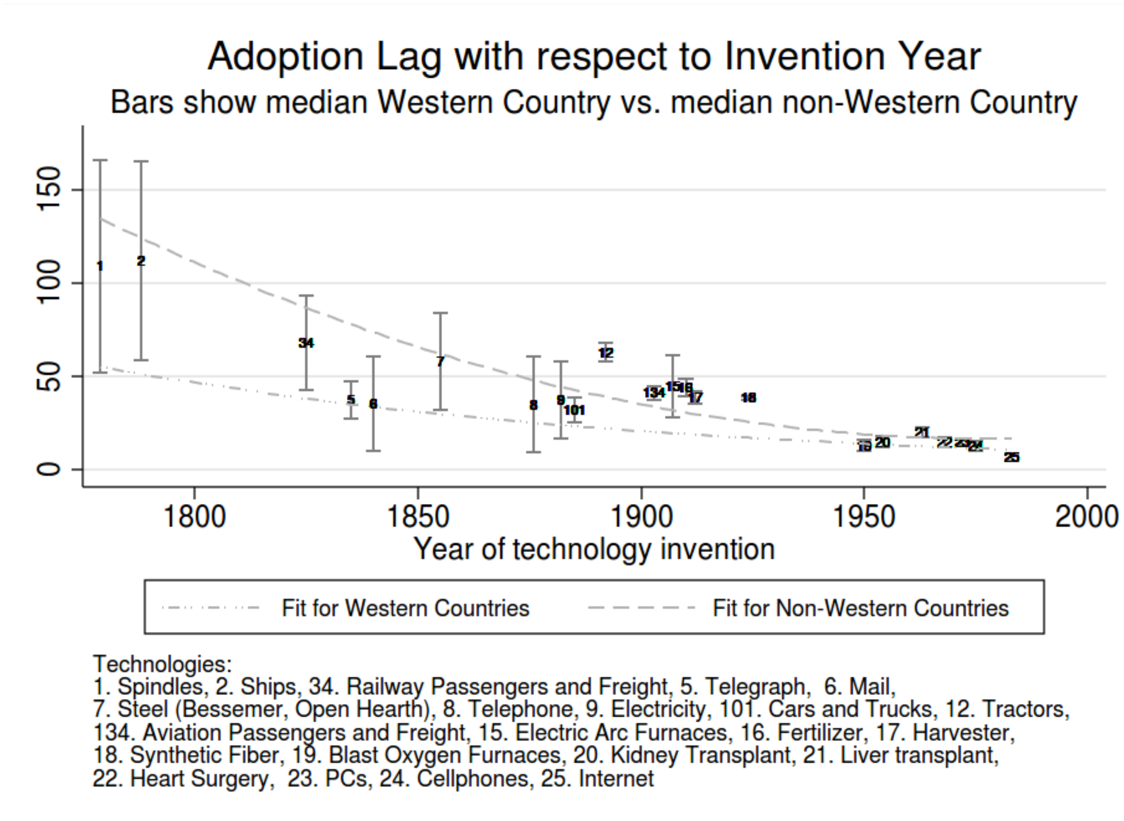
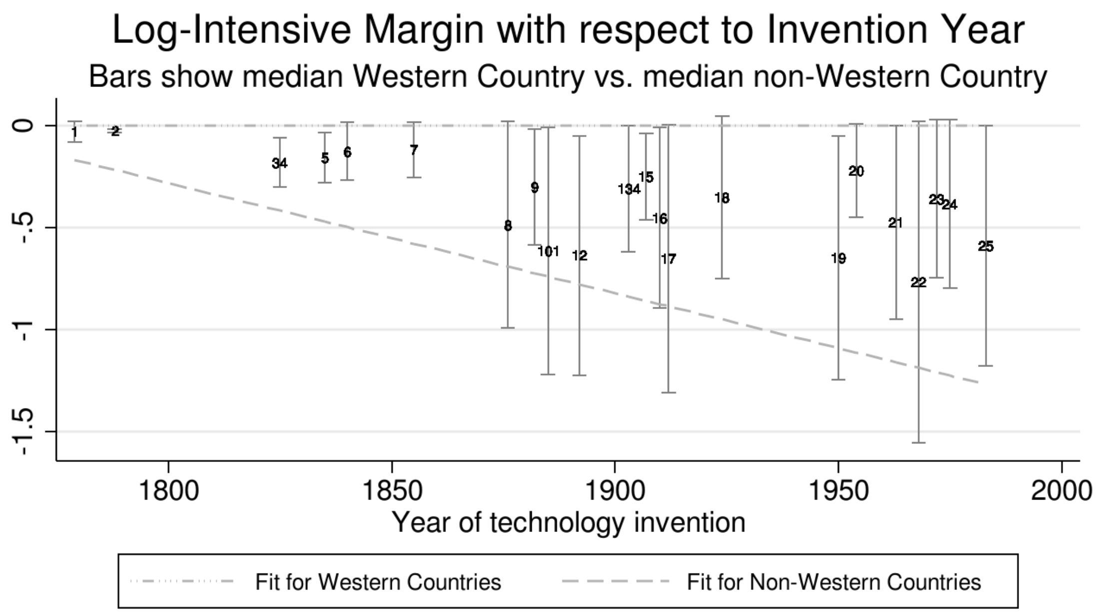


Table 3: Evolution of the Adoption Lag

	(1)	(2)	(3)
Dependent Variable is:	Log(Lag) World	Log(Lag) Western Countries	Log(Lag) Rest of the World
Year-1820	-0.0106	-0.0081	-0.0112
	(0.0004)	(0.0006)	(0.0004)
Constant	4.27	3.67	4.48
	(0.06)	(0.07)	(0.05)
Observations	1274	336	938
R-squared	0.45	0.34	0.53

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

Figure 3: Divergence of the Intensive Margin



Technologies:
1. Spindles, 2. Ships, 34. Railway Passengers and Freight, 5. Telegraph, 6. Mail,
7. Steel (Bessemer, Open Hearth), 8. Telephone, 9. Electricity, 101. Cars and Trucks, 12. Tractors,
134. Aviation Passengers and Freight, 15. Electric Arc Furnaces, 16. Fertilizer, 17. Harvester,
18. Synthetic Fiber, 19. Blast Oxygen Furnaces, 20. Kidney Transplant, 21. Liver transplant,
22. Heart Surgery, 23. PCs, 24. Cellphones, 25. Internet

Table 4: Evolution of the Intensive Margin

	(1)	(2)	(3)
Dependent Variable is:	Intensive World	Intensive Western Countries	Intensive Rest of the World
Year-1820	-0.0029	0.0000	-0.0054
	(0.0005)	(0.0002)	(0.0005)
Constant	-0.32	-0.00	-0.39
	(0.05)	(0.06)	(0.07)
Observations	1306	350	956
R-squared	0.042	0	0.13

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

4.1 Balanced Growth

As described in Section 2, the first vintage of a new technology and a new vintage for all past technologies appear at each instant of time. Thus, the set of technologies available at time t is given by $[-\infty, t - D_t)$, and 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 corresponding adoption lag. Let a dot and the letter g denote time derivatives and growth rates, respectively. Taking the time derivative of (17) and using (15) and (18), we find that

$$(1-\alpha) g_Y = \underbrace{(\theta - 1)\left(\frac{Z_{t-D_t}}{Y}\right)^{\frac{1}{\theta-1}} (1 - \dot{D}_t)}_{\text{New Technology}} + \underbrace{\int_{-\infty}^{t-D_t} \left(\frac{Z_\tau}{Y}\right)^{\frac{1}{\theta-1}} g_{Z_\tau} d\tau}_{\text{Old Technologies}}, \tag{31}$$

where

$$g_{Z_\tau} = \gamma \left(1 + \frac{e^{\frac{-\gamma}{\mu-1}(t-\tau-D_\tau)}}{1 - e^{\frac{-\gamma}{\mu-1}(t-\tau-D_\tau)}}\right). \tag{32}$$

The first term in (31) captures the growth imputable to a new technology being introduced in the economy. This term has three parts. $(1 - \dot{D}_t)$ captures the rate at which new technologies are 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 of new technologies depends on two factors. The first is the inverse of the 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. The second is the share of the new technologies in output, i.e., $(Z_{t-D_t}/Y)^{1/(\theta-1)}$.[22] The higher the productivity embodied in a technology, the larger the impact of its arrival on GDP growth. Note 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 (31) captures the increases of productivity due to the introduction of new vintages of already adopted 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 (32) that the productivity of new technologies grows faster than for older ones because of the larger gains from variety when fewer vintages of a technology 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.

[22] Recall from (17) and (18) that $Y_t = \alpha^{\frac{\alpha}{1-\alpha}} \left(\int_{-\infty}^{t-D_t} Z_\tau^{\frac{1}{\theta-1}} d\tau\right)^{\frac{\theta-1}{1-\alpha}}$.

4.2 Transitional dynamics after an acceleration in frontier growth

How does the economy respond to a permanent, instantaneous increase in the growth rate of the technology frontier? Let's suppose the growth rate of the frontier increases at time T from g_{Old} to $\gamma + \chi$. According to Mokyr (1990) and Crafts (1997), this acceleration in the technology frontier growth captures the Industrial Revolution. For the time being, we keep the adoption margins constant, and denote them, respectively, by D and a.

Let's decompose output and output growth as follows

$$\begin{aligned} Y(t) &= \alpha^{\frac{\alpha}{1-\alpha}} \left(\int_{-\infty}^{T} Z_\tau^{\frac{1}{\theta-1}} + \int_{T}^{t-D} Z_\tau^{\frac{1}{\theta-1}} \right)^{\frac{\theta-1}{1-\alpha}} \equiv \alpha^{\frac{\alpha}{1-\alpha}} \left(Y_{Old}^{\frac{1}{\theta-1}} + Y_{Modern}^{\frac{1}{\theta-1}} \right)^{\frac{\theta-1}{1-\alpha}}, && (33) \\ (1-\alpha) g_Y &= (1-s)\ g_{Old} + s\ g_{Modern}, && (34) \end{aligned}$$

where Y_{Old} is the output produced with pre-Industrial technologies, and Y_{Modern} is the output produced with Modern technologies. Let's denote by s the output share of Modern technologies $\left(\frac{Y_{Modern}}{Y}\right)^{\frac{1}{\theta-1}}$, and by g_i the growth rate of output produced with technologies of type $i = \{Old,\ Modern\}$.

It is clear from (34) that changes in growth may 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 *Modern and pre-Industrial output are given by*

$$\begin{aligned} Y_{Old}(t) &= aA_{Old} e^{g_{Old}(t-D)} && \textit{for all } t, && (35) \\ Y_{Modern}(t) &= \begin{cases} 0 & \textit{for } t < T + D, \\ aA_{Modern} e^{(\chi+\gamma)(t-D)} h(t)^{\theta-1} & \textit{for } t \geq T + D, \end{cases} && && (36) \end{aligned}$$

where D denotes the adoption lag, a is the intensive margin, and A_{Old}, A_{Modern} are positive constants. $h(t)$ 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, $\lim_{t\to\infty} h(t) = 1$. *Moreover, it approaches smoothly to its minimum and maximum values,* $h'(T+D) = \lim_{t\to\infty} h'(t) = \lim_{t\to\infty} h''(t) = 0$.[23]

The output produced using Old technologies grows at rate g_{Old}.[24] Modern output, in-

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

$$h(t) = \frac{\chi(\chi+\gamma)}{\gamma} \left(\frac{1}{\chi}\left(1 - e^{-\frac{\chi \Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right), \qquad (37)$$

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

[24] Note that here we are assuming that output produced with pre-Modern technologies keeps increasing

stead, has two components that change over time: a log-linear trend, $(\chi+\gamma)t$, and a transient source of growth, $h(t)$. The log-linear trend captures the higher productivity embodied in Modern 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 Modern vintages is larger due to gains from variety. Hence, the initial convexity of $h(t)$. At some point, though, the decreasing marginal gains from variety strike 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 of aggregate output from the pre-Industrial balanced growth path to the Modern balanced growth path is S-shaped. The growth rate starts the transition from its initial value g_{Old}. It is increasing and convex first, then concave. It approaches asymptotically the long-run growth rate $(\chi+\gamma)/(1-\alpha)$ from above, thus declining in a convex manner. In the case that $\gamma = \chi \gg g_{Old}$, the growth rate is increasing for $t < t^*$ and decreasing thereafter with inflexion points t_{i1} and t_{i2}, such that $t_{i1} < t^* < t_{i2}$, where*

$$t^* = T + D + \frac{(\theta-1)}{\chi}\log\left(\sqrt{2}\sqrt{2\kappa^2\chi^2+\kappa\chi}+2\kappa\chi+1\right), \tag{38}$$

$$t_{i1} = T + D + \frac{(\theta-1)}{2\chi}\log(2\kappa\chi+1), \tag{39}$$

$$t_{i2} = T + D + \frac{(\theta-1)}{\chi}\log\left(\sqrt{2}\sqrt{8\kappa^2\chi^2+3\kappa\chi}+4\kappa\chi+1\right), \tag{40}$$

with $\kappa = 2\chi\left(\frac{A_{Old}}{A_{Modern}}\right)^{\frac{1}{\theta-1}} e^{\frac{(2\chi-g_{Old})D}{\theta-1}}$.

From (34), 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. Hence, the dynamics of the growth rate are pinned down by the behavior of $s(t)(g_{Modern}(t) - g_{Old})$. The intuition for the result is that the share of the Modern sector $s(t)$ inherits the properties of the transient component $h(t)$, so that the weight on Modern output $s(t)$ is increasing and has an S-shape.

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 $(1-\alpha)g_Y = (1-s)g_{Old} + s(\chi+\gamma)$. It follows from this expression that aggregate output growth would be increasing over time reaching $(\chi+\gamma)/(1-\alpha)$ asymptotically.

independently from the advent of the Industrial Revolution. Thus, we are assuming that productivity of Old vintages does not increase with the Industrial Revolution. In Appendix D, we analyze the case in which new vintages of Old technologies become also more productive with the Industrial Revolution. The differences that we obtain are qualitatively minor, and quantitatively insignificant for the relevant parameter range.

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 because of the transient growth component $h(t)$. Thus, aggregate output growth will overshoot its long-run level $(\chi+\gamma)/(1-\alpha)$. 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 $(\chi+\gamma)/(1-\alpha)$. Our simulations, e.g. Figure 4b, suggest that this effect is not substantial.

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

Proposition 3 *Approximating the transient term $h(t)$ by its long-run value, the half-life of the output gap and the growth rate are*

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

$$t_{1/2}^{growth} \simeq D + \frac{1}{\chi+\gamma-g_{Old}} \ln\left(\frac{A_{Old}}{A_{Modern}}\right), \tag{42}$$

where the output gap is defined as the ratio of output in the Modern balanced growth path relative to current output.

The first term in both equations captures the fact that there is a lag between the advent of the Industrial Revolution and when a country starts adopting Modern technologies (i.e., the extensive margin). The second term captures 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. On the other hand, this effect is mitigated if the difference between the new and old growth rates $\chi+\gamma-g_{Old}$ is large.

4.3 Changes in Adoption Margins

Next, we study how changes in the adoption lag and intensive margin affect the transitional dynamics. Perhaps surprisingly, we show that qualitatively our analytic results continue to hold.

We consider a one-period, permanent change of adoption lags and the intensive margin from its pre-Modern levels to their average Modern levels.[25] Formally, we pose the following

[25] In Appendix B, we extend the results to the case where these margins evolve linearly over time, and show that the same qualitative results remain.

one-shot changes,

$$D_\tau = \begin{cases} D_{Old} & \text{for } \tau < T \\ D_{Modern} & \text{for } \tau \geq T \end{cases} \qquad a_\tau = \begin{cases} a_{Old} & \text{for } \tau < T \\ a_{Modern} & \text{for } \tau \geq T \end{cases} \tag{43}$$

where T denotes the time when the first Modern technology appears.

Proposition 4 *Let the evolution of the adoption lag and the intensive margin be given by (43), then pre-Industrial and Modern output are*

$$\begin{aligned} Y_{Old}(t) &= a_{Old} A_{Old} e^{g_{Old}(t-D_{Old})} && \textit{for all } t, && (44)\\ Y_{Modern}(t) &= \begin{cases} 0 & \textit{for } t < T + D_{Modern}, \\ a_{Modern} A_{Modern} e^{(\chi+\gamma)(t-D_{Modern})} h(t)^{\theta-1} & \textit{for } t \geq T + D_{Modern}, \end{cases} && && (45) \end{aligned}$$

where A_{Old}, A_{Modern} and $h(t)$ are as in in Proposition 1. In particular, $h(t)$ is S-shaped: increasing, initially convex and concave thereafter, reaching a plateau.[26] *The transition of the growth rate of aggregate output from the pre-Industrial balanced growth path to the Modern growth path is S-shaped as in Proposition 2. The characterization of the S-shape transition from equations (38), (39) and (40) holds with*

$$\kappa = 2\chi \left(\frac{a_{Old} A_{Old}}{a_{Modern} A_{Modern}} \right)^{\frac{1}{\theta-1}} e^{\frac{2\chi D_{Modern} - g_{Old} D_{Old}}{\theta - 1}}. \tag{46}$$

Approximating the transient term $h(t)$ by its long-run value, the half-lives of the system's output gap and growth rate are

$$\begin{aligned} t_{1/2}^{gap} &\simeq D_{Modern} - \frac{g_{Old} D_{Old}}{\chi+\gamma-g_{Old}} + \frac{1}{\chi+\gamma-g_{Old}} \ln \left(\frac{1}{2^{1-\alpha}} \frac{a_{Old} A_{Old}}{a_{Modern} A_{Modern}} \right), && (47)\\ t_{1/2}^{growth} &\simeq D_{Modern} - \frac{g_{Old} D_{Old}}{\chi+\gamma-g_{Old}} + \frac{1}{\chi+\gamma-g_{Old}} \ln \left(\frac{a_{Old} A_{Old}}{a_{Modern} A_{Modern}} \right). && (48) \end{aligned}$$

Note the similarity of these results with the previous propositions, where we only allowed for an acceleration of the frontier growth. This is the case for two reasons. First, changes in the adoption margins do not affect the pre-Modern sector. Second, up to a re-scaling of adoption lags, 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 (43). However, the changes in the adoption margins have a quantitative impact on the transitional dynamics.

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

This can be seen in expressions (47) and (48) for the half-lives, which depend on both Old and Modern levels of adoption.

To assess the protractedness of the model's transitional dynamics, we use Proposition 4 to calculate the half-lives during the transition to the new steady state. We calibrate D and a using information on the averages of both margins from our estimates from Tables 3 and 4.[27] The first two terms in the expressions for the half-lives (47) and (48) capture the extensive margin of adoption. The sum of these terms is approximately 40 years in our calculation (which almost coincides with the sample average lag of 44 years). The last term in the half-life expressions reflects output dynamics once adoption has started. Our calculations yield a value for this last term of 86 years for the half life of the output gap and 124 years for the half life of the growth rate. Hence, the resulting half-lives are 125 years for the half-life of the output gap and 163 for the half-life of the growth rate, showing how protracted dynamics are.

5 Income Dynamics: Simulation Results

We use our model to evaluate quantitatively the effects of dynamics in technology diffusion on the cross-country evolution of economic growth. We explore three questions: *(i)* The model's ability to generate pre-industrial income differences, *(ii)* the protractedness of the model's transitional dynamics, and *(iii)* the model's account of the Great Divergence.

Calibration.– To simulate the model we need to calibrate four parameters. First, we take $\alpha = .3$ to match the labor income share. Second, 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), we assume that the technology frontier grew at 0.2%. This is the growth rate of Western Europe according to Maddison (2004) from 1500 to 1800.[28] After 1765, the frontier growth rate, $\chi + \gamma$, equals $(1 - \alpha) \cdot 2\%$ per year. This ensures that Modern growth along the balanced growth path is 2%. The literature has not determined what fraction of frontier growth comes from increases in productivity due to new technologies (χ) and new vintages (γ). 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.

Finally, we need to calibrate the elasticities of substitution 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 average value we estimate for θ is 1.28, which is very similar to the values implied by the estimates of price markups from Basu and Fernald (1997) and Norbin (1993). Thus, we set $\theta = 1.28$.

[27] The rest of the parameters we take from our baseline calibration, which is explained in Section 5. These parameters are $\chi = \gamma = 1\%$, $\theta = 1.28$. We take the estimated adoption levels in 1820 as proxies for the Old margins of adoption, and the Old growth $g_{Old} = .2\%$. In Appendix F, we show the details of the calculation.

[28] 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.

Initial income differences.– In our model, differences in productivity along the pre-industrial balanced growth path are due to differences in adoption lags and in the intensive margin in the pre-Modern era. Since pre-Modern growth was roughly constant (e.g., Maddison, 2004), we assume that pre-Modern levels of adoption were constant and coincide with th adoption levels that we estimate for the beginning of 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 the rest of the world is 49 years in 1820. The average gap in the (log) intensive margin is 0.39. With this assumption and using Maddison's estimates of pre-industrial growth in Western Europe (0.2%) to calibrate the pre-industrial growth rate of the world technology frontier, equation (24) implies an income gap between Western countries and the rest of the world of 90%.[29] This is in line with the results from Maddison (2004), who reports an income gap of the same magnitude. Hence, the pre-industrial income differences generated by our model account for those observed in the data.

Protracted dynamics.– Next, we explore the protractedness of the model transitional dynamics. To this end, we consider the average country in our sample, and suppose that there is a one time permanent increase in the growth of the world technology frontier $(\gamma + \chi)$ like the one we observed in the Industrial Revolution (so that the balanced growth rate increases from 0.2% to 2%). The average country is parametrized so that its adoption lag and its degree of penetration rate are constant and equal to the average adoption lag and intensive margin across countries over our sample of technologies. In particular, the resulting D is 44 years and the intensive margin is 54% of the Western level. Figure 4 plots the transition of the output gap and income growth rate for this representative economy. The output gap is defined as the ratio of output in the Modern balanced growth path relative to current output. 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 117 years, while for output growth it is 145 years. These half-lives are 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 (44 years) imply that it takes this amount of time for the new technologies (which embody the higher productivity gains) to arrive to 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 (see equation 34). Since the Modern sector's share increases slowly, so does aggregate output. Third, the growth rate of the Modern sector is initially very small and grows progressively (see Proposition 1).

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

[29] That is, $\exp(.2\% \cdot 49 + .39/(1-\alpha)) = 1.9$.

Table 5: Growth rates of GDP per capita

		Time Period		
		1820-2000	1820-1913	1913-2000
Simulation	Western Countries	1.47%	.84%	2.15%
	Rest of the World	.82%	.35%	1.31%
	Difference West-Rest	.65%	.49%	.84%
Maddison	Western Countries	1.61%	1.21%	1.95%
	Rest of the World	.86%	.63%	1.02%
	Difference West-Rest	.75%	.58%	.93%

Notes: Simulation results and median growth rates from Maddison (2004). We use 1913 instead of 1900 to divide the sample because there are more country observations in Maddison (2004). The growth rates reported from Maddison for the period 1820-1913 for non-Western countries are computed imputing the median per capita income in 1820 for those countries with income data in 1913 but missing observations in 1820. These represent 11 observations out of the total 50. We do the same imputation for computing the growth rate for non-Western countries for 1820-2000. This represents 106 observations out of 145. For the 1913-2000 growth rate of non-Western countries, we impute the median per capita income in 1913 to those countries with income per capita data in 2000 but missing observations in 1913. These represent 67 observations out of the total 145.

of the world after feeding in a (common) one time permanent increase in frontier growth and the estimated evolutions for adoption lags and the intensive margin for each group of countries reported in Tables 3 and 4.[30] The results from this exercise are reported in Figure 5 and Table 5.

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, converging to the steady state growth of 2% in the early twentieth century. 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 non-Western country slowly accelerates, but it is still around 1.5% by year 2000. The gap in growth between both countries is considerable. Annual growth rates differ by more than 0.7% for over 100 years. The peak gap is reached around 1915 at 1.1%. From then, the gap declines monotonically until reaching around 0.6% by 2000. Table 5 reports the average growth and growth gaps of our simulation comparing it to Maddison (2004). The patterns and levels in our data trace quite well Maddison'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 3.2 income gap between the Western countries and the rest of the world. Maddison (2004) reports an actual income widening by a factor of 3.9 between Western countries and the rest of the world since the

[30] We assume that after the last technology invented in our sample (the internet, in 1983), the estimated margins remain constant at their 1983 values. This ensures that both groups of countries exhibit the same long-run growth. This assumption is quantitatively inconsequential, as it only affects the dynamics of the last ten years of our simulations. If anything, it tends to understate the effect of technology dynamics.

Industrial Revolution. Hence, most of the variation (82%) in the income gap between Western and non-Western countries in the last two centuries is accounted for.

The simulation does also well in replicating the time series income evolution of each country group separately. For Western countries, Maddison (2004) reports a 18.5-fold increase in income per capita between 1820 and 2000. Approximately 19% of this increase occurred prior to 1913. In our simulation, we generate a 14-fold increase over the same period, and 16% of this increase is generated prior to 1913. For non-Western countries, Maddison (2004) reports an almost 5-fold increase, with around 37% of the increase being generated prior to 1913. Our simulation generates a 4.3-fold increase in the 1820-2000 period with 32% of this increase occurring in pre-1913. The fact that we underpredict the time series increase in output reflects, in our view, that we are not accounting for the accumulation of factors of production over this time period (e.g., human capital), which also contributed to income growth.

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, we dissect 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 6 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 2.3 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 evolution of adoption lags that we estimated from the data. The intensive margins are kept constant at pre-industrial levels. Figure 7a 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 countries starting to adopt Modern technologies, the income gap reaches a level of 2.1. However, after that, the faster reduction in adoption lags in non-Western countries induces higher growth rates in this group of countries than in Western countries during the twentieth century. As a result, during the twentieth century income converges, and the relative income between the two countries is almost equalized at pre-industrial levels 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 study this hypothesis, 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 their pre-industrial levels.

Figure 7b 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 Western countries and the rest of the World. In this simulation, the growth acceleration in non-Western countries starts much later than in the baseline (Figure 5). This is a consequence of omitting the productivity gains from a reduction in adoption lags in non-Western countries. 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 non-Western countries.

We also see in Figure 7b that Western countries grow less than in the baseline, especially during the nineteenth century. This is a reflection of the productivity gains brought by the reduction in adoption lags for Western countries. Furthermore, as shown in the bottom panel of Figure 7b, the growth gap between the two groups of countries during the nineteenth century is smaller when we omit the evolution of the adoption lags. Despite that, the growth rate of non-Western countries falls behind during the first half of the twentieth century. This gap does not begin to close until the second half of the twentieth century. In this simulation, the income gap between Western countries and the rest of the world has increased by a factor of 3.8 by year 2000.

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

1. The model is capable of generating a Great Divergence. Income per capita of Western countries relative to the rest of the world increases by a factor of 3.2 over the last 200 years. This represents 80% of the actual increase in the income gap observed in the data.

2. Our model generates very protracted transitional dynamics due to the length of the adoption lags and to the time it takes for new technologies to become significant in aggregate output.

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

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

6 Extensions

Next, we show that the findings from the previous section are robust to alternative *(i)* calibrations, *(ii)* assumptions about the income elasticity of technology, and *(iii)* definitions of the samples of rich and poor countries.

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, we study the robustness of our findings to the relative contributions of new technologies and new varieties to balanced growth. To this end, we redo our baseline simulation under two polar assumptions. Figure 8a depicts the dynamics of productivity growth when balanced growth comes only from the development of better vintages (i.e., $\chi = 0$), while 8b shows the polar case, in which all productivity growth comes from the adoption of new technologies (i.e., $\gamma = 0$).

We draw two conclusions from this exercise. First, the main findings of the paper are robust quantitatively and qualitatively to the source of long-run growth. In particular, the income gap between Western and non-Western increases by a similar magnitude as in the benchmark (2.9 when growth comes from γ, 3.6 when it comes from χ, and 3.2 in the benchmark). Second, the income gap between Western and non-Western countries is larger when growth comes only from the adoption of new technologies. Intuitively, in this case, for a given technology, all vintages have the same productivity. 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 Western and non-Western countries) are less important in this case.

6.2 Non-homotheticities in production

Next, we explore the robustness of the dynamics of adoption margins once we allow for non-homotheticities in the production function. Non-homotheticities alter our baseline estimating equation (27) by introducing an income elasticity in the demand for technology, $\beta_{\tau y}$, potentially different from one,

$$y^c_{\tau t} = \beta^c_{\tau 1} + \beta_{\tau y} y^c_t + \beta_{\tau 2} t + \beta_{\tau 3} \left((\mu - 1) \ln(t - D^c_\tau - \tau) - (1 - \alpha)(y^c_t - l^c_t) \right) + \varepsilon^c_{\tau t}. \qquad (49)$$

To estimate (49), we use the same strategy as in the main specification. We first use the time series variation in the diffusion curves of the United States, the United Kingdom and France to estimate the elasticity parameter $\beta_{\tau y}$ (jointly with $\beta_{\tau 2},\ \beta_{\tau 3}$). However, by relaxing

Table 6: Income Elasticities by Period

	Period			
	pre-1850	1850-1900	1900-1950	post-1950
β_{Ty}	1.58	1.99	1.88	1.75
	(.02)	(.02)	(.02)	(.02)

Note: Standard errors are shown in parenthesis.

the theoretical constraint that $\beta_{\tau y} \neq 1$, we encounter a potential collinearity of income and the time trend. In practice, we can identify the income elasticity if we group technologies by their invention date in four groups: pre-1850, 1850-1900, 1900-1950 and post-1950. Then, we perform a non-linear version of seemingly unrelated regressions. We estimate jointly (49) for all technologies in a given time period with the restriction that the income elasticity has to be common across technologies of the same period. This allows us to identify four elasticity terms, β_{Ty}, where T indexes the four periods in which we have divided the sample. Table 6 reports the elasticities we find. They range from 1.58 in the pre-1850 period to 1.99 in the 1850-1900 period.

Once we have obtained the estimates for the income elasticity, we proceed as in the baseline estimation, but instead of using an income elasticity of one, we use the income elasticity that we have estimated.[31] We estimate for each country

$$y^c_{\tau t} = \beta^c_{\tau 1} + \bar{\beta}_{Ty} y^c_t + \bar{\beta}_{\tau 2} t + \bar{\beta}_{\tau 3} \left((\mu - 1)\ln(t - D^c_\tau - \tau) - (1-\alpha)(y^c_t - l^c_t)\right)) + \varepsilon^c_{\tau t}. \tag{50}$$

where $\bar{\beta}_{\tau 2}$, $\bar{\beta}_{\tau 3}$ and $\bar{\beta}_{\tau y}$ are the values of $\beta_{\tau 2}$, $\beta_{\tau 3}$ and $\beta_{\tau y}$ estimated in the first step.

To compute the intensive margin, we proceed as in the baseline model and calibrate the intermediate goods share to $\alpha = .3$ and the embodied productivity growth to be $\gamma = (1-\alpha) \cdot 1\%$ (i.e., half of the long-run productivity growth). Finally, for the elasticity of substitution across technologies, we use the average value we estimated, $\theta = 1.24$.

The summary of our estimates by technology of the intensive and the extensive margin are reported in Tables 10 and 11. The estimates obtained allowing for non-homotheticities are similar to our baseline estimates. The correlation between the adoption lags estimated allowing for non-homotheticities and the baseline model is 98%. For the intensive margin, the correlation is 89%. More importantly, the patterns of convergence of adoption lags and divergence of the intensive margin remain. Tables 12 and 13 show the evolution of adoption lags and intensive margin. We see that the convergence rate (measured as the difference between the coefficients of Western and non-Western countries) is -.44% per year, while in the baseline

[31] That is, given the income elasticity of a technology β_{Ty}, we first estimate the common technological parameters of 49 for each technology using the U.S., the U.K. and France. Note that this means that we re-estimate $\beta_{\tau 2}$, $\beta_{\tau 3}$ for each technology instead of using the estimates from the joint estimation. This is to allow the estimation to be the most flexible possible.

case it was -.31% per year. Hence, if anything, adoption lags close faster. The divergence rate in the intensive margin is reduced by 19% when allowing for non-homotheticities relative to the baseline model (the regression coefficient estimated on the time trend is $-.44\%$, versus $-.54\%$ in the baseline model). However, the divergence is still very significant. After feeding these trends, we observe that the productivity gap between Western and non-Western countries increases by a factor of 2.6 over the last 200 years, which accounts for 67% of the Great Divergence. We conclude that our findings are robust to allowing for non-homotheticities in demand.[32]

6.3 Alternative definitions of non-Western

Finally, we extend the analysis to alternative country classifications to Maddison's Western/non-Western. We conduct two exercises. First, we focus on the countries that were in the bottom 25th and 10th percentile of the income distribution in year 2000. We also report the model predictions for the evolution of productivity growth for various continents.

Table 7 shows that initial technology adoption patterns accounted for cross-country differences in productivity circa 1820 in the world's poorest economies (see the first row of Table 7). For example, in 1820, based on the observed differences in technology adoption between Western countries and those in the bottom 10%, the model predicts that the average income in Western countries should be 2.9 times the average income of the bottom 10% countries. In Maddison's data we observe that the actual income ratio is 2.8.

Technology dynamics are also important to account for the evolution of income between 1820 and 2000 for the two new country groupings (see the second row Table 7). For example, our model predicts an accumulated income gap between Western countries and the bottom 10th of 17.2, while in the data the actual gap was 17.6.

Table 8 reports the results from a similar exercise after grouping countries by continent. This table shows that technology dynamics in each continent have induced income dynamics that resemble closely those observed in the data over the last two centuries. As in the baseline case, the model tends to underpredict the average growth rates during the nineteenth century. However, the correlation between the actual growth rates across continents between 1820 and 1913 and those predicted by the model is 0.99. For the period 1913-2000, the correlation between actual and predicted income growth is 0.94. We conclude that technology dynamics account well for income dynamics when looking at narrower country groupings either based on income levels or on geographical locations.

[32] Another robustness check that we have performed is to group technologies by sector rather than invention dates to compute the income elasticity. These categories are listed in Appendix A. The results in this case are also very similar to our baseline results and we do not report them.

Table 7: Evolution of the Income Gap for different groups of countries

Period	Income gap of the West relative to ...					
	Non-Western		Bottom 25th		Bottom 10th	
	Maddison	Simulation	Maddison	Simulation	Maddison	Simulation
Pre 1820	1.9	1.9	2	2.3	2.8	2.9
1820-2000	3.9	3.2	5.4	4.5	6.3	5.8
Cumulative	7.2	6.2	10.3	10.3	17.6	17.2

Note: The Maddison column refers to the raw data provided in Maddison (2004). The data reported for each time period is done using the maximal amount of data available to compute it. The only exception is for the calculation of the pre-1820 bottom 10th countries, for which we have imputed the minimum subsistence level estimated by Maddison to missing income per capita data in 1820. This means that the line 1820-2000 reports relative income gap only for countries with 1820 data, while the Cumulative line uses data for all countries with data in 2000. If Maddison's data were balanced, the accumulated value in 2000 should equal to the product of the gap generated prior to 1820 and from 1820 to 2000.

Table 8: Annual Growth rates of GDP per capita by regions.

	Simulation		Maddison	
	1820-1913	1913-2000	1820-1913	1913-2000
USA & Canada	.77%	2.05%	1.63%	1.90%
Western Europe	.62%	1.91%	1.29%	2.16%
Africa	.26%	.75%	.36%	.90%
Asia	.34%	1.37%	.49%	1.70%
Latin America	.37%	1.28%	.59%	1.50%

Note: Annual Growth rates of GDP per capita by regions. Simulation results and growth rates from Maddison (2004). We use 1913 instead of 1900 because there are more country observations in Maddison (2004). For the missing income per capita values in 1820 and 1913, we have imputed the minimal within group income reported in that year.

7 Conclusions

In what has now become a classic paper, Klenow and Rodríguez-Clare (1997) show that factor accumulation accounts for 10% of cross-country variation in productivity growth between 1960 and 1985 leaving for the TFP residual a staggering 90% of the variation in income growth. What drives variation in income growth over the long-term?

In this paper we have explored one potential driver: the dynamics of technology adoption. Using a stylized model of adoption that fits well diffusion curves for individual technologies, we have identified two margins of adoption: adoption lags and the penetration rates. Analyzing the panel of adoption lags and penetration rates for 25 technologies and 132 countries, we have uncovered two new facts. Adoption lags have converged across countries over the last 200 years, while penetration rates have diverged. Feeding in these patterns into the aggregate representation of our model economy we have evaluated the effects of the cross-country evolution of adoption patterns on the cross-country evolution of income growth.

The main finding of the paper is that the evolution of adoption patterns accounts for the vast majority of cross-country evolution of income growth for many country groupings. Hence, this shows that adoption dynamics are at the core of the Great Divergence that has taken place over the last two centuries.

Our findings motivate some new questions that we plan to pursue in future research. Probably the main one is why has the intensive margin of adoption diverged. Future work shall formulate hypotheses about the nature, drivers and sources of dynamics for the intensive margin of adoption. These explorations will complement our analysis towards a fuller understanding of cross-country income dynamics.

References

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

— and Sala-i-Martin, X. (2003). *Economic Growth, 2nd Edition*. 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, National Bureau of Economic Research, Inc, pp. 277–322.

Comin, D. and Hobijn, B. (2004). Cross-country technology adoption: making the theories face the facts. *Journal of Monetary Economics*, **51** (1), 39–83.

— and — (2009). *The CHAT Dataset*. Working Paper 15319, National Bureau of Economic Research.

— and — (2010). An exploration of technology diffusion. *American Economic Review*, **100** (5), 2031–59.

—, — and Rovito, E. (2008). A new approach to measuring technology with an application to the shape of the diffusion curves. *The Journal of Technology Transfer*, **33** (2), 187–207.

— and Mestieri, M. (2010). An intensive exploration of technology adoption. *NBER Working Paper 16379*.

Crafts, N. (1997). Endogenous growth: Lessons for and from economic history. In D. Kreps and K. F. Wallis (eds.), *Advances in Economics and Econometrics: Theory and Applications*, vol. 2, CUP, pp. 38–78.

Dholakia, R. D., N and Kshetri, N. (2003). Internet diffusion. In H. Bidgoli (ed.), *The Internet Encyclopedia*, Wiley.

Easterly, W. and Levine, R. (2002). Its not factor accumulation: Stylized facts and growth models. In N. Loayza, R. Soto, N. L. S. Editor) and K. S.-H. S. Editor) (eds.), *Economic Growth: Sources, Trends, and Cycles*, *Central Banking, Analysis, and Economic Policies Book Series*, vol. 6, *3*, Central Bank of Chile, pp. 061–114.

Galor, O. and Mountford, A. (2006). Trade and the great divergence: The family connection. *American Economic Review*, **96** (2), 299–303.

— and WEIL, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, **90** (4), 806–828.

HANSEN, G. D. and PRESCOTT, E. C. (2002). Malthus to solow. *American Economic Review*, **92** (4), 1205–1217.

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

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, National Bureau of Economic Research, Inc, pp. 73–114.

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

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

MOKYR, J. (1990). *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford paperbacks, Oxford University Press, USA.

— (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press.

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

O'ROURKE, K. H., RAHMAN, A. S. and TAYLOR, A. M. (2012). *Trade, Technology and the Great Divergence*. Departmental Working Papers 35, United States Naval Academy Department of Economics.

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

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

TAMURA, R. (2002). Human capital and the switch from agriculture to industry. *Journal of Economic Dynamics and Control*, **27** (2), 207–242.

Table 9: Quality of the Estimates

Technology	Invention						Detrended R^2		
	Year	Total	Implausible	Imprecise	Precise	% Precise	$R^2 > 0$	Mean	SD
Spindles	1779	34	3	0	31	91	26	0.57	0.25
Steam and Motor Ships	1788	61	16	0	45	74	45	0.84	0.10
Railways Freight	1825	85	39	0	46	54	46	0.78	0.15
Railways Passengers	1825	80	41	0	39	49	39	0.79	0.14
Telegraph	1835	62	19	0	43	69	30	0.58	0.25
Mail	1840	67	20	0	47	70	47	0.87	0.10
Steel (Bessemer, Open Hearth)	1855	52	11	0	41	79	41	0.74	0.17
Telephone	1876	139	84	0	55	40	54	0.88	0.15
Electricity	1882	134	52	0	82	61	82	0.91	0.12
Cars	1885	124	54	0	70	56	68	0.76	0.22
Trucks	1885	108	46	0	62	57	62	0.78	0.20
Tractor	1892	135	45	2	88	65	81	0.74	0.20
Aviation Freight	1903	93	50	0	43	46	43	0.88	0.10
Aviation Passengers	1903	96	52	0	44	46	44	0.90	0.07
Electric Arc Furnace	1907	75	22	0	53	71	46	0.62	0.24
Fertilizer	1910	132	39	4	89	67	76	0.62	0.25
Harvester	1912	104	32	2	70	67	59	0.68	0.24
Synthetic Fiber	1924	49	1	0	48	98	45	0.69	0.24
Blast Oxygen Furnace	1950	49	10	0	39	80	30	0.62	0.29
Kidney Transplant	1954	27	3	0	24	89	24	0.82	0.17
Liver Transplant	1963	21	0	0	21	100	20	0.81	0.16
Heart Surgery	1968	18	0	0	18	100	17	0.63	0.20
Cellphones	1973	84	2	0	82	98	82	0.91	0.07
PCs	1973	69	1	0	68	99	68	0.93	0.07
Internet	1983	59	1	0	58	98	58	0.96	0.04
All Technologies		1957	643	8	1306	67	1233	0.79	0.21

Table 10: Adoption Lags with Non-homotheticities

Technology Name	Invention Year	Obs.	Mean	Sd	p10	p50	p90	iqr
Spindles	1779	28	114	49	49	108	170	94
Steam and Motor Ships	1788	36	121	56	47	145	179	111
Railways Freight	1825	34	71	30	31	74	117	47
Railways Passengers	1825	23	64	37	15	63	116	61
Telegraph	1835	32	48	30	18	39	94	33
Mail	1840	31	44	39	6	33	104	78
Steel (Bessemer, Open Hearth)	1855	40	61	36	13	61	106	57
Telephone	1876	35	49	33	6	47	91	56
Electricity	1882	56	44	25	12	41	69	37
Cars	1885	47	40	23	14	33	64	33
Trucks	1885	53	36	20	14	33	63	29
Tractor	1892	55	57	20	29	63	79	15
Aviation Freight	1903	32	47	14	32	47	65	22
Aviation Passengers	1903	30	31	16	16	26	53	24
Electric Arc Furnace	1907	28	53	21	22	59	78	37
Fertilizer	1910	74	42	11	26	43	52	13
Harvester	1912	67	37	17	17	42	50	23
Synthetic Fiber	1924	46	38	4	33	39	41	2
Blast Oxygen Furnace	1950	37	15	8	8	13	28	10
Kidney Transplant	1954	24	13	7	4	13	25	4
Liver Transplant	1963	18	18	3	15	18	24	3
Heart Surgery	1968	16	12	3	9	12	17	3
PCs	1971	69	16	3	12	16	19	3
Cellphones	1975	82	13	4	9	14	17	5
Internet	1983	50	7	3	2	7	10	4
All Technologies		1043	41	35	10	34	82	41

Table 11: Intensive Margin with Non-homotheticities

Technology Name	Invention Year	Obs.	Mean	sd	p10	p50	p90	iqr
Spindles	1779	28	0.1	0.8	-1.6	0.1	1.6	1.0
Steam and Motor Ships	1788	36	0.0	0.8	-2.1	0.1	1.6	0.8
Railways Freight	1825	34	0.0	0.4	-0.7	-0.1	0.8	0.5
Railways Passengers	1825	23	0.1	0.4	-0.6	0.1	0.6	0.5
Telegraph	1835	32	-0.1	0.5	-1.2	-0.1	0.8	0.5
Mail	1840	31	-0.1	0.4	-1.1	0.0	0.6	0.5
Steel (Bessemer, Open Hearth)	1855	40	-0.1	0.6	-1.3	0.0	0.6	0.7
Telephone	1876	35	-0.2	0.9	-1.9	-0.1	0.9	1.0
Electricity	1882	56	-0.2	0.6	-1.4	-0.2	0.8	0.9
Cars	1885	47	-0.4	0.8	-1.7	-0.4	0.5	1.2
Trucks	1885	53	-0.3	0.7	-1.6	-0.2	0.7	0.7
tractor	1892	55	-0.6	0.9	-1.9	-0.7	0.7	1.3
Aviation Freight	1903	32	0.0	0.4	-0.6	0.0	0.8	0.5
Aviation Passengers	1903	30	-0.1	0.6	-0.9	0.0	1.0	0.7
Electric Arc Furnace	1907	28	0.1	0.4	-0.6	0.2	0.8	0.6
fertilizer	1910	74	-0.5	0.7	-1.4	-0.5	0.4	0.9
harvester	1912	67	-0.6	0.8	-1.8	-0.6	0.8	1.0
Synthetic	1924	46	-0.2	0.6	-1.1	-0.1	0.5	0.6
Blast Oxygen Furnace	1950	37	-0.6	0.8	-2.1	-0.3	0.3	1.1
Kidney Transplant	1954	24	-0.1	0.4	-0.8	-0.1	0.3	0.5
Liver Transplant	1963	18	-0.2	0.5	-1.6	-0.1	0.2	0.3
Heart Surgery	1968	16	-0.3	0.7	-2.1	-0.1	0.3	0.4
PCs	1971	69	-0.3	0.5	-1.2	-0.3	0.4	0.7
Cellphones	1975	82	-0.5	0.6	-1.6	-0.3	0.4	0.9
Internet	1983	50	-0.7	0.9	-2.6	-0.4	0.4	1.1
All Technologies		1043	-0.3	0.7	-1.6	-0.2	0.6	0.8

Table 12: Evolution of Adoption Lags with Non-homotheticities

	(1)	(2)	(3)
Dependent Variable is:	Intensive World	Intensive Western Countries	Intensive Rest of the World
Year-1820	-0.011***	-0.0076***	-0.012***
	(0.0005)	(0.0007)	(0.0004)
Constant	4.23***	3.58***	4.53***
	(0.08)	(0.08)	(0.06)
Observations	1027	314	713
R-squared	0.43	0.28	0.58

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

Table 13: Evolution of the Intensive Margin with Non-homotheticities

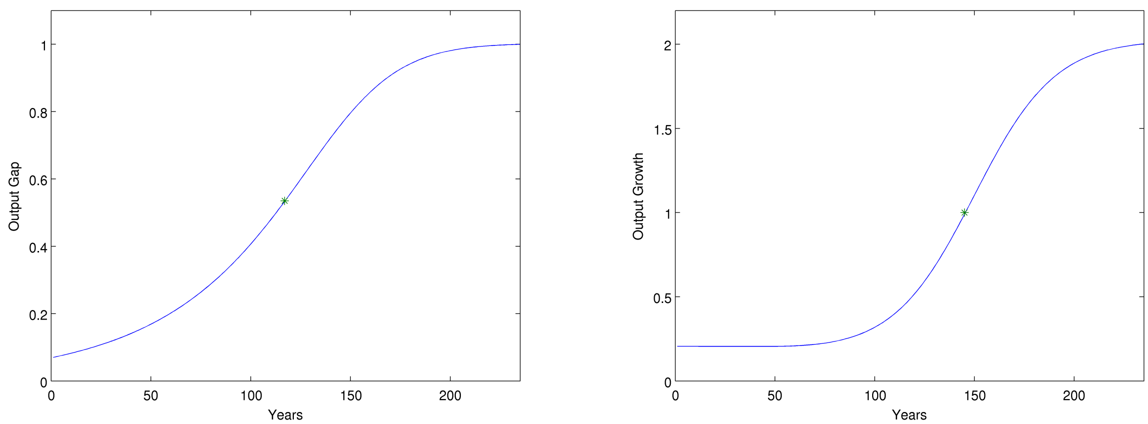
	(1)	(2)	(3)
Dependent Variable is:	Intensive World	Intensive Western Countries	Intensive Rest of the World
Year-1820	-0.0025***	0	-0.0044***
	(0.0005)	(0.0002)	(0.0006)
Constant	-0.04	0	-0.04
	(0.08)	(0.07)	(0.1)
Observations	1043	323	720
R-squared	0.04	0	0.10

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

Figure 4: 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 (44 years) and average intensive margin (54% of the Western productivity level). The star * denotes the half-life.

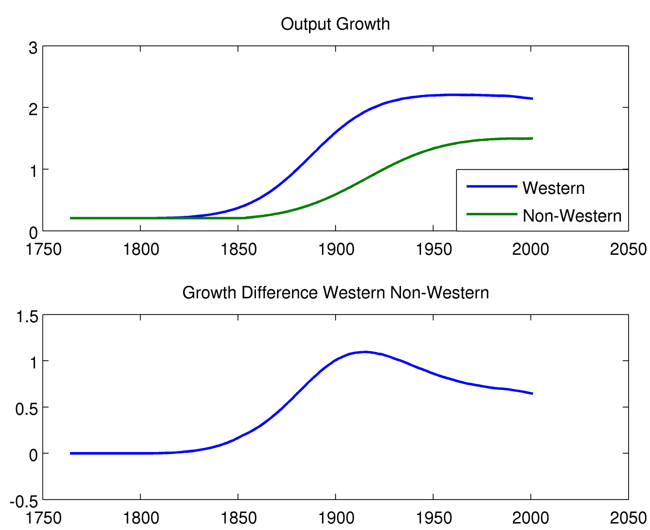


Figure 5: Growth of Western and non-Western countries imputing the estimated evolution of the intensive and extensive margins.

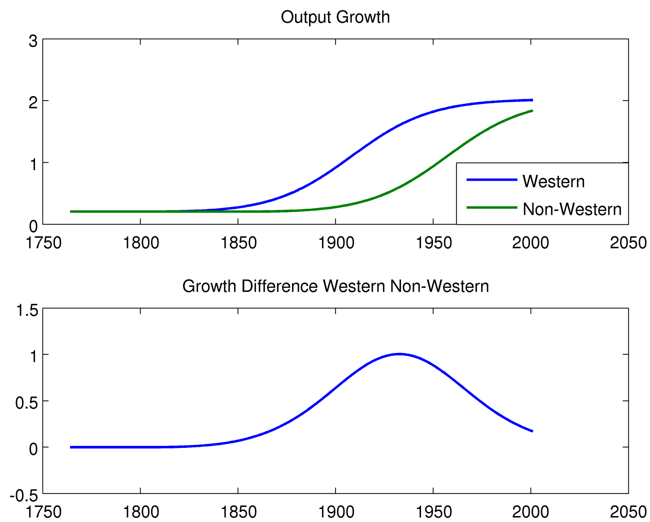
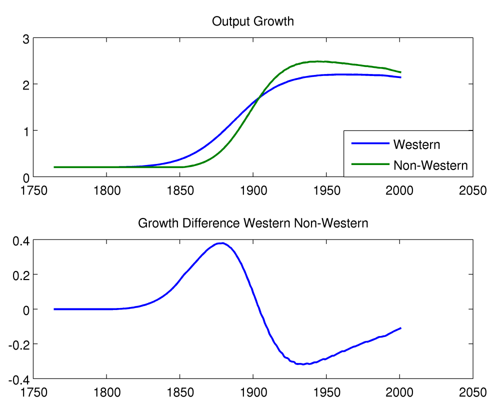
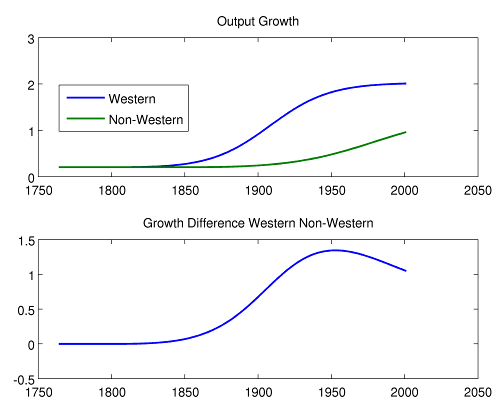


Figure 6: Growth of Western and non-Western countries with *only* an acceleration of the technology frontier. Both margins of adoption are held constant.

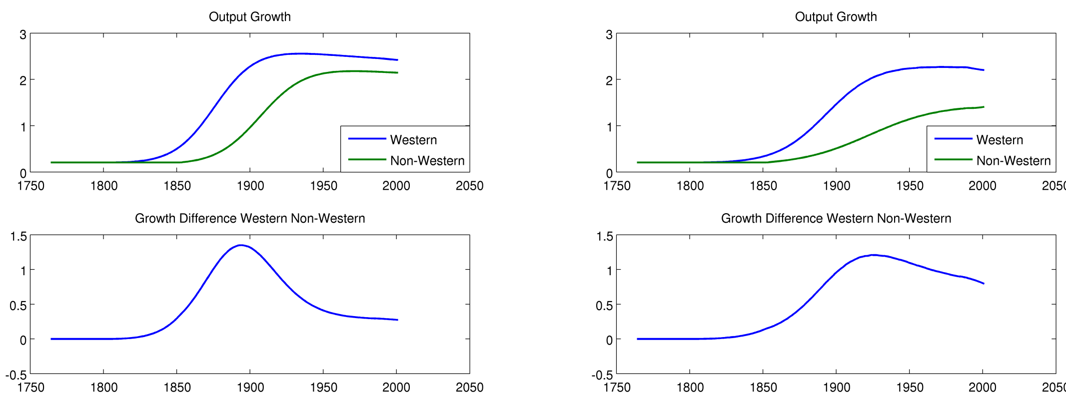


(a) Dynamics due only to a decline in lags.



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

Figure 7: 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 8: Role played by the different margins of productivity gains

A Data Description

The twenty-five particular technology measures used in the paper, organized by broad category (transportation, communication, IT, industrial, agricultural and medical) are described below.

Transportation

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.

Communication and IT

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 cross-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.

Industrial

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 the 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.

Agricultural

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

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

Medical

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.

B Transitional Dynamics with Linear Trends in Adoption

In this appendix, we extend the analysis of the model's transitional dynamics to the case where the adoption margins change continuously. As shown in Section 3, the evolution of the intensive margin and adoption lags has been smoother than in (43). A more realistic characterization of the evolutions of the extensive and intensive margins is given by:[33]

$$D_\tau = \begin{cases} d_o & \text{for } \tau < T, \\ d_o - d_1\tau & \text{for } \tau \in [T, \bar{T}], \\ d_m & \text{for } \tau > \bar{T}, \end{cases} \qquad \ln a_\tau = \begin{cases} a_o & \text{for } \tau < T, \\ a_o - a_1\tau & \text{for } \tau \in [T, \bar{T}], \\ a_m & \text{for } \tau > \bar{T}, \end{cases} \tag{51}$$

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.

[33] The specification we have estimated in Section 3.2 differs slightly from (43) in that in Section 3.2 we fit a linear trend to the log adoption lag while in (43) the trend is fit to the level. Both approaches seem sensible to us and, quantitatively, there are no significant differences between them. The linear trend is more amenable to analytical manipulation.

Proposition 5 *Let the evolution of adoption margins be given by (51). Pre-Industrial productivity is described by $Y_{Old}(t) = A_{Old}e^{g_{Old}t}$. Modern productivity is a continuous, increasing function,*

$$Y_{Modern}(t) = \begin{cases} 0 & \text{for } t < T + d_0, \\ A_0 \; e^{(\chi+\gamma+g_a)t} \; h_0(t)^{\theta-1} & \text{for } t \in [T + d_0, d_0 + \bar{T}/(1 + d_1)], \\ A_1 \; e^{(\chi+\gamma)t} \; h_1(t)^{\theta-1} & \text{for } t > d_0 + \bar{T}/(1 + d_1), \end{cases} \tag{52}$$

where A_0, A_1 are positive constants, $g_a = d_1(\chi+(1+d_1)\gamma)-(1+d_1)a_1$, $h_0(t)$ 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.[34]

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 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 $-(1 + d_1)a_1$.

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.

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

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

$$h_0(t) = \frac{1}{\chi + \gamma d_1 - a_1}\left[1 - e^{-\frac{\chi+\gamma d_1 - a_1}{\theta-1}(1+d_1)(t-d_0)}\right] - \frac{e^{-\frac{d_1^2\gamma}{\theta-1}(t-d_0)}}{\chi+\gamma-a_1}\left[1 - e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]. \tag{53}$$

See Appendix E 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.

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.[35]

C Derivation of Equilibrium Conditions

Derivation of equation (15) It follows from

$$Z_\tau = \left(\int_\tau^{\max\{t-D_\tau,\tau\}} Z(\tau,v)^{\frac{1}{\mu-1}} dv\right)^{\mu-1} \tag{54}$$

$$= a_\tau e^{(\chi+\gamma)\tau}\left(\int_\tau^{t-D_\tau} e^{\frac{\gamma}{\mu-1}(v-\tau)} dv\right)^{\mu-1}$$

$$= \left(\frac{\mu-1}{\gamma}\right)^{\mu-1} a_\tau e^{(\chi+\gamma)\tau}\left(e^{\frac{\gamma}{\mu-1}(t-D_\tau-\tau)}-1\right)^{\mu-1}. \tag{55}$$

Derivation of equation (24) Using the definition of the production function and integrating, we have that

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

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

With a constant D and a, we find

$$A = a\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\frac{\theta-1}{\chi}-\frac{\theta-1}{\chi+\gamma}\right)^{\theta-1} e^{(\chi+\gamma)(t-D)}, \tag{56}$$

after rearranging, we obtain (24).

Derivation of equation (26) Start considering a second order approximation of Z_τ around $\Delta t \equiv t-D_\tau-\tau=0$,

$$Z_\tau \simeq a_\tau e^{(\chi+\gamma)\tau}\left[\Delta t\left(1+\frac{1}{2}\frac{\gamma}{\mu-1}\Delta t\right)\right]^{\mu-1}. \tag{57}$$

[35] We could also characterize the half-lives of this system in a similar manner to the main text. We obtain similar expressions when assuming that $t_{1/2} \in [T+d_0, d_0+\frac{\bar{T}}{1+d_1}]$.

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

$$\ln Z_\tau \simeq \ln a_\tau + (\chi + \gamma)\tau + (\mu - 1)\ln \Delta t + \frac{\gamma}{2}\Delta t. \tag{58}$$

Equation (26) is obtained then by direct substitution.

D Evolution of the Output Produced with Old Technologies when Old Vintages Become more Productive after the Industrial Revolution

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. To lighten up notation, we make the innocuous simplifying assumption that $\alpha = 0$.

Proposition 6 *Suppose that before the adoption of Modern technologies, the economy is in a balanced growth path with growth rate $\chi_o + \gamma_o$, so that output is*

$$Y(t) = Ae^{(\chi_o+\gamma_o)t}. \tag{59}$$

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

$$Y_{Old}(t) \quad = \quad A_o e^{\gamma_o t} d(t)^{\theta-1}, \tag{60}$$

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 $\lim_{t\to\infty} d(t) = 1$. The output produced with Modern technologies is as in Proposition 1.

Proof For output produced before a country starts adopting Modern technologies, $\tau < T$, we have that equation (24) holds, and hence

$$Y_{Old} = a_o \left(\frac{(\theta-1)^2}{\chi_o(\chi_o+\gamma_o)} \right)^{\theta-1} e^{(\chi_o+\gamma_o)(t-D_o)}. \tag{61}$$

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

$$\begin{aligned} Y_{Old} &= a\left(\int_{-\infty}^{T} d\tau \int_{\tau}^{t-D_o} dv e^{\frac{\chi_o\tau+\gamma_o v}{\theta-1}}\right)^{\theta-1} && (62)\\ &= a\left(\frac{\theta-1}{\gamma_o}\int_{-\infty}^{T} d\tau\left(e^{\frac{\chi_o\tau+\gamma_o(t-D_o)}{\theta-1}} - e^{\frac{(\chi_o+\gamma_o)\tau}{\theta-1}}\right)\right)^{\theta-1} && (63)\\ &= a\left(\frac{(\theta-1)^2}{\gamma_o\chi_o}\right)^{\theta-1} e^{(\chi_o+\gamma_o)T} e^{\gamma_o(t-D_o-T)}\left(1-\frac{\chi_o}{\chi_o+\gamma_o}e^{\frac{-\gamma_o(t-D_o-T)}{\theta-1}}\right)^{\theta-1}. && (64)\end{aligned}$$

Equation (60) 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 (35). 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.

E Proofs and Derivations of Section 4 and Appendix B

Proof of Proposition 1 For the output produced with Modern technologies, applying (18), we have that

$$\begin{aligned} Y_{Modern} &= a\left(\int_{T}^{t-D} d\tau \int_{\tau}^{t-D} dv e^{\frac{\chi\tau+\gamma v}{\theta-1}}\right)^{\theta-1} && (65)\\ &= a\left(\frac{\theta-1}{\gamma}\int_{T}^{t-D} d\tau\left(e^{\frac{\chi\tau+\gamma(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)\tau}{\theta-1}}\right)\right)^{\theta-1} && (66)\\ &= a\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(e^{\frac{\chi(t-D)+\gamma(t-D)}{\theta-1}} - e^{\frac{\chi T+\gamma(t-D)}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(e^{\frac{(\chi+\gamma)(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)T}{\theta-1}}\right)\right\}\right]^{\theta-1} && (67)\\ &= ae^{(\chi+\gamma)(t-D)}\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(1-e^{\frac{\chi(T-(t-D))}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(1-e^{\frac{(\chi+\gamma)(T-(t-D))}{\theta-1}}\right)\right\}\right]^{\theta-1} && (68)\\ &= ae^{(\chi+\gamma)(t-D)}\left[\frac{(\theta-1)^2}{\gamma}\left\{\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right\}\right]^{\theta-1} && (69)\end{aligned}$$

where $\Delta t \equiv t - D - T$. This last expression can be identified with (36), where

$$h(t) = \frac{\chi(\chi+\gamma)}{\gamma}\left(\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right). \quad (70)$$

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

$$\frac{\gamma(\theta-1)}{\chi(\chi+\gamma)} h'^{-\frac{\chi\Delta t}{\theta-1}} - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}, \tag{71}$$

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

$$\frac{\gamma(\theta-1)^2}{\chi(\chi+\gamma)} h'^{-\frac{\chi\Delta t}{\theta-1}} + (\chi+\gamma)e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}. \tag{72}$$

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

For the output produced with pre-Modern technologies, we have that by assumption, it grows at rate g_{Old}. In other words, as new vintages of Old technologies do not experience an increase in productivity (and there keep appearing Old technologies at the g_{Old} rate). Letting $(1-\alpha)g_{Old} = (\gamma_o+\chi_o)$, we can repeat the previous steps for the output produced with old technologies. Note that, by assumption, the pre-Modern economy was in a balanced growth path before the Industrial Revolution. This implies that the transient factor $h_{Old}(t)=1$. Which gives,

$$Y_{Old} = a_o\left(\frac{(\theta-1)^2}{\chi_o(\chi_o+\gamma_o)}\right)^{\theta-1} e^{(\chi_o+\gamma_o)(t-D_o)}. \tag{73}$$

Alternatively, this could have been derived by noting that $dY/dt = g_{Old}$ with the appropriate boundary condition. $\blacksquare$

Proof of Proposition 2 The shape of the growth rate is determined by the time dependent parameters of the growth rates.

$$s(t)(g_{Modern}(t) - g_{Old}). \tag{74}$$

The growth rate of the Modern sector is

$$g_{Modern} = \frac{\gamma+\chi}{1 - \frac{\chi e^{-\frac{(\chi+\gamma)t}{\theta-1}}\left(e^{\frac{\gamma t}{\theta-1}}-1\right)}{\gamma\left(1-e^{-\frac{\chi t}{\theta-1}}\right)}}. \tag{75}$$

Note that the growth rate is $\chi+\gamma$ at time zero and as $t\to\infty$.

The share $s(t)$ is, using its definition,

$$s(t) = \frac{e^{\frac{t(\chi+\gamma)}{\theta-1}}\left(\frac{1-e^{-\frac{\chi t}{\theta-1}}}{\chi} - \frac{1-e^{-\frac{t(\chi+\gamma)}{\theta-1}}}{\chi+\gamma}\right)}{\kappa e^{\frac{g_{Old}t}{\theta-1}} + e^{\frac{t(\chi+\gamma)}{\theta-1}}\left(\frac{1-e^{-\frac{\chi t}{\theta-1}}}{\chi} - \frac{1-e^{-\frac{t(\chi+\gamma)}{\theta-1}}}{\chi+\gamma}\right)}, \tag{76}$$

where using (69) and (73) and using O and M to denote Old and Modern, respectively, we have that

$$\kappa = \left(\frac{a_O}{a_M}\right)^{\frac{1}{\theta-1}} e^{\frac{(\chi+\gamma)D_M - g_O D_O}{\theta-1}} \frac{\gamma_M}{\chi_O(\gamma_O + \gamma_O)}. \tag{77}$$

Indeed, in this case adoption margins do not change and $D_M = D_O$, $a_M = a_O$. This is noted for future reference. Inspection of this term reveals that the last fraction has a very high value if we assume $\gamma_M = 1\%$ and $\gamma_O = \chi_O = .1\%$ (its numerical value is 5000). The other two terms are also greater or equal than one. So, we have that $\kappa \gg 1$.

The time derivative of (74) with respect to time is

$$\frac{\chi(\gamma+\chi)e^{\frac{(\gamma+g_O)t}{\theta-1}}}{(\theta-1)\left(\kappa\chi(\gamma+\chi)e^{\frac{g_O t}{\theta-1}} + \left(\gamma e^{\frac{(\gamma+\chi)t}{\theta-1}} - (\gamma+\chi)e^{\frac{\gamma t}{\theta-1}} + \chi\right)\right)^2}. \tag{78}$$

$$\left\{\kappa\left(g_O^2\chi e^{-\frac{\gamma t}{\theta-1}} - (\gamma+\chi)(\gamma-g_O)^2 + \gamma(\gamma+\chi-g_O)^2 e^{\frac{\chi t}{\theta-1}}\right) + \gamma\left(\gamma\left(e^{\frac{\chi t}{\theta-1}} - 1\right) + \chi e^{\frac{(\gamma+\chi)t}{\theta-1}}\left(e^{\frac{-\gamma t}{\theta-1}} - 1\right)\right) e^{-\frac{g_O t}{\theta-1}}\right\}$$

It can be verified that the value of (78) at $t = 0$ is always positive (it is $(\kappa(\theta-1))^{-1}$ and that it is zero as $t \to \infty$. Next, we show that (78) approaches zero from below as $t \to \infty$. Note that the sign of the derivative depends on the sign of the second line of equation (78), as the first line is always positive. Moreover, note that the fastest growing term in the second line is $-\chi\gamma e^{\frac{(\gamma+\chi-g_O)t}{\theta-1}}$. Thus it is immediate to verify that for t high enough, this term will dominate in the second line of (78) and hence the derivative will be negative. Note moreover that we have that there exists a t^* such that for $t < t^*$, equation (78) is increasing and it is decreasing thereafter. We have shown already that at $t = 0$, the second line is positive.

Next, we decompose the second line of (78) as the sum of two functions, one positive and another negative. We show that the negative function decreases faster than the rate at which the positive function increases. Thus, there is only one change of sign of the derivative at t^*. The second line of (78) can be written as,

$$\text{Positive Terms:} \quad \kappa\left(g_O^2\chi e^{-\frac{\gamma t}{\theta-1}} + \gamma(\gamma+\chi-g_O)^2 e^{\frac{\chi t}{\theta-1}}\right) + \gamma e^{\frac{(\chi-g_O)t}{\theta-1}}(\gamma+\chi) \tag{79}$$

$$\text{Constant:} \quad -\kappa(\gamma+\chi)(\gamma-g_O)^2 \tag{80}$$

$$\text{Negative Terms:} \quad -\gamma\left(\gamma + \chi e^{\frac{(\gamma+\chi)t}{\theta-1}}\right) e^{-\frac{g_O t}{\theta-1}} \tag{81}$$

To avoid a taxonomical analysis, we discuss only the case in point which is $\gamma, \chi \gg g_O$. In this case the derivative of the positive terms is

$$\frac{\chi\gamma(\chi+\gamma)}{\theta-1}(1+\kappa(\chi+\gamma))e^{\frac{\chi t}{\theta-1}}, \tag{82}$$

while the derivative of the negative terms is

$$-\frac{\chi\gamma(\chi+\gamma)}{\theta-1}e^{\frac{(\chi+\gamma)t}{\theta-1}}. \tag{83}$$

Looking at the ratio (in absolute values) of the two, it is clear that initially the positive terms grow faster (recall that $\kappa \gg 1$). Thus, the term in the second line of (78) initially increases. Eventually, it is the negative terms that grow faster, making the second line of (78) to decline and eventually become negative. Note however that the first term of (78) is always decreasing. Thus the magnitude of the negative slope will be attenuated by this first term.

To gain some quantitative insight on when (78) ceases to be positive, assume that $\gamma = \chi \gg g_O$. In this case, (78) simplifies to

$$\frac{2\chi^2 e^{\frac{t\chi}{\theta-1}}\left(2\chi\left(2\kappa e^{\frac{t\chi}{\theta-1}}-1\right)-\left(e^{\frac{t\chi}{\theta-1}}-1\right)^2\right)}{(\theta-1)\left(2\kappa\chi+\left(e^{\frac{t\chi}{\theta-1}}-1\right)^2\right)^2}. \tag{84}$$

In this case, we can solve explicitly for the time in which there is a change of sign,

$$t^* = \frac{(\theta-1)\log\left(\sqrt{2}\sqrt{2\kappa^2\chi^2+\kappa\chi}+2\kappa\chi+1\right)}{\chi}. \tag{85}$$

Assuming that $\chi = \gamma = 1\%$ and $D = 45$ years, this gives a $t^* = 240$ years. At this point, the value of the growth rate is $g(t = 240) = 2\%$, so the decreasing part of the growth rate is of no numerical significance.

Next we study the second derivative of the system. Its expression is

$$\frac{\chi(\gamma+\chi)e^{\frac{(2\gamma+g_O+\chi)t}{\theta-1}}}{(\theta-1)^2\left(\kappa\chi(\gamma+\chi)e^{\frac{g_O t}{\theta-1}}+\gamma e^{\frac{(\gamma+\chi)t}{\theta-1}}-(\gamma+\chi)e^{\frac{\gamma t}{\theta-1}}+\chi\right)^3}f(t) \tag{86}$$

with

$$
\begin{aligned}
f(t) \quad = \quad & \gamma\kappa(\gamma+\chi)(\gamma-g_O-\chi)(2\gamma-2g_O+\chi)(\gamma-g_O+2\chi) && (87)\\
& -g_O^3\kappa\chi^2\left(\kappa(\chi+\gamma)e^{\frac{g_O t}{\theta-1}}-1\right)e^{-\frac{t(2\gamma+\chi)}{\theta-1}} && (88)\\
& -\gamma^3\chi e^{-\frac{(\gamma+\chi+g_O)t}{\theta-1}}-\gamma^3(\gamma+\chi)e^{-\frac{(\chi+g_O)t}{\theta-1}} && (89)\\
& -\kappa\chi(\gamma+\chi)(\gamma-2g_O)(\gamma+g_O)(2\gamma-g_O)e^{-\frac{(\gamma+\chi)t}{\theta-1}} && (90)\\
& -\gamma^2\kappa(\gamma-g_O+\chi)^3e^{\frac{\chi t}{\theta-1}}-\gamma^2(\gamma+\chi)^2e^{\frac{(\chi-g_O)t}{\theta-1}} && (91)\\
& +\gamma\kappa^2\chi(\gamma+\chi)(\gamma-g_O+\chi)^3e^{-\frac{(\gamma-g_O)t}{\theta-1}}+\kappa^2\chi(\gamma+\chi)^2(g_O-\gamma)^3e^{-\frac{(\gamma-g_O+\chi)t}{\theta-1}} && (92)\\
& +\kappa(\gamma+\chi)^2(g_O-\gamma)^3e^{-\frac{\chi t}{\theta-1}} && (93)\\
& +\gamma\kappa\chi(\gamma-2g_O+\chi)(\gamma+g_O+\chi)(2(\gamma+\chi)-g_O)e^{-\frac{\gamma t}{\theta-1}} && (94)\\
& +\gamma\chi(\gamma+\chi)^2e^{-\frac{(\gamma+g_O)t}{\theta-1}}+\gamma(\gamma-\chi)(2\gamma+\chi)(\gamma+2\chi)e^{-\frac{g_O t}{\theta-1}} && (95)\\
& +\gamma\chi^2(\gamma+\chi)e^{\frac{(\gamma-g_O)t}{\theta-1}}+\gamma^2\chi^2e^{\frac{(\gamma+\chi-g_O)t}{\theta-1}} && (96)
\end{aligned}
$$

First, it can be verified that the value of the second derivative at $t=0$ is $\frac{\gamma(2\gamma-3g_O+\chi)}{\kappa(\theta-1)^2}$. Hence, as long as $2\gamma+\chi>3g_O$, the curve is initially convex (which is indeed true as $\chi+\gamma\sim 2\%$ and $g_O\sim 0.2\%$. Second, asymptotically, the term that dominates is the last term of $f(t)$, (96), because it is the one that grows at the fastest rate, as long as $\gamma>g_O$. This implies that the curve is asymptotically convex.

Next we are interested in whether the terms in (91) can make $f(t)$ negative, and hence the growth rate of the economy concave. Using basic results from calculus we can guarantee that there exists a concave region. Recall that the growth rate is twice-continuously differentiable. It asymptotes continuously a constant value as $t\to\infty$ and approaches it from above. Moreover we have shown that it is initially convex and asymptotically convex. Hence, the growth rate has to be concave for some range of t. For one thing, it is impossible to have an increasing, convex and bounded function over $\mathbb{R}_+$. For another, we can apply Rolle's theorem on the derivative of the growth rate $g'(t)$. We have already shown that $g'(t)$ takes at least two times the initial value (at $t=0$ and some time after, as we have shown that $g''(t=0)>0$ and $g'(t^*)=0$). So given that $g'(t)$ is a continuous function, it is decreasing in some region.

To have a sharper characterization, we study again the case $\chi=\gamma\gg g_O$. In this case

$$
f(t)=\chi^3\left(e^{\frac{2t\chi}{\theta-1}}-2\kappa\chi-1\right)\left(e^{\frac{t\chi}{\theta-1}}\left(e^{\frac{t\chi}{\theta-1}}-8\kappa\chi-2\right)+2\kappa\chi+1\right). \tag{97}
$$

We can thus solve analytically for the region in which the curve become concave by solving for the zeros of $f(t)$. Which yields, the following two solutions

$$t_{i1} = \frac{(\theta-1)\log(2\kappa\chi+1)}{2\chi}, \tag{98}$$

$$t_{i2} = \frac{(\theta-1)\log\left(\sqrt{2}\sqrt{8\kappa^2\chi^2+3\kappa\chi}+4\kappa\chi+1\right)}{\chi}. \tag{99}$$

Quantitatively, using the same parametrization as before, we have that $t_{i1} = 110$ and $t_{i2} = 261$. Inspection of equations (85), (99), and (99) shows that

$$t_1 < t^* < t_2. \tag{100}$$

Thus we have an S-shape in the sense that growth rate is increasing and convex for $t < t_1$, then increasing and concave $t \in [t_1, t^*]$, then decreasing and concave for $t \in [t^*, t_2]$ and finally, convex and decreasing for $t > t_2$, asymptotically converging to $(\gamma+\chi)/(1-\alpha)$. Quantitatively we have found that $g(t^*) - (\gamma+\chi)/(1-\alpha)$ is very small. Thus, in this case the "overshooting" above the long-run growth rate is not quantitatively important.[36] ∎

Proof of Proposition 3 We start with the half-life of the growth rate. The definition of the half-life is

[36] In the explanation in the text we refer to an interpretation that claims that the evolution of the share $s(t)$ is S-shaped. Here we outline part of the analysis. Note that

$$\left(\frac{Y_{Old}}{Y}\right)^{\frac{1}{\theta-1}} = \frac{1}{1+\left(\frac{Y_1}{Y_0}\right)^{\frac{1}{\theta-1}}}, \tag{101}$$

$$\frac{Y_{Modern}}{Y_{Old}} \propto e^{(\chi+\gamma-g_o)t}\left[\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right]^{\theta-1}. \tag{102}$$

Taking the time derivative of (101), it can be verified that this share declines over time. Moreover, the sign of the second derivative of (101) coincides with the sign of

$$\left(\frac{d}{dt}\left(\frac{Y_{Modern}}{Y_{Old}}\right)^{\frac{1}{\theta-1}}\right)^2 - \left(1+\left(\frac{Y_{Modern}}{Y_{\text{Old}}}\right)^{\frac{1}{\theta-1}}\right)\frac{d^2}{dt^2}\left(\frac{Y_{Modern}}{Y_{Old}}\right)^{\frac{1}{\theta-1}}. \tag{103}$$

Note that in the case that $\left(\frac{Y_{Modern}}{Y_{Old}}\right)^{\frac{1}{\theta-1}}$ is concave, the share is unambiguously convex. As we previously discussed for the transient factor $h(t)$, 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 (103) coincides with the sign of

$$e^{gt}\left(h'^2 - h''(t)h(t)\right) - \left(g^2h(t) + 2gh'(t) + h''(t)\right). \tag{104}$$

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 (104) is positive. Similarly, when a country starts to adopt technologies of the Industrial Revolution (for $t = 0$ after the change of variables), equation (104) 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.

$$(1-\alpha)\frac{\chi+\gamma}{2(1-\alpha)} = \left(\frac{Y_M(t_{1/2})}{Y(t_{1/2})}\right)^{\frac{1}{\theta-1}} g_M(t_{1/2}) + \left(\frac{Y_O(t_{1/2})}{Y(t_{1/2})}\right)^{\frac{1}{\theta-1}} g_O, \tag{105}$$

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

$$g_M = \chi + \gamma + (\theta - 1)\frac{h'(t)}{h(t)}. \tag{106}$$

Rearranging, equation (105) becomes

$$(\chi + \gamma + 2(\theta - 1)g_h(t_{1/2}))Y_M(t_{1/2})^{\frac{1}{\theta-1}} = (\chi + \gamma - 2g_O)Y_O(t_{1/2})^{\frac{1}{\theta-1}}. \tag{107}$$

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

$$< e^{-\beta t} >= \frac{1}{T}\int_0^T e^{-\beta t} dt = \frac{1 - e^{-\beta T}}{\beta T}. \tag{108}$$

We proceed by averaging $h(t)$ and $h'(t)$ to make (107) analytically solvable,

$$(\chi + \gamma + 2(\theta - 1) < g_h >)Y_M(t_{1/2})^{\frac{1}{\theta-1}} = (\chi + \gamma - 2g_O)Y_O(t_{1/2})^{\frac{1}{\theta-1}}. \tag{109}$$

Denoting by

$$\alpha \equiv \left(\frac{(\chi + \gamma + 2(\theta - 1) < g_h >}{\chi + \gamma - 2g_O}\right)^{\theta-1}, \tag{110}$$

equation (109) is

$$\alpha A_M e^{(\chi+\gamma)(t-D)} h(t)^{\theta-1} = A_O e^{g_O(t-D)}, \tag{111}$$

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,

$$t = D + \frac{1}{\chi + \gamma - g_O} \ln\left(\frac{A_O}{\alpha A_M < h(t) >^{\theta-1}}\right). \tag{112}$$

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.

Derivation of the half-life for the output gap The definition of the output gap relative to the Modern balanced growth path is

$$\tilde{Y}(t) = \frac{\left(a_M A_M e^{(\chi+\gamma)(t-D)}\right)^{\frac{1}{1-\alpha}}}{Y(t)}. \tag{113}$$

Qualitatively, note that the output gap will be growing over time until $Y(t)$ reaches the Modern balance growth path, in which case the output gap becomes constant. In fact, by construction, $\lim_{t\to\infty} \tilde{Y}(t) = 1$. Moreover, as discussed $\tilde{Y}(t)$ is an increasing function.[37] The definition of the half-life is

$$\tilde{Y}(0) + \frac{1}{2}(1 - \tilde{Y}(0)) = \tilde{Y}(t_{1/2}). \tag{114}$$

We know that real income per capita in the world since the Industrial Revolution until 2000 has increased on average 15-fold. Supposing that the Industrial Revolution started in 1765 with James Watt's steam engine invention, we have that in this time period, the income per capita of a fictional country on the technological frontier would have increased $\exp(2\% \cdot 235) = 108.4$-fold by 2000. Suppose further that by year 2000, $\tilde{Y}(t = 2000) \simeq 1$. This means that $\tilde{Y}(0) \simeq \frac{15}{108.4} = 0.14$. In interest of having a simple expression, we proceed by assuming that $1 \gg .14$, so that the left hand side of (114) is approximately $1/2$. In this case, the half-life definition simplifies to

$$\frac{1}{2} \simeq \tilde{Y}(t_{1/2}). \tag{115}$$

Following the same steps as in the derivation of the half-life of the growth rate (i.e., approximating the average of $h(t)$ by its long-run level), one obtains that

$$\frac{A_O}{A_M \left(2^{\frac{1-\alpha}{\theta-1}} - 1\right)^{\theta-1}} = e^{(\chi+\gamma-g_O)(t_{1/2}-D)}. \tag{116}$$

Using that $\theta = 1.3$, we can further simplify by using that $2^{\frac{1}{\theta-1}} \gg 1$. Taking logs and solving for $t_{1/2}$ we obtain (41). ∎

Proof of Proposition 4 With the definition of the evolution of the intensive and extensive margins (43), 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

[37] As we discussed, the transitional dynamics can feature an overshot of the long-run growth, in which case the output gap can be decreasing after surpassing $\tilde{Y} = 1$. However, for purposes of our calculation this is not a relevant range.

claims in the proposition can be derived analogously to Propositions 1, 2 and 3 replacing D for D_{Modern}. An additional correction appears linearly, $t_{1/2} = D_{Modern} - \frac{g_O}{\chi+\gamma} D_{Old} + \ldots$ 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, $v_\tau \in [\tau, t - D_\tau]$. The output produced using modern technologies can be written as

$$\begin{aligned}
Y_m &= \left(\int_0^{t-D_t} d\tau \int_\tau^{t-D_\tau} dv \left[a_\tau Z(\tau, v) \right]^{\frac{1}{\theta-1}} \right)^{\theta-1} && (117)\\
&= e^{a_0} \left(\int_0^{(1+d_1)(t-d_0)} d\tau \left[e^{\frac{(\chi-a_1)\tau}{\theta-1}} \int_\tau^{t-(d_0-d_1\tau)} dv e^{\frac{\gamma v}{\theta-1}} \right] \right)^{\theta-1} && (118)\\
&= e^{a_0} \left(\frac{\theta-1}{\gamma} \right)^{\theta-1} \left(\int_0^{(1+d_1)(t-d_0)} d\tau e^{\frac{(\chi-a_1)\tau}{\theta-1}} \left[e^{\frac{\gamma(t-(d_0-d_1\tau))}{\theta-1}} - e^{\frac{\gamma\tau}{\theta-1}} \right] \right)^{\theta-1} && (119)\\
&= e^{a_0} \left(\frac{\theta-1}{\gamma} \right)^{\theta-1} \left(\int_0^{(1+d_1)(t-d_0)} d\tau \left[e^{\frac{(\chi+d_1\gamma-a_1)\tau+\gamma(t-d_0)}{\theta-1}} - e^{\frac{(\gamma+\chi-a_1)\tau}{\theta-1}} \right] \right)^{\theta-1} && (120)\\
&= e^{a_0} \left(\frac{\theta-1}{\gamma} \right)^{\theta-1} \left(^{\ldots\theta-1} e^{\frac{\gamma(t-d_0)}{\theta-1}} \left(\frac{\theta-1}{\chi+\gamma d_1 - a_1} \right) \left[e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}((1+d_1)(t-d_0))} - 1 \right] \right) - \ldots \\
&\quad \left(^{\ldots\theta-1} \ldots - \frac{\theta-1}{\chi+\gamma-a_1} \left[e^{\frac{\chi+\gamma-a_1}{\theta-1}((1+d_1)(t-d_0))} - 1 \right] \right)^{\theta-1} && (121)\\
&= e^{a_0} \left(\frac{(\theta-1)^2}{\gamma} \right)^{\theta-1} \exp\left[\left((\gamma+\chi-a_1)(1+d_1) + d_1^2\gamma \right)(t-d_0) \right] && (122)\\
&\quad \left(\frac{1}{\chi+\gamma d_1-a_1} \left[1 - e^{-\frac{\chi+\gamma d_1-a_1}{\theta-1}(1+d_1)(t-d_0)} \right] - \frac{e^{-\frac{d_1^2\gamma}{\theta-1}(t-d_0)}}{\chi+\gamma-a_1} \left[1 - e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)} \right] \right)^{\theta-1}.
\end{aligned}$$

This last expression can be rewritten as

$$Y_m(t) = A_m e^{g_m t} f(t) \tag{123}$$

where

$$
\begin{aligned}
A_m &= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1} e^{-d_0\left((\gamma+\chi-a_1)(1+d_1)+d_1^2\gamma\right)} &(124)\\
g_m &= (\gamma+\chi-a_1)(1+d_1)+d_1^2\gamma &(125)\\
f(t) &= \left(\frac{1}{\chi+\gamma d_1-a_1}\left[1-e^{-\frac{\chi+\gamma d_1-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]-\frac{e^{-\frac{d_1^2\gamma}{\theta-1}(t-d_0)}}{\chi+\gamma-a_1}\left[1-e^{-\frac{\chi+\gamma-a_1}{\theta-1}(1+d_1)(t-d_0)}\right]\right)^{\theta-1}
\end{aligned}
$$

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 behavior of the damp factor $f(t)$ it is useful to rewrite it as $f(t)=h(t)^{\theta-1}$, note that

$$
\begin{aligned}
& f'^{\theta-2}h'(t), &(126)\\
& f''^{\theta-2}\left[(\theta-2)h(t)^{-1}h'^2+h''(t)\right], &(127)\\
g_f \equiv (\ln f(t))' &= (\theta-1)\frac{h'(t)}{h(t)}, &(128)\\
g_f' &= (\theta-1)\frac{h''(t)h(t)-h'^2}{h(t)^2}. &(129)
\end{aligned}
$$

The time derivative of $h(t)$ is

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

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,

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

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 $(\theta-1)^2 h''(t)$ 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

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

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 behavior of $h(t)$ can inform our analysis on $f(t)$ and its derivatives. Using equation (126) 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 (129), 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,

$$
\begin{aligned}
Y_{m0}(t) &= \left(\int_0^T d\tau \int_\tau^{t-D_\tau} dv\, [a_\tau Z(\tau,v)]^{\frac{1}{\theta-1}}\right)^{\theta-1} && (130)\\
&= e^{a_0}\left(\int_0^T d\tau \left[e^{\frac{(\chi-a_1)\tau}{\theta-1}} \int_\tau^{t-(d_0-d_1\tau)} dv e^{\frac{\gamma v}{\theta-1}}\right]\right)^{\theta-1} && (131)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\int_0^T d\tau e^{\frac{(\chi-a_1)\tau}{\theta-1}}\left[e^{\frac{\gamma(t-(d_0-d_1\tau))}{\theta-1}} - e^{\frac{\gamma\tau}{\theta-1}}\right]\right)^{\theta-1} && (132)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(\int_0^T d\tau \left[e^{\frac{(\chi+d_1\gamma-a_1)\tau+\gamma(t-d_0)}{\theta-1}} - e^{\frac{(\gamma+\chi-a_1)\tau}{\theta-1}}\right]\right)^{\theta-1} && (133)\\
&= e^{a_0}\left(\frac{\theta-1}{\gamma}\right)^{\theta-1}\left(e^{\frac{\gamma(t-d_0)}{\theta-1}}\left(\frac{\theta-1}{\chi+\gamma d_1-a_1}\right)\left[e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}T}-1\right] - \frac{\theta-1}{\chi+\gamma-a_1}\left[e^{\frac{\chi+\gamma-a_1}{\theta-1}T}-1\right]\right)^{\theta-1} && (134)\\
&= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1}\exp\left[\gamma(t-d_0)\right] && (135)\\
&\quad \left(\left(\frac{\theta-1}{\chi+\gamma d_1-a_1}\right)\left[e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}T}-1\right] - e^{-\frac{\gamma}{\theta-1}(t-d0)}\frac{\theta-1}{\chi+\gamma-a_1}\left[e^{\frac{\chi+\gamma-a_1}{\theta-1}T}-1\right]\right)^{\theta-1}
\end{aligned}
$$

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

$$Y'_{m0}(t) = Age^{gt}(\theta-1)(Ae^{gt}-B)^{\theta-2} > 0. \tag{136}$$

The second time derivative is

$$Y''_{m0}(t) = (\theta-1)(Ae^{gt}-B)^{\theta-2}g^2A\left(A(\theta-2)(Ae^{gt}-B)^{-1}+1\right). \tag{137}$$

It is clear that (137) 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 (65) in which case the adoption margins are constants and we denote $T \equiv t_i$

$$
\begin{aligned}
Y_{m1} &= a\left(\int_{t_i}^{t-D} d\tau \int_{\tau}^{t-D} dv e^{\frac{\chi\tau+\gamma v}{\theta-1}}\right)^{\theta-1} && (138)\\
&= a\left(\frac{\theta-1}{\gamma}\int_{t_i}^{t-D} d\tau \left(e^{\frac{\chi\tau+\gamma(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)\tau}{\theta-1}}\right)\right)^{\theta-1} && (139)\\
&= a\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(e^{\frac{\chi(t-D)+\gamma(t-D)}{\theta-1}} - e^{\frac{\chi t_i+\gamma(t-D)}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(e^{\frac{(\chi+\gamma)(t-D)}{\theta-1}} - e^{\frac{(\chi+\gamma)t_i}{\theta-1}}\right)\right\}\right]^{\theta-1} && (140)\\
&= ae^{(\chi+\gamma)(t-D)}\left[\frac{\theta-1}{\gamma}\left\{\frac{\theta-1}{\chi}\left(1-e^{\frac{\chi(t_i-(t-D))}{\theta-1}}\right) - \frac{\theta-1}{\chi+\gamma}\left(1-e^{\frac{(\chi+\gamma)(t_i-(t-D))}{\theta-1}}\right)\right\}\right]^{\theta-1} && (141)\\
&= ae^{(\chi+\gamma)(t-D)}\left[\frac{(\theta-1)^2}{\gamma}\left\{\frac{1}{\chi}\left(1-e^{-\frac{\chi\Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1-e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right)\right\}\right]^{\theta-1} && (142)
\end{aligned}
$$

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 Y_{m1} because the transient part of (142) 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

$$
Y_m = \left(Y_{m0}^{\frac{1}{\theta-1}} + Y_{m1}^{\frac{1}{\theta-1}}\right)^{\theta-1}. \quad (143)
$$

As both Y_{m0} and Y_{m1} are increasing functions of time, it is immediate to verify that Y_m is increasing over time. To gain further insight on its behavior, note that (135) and (142) can be written as

$$
\begin{aligned}
Y_{m0} &= Ae^{(\chi+\gamma)t}\left(Be^{-\frac{\chi}{\theta-1}t} - Ce^{-\frac{\chi+\gamma}{\theta-1}t}\right)^{\theta-1}, && (144)\\
Y_{m1} &= Ae^{(\chi+\gamma)t}\left(De^{-\frac{\chi+\gamma}{\theta-1}t} - Ee^{-\frac{\chi}{\theta-1}t} + F\right)^{\theta-1}, && (145)\\
A &= e^{a_0}\left(\frac{(\theta-1)^2}{\gamma}\right)^{\theta-1}, && (146)\\
B &= \frac{e^{-\frac{\gamma}{\theta-1}d0}}{\chi+\gamma d_1 - a_1}\left(e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}T} - 1\right), && (147)\\
C &= \frac{1}{\chi+\gamma-a_1}\left(e^{\frac{\chi+\gamma-a_1}{\theta-1}T} - 1\right), && (148)\\
D &= \frac{e^{\frac{\chi+\gamma}{\theta-1}T}}{\chi+\gamma}, && (149)\\
E &= \frac{e^{\frac{\chi T-\gamma d_m}{\theta-1}}}{\chi}, && (150)\\
F &= \frac{\gamma e^{-\frac{\chi+\gamma}{\theta-1}d_m}}{\chi(\chi+\gamma)}. && (151)
\end{aligned}
$$

Using (143), we have that

$$Y_m = Ae^{(\chi+\gamma)t}\left(\underbrace{(D-C)e^{-\frac{\chi+\gamma}{\theta-1}t} - (E-B)e^{-\frac{\chi}{\theta-1}t} + F}_{h(t)}\right)^{\theta-1} \tag{152}$$

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

$$(\theta-1)h'^{-\frac{\chi+\gamma}{\theta-1}t} + \chi(E-B)e^{-\frac{\chi}{\theta-1}t}, \tag{153}$$

$$(\theta-1)^2h''^2(D-C)e^{-\frac{\chi+\gamma}{\theta-1}t} - \chi^2(E-B)e^{-\frac{\chi}{\theta-1}t}, \tag{154}$$

where

$$\begin{aligned}
D-C &= \frac{e^{\frac{\chi+\gamma}{\theta-1}T}}{\chi+\gamma} - \frac{1}{\chi+\gamma-a_1}\left(e^{\frac{\chi+\gamma-a_1}{\theta-1}T}-1\right), && (155)\\
&= \frac{1}{\chi+\gamma-a_1}\left[1+e^{\frac{\chi+\gamma}{\theta-1}T}\left(1-e^{-\frac{a_1}{\theta-1}T}\right) - \frac{a_1e^{\frac{\chi+\gamma}{\theta-1}T}}{\chi+\gamma}\right], && (156)\\
E-B &= \frac{e^{\frac{\chi T-\gamma d_m}{\theta-1}}}{\chi} - \frac{e^{-\frac{\gamma}{\theta-1}d0}}{\chi+\gamma d_1-a_1}\left(e^{\frac{\chi+\gamma d_1-a_1}{\theta-1}T}-1\right), && (157)
\end{aligned}$$

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 $\chi, \gamma \gg a_1, d_1$,

$$\begin{aligned}
D-C &= \frac{1}{\chi+\gamma}, && (158)\\
E-B &= \frac{1}{\chi}\left(e^{-\frac{\gamma d_0}{\theta-1}} + e^{\frac{\chi T}{\theta-1}}\left(e^{-\frac{\gamma d_m}{\theta-1}} - e^{-\frac{\gamma d_0}{\theta-1}}\right)\right). && (159)
\end{aligned}$$

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

$$-e^{-\frac{\chi+\gamma}{\theta-1}T} + e^{-\frac{\chi T+\gamma d_0}{\theta-1}} + e^{-\frac{\gamma d_m}{\theta-1}}\left(1-e^{-\frac{\gamma}{\theta-1}(d_o-d_m)}\right). \tag{160}$$

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

$$(\chi+\gamma)e^{-\frac{\chi+\gamma}{\theta-1}t} - \chi e^{-\frac{\chi}{\theta-1}t}\left(e^{-\frac{\gamma d_0}{\theta-1}} + e^{\frac{\chi T-\gamma d_m}{\theta-1}}\left(1-e^{-\frac{\gamma}{\theta-1}(d_o-d_m)}\right)\right). \tag{161}$$

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 (161) to be positive is $d_0 < T$. In this case we would have an S-shape.

Next, we study the behavior of the share

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

The quotient in the previous expression can be written as,

$$y \equiv \left(\frac{Y_m}{Y_o}\right)^{\frac{1}{\theta-1}} = \frac{Y_{m0}^{\frac{1}{\theta-1}} + Y_{m1}^{\frac{1}{\theta-1}}}{Y_o^{\frac{1}{\theta-1}}} = C_0 e^{g_0 t} + C_1 e^{g_1 t} h(t) \tag{162}$$

where C_0, C_1 are two constants, $g_0 < g_1$, and $h(t)$ is given by equation (142),

$$\frac{1}{\chi}\left(1 - e^{-\frac{\chi \Delta t}{\theta-1}}\right) - \frac{1}{\chi+\gamma}\left(1 - e^{-\frac{(\chi+\gamma)\Delta t}{\theta-1}}\right) \tag{163}$$

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

$$\begin{aligned} \dot{y} &= g_0 C_0 e^{g_0 t} + g_1 C_1 e^{g_1 t} h + C_1 e^{g_1 t} \dot{h} && (164)\\ \ddot{y} &= g_0^2 C_0 e^{g_0 t} + g_1^2 C_1 e^{g_1 t} h + 2 g_1 C_1 e^{g_1 t} \dot{h} + C_1 e^{g_1 t} \ddot{h} && (165) \end{aligned}$$

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

$$g_0^2 C_0^2 e^{2 g_0 T} - (1 + C_0 e^{g_0 t})(g_0^2 C_0 e^{g_0 T} + C_1 \gamma) < 0. \tag{166}$$

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

$$\lim_{t\to\infty} \dot{y}(t)^2 - (1+y(t))\ddot{y}(t)^2 \sim \lim_{t\to\infty} \left(\frac{\chi}{\gamma(\chi+\gamma)} - 1\right) e^{2 g_1 t}. \tag{167}$$

Hence, the asymptotic behavior depends on whether $\chi \lessgtr \gamma(\chi+\gamma)$. Note that given that both χ and γ are on the order of $1/100$, we have that $\chi > \gamma(\chi+\gamma)$, and hence s is asymptotically convex.[38] ∎

F Computation of the Half-Lives Reported in the Main Text

We use the microfoundations for A_{Modern} and A_{Old} derived in (69) and (73), respectively. As in the baseline model, we assume that growth is evenly split between χ and γ, both for Modern and pre-Modern growth. This implies that to match long-run growth rates of 2% and .2% for Modern and pre-Modern growth, we have that $\chi = \gamma = (1-\alpha)1\%$ and $\chi_o = \gamma_o = (1-\alpha).1\%$. Using that $\theta = 1.3$, we have that

$$\frac{A_{Old}}{A_{Modern}} = \left(\frac{\chi(\chi+\gamma)}{\chi_o(\chi_o+\gamma_o)}\right)^{\theta-1} = 4.0 \tag{168}$$

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

Taking the ratio of intensive margins of Modern and pre-Modern as the ratio of the intercept of our estimated diffusion regression to the post-1820 average (see regression (1) in Table 4), we have that $\frac{a_{Old}}{a_{Modern}} = \exp(-.24 - \frac{-.24-180\cdot.0034}{2}) = 1.2$. Finally, for adoption lags, we take D_{Old} from the intercept of regression (1) in Table 3 and D_{Modern} as the post-1820 average. We have $D_{Old} = \exp(4.37) = 80$ years and $D_{Modern} = \exp\left(\frac{4.37+4.37-180\cdot.011}{2}\right) = 30$.

Using equations (47) and (48) and the values we have calculated we can proceed and calculate the half-lives. The total we find that

$$
\begin{aligned}
t_{1/2}^{\text{gap}} &= 39 + 86 = 125, && (169)\\
t_{1/2}^{\text{growth}} &= 39 + 124 = 163. && (170)
\end{aligned}
$$

Finally, note that the contribution of adoption lags, is approximately the average adoption lag in the sample,

$$D_{Modern} - \frac{g_{Old} D_{Old}}{\chi + \gamma - g_{Old}} = 30 + \frac{.2\%}{1.8\%} 80 = 39 \text{ years.} \tag{171}$$