If Technology Has Arrived Everywhere, Why Has Income Diverged?

Diego Comin Martí Mestieri*

July 20, 2017

Abstract

We study the cross-country evolution of technology diffusion over the last two centuries. We document that adoption lags between poor and rich countries have converged, while the intensity of use of adopted technologies of poor countries relative to rich countries has diverged. The evolution of aggregate productivity implied by these trends in technology diffusion resembles the actual evolution of the world income distribution in the last two centuries. Cross-country differences in adoption lags account for a significant part of the cross-country income divergence in the nineteenth century. The divergence in intensity of use accounts for the divergence during the twentieth century.

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

*Comin: Dartmouth College, NBER and CEPR, Department of Economics, Dartmouth College, 6106 Rockefeller Hall, Hanover, NH 03755, diego.comin@dartmouth.edu; Mestieri: Northwestern University and CEPR, Department of Economics, Northwestern University, 2211 Campus Dr., Evanston, IL 60208, marti.mestieri@northwestern.edu. This paper subsumes the working paper "An Intensive Exploration of Technology Adoption." We are grateful to Richard Rogerson (the Editor) and two anonymous referees for comments and suggestions. We have also benefitted from comments of Thomas Chaney, Gino Gancia, Christian Hellwig, Chad Jones, Pete Klenow, Franck Portier, Mar Reguant, Doug Staiger, Jon Van Reenen and seminar participants at Aix-Marseille, Boston University, Brown, CERGE-EI, Dartmouth, Edinburgh, HBS, Harvard, LBS, Minnesota Fed, OECD, Pompeu Fabra, Stanford GSB, SIEPR, Autònoma de Barcelona, University of Toronto, TSE and the World Bank for useful comments and suggestions. Comin acknowledges the generous support of the National Science Foundation and of the Institute for New Economic Thinking. Mestieri acknowledges the generous support of the Agence Nationale de la Recherche while at TSE. All remaining errors are our own.

1 Introduction

Cross-country differences in productivity have increased dramatically over the last 200 years. The ratio between the average (per capita) income in Maddion's "Western countries" and the average income among the non-Western countries was 1.9 in 1820. By 2000, the average income ratio was 7.2.[1]

Observed differences in productivity growth can be driven by differences in the technology used in production or by other non-technological factors. This paper studies the role played by technology diffusion on the evolution of productivity growth over the last two centuries. In particular, we study *(i)* how the technology diffusion processes have evolved over time in rich and poor countries and *(ii)* how potential differences in the evolution of the diffusion process may have contributed to the dynamics of productivity growth.

We study the evolution of technology diffusion by taking advantage of a property of diffusion curves identified by Comin and Hobijn (2010). To illustrate it, consider Figure 1 which represents the evolution of the log of the number of telephone lines over GDP in various countries. These diffusion curves can be well approximated by a common curve plus country-specific vertical and horizontal shifters. Comin and Hobijn (2010) use a model to relate the horizontal shifters to the time it takes for the technology to arrive to each country (i.e. the adoption lags). We extend the model in Comin and Hobijn (2010) by relating the vertical shifters to country-technology specific factors such as sectoral distortions (e.g., Restuccia and Rogerson, 2008) and barriers to adoption (e.g., Parente and Prescott, 1994) that impact the intensity with which the technology is used. We denote the technology-country specific parameters that captures the vertical shifters in the diffusion curves the intensity of adoption parameter.[2]

We estimate our model of diffusion using data for 25 major technologies invented over the last 200 years for a sample of 139 countries. Our estimation strategy takes care of the effect of income on the demand for technology,[3] and of the effect on production costs of economy-wide factors that have a symmetric effect across technologies.[4] The average estimated adoption lag is 42 years. The average intensity of adoption in Non-Western countries is half (47%) of the level of Western countries. We document significant dispersion on both margins of adoption

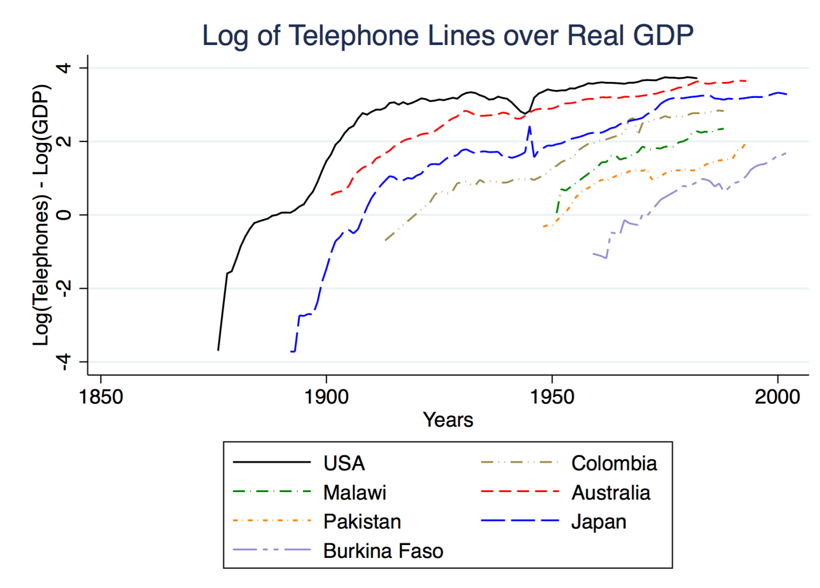
[1]These differences in income growth can be attributed to differences in total factor productivity (TFP) growth (Klenow and Rodríguez-Clare (1997) and Clark and Feenstra (2003)). Maddison (2004) classifies as "Western countries" Austria, Belgium, Denmark, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland, Untied Kingdom, Japan, Australia, New Zealand, Canada and the United States.

[2]See Clark (1987) for a detailed account of the importance of this margin in the case of spinning spindles, which is one of the technologies in our data set.

[3]In the baseline calculations, the restriction that our model has a balanced growth path implies that the income elasticity of technology demand is equal to one. In section 5.3, we show that our findings are robust to allowing for non-homotheticities in the demand for technology.

[4]As discussed in section 3.1, we can do this by taking advantage of the three dimensional nature of our data.

Figure 1: Diffusion Curves of Telephone Lines



across countries and technologies. For example, comparing the 90th to the 10th percentiles of our estimates, we find 10-fold differences in adoption lags and 8-fold differences in the intensity of adoption.

To study how technology diffusion has evolved over the last two centuries, we take advantage of the variation in invention dates across the technologies in our sample. We document two new facts: *(i)* cross-country differences in adoption lags have narrowed over the last 200 years. That is, adoption lags have declined more over time in poor/slow adopter countries than in rich/fast adopter countries, *(ii)* the intensity of use of technology has diverged. That is, the gap in the use of individual technologies between rich and poor countries has widened and it is larger for recent technologies than for those invented 200 years ago.

After documenting these cross-country trends in technology diffusion processes, we evaluate their effect on the evolution of productivity growth across countries. We take advantage of the fact that our model has a parsimonious aggregate representation. This allows us to feed in the estimated trends in adoption and compute the implied evolution of productivity growth.[5]

Cross-country differences in the diffusion margins in 1820 induce a world income distribution that is similar to the actual distribution. Our main result is that cross-country differences in the evolution of technology diffusion imply an evolution of aggregate TFP that accounts well for the evolution of the world income distribution over the last 200 years. For example, we find that differences in the evolution of technology diffusion increase by a factor of 2.9 the

[5]By using a model to compute the effect of technology diffusion to productivity growth, we avoid the biases from the endogeneity of technology adoption that would plague standard regression estimates.

income gap between Western and non-Western countries during the period 1820-2000. This represents 75% of the actual increase in the income gap, which grew by a factor of 3.9. More generally, the aggregate representation implied by our country-technology estimates can reproduce the evolution of the world income distribution when we look at different time periods, or when we split the countries by income quintiles or continents.

This paper is related to the literature analyzing the channels driving the Great Divergence. One strand of the literature has related the divergerce to the evolution of Solow residuals which are taken as exogenous processes (Easterly and Levine, 2001, Clark and Feenstra, 2003, Lucas (2000), and Gancia *et al.* (2013)).[6] Another strand of the literature uses aggregate production functions with biased technical change to study whether the divergence is due to efficiency differences or to appropriate technology considerations (Allen (2012) and Jerzmanowski (2007)). Often theories that emphasize the importance of the direction of technical change have been framed in the context of trade models (see, e.g., O'Rourke *et al.* (2012)).[7] Our focus differs from these studies in that we measure the relevant dimensions of technology diffusion directly from diffusion curves and analyze the implications for productivity growth. Unlike these studies we do not intend to explain why the use of technology diverged between rich and poor countries.[8]

The rest of the paper is organized as follows. Section 2 develops the model of technology diffusion and productivity growth. Section 3 presents the estimation strategy. Section 4 estimates the two margins of adoption and documents their cross-country evolution. Section 5 quantifies the effect of the technology dynamics on the evolution of the world income distribution. Section 6 concludes.

2 Model

We present a model of technology adoption and growth. The goals of the model are two. First, to derive a functional form for the shape of diffusion curves for specific technologies. Second, to provide an aggregate represnetation that allows us to explore how the adoption margins affect productivity growth.

[6]Lucas (2000) studies the evolution of the world income distribution using a model that assumes a negative relationship between the time a country takes off and the TFP growth it experiences during the transition. He predicts either no growth or a strong convergence. Instead, our model has more nuanced predictions and captures well the entire world income distribution.

[7]Galor and Mountford (2006) develop a model where trade affects asymmetrically fertility decisions in developed and developing economies, due to different initial endowments of human capital, leading to different evolutions of productivity growth.

[8]Trade-based theories of the Great Divergence, need to confront two facts. Prior to 1850, the technologies brought by the Industrial Revolution were unskilled-biased rather than skilled-biased (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.

We follow different approaches to model the two adoption margins. For simplicity, we assume an exogenous adoption lag for each technology-country pair. To model the intensity of use, we introduce various frictions that differ in their scope (economy-wide vs. technology-specific) and nature (taxes on labor income vs. intermediate goods prices). Once we derive the expression for the diffusion of a specific technology, we define formally the intensity of use parameter and study which of the factors introduced in the model affect it.

2.1 Preferences and endowments

There is a unit measure of identical households in the economy.[9] Each household supplies inelastically one unit of labor, for which they earn a wage w_t at time t. Households can save in domestic bonds which are in zero net supply. The utility of the representative household is given by

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

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

$$\dot{B}_t + C_t = w_t + r_t B_t + T_t, \tag{2}$$

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

where T_t denotes government transfers to consumers, 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

At a given instant of time t, the world technology frontier is characterized by a set of technologies and a set of vintages specific to each technology. Each instant, a new technology τ exogenously appears. We denote a technology by the time it was invented, τ. 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 v_τ of technology τ available at time t ($> \tau$) is $[\tau, t]$. That is, $v_\tau \in [\tau, t]$, and v_τ takes the value of the time period in which vintage v_τ appears in the technology frontier.

The productivity of a technology-vintage pair, $Z(\tau, v_\tau)$, is common across countries and

[9] We identify an economy with a country in our empirical application. We refer interchangeably to economy and country in what follows.

is given by

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

where $(\sigma+\gamma)\tau$ is the productivity level associated with the first vintage of technology τ and $\gamma(v_\tau - \tau)$ captures the productivity gains associated with the subsequent introduction of new vintages $v_\tau \geq \tau$. By allowing subsequent vintages of the same technology to be more productive, we want to capture the idea that productivity not only increases through the introduction of new technologies, but also that technologies become more productive over time (e.g., Aghion and Howitt, 1992, Klette and Kortum, 2004).

Economies are typically below the world technology frontier. Let $D_{\tau c}$ denote the age of the best vintage available of technology τ used in production in country c. $D_{\tau c}$ reflects the time lag between the best vintage in use and the world frontier for technology τ; that is, the *adoption lag*.[10] For technology τ, the set of vintages available in economy c is $v_{\tau c} = [\tau, t - D_{\tau c}]$.[11]

2.3 Production

Now that we have defined the notion of technology, we can specify how technology is used in production. To simplify notation, we try to eliminate country subindeces. Technologies are embodied in intermediate goods. It takes one unit of aggregate output to produce one unit of intermediate good. Associated with each intermediate good, there is a service, $Y_{\tau,v,t}$, that results from combining units of the vintage-specific intermediate good, $X_{\tau,v,t}$, with labor, $L_{\tau,v,t}$.[12] The productivity in the production of these services depends on two variables: the productivity of the technology-vintage, $Z(\tau, v)$, and the economy-wide exogenous TFP level, e^{χ_t}. Equation (5) formalizes the production function for vintage-level services,

$$Y_{\tau,v,t} = e^{\chi_t} Z(\tau, v) X^{\alpha}_{\tau,v,t} L^{1-\alpha}_{\tau,v,t}, \tag{5}$$

with $\alpha \in (0, 1)$.

Vintage-level services are used to produce differentiated outputs associated with each technology τ, indexed by $i \in [0, N_\tau]$. In particular, differentiated output i, denoted $Y^i_{\tau,t}$, is

[10]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).

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

[12]See Comin and Hobijn (2010) for a formulation where vintage-specific technologies are embodied in investment. The key difference between these two approaches is that, in ours, intermediate goods depreciate instantaneously. This makes their demand a static problem. Otherwise, the two formulation are isomorphic. We use below this similarity to calibrate α to match the capital share as it is usually done in endogenous growth models with expanding varieties (e.g., Jones and Williams, 2000).

produced combining vintage-level services according to

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

where $Y^i_{\tau,v,t}$ denotes the amount of vintage v services used by the i^{th} differentiated output producer associated with technology τ at time t. Note that, since we are assuming a constant elasticity of substitution production function in (6), it is optimal for every producer of differentiated output to use all available vintage-level services. There are N_τ differentiated outputs associated with technology τ. Therefore, the total demand for services associated with a vintage (v, τ) is equal to its supply:

$$Y_{\tau,v,t} = \int_0^{N_\tau} Y^i_{\tau,v,t} di. \tag{7}$$

We can think of these as corresponding to the number of companies that use a given technology to produce a good or a service. For simplicity, we take N_τ as a parameter and allow it to vary across countries.[13]

The final output associated with technology τ, $Y_{\tau,t}$, is produced by combining the available differentiated outputs

$$Y_{\tau,t} = N_\tau^{-(\psi'-\psi)} \left(\int_0^{N_\tau} \left(Y^i_{\tau,t}\right)^{1/\psi'} di \right)^{\psi'}. \tag{8}$$

Note that in this expression, there may be productivity gains from the number of differentiated outputs, N_τ, as in Benassy (1996).

Aggregate output, Y_t, results from aggregating the available technology-specific outputs, $Y_{\tau,t}$, as follows:

$$Y_t = \left(\int_{-\infty}^{\bar{\tau}} Y_{\tau,t}^{\frac{1}{\theta}} d\tau \right)^{\theta}, \qquad \text{with} \quad \theta > 1 \tag{9}$$

where $\bar{\tau}$ is the last technology adopted in the country.

Taxes We introduce two frictions in the form of proportional taxes: one to labor, $\zeta_{L,t}$, and another to intermediate goods, $\zeta_{\tau x}$. We assume that the labor tax $\zeta_{L,t}$ faced by producers is the same across the economy (but may change over time). However, $\zeta_{\tau x}$ can be technology-specific.[14] This feature reflects interest groups capturing the political system to limit the use

[13] See the working paper version, Comin and Mestieri (2010) for a simple approach to make it endogenous.

[14] For simplicity, we assume $\zeta_{\tau x}$ to be constant across all vintages of technology τ. Simplicity also motivated our choice of having only two taxes, one economy-wide and another sector-specific. The logic for sector-specific and economy-wide taxes carries over to a tax on sectoral output or interchanging the role of economy-wide and sector-specific taxes to the labor and intermediate tax.

of a given technology[15] as well as other technology-specific frictions that affect the intensity of technology use.[16] In this paper, we do not take a stand on the nature and relative importance of these factors. The proceeds from taxes are returned to consumers through lump sum transfers.

2.4 Equilibrium

The numeraire in the economy is aggregate output. To save on notation, we focus on the case in which all goods and services are produced competitively.[17]

Given a sequence of adoption lags and number of differentiated outputs $\{D_\tau, N_\tau\}_{\tau=-\infty}^{\infty}$, a sequence of exogenous TFP $\{e^{\chi_t}\}_{t=-\infty}^{\infty}$, and tax rates $\left\{\zeta_{L,t}\right\}_{t=t_0}^{\infty}$, $\{\zeta_{\tau x}\}_{\tau=-\infty}^{\infty}$, a competitive equilibrium in this economy is defined by consumption, output, and labor allocations paths $\{C_t, L_{\tau,v,t}, Y_{\tau,v,t}, X_{\tau,v,t}, Y^i_{\tau,t}, Y^i_{\tau,v,t}, Y_{\tau,t}, Y_t, T_t\}_{t=t_0}^{\infty}$ and prices $\{P_{\tau,t}, P^i_{\tau,t}, P_{\tau,v,t}, W_t, R_{x,t}, r_t\}_{t=t_0}^{\infty}$, such that

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

$$\frac{\dot{C}_t}{C_t} = r_t - \rho, \tag{10}$$

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

2. Firms maximize profits taking prices as given. The resulting optimality conditions yield the demand for labor and intermediate goods for each technology and vintage, equations (B.1) and (B.2), for the output produced with a vintage (equation B.4), for differentiated output (equation B.14) and for the output produced with a technology (equation B.13). Furthermore the prices of technology-vintage output, differentiated output and technology-levl output are respectively determined by equations (B.3), (B.6) and (B.7).

3. Market clearing. Labor and intermediates market clearing conditions are

$$L_t = \int_{-\infty}^{\bar{\tau}} \int_{\tau}^{t-D_\tau} L_{\tau,v,t} dv_\tau d\tau = 1, \tag{11}$$

$$X_t = \int_{-\infty}^{\bar{\tau}} \int_{\tau}^{t-D_\tau} X_{\tau,v,t} dv_\tau d\tau. \tag{12}$$

[15]See Olson (1982), Parente and Prescott (1994), Acemoglu *et al.* (2004) and Rios-Rull and Krusell (1999) for theoretical arguments and Comin and Hobijn (2009b) and Mokyr (2000) for a systematic empirical analysis.

[16]For example, variation in the abundance of some complementary factors (e.g., roads for cars or trucks), or on the relevance of access to credit to finance the technologies (e.g., Comin and Nanda, 2014)

[17]Our results hold in the case of a monopolistic competitive environment, see Comin and Mestieri (2010).

Additionally, total demand for services associated with a vintage $(v,\ \tau)$ is equal to its supply, (7). The output associated with a technology $Y_{\tau,t}$ equals its demand from the final good producer.

4. The resource constraint holds, $Y_t = C_t + X_t$.
5. The government balances its budget constraint period-by-period by setting transfers T_t as follows:

$$T_t = \int_{-\infty}^{\bar{\tau}} \frac{\zeta_{\tau x} P_{\tau,t} Y_{\tau,t}}{1+\zeta_{\tau x}} d\tau + \frac{\zeta_{L,t}(1-\alpha)Y_t}{1+\zeta_{L,t}}. \tag{13}$$

We note that combining (11) and the labor demands from the producers of vintage-level services assicated to all technologies, (B.2), it follows that the wage rate is given by

$$W_t = \frac{(1-\alpha)Y_t}{(1+\zeta_{L,t})L}. \tag{14}$$

Finally, in what follows, we omit time subscripts and make the time dependence of variables implicit to ease notation (e.g., Y_t becomes Y) .

2.5 Aggregate representation

We show that there exists an aggregate production function representation for both final output associated with a technology, Y_τ, and aggregate output, Y. The aggregate representation of final output associated with a technology, Y_τ, provides the structural diffusion curve that we use in the estimation of the adoption margins. The aggregate representation of aggregate output, Y, provides the theoretical foundation for linking technology diffusion patterns to the evolution of aggregate output. We use this result in our quantitative exercise in Section 5.

Proposition 1 (Technology-level representation): *(i) The level of output associated with technology τ, Y_τ can be represented by the Cobb-Douglas function*

$$Y_\tau = A_\tau X_\tau^\alpha L_\tau^{1-\alpha}, \tag{15}$$

where $L_\tau = \int_\tau^{t-D_\tau} L_{\tau,v} dv$, $X_\tau = \int_\tau^{t-D_\tau} X_{\tau,v} dv$, and technology level TFP is equal to

$$A_\tau = e^{\chi_t} N_\tau^{(\psi-1)} \left(\int_\tau^{t-D_\tau} Z(\tau,v)^{1/(\mu-1)} dv \right)^{\mu-1} \tag{16}$$

$$= \left(\frac{\mu-1}{\gamma}\right)^{\mu-1} \underbrace{e^{\chi_t}}_{\text{Exog. TFP}} \underbrace{N_\tau^{(\psi-1)}}_{\text{Variety of services}} \underbrace{e^{\sigma\tau+\gamma(t-D\tau)}}_{\text{Embodiment}} \underbrace{\left(1-e^{\frac{-\gamma}{\mu-1}(t-D_\tau-\tau)}\right)^{\mu-1}}_{\text{Variety of vintages}} \tag{17}$$

(ii) The demand for output associated with technology τ is given by

$$Y_\tau = Y P_\tau^{-\theta/(\theta-1)}, \tag{18}$$

where

$$P_\tau = \frac{\left(\frac{(1+\zeta_{\tau x})}{\alpha}\right)^\alpha \left(\frac{(1+\zeta_L)W}{1-\alpha}\right)^{1-\alpha}}{A_\tau}. \tag{19}$$

(iii) The demand for technology τ intermediate goods and labor used in the production of technology τ services are respectively

$$X_\tau = \frac{\alpha Y_\tau P_\tau}{(1+\zeta_{\tau x})} = \frac{\alpha Y P_\tau^{-1/(\theta-1)}}{(1+\zeta_{\tau x})}, \tag{20}$$

$$L_\tau = \frac{(1-\alpha) Y_\tau P_\tau}{(1+\zeta_L)W} = \frac{(1-\alpha) Y P_\tau^{-1/(\theta-1)}}{(1+\zeta_L)W}. \tag{21}$$

□

The first part of Proposition 1 defines technology-level TFP, A_τ. A_τ has four components. The first is the standard economy-wide exogenous TFP level. The second captures the productivity gains associated with how many producers are using technology τ to produce their differentiated services. The third term captures the productivity embodied in the most advanced vintage adopted. The final term captures the productivity gains associated with the range of vintages used in production. Because the range of vintages used diminishes with the adoption lag, so does the variety effect. The first three terms are log-linear. The variety effect generates the log-concavity of A_τ.

The second part of Proposition 1 provides an alternative way to characterize the output associated with technology τ. Equation (18) implies that Y_τ depends only on aggregate output, Y and of the factors that affect the price of technology-level output, P_τ. These are the technology-level TFP, A_τ, the economy-wide wage rate, W, and the taxes on intermediate goods and labor, $\zeta_{\tau x}$ and ζ_L.

The third part of the proposition characterizes the factors used to produce services associated with technology-τ. These predictions are relevant in the estimation because for some technologies we have measures of diffusion that capture the units of output produced, while for others we have information on the intermediate goods that embody it, which corresponds to (20).

Diffusion curves and adoption margins Equations (18) and (20) characterize the evolution of output (or intermediate inputs) associated with a technology. These are the diffusion measures that we observe in our data (e.g., electricity produced or number of spindles). Next, we show how the adoption margins affect the diffusion curves. For illustration purposes, we focus on the case of output, Y_τ.

Combining the demand for technology output, (18), with the definition of sectoral aggregate productivity and equilibrium prices, (17) and (19), we obtain the following expression for the output of technology τ, Y_τ^c, scaled by total country income Y^c

$$\frac{Y_\tau^c}{Y^c} = \left(\xi^c e^{\sigma\tau} \underbrace{(1+\zeta_{\tau x}^c)^{-\alpha} N_\tau^{c(\psi-1)}}_{\text{Intensity Parameter } I_\tau^c} e^{\gamma(t-D_\tau^c)} \left(1 - e^{\frac{-\gamma}{\mu-1}(t-D_\tau^c-\tau)}\right)^{\mu-1} \right)^{\frac{\theta}{\theta-1}}, \tag{22}$$

where the superscript c denotes country-specific variables and ξ^c represents country-wide factors such as the labor tax or the exogenous TFP component.[18]

In addition to ξ^c and the technology-specific productivity term, $e^{\sigma\tau}$, technology-τ output depends on the intensity of use parameter, $I_\tau^c \equiv (1+\zeta_{\tau x}^c)^{-\alpha} N_\tau^{c(\psi-1)}$, and the adoption lag, D_τ^c. The intensity of use term I_τ^c in (22) captures the effect of technology-specific distortions $\zeta_{\tau x}^c$ and of the number of differentiated users of the technology, N_τ^c, on the level of output associated with a technology.[19] The intensity of use parameter affects the level of technology-level output. That is, it is a vertical shifter of the diffusion curve (22).[20] The parameter D_τ^c shifts the diffusion curve horizontally. If we increase D_τ^c by T years, it will take T additional years to reach any given level of technology-τ output. Note also that Y_τ^c inherits the non-linearities of technology-τ productivity, A_τ, in D_τ^c that we have already discussed above.

After studying the implications of our model economy at the technology level, next we characterize the level of aggregate output and productivity (we drop the superscript c in the exposition of the result).

Proposition 2 (Aggregate representation): *Suppose that $\zeta_{\tau x} \equiv \zeta_x$ is the same across technologies. Then, the economy has an aggregate representation where output is given by*

$$Y = AX^\alpha L^{1-\alpha} = \left(\frac{\alpha}{1+\zeta_x}\right)^{\frac{1}{1-\alpha}} A^{\frac{1}{1-\alpha}} \tag{23}$$

with $X = \int_\tau X_\tau d\tau$, $L = \int_\tau L_\tau d\tau$, and

$$A = \left(\int_\tau^{t-D_\tau} A_\tau^{1/(\theta-1)} d\tau\right)^{\theta-1}. \tag{24}$$

[18] In particular, $\xi^c = e^{\chi_t^c} \alpha^{-\alpha} \left(\frac{(1+\zeta_L^c)W^c}{1-\alpha}\right)^{-(1-\alpha)}$.

[19] Our model introduces two different mechanisms as sources of the intensity of use for two reasons. First, to reflect the two most prominent strands in the literture: the barriers to adoption (Parente and Prescott, 1994) and the sectoral distortions (Restuccia and Rogerson, 2008). Second, we want to stress that the intensity of use may impact sectoral TFP as it is the case with the number of differentiated users of tecnology, N_τ.

[20] Increasing I_τ by one percent, increases Y_τ by $\frac{\theta}{\theta-1}$ percent.

□

Proposition 2 implies that the aggregate production function in our economy takes a standard Cobb-Douglas form, where X and L denote the aggregate number of intermediate goods and labor, respectively, and A is the aggregate TFP. Moreover, we show that total output Y is proportional to $A^{\frac{1}{1-\alpha}}$. This implies that output dynamics are completely determined by the dynamics of aggregate TFP, A.[21] We use this property extensively to analyze the output dynamics implied by our measures of technology diffusion in Section 5.[22] A key difference of our model from the traditional one-sector neoclassical models is that our model provides a theory of aggregate TFP in the sense that aggregate TFP depends on technology-level TFPs, A_τ according to (24). In turn, technology-level TFPs, A_τ, depend on the adoption lags, D_τ, the intensity of use of technologies, I^c_τ, and the economy-wide exogenous TFP level, χ_t. Thus our theory relates measures of adoption at the technology level to aggregate TFP.

Finally, we note that to guarantee the existence of a balanced growth path, a sufficient condition, which we take as a benchmark, is that both adoption margins are constant across technologies, $D_\tau = D$ and $N_\tau = N$.[23] If we make the simplifying assumption that $\theta = \mu$,[24] aggregate TFP can be computed in closed form,

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

Naturally, a higher intensity of adoption, N, a shorter adoption lag, D, and a higher exogenous TFP component lead to higher aggregate TFP. Along this balanced growth path, productivity grows at rate $\sigma + \gamma$, and output grows at rate $(\sigma+\gamma)/(1-\alpha)$.[25]

[21] The literature on growth and development accounting (e.g., Klenow and Rodríguez-Clare, 1997) has focused on whether cross-country differences in productivity arise from differences in TFP or in capital per capita (or their growth rates). In light of that literature, one might wonder whether cross-country differences in adoption margins show up in aggregate TFP or in aggregate capital, here proxied by X. The answer to this question depends on *(i)* whether technologies are embodied in capital or disembodied, and *(ii)* whether price deflators adequately correct for the quality of capital which captures the type of capital good and the productivity embodied in it. In our empirical analysis of the drivers of productivity growth we abstract from this question by computing the model predictions for productivity growth (as opposed to TFP growth).

[22] We use as a benchmark for our calibration the case with constant taxes because, in this case, the growth rate of the economy is solely pinned down by the evolution of sectoral TFPs, A_τ, through the aggregation equation, (24). If, instead, we allowed for technology-specific taxes, there would be a direct effect of taxes on aggregate output that is not related to technology. Proposition 1 in the online appendix presents the aggregation results with technology-specific taxes.

[23] Comin and Mestieri (2010) show in their micro-founded model of adoption that this is a necessary and sufficient condition for a balanced growth path.

[24] Our empirical analysis in Section 4 suggests that this is a reasonable approximation.

[25] For discounted utility to be bounded, it is required that $(\sigma+\gamma)/(1-\alpha) < \rho$.

3 Estimation Strategy

In this section, we describe the estimation procedure used to measure the adoption margins for each technology-country pair. Our estimation strategy relies on using the diffusion equations derived in the theoretical section to identify the adoption margins for each country-technology pairs. In a nutshell, we use the functional form of the diffusion curve derived in equation (22) to estimate the diffusion of different technologies across countries allowing for several parameters in the diffusion curve to be country-technology specific.

3.1 Estimating Equations

We derive our estimating equation from a log-linearized version of the diffusion curve (22) derived in Section 2.5. Denoting by lowercase the logs of uppercase variables, combining the demand for sector τ output (18), the sectoral price deflator (19), the expression for the equilibrium wage rate (14), and the expression for sectoral TFP, A_τ, (17), we obtain the demand equation

$$y_\tau = y + \frac{\theta}{\theta - 1}\left[a_\tau - (1-\alpha)(y-l) - \alpha \ln\left((1+\zeta_x)/\alpha\right)\right]. \tag{26}$$

The terms $(1-\alpha)(y-l)$ and $\alpha \ln\left((1+\zeta_x)/\alpha\right)$ correspond to the logarithm of the sectoral price index P_τ. We then use the fact that γ takes small values to simplify the expression of sectoral TFP a_τ to its first order approximation in γ (see Appendix B for calculation details),

$$a_\tau \simeq \chi_t + (\psi-1)n_\tau + (\sigma+\gamma)\tau + (\mu-1)\ln(t-\tau-D_\tau) + \frac{\gamma}{2}(t-\tau-D_\tau). \tag{27}$$

Substituting (27) in (26) gives our main estimating equation. Explicitly indexing country-specific variables with superscript c, technology-specific by τ, and denoting time dependence by a subindex t, the estimating equation is

$$y^c_{\tau t} = \beta^c_{t0} + \beta^c_{\tau 1} + y^c_t + \beta_{\tau 2} t + \beta_{\tau 3}\left((\mu-1)\ln(t-D^c_\tau-\tau) - (1-\alpha)(y^c_t - l^c_t)\right) + \varepsilon^c_{\tau t}, \tag{28}$$

where $\varepsilon^c_{\tau t}$ is a country c, technology τ and time t error term that we introduce to account for potential discrepancies between our model and data.[26] Equation (28) shows that we can express the (log of) output produced with technology τ, $y^c_{\tau t}$, as the summation of a country time-varying term, β^c_{t0}, a country-technology specific constant, $\beta^c_{\tau 1}$, a log-linear term in time with coefficient $\beta_{\tau 2} = \gamma/2$ that captures technology productivity growth, a log-linear term in income y^c_t with coefficient equal to one, and a non-linear function of the adoption lag with

[26]This error term could be rationalized in our theory as time-dependent shocks to the number of users or to technology-specific taxes.

coefficient $\beta_{\tau 3} = \frac{\theta}{\theta - 1}$.[27]

The country-technology intercept, $\beta^c_{\tau 1}$, is given by

$$\beta^c_{\tau 1} = \beta_{\tau 3} \left(\underbrace{(\psi - 1)\, n^c_\tau - \ln\left((1 + \zeta^c_{\tau x})\right)}_{\text{Log-Intensity Of Use, } \ln I^c_\tau} + \alpha \ln \alpha + \left(\sigma + \frac{\gamma}{2}\right)\tau - \frac{\gamma}{2} D^c_\tau \right). \tag{29}$$

The first two terms of expression (29) capture variation in technology-τ output associated with the number of producers use the technology to produce differentiated services and the number of units of technology used per producer. These two components define the intensity of use parameter, I_τ. The last two terms in (29) reflect the effect on the intercept of the initial level of productivity embodied in the technology.

The term β^c_{t0} captures the variation in $y^c_{\tau t}$ induced by country-wide factors such as exogenous TFP. Specifically,

$$\beta^c_{t0} = \beta_{\tau 3}(\chi^c_t - (1 - \alpha) \ln\left((1 + \zeta^c_L)/(1 - \alpha)\right)). \tag{30}$$

Aggregate output, y^c_t, enters in the estimation equation (28) because the level of aggregate demand affects the demand for technology. The coefficient on aggregate output in the estimating equation (28) is one (and it is not estimated). In Section 4.4 we relax this assumption and estimate directly the Engel curve from the data to assess the robustness of our estimates and findings.

Some of the variables in our data set measure the number of units of the input that embody the technology (e.g. number of computers) rather than output. For this case, we derive an estimating equation for input measures. We take the logarithm of technology-τ intermediate goods, equation (20), and combine it with the sectoral price deflator (19), the equilibrium wage rate (14) and the expression for sectoral TFP A_τ, (17). Approximating A_τ using (27), we obtain the following expression, which inherits the properties from (28),[28]

$$x^c_{\tau t} = \beta^c_{t0} + \beta^c_{\tau 1} + y^c_t + \beta_{\tau 2} t + \beta_{\tau 3}\left((\mu - 1)\ln(t - D^c_\tau - \tau) - (1 - \alpha)(y^c_t - l^c_t)\right) + \varepsilon^c_{\tau t}. \tag{31}$$

3.2 Identification

The goal of the estimation is to measure the adoption lags and the intensity of use parameters for each technology-country pair, $\{I^c_\tau, D^c_\tau\}_{c,\tau}$, using the structural diffusion curves derived

[27] According to our theory, $\beta_{\tau 3}$ should coincide across technologies. In the estimation, however, we let $\beta_{\tau 3}$ vary across technologies to obtain a better fit of the diffusion curve (as in Comin and Hobijn, 2010). We check ex-post that the estimates do not vary significantly across technologies and that our results are robust to imposing a common $\beta_{\tau 3}$.

[28] Note that there are two minor differences between (28) and (31). 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 $\beta^c_{\tau 1}$ has an extra term equal to $\beta_{\tau 3} \ln \alpha$.

from our theory, (28) and (31). To this end, we assume that the parameters that govern the growth in the technology frontier (γ and σ), and the inverse of the elasticity of demand (θ) are the same across countries, for any given technology. These restrictions imply that the technology-specific coefficients of the time-trend, $\beta_{\tau 2}$, and of the non-linear term, $\beta_{\tau 3}$, in (28) and (31) are common across countries. In addition, in our baseline estimation, we calibrate α, μ, and the invention date, τ. We infer the elasticity of substitution across vintages of the same technology $\mu = 1.3$ by using the price markups from Basu and Fernald (1997) and Norbin (1993),[29] $1 - \alpha = .7$ to match the labor share in the U.S.,[30] and τ to the invention date of each technology. Invention dates are detailed in Appendix A.

These parameter restrictions imply that, conditional on $\beta_{\tau 3}$, the term β_{t0}^c has a symmetric effect across all technologies in a country, as $\beta_{t0}^c = \beta_{\tau 3} \cdot \left(\chi_t^c - (1-\alpha)\ln\left((1+\zeta_L^c)/(1-\alpha)\right)\right)$. One implication of this observation is that the term β_{t0}^c can be absorbed by a full set of country-specific time-dummies (interacted with $\beta_{\tau 3}$) that are restricted to be the same across technologies. That is, once we introduce these time-varying country dummies, the rest of coefficients in (28) and (31) can be identified as deviations in technology-level diffusion from the country-dummies.

The parameter restrictions also imply that cross-country variation in the curvature of (28) and (31) is driven by variation in adoption lags after purging the effect of income and the exogenous TFP process. Specifically, D_τ^c causes the slopes in y_τ^c and x_τ^c with respect to time to monotonically decline in time since adoption. Consider two identical economies except for their adoption of technology τ. If at a given moment in time, we observe that the slope of the diffusion curve y_τ^c (or x_τ^c) is diminishing faster over time in the first country than the second, the concavity of the diffusion curve implies that this must be because the former country has started adopting the technology more recently than the latter (i.e., it is in a more "concave" region of the diffusion curve). This is the basis of our empirical identification strategy for D_τ^c. Equivalently, a higher adoption lag D_τ^c shifts the diffusion curve (28) to the right. Thus, countries that for the same income levels have their diffusion curves "shifted to the right" have a longer adoption lag. This is illustrated by the horizontal shifts in the diffusion curves in Figure 2.[31]

When the number of adopted vintages becomes sufficiently large (i.e., $t \gg D_\tau + \tau$), the effect of D_τ^c in the diffusion curve vanishes because the gains from additional varieties

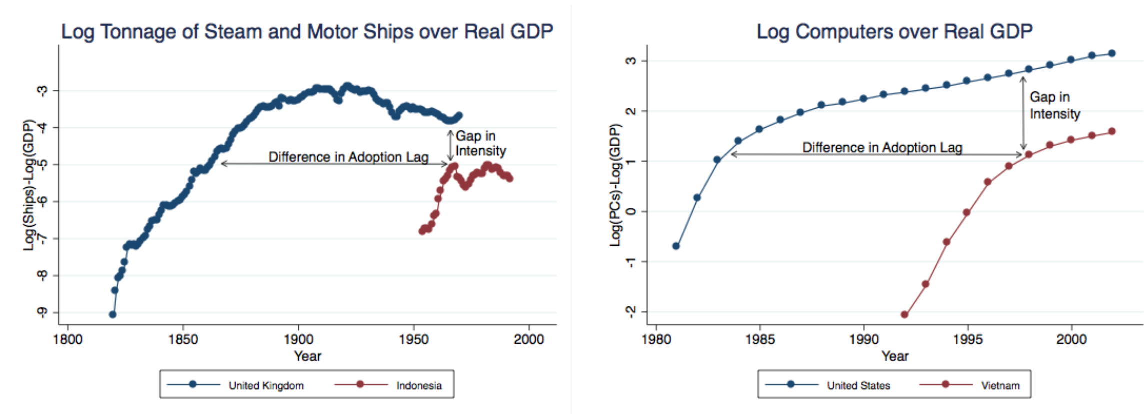
[29] Note that, as we have previously indicated, our model is perfectly competitive and abstracts from markups. However, it is possible to rewrite our model using monopolistic competition, in which case there would exist constant markups that would be solely a function of the elasticity of substitution. It is in this sense that we can infer the elasticity of substitution from markups.

[30] Note that this implies that $(1-\alpha)$ in the term $(1-\alpha)(y_t^c - l_t^c)$ of the estimating equations (28) and (31) are calibrated rather than estimated.

[31] We can identify adoption lags even if we do not have data starting at the exact adoption date. However, to separately identify the adoption lag from the log-linear trends, it is necessary that the data covers some of the non-linear segment of the diffusion curve.

Figure 2: Examples of diffusion curves

(a) Diffusion of Steam and Motor ships (b) Diffusion of PCs



become negligible. At this point, y^c_τ asymptotes to the common linear trend in time and (log) income plus the country-specific intercept, $\beta^c_{\tau 1}$. Therefore, after filtering the country-time varying effects, differences in aggregate demand, and technology-specific time trends, asymptotic cross-country differences in technology are fully captured by the intercept, $\beta^c_{\tau 1}$. These differences in $\beta^c_{\tau 1}$ are captured by vertical shifts in the diffusion curves, as illustrated Figure 2.

In our model, $\beta^c_{\tau 1}$ reflects the intensity of use, I^c_τ, and differences in the average productivity of the technology due to differences in D^c_τ and the invention year. The latter effect can be subtracted from $\beta^c_{\tau 1}$ using the estimated adoption lag D^c_τ in equation (29), to obtain an expression for I^c_τ as

$$\ln I^c_\tau = \frac{\beta^c_{\tau 1}}{\beta_{\tau 3}} + \frac{\gamma}{2} D^c_\tau - \left(\sigma + \frac{\gamma}{2}\right)\tau. \tag{32}$$

In order to difference out the technology-specific term $\left(\sigma + \frac{\gamma}{2}\right)\tau$ and make the estimates of the intensity of use parameters comparable across technologies (which are measured in different units), we define the intensity of use parameter of technology τ in country c, $\hat{I}^c_\tau$, relative to the average value of the intensity of use parameter in technology τ for the seventeen Western countries defined in Maddison (2004),

$$\ln \hat{I}^c_\tau \equiv \ln I^c_\tau - \ln I^{\text{Western}}_\tau = \frac{\beta^c_{\tau 1} - \beta^{\text{Western}}_{\tau 1}}{\beta_{\tau 3}} + \frac{\gamma}{2}(D^c_\tau - D^{\text{Western}}_\tau). \tag{33}$$

Before turning to the implementation of the estimation, we discuss possible sources of bias in the estimates of the adoption margins. As we have already discussed, we achieve identification through the functional form of diffusion curves derived from our theory. One important concern is whether this functional form provides a reasonably good description of diffusion curves for different technologies and countries. We show that our theory provides a good fit of diffusion curves in Section 4.2. After partialling out technology time-trends $\beta_2 t$ and country-time fixed effects β_{t0}^c from the estimated diffusion curves (28) and (31), the average detrended R^2 is .65 (it would be significantly higher without detrending). We take from this result that our framework can capture a very significant amount of the observed variation in technology adoption. This result provides a basis to take our framework as a reasonable approximation of the true diffusion processes, and, under the null that our model is well-specified, then analyze how the different adoption margins contribute to the fit of the diffusion curves.

Even under the assumption that our model provides a good representation of the diffusion curves, there is still the concern of whether the estimates of adoption margins are well-identified in our empirical exercise due to omitted variable bias. We start by noting that the estimating equations, (28) and (31), control for confounding factors that are *(i)* country-time specific through β_{t0}^c (e.g., residual TFP) and aggregate demand, and *(ii)* technology-time specific through technology-specific time trends $\beta_{\tau 2} t$. We obtain the intensity of use parameter from the intercept $\beta_{\tau 1}^c$ in the estimation equation. The intensity of use parameter captures, by definition, the average level of adoption at the country-technology level after filtering out country-time dummies, aggregate demand effects, and time trends. Since it is the collective projection of all country-technology specific factors affecting adoption and we do not attempt to separate out the causal effect of different factors affecting it, there is no concern of omitted variable bias other than model mis-specification. In Section 4.4 we conduct a number of robustness checks on the model specification. For example, we allow a more flexible (non-unitary) effect of aggregate income on the demand for technology and we find similar estimates.

Our estimates of adoption lags D_τ^c come from the non-linear component of the estimating equation, $\beta_{\tau 3} \ln(t - D_\tau^c - \tau)$. Thus, our estimated adoption lags may be biased if there is a technology-country-time varying omitted variable that affects the use of a technology in a way that is correlated with this non-linear term. This omitted variable would bias the estimates of $\beta_{\tau 3}$ and D_τ^c. We cannot rule out that this is occurring in our estimation. However, in the robustness Section 4.4, we find that when we allow for a more flexible non-linear estimation with country-specific curvature $\beta_{\tau 3}^c$, we cannot rule out the null that the diffusion curvature $\beta_{\tau 3}$ is common across countries for 94% of the country-technology pairs. Moreover, the correlation between adoption lags in the two estimations is very high, .93. Thus, it does not appear that allowing for additional flexibility in the curvature of this non-linear term changes substantially our estimates, alleviating the concern of omitted variable bias.

3.3 Implementation of estimation

Next, we discuss the baseline estimation procedure for the diffusion equations (28) and (31).[32] We estimate (28) and (31) in two stages. For each technology, we first estimate the corresponding diffusion equations jointly for the U.S., the U.K. and France, which are the countries for which we have the longest time series and, arguably, the data with the least measurement error.[33] From this estimation, we obtain the technology-specific parameters $\hat{\beta}_{\tau 2}$ and $\hat{\beta}_{\tau 3}$.[34] Then, in the second stage, we jointly estimate the system of equations (28) and (31) for all technology-country pairs, fixing the values of of $\beta_{\tau 2}$ and $\beta_{\tau 3}$ to the estimated $\{\hat{\beta}_{\tau 2}, \hat{\beta}_{\tau 3}\}_{\tau}$ obtained in the first stage.[35] We obtain estimates for $\{\hat{\beta}^c_{t0}, \hat{\beta}^c_{\tau 1}, \hat{D}^c_{\tau}\}_{c,\tau,t}$ in the second stage. Both of these estimations are conducted using non-linear least squares.

Recall from equation (30) that $\beta^c_{t0} = \beta_{\tau 3} \cdot (\chi^c_t - \alpha \ln((1 + \zeta^c_{\tau x})/\alpha)) \equiv \beta_{\tau 3} \cdot \delta^c_t$, where δ^c_t denotes a country-time varying term. Thus, our theory implies that the term β^c_{t0} in each technology diffusion equation can be decomposed between the technology specific term $\beta_{\tau 3}$ and a country-time-varying parameter that is common across technologies. We estimate this latter term δ^c_t in a flexible form using decade fixed effects.

In practice, to separately identify the country-time-varying parameter, δ^c_t, from country-technology specific terms, $\beta^c_{\tau 1}$, requires a balanced panel of technologies for each country. This is not the case for the vast majority of our countries.[36] Thus, rather than estimating country-specific aggregate time trends, we group countries by income groups, and estimate aggregate time trends β^c_{0t} for each income group. This allows us to have long time series of technology diffusion that span the entire period of interest for all technologies. In our baseline estimation, we only use two country groupings, Western and non-Western. This boils down to including two full sets of decade dummies, one for the Western countries and another for the non-Western countries.[37] Finally, we use the estimates for $\{\hat{\beta}^c_{\tau 1}, \hat{\beta}_{\tau 3}, \hat{D}^c_{\tau}\}_{c,\tau}$, in equation (33) to obtain the estimate of the intensity of use parameter for each country-technology pair.[38]

Next, we discuss some of the features of the estimation procedure. The second stage requires the joint estimation of all country-technology pairs because of the presence of the aggregate time effect β^c_{t0}, which is common to all technology diffusion curves for any given country. Otherwise, if β^c_{t0} were absent in the estimating equation, we could obtain consistent

[32] In Section 4.4, we discuss alternative approaches, their rationale and the robustness of our baseline estimates.

[33] In the case of railways, we substitute Germany for the U.K. because we lack the initial phase of diffusion of railways for the UK. In the case of tractors, we also replace the U.S. with Germany for the same reason.

[34] Note that the coefficients $\beta_{\tau 2}$ and $\beta_{\tau 3}$ in (28) 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.

[35] In Section 4.4 we study how sensitive our results are to assume that $\beta_{\tau 3}$ is common across countries.

[36] We have data for all 25 technologies for 9 countries in our sample.

[37] We show in Section 4.4 that our results are robust to a finer division of income groups into quintiles.

[38] Consistent with our calibration below, we compute the intensity of use parameter using a value for γ in (33) of $2/3 \cdot 1\%$. In Section 4.4 we conduct robustness analysis of this parametrization.

estimates of all other parameters by estimating each country-technology pair separately.

The use of just two full sets of decade dummies (one for Western countries and another for the rest) significantly simplifies the estimation and still allows us to control for the possibility that exogenous TFP or other country-level characteristics have diverged in Western vs. non-Western countries.[39]

Third, the use of decade-dummies instead of year dummies also simplifies the computational complexity of the estimation and, given that our interest lies in long-run phenomena, this modification should be inconsequential for our findings.

Fourth, the joint estimation requires estimating non-linearly around 2500 parameters. We have used a nonlinear commercial solver to estimate the nonlinear problem. As a robustness check, we have used an iterative method that exploits the fact that all terms in the joint estimation are log-linear except for the adoption lag. This method proceeded by estimating first the linear part jointly for all country-technology pairs. Then, we used the fact that conditional on the linear terms, the adoption lag can be estimated independently for each country-technology pair. We iterated the estimation until the estimates converged. Both methods yield very similar estimates. The codes and estimation results are available on the authors' websites.[40]

4 Estimation Results

4.1 Data Description

We implement our estimation procedure using data on the diffusion of technologies from the CHAT data set (Comin and Hobijn, 2009a), and data on income and population from Maddison (2010). 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 major technologies that have a broad coverage over rich and poor countries and for which the data captures the initial phases of diffusion. The twenty-five technologies that meet these criteria cover a wide range of sectors in the economy (transportation, communication and IT, industrial, agricultural and medical sectors) as well as 139 countries. They are listed and briefly described in the online Appendix. The invention dates are spread quite evenly throughout the two hundred year period being studied. Our analysis proceeds under the assumption that these technologies are a representative sample of all technologies invented

[39]To study the evolution of adoption margins by continent or income quintiles, we include one full set of decade dummies for each group of countries. Section 5.2 reports the implied and observed income growth patterns for these alternative divisions of countries.

[40]For example, the correlation in estimated adoption lags and technology specific intercepts is over .99. To compute the solution in the nonlinear joint estimation, we have used the Knitro nonlinear solver with the AMPL optimization language. The alternative iterative method has the advantage that can be implemented in more standard econometrics software, e.g., Stata.

during this period. To the best of our knowledge, there is no more comprehensive data available to conduct this analysis than our data.

Overall, we have time series of technology adoption for 1841 country-technology pairs. 57% of these observations correspond to technologies invented prior to 1900. Country-technology pairs corresponding to Western countries represent 22% of the sample, while 66% of the country-technology pairs correspond to countries in the bottom third of the world income per capita distribution in year 2000.[41]

The specific measures of technology diffusion in CHAT match the dependent variables in specification (28) or (31). 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).

4.2 Estimates

Following Comin and Hobijn (2010), we only use the estimates of technology-country pairs that satisfy plausibility and precision conditions. Estimated adoption lags are plausible when they imply an adoption date after the invention year (allowing for some inference error).[42] The majority of the implausible estimates correspond to technology-country pairs for which we do not have data on the concave portion of the curve. With only the log-linear part of the diffusion curve, it is impossible to infer adoption lags since the time trend suffices to fit the diffusion curve. We define estimates as precise if the estimate of the adoption lag is significant at a 5% level. The plausible and precise criteria are met for the majority of the technology country-pairs (65%).

For the plausible and precise technology country-pairs, we find that our estimating equations provide a good fit with an average detrended R^2 of 0.65 across countries and technologies (in the online appendix, Table C.1 reports summary statistics by technology and Figure C.4 shows examples of fit for two technologies).[43] The fit of the model indicates that the restriction that adoption lags and the intensity of use parameter are constant for each technology-country pair and that the curvature of diffusion is the same across countries are not a bad approximation to the data.

Table 1 reports summary statistics of the estimates of the adoption lags for each technology using the estimation procedure described in Section 3.3. The average adoption lag across all technologies and countries is 42 years. We find significant variation in average adoption lags

[41]For technologies invented prior to 1900, 19% of the country-technology pairs belongs to Western countries and 70% belongs to countries in the bottom third. These numbers for technologies invented post 1900 are 26% and 60%, respectively.

[42]We parametrize the error margin such that it is 5 years for a technology invented in year 2000 and 20 years for a technology invented in 1800 (and a linear interpolation for years in between). The results are robust to other definitions, e.g., a flat ten-year window across all technologies.

[43]To compute the detrended R^2, we partial out the linear trend component, γt, of the estimation equation and the country group specific trend, β_{t0}^{c}. We compute the R^2 for the detrended data.

Table 1: Estimates of Adoption Lags

	Invention Year	Obs.	Mean	SD	P10	P50	P90	IQR
Spindles	1779	23	130	49	58	167	171	96
Ships	1788	40	110	65	21	107	180	120
Railway Passengers	1825	36	73	35	26	78	120	47
Railway Freight	1825	43	69	38	12	79	114	56
Telegraph	1835	30	46	36	-1	45	91	56
Mail	1840	45	43	35	-5	45	86	54
Steel	1855	41	68	32	12	69	105	37
Telephone	1876	54	49	33	4	52	92	55
Electricity	1882	75	47	21	16	49	71	33
Cars	1885	61	36	24	6	33	65	30
Trucks	1885	52	37	22	10	31	62	32
Tractor	1892	87	57	17	24	65	69	17
Aviation Passengers	1903	40	26	18	2	24	52	19
Aviation Freight	1903	39	42	17	19	42	63	27
Electric Furnace	1907	46	48	18	25	54	65	33
Fertilizer	1910	92	43	13	29	47	51	9
Harvester	1912	70	36	14	14	39	48	15
Synthetic Fiber	1931	45	29	7	23	31	33	3
Oxygen Furnace	1950	36	13	8	5	13	24	10
Kidney Transplant	1954	24	14	6	6	14	24	5
Liver Transplant	1963	19	19	4	14	18	25	3
Heart Surgery	1968	16	13	3	9	13	19	4
Pcs	1973	62	14	2	11	14	17	2
Cellphones	1973	71	14	5	11	16	19	6
Internet	1983	42	6	3	3	7	9	3
All Technologies		1189	42	35	8	37	82	44

Note: SD denotes Standard Deviation. P10, P50 and P90 refer to the tenth percentile, the median and the ninetieth percentile, respectively. IQR refers to the Interquartile Range, defined as the difference between the third and first quartiles.

across technologies. The range of average adoption lags by technology goes from 6 years for the internet to 130 years for spindles. There is also considerable cross-country variation in adoption lags for any given technology. The range for the cross-country standard deviations goes from 2 years for PCs to 65 years for steam and motor ships.

To compute the intensity of use parameter $\ln \hat{I}^c_\tau$ (equation 33), we calibrate $\gamma = (1-\alpha)\cdot 1\%$, with $\alpha = .3$. We choose this calibration so that half of the 2% long run growth rate of Western countries comes from productivity improvements within a technology (γ) and the other half comes from new technologies being more productive (σ). Section 4.4 conducts the robustness

Table 2: Estimates of the Log-Intensity of Use Parameter relative to Western Countries

	Invention Year	Obs.	Mean	SD	P10	P50	P90	IQR
Spindles	1779	23	-0.19	0.64	-1.24	-0.14	0.64	0.96
Ships	1788	40	-0.29	0.63	-1.24	-0.22	0.44	0.74
Railway Passengers	1825	36	-0.46	0.49	-1.05	-0.43	0.20	0.84
Railway Freight	1825	43	-0.33	0.49	-0.95	-0.33	0.32	0.65
Telegraph	1835	30	-0.34	0.64	-1.21	-0.23	0.32	0.79
Mail	1840	45	-0.31	0.35	-0.79	-0.30	0.15	0.54
Steel	1855	41	-0.37	0.46	-0.88	-0.31	0.19	0.66
Telephone	1876	54	-1.09	0.85	-2.15	-1.05	-0.10	1.17
Electricity	1882	75	-0.74	0.59	-1.49	-0.65	-0.04	0.90
Cars	1885	61	-1.31	1.23	-2.61	-1.20	0.01	1.84
Trucks	1885	52	-1.10	1.04	-2.20	-1.15	0.11	1.17
Tractor	1892	87	-1.20	0.89	-2.43	-1.19	-0.06	1.24
Aviation Passengers	1903	40	-0.49	0.73	-1.45	-0.34	0.26	0.95
Aviation Freight	1903	39	-0.50	0.64	-1.46	-0.36	0.25	0.99
Electric Furnace	1907	46	-0.33	0.51	-1.00	-0.18	0.25	0.70
Fertilizer	1910	92	-0.97	0.78	-1.93	-0.91	-0.03	1.20
Harvester	1912	70	-1.44	1.13	-3.17	-1.36	0.02	1.66
Synthetic Fiber	1931	45	-0.76	0.78	-1.93	-0.69	0.18	1.09
Oxygen Furnace	1950	36	-1.00	1.01	-2.43	-0.80	0.09	1.97
Kidney Transplant	1954	24	-0.25	0.42	-0.99	-0.06	0.11	0.56
Liver Transplant	1963	19	-0.46	0.80	-1.98	-0.09	0.15	0.95
Heart Surgery	1968	16	-0.48	0.88	-1.96	-0.10	0.21	0.67
Pcs	1973	62	-0.59	0.56	-1.42	-0.61	0.05	0.84
Cellphones	1973	71	-0.74	0.68	-1.82	-0.57	0.05	0.98
Internet	1983	42	-0.76	0.79	-1.79	-0.63	0.07	1.29
All Technologies		1189	-0.76	0.85	-1.94	-0.59	0.14	1.13

Note: SD denotes Standard Deviation. P10, P50 and P90 refer to the tenth percentile, the median and the ninetieth percentile, respectively. IQR refers to the Interquartile Range, defined as the difference between the third and first quartiles.

checks of this calibration. We use the value of $\beta_{\tau 3}$ that results from setting the elasticity across technologies, θ, to be the mean across our estimates, which is $\theta = 1.28$.[44]

Table 2 reports the summary statistics for the estimates of the intensity of use parameters. The average intensity of use parameter is -.76. This implies that the intensity of use in the average country is 47% (i.e., $\exp(-.76)$) of the Western countries. There is significant cross-

[44] This value is very similar to the estimates of demand elasticity found by Basu and Fernald (1997) and Norbin (1993) despite using a different identification strategy.

country variation in the intensity of use parameter. The technology average ranges from $-.19$ for spindles, which implies being 18% less productive relative to the benchmark, to -1.00 for blast oxygen furnaces, which implies being 63% less productive than the average. The range of the cross-country standard deviation goes from 0.35 for mail to 1.23 for cars. The average 10-90 percentile range in the (log) intensity of use parameter is 2.08. This gap implies output differences along a balanced growth path of a factor of 11.4.[45]

4.3 Cross-country evolution of the diffusion process

To analyze the cross-country *evolution* of the adoption margins, we divide the countries in our data set into two groups defined by Maddison (2004): Western countries, and the rest of the world, labeled "Rest of the World" or, simply, non-Western.

Figure 3 plots, the median adoption lag for each technology among Western countries and the rest of the world. This figure illustrates that adoption lags have declined over time, and that cross-country differences in adoption lags have narrowed. Table 3 formalizes these impressions by regressing (log) adoption lags on their year of invention (and a constant),

$$\ln D_\tau^c = \rho + \omega \cdot (\text{Invention Year}_\tau - 1820) + \varepsilon_\tau^c, \tag{34}$$

where ε_τ^c denotes an error term. Column (1) reports this regression for the whole sample of countries showing that adoption lags have declined with the invention date. Columns (2) and (3) report the same regression separately for Western and non-Western countries, respectively. We find that the rate of decline in adoption lags is almost 40% higher in non-Western than in Western countries (-1.106% vs. -.76%). Hence, there has been *convergence* in the evolution of adoption lags between Western and non-Western countries. The yearly convergence rate between Western and non-Western countries we find is $-.76 - (-1.106) = .346\%$.

Do we observe a similar pattern for the intensity of use? Figure 3b plots, for each technology and country group the median intensity of use parameter. This figure shows that the gap in the intensity of use parameter between Western countries and the rest of the world is larger for newer than for older technologies. In other words, the gap in the intensity of use has widened over time. Table 3 studies econometrically this question by regressing the intensity of use parameter on the invention year and a constant,

$$\ln \hat{I}_\tau^c = \rho + \omega \cdot (\text{Invention Year}_\tau - 1820) + \varepsilon_\tau^c. \tag{35}$$

Column (6) shows that, for non-Western countries, the intensity of use parameter has declined at an annual rate of .50%.[46] Since the intensity of use parameter is measured relative to the

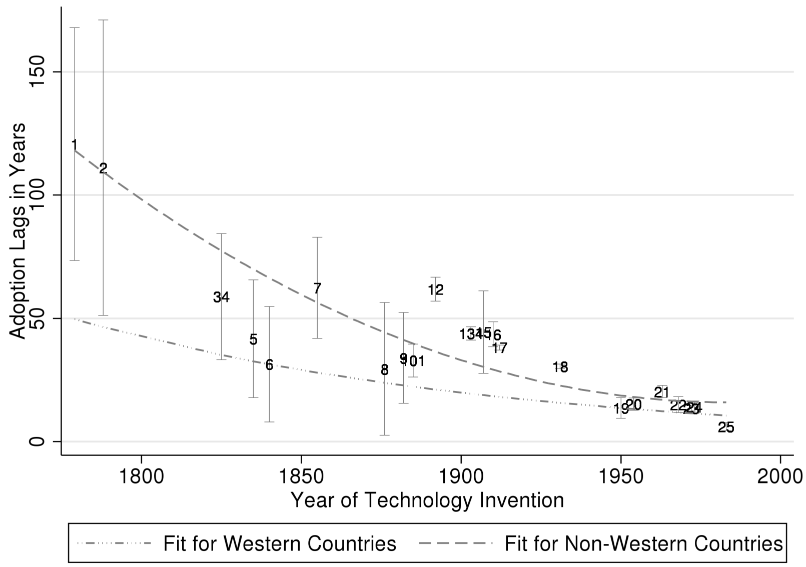
[45] To see this, use equation (25) with our calibration (discussed below) of $1 - \alpha = .7$, to find $\exp(2.08)/.7 = 11.4$.

[46] For Western countries, column (5) shows that there is no trend in the intensity of use parameter, as by

Figure 3: Evolution of Adoption Margins

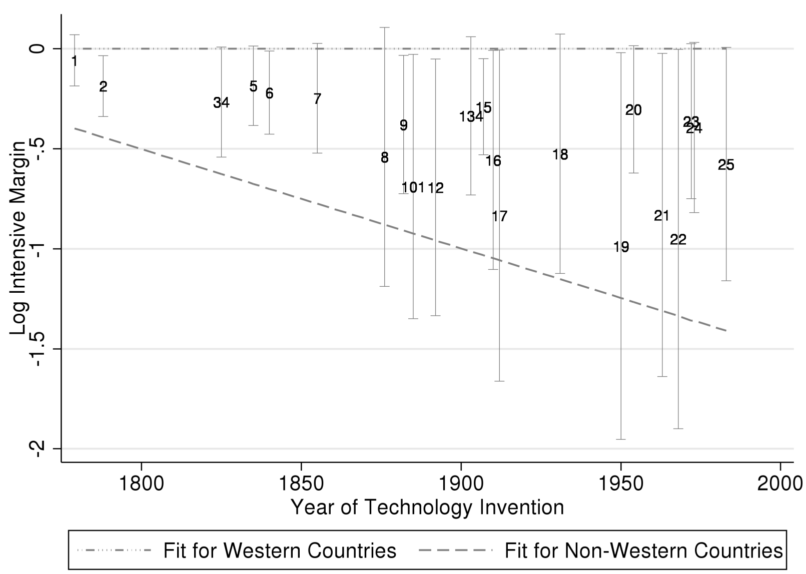
(a) Convergence of Adoption Lags

The gap between Western and non-Western countries decreases for newer technologies



(b) Divergence of the Intensity of Use

The gap between Western and non-Western countries increases for newer technologies



Note: Bars show median margins of adoption for Western vs. non-Western countries. 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 3: Evolution of the Adoption Lag and Intensity of Use

Dep. Var.:	Log(Lag)			Log(Intensity)		
	World	Western	Rest	World	Western	Rest
	(1)	(2)	(3)	(4)	(5)	(6)
Year -1820	-0.0106	-0.0076	-0.0106	-0.0018	0.0000	-0.0050
	(0.0005)	(0.0009)	(0.0005)	(0.0006)	(0.0002)	(0.0005)
Constant	4.23	3.60	4.37	-0.52	-0.00	-0.60
	(0.07)	(0.12)	(0.07)	(0.06)	(0.04)	(0.07)
Obs.	1151	288	863	1189	306	883
R^2	0.40	0.27	0.41	0.02	0.00	0.11

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

mean of Western countries, this estimate implies a *divergence* in the intensity of use of new technologies between Western and non-Western countries over the last 200 years.[47]

Figure 2 is consistent with our findings on the evolution of technology diffusion. The horizontal gap in diffusion between the rich and developed economies is much larger for ships than for personal computers. This is consitent with the closing of the gap in adoption lags between Western and non-Western countries. Conversely, the vertical gap in the diffusion curves evaluated in the “long-run” is larger for personal computers than for ships. This is consistent with the widening of the intensity of use between Western and non-Western countries.

4.4 Robustness

In this section, we assess the robustness of our estimates and the two cross-country trends in technology diffusion to alternative measurement and estimation approaches.

Curvature The main assumption used in the identification of the adoption margins is that the curvature of the diffusion curve is the same across countries. In our model, this property follows from the common elasticity of substitution between sectoral outputs across countries (i.e., $1/(\theta-1)$). To explore the empirical validity of this assumption, we re-estimate equation

construction the intensity of use parameter is defined relative to Western countries.

[47]One alternative interpretation of Figure 3b is that, rather than a continuous decline in the intensity of use parameter in non-Western countries, there was a structural break around 1860. We find that the linear model provides a better statistical fit as measured by the R^2. In Section 5.3, we examine the implications for income dynamics of modeling the divergence in the intensity of use parameter as a continuous or as a discrete process and show that our results are robust to this modeling choice.

(28) allowing $\beta_{\tau 3}$ to differ across countries. Thus, we obtain an estimate $\hat{\beta}^{c}_{\tau 3}$ for each country-technology pair. Then, we test whether $\hat{\beta}^{c}_{\tau 3}$ in each country is equal to the baseline estimate $\hat{\beta}_{\tau 3}$. We find that in 94% of the cases, we cannot reject the null that the curvature is the same as for the baseline countries at a 5 percent significance level.

Table 4 documents the robustness of our estimates to relaxing the restriction that $\beta_{\tau 3}$ is the same across countries. The first row reports the correlation between the estimates of the diffusion margins in the baseline and in the unrestricted estimations. The unconditional correlation of adoption lags across all technologies is .931 (column 1). The median correlation when we compute it technology by technology is is .78 (column 2). We also report the 25th and 75th percentiles of the within technology correlation which are .68 and .86, respectively. For the intensity of use, the unconditional correlation is .80 (column 3) and the median correlation within technologies is .78 (column 4). Therefore, we conclude that the adoption lags and intensity of use that arise under the unrestricted estimation are highly correlated with the baseline estimates.

Table 5 studies the robustness of the patterns uncovered for the evolution of the adoption margins. Columns (1) and (2) report the time-trend coefficient of the (log) adoption lag with respect to the invention date for Western and non-Western countries. Column (3) reports the time-trend coefficient for the intensity of use parameter in non-Western countries. For comparison purposes, we report the baseline estimates in the first row of the table. The new estimates confirm that both the convergence of adoption lags and the divergence in the intensity of use parameter are robust to relaxing the restriction of a common curvature across countries. If anything, the new estimates suggest stronger convergence and divergence patterns than the ones reported in the baseline.

Division in Quintiles One possible concern is that, when computing the country-group decade fixed effects, β^{c}_{t0}, dividing the sample between West and Rest is too coarse. There may be substantial heterogeneity in the aggregate country trends of non-Western countries that is not well captured by the common trend for non-Western countries. To address this concern, we group the non-Western countries into quartiles according to their income per capita in year 2000 and allow for a specific country-group time trend in our estimation of the diffusion curve. Thus, we estimate five different sets of decade dummies β^{c}_{t0}s (one for Western and four for non-Western), rather than just the two from the baseline estimation. We find that the estimates we obtain for both adoption margins hardly change (See second row of Table 4.) Furthermore, the resulting trends in the adoption margins by income group are very similar to those uncovered in the baseline specification (See the third line in Table 5.)[48]

[48] We have also experimented with other robustness checks on how to specify β^{c}_{t0}. For example, the results are robust to eliminating the term $\beta_{\tau 3}$ from the estimation such that the effect of the aggregate trends is symmetric across technologies. The results are also robust to including an additional group for the East Asian Tigers.

Table 4: Correlation of Baseline Estimates with Alternative Specifications

Alternative Specification	Adoption Lags		Intensity of Use	
	Overall	Within Tech.	Overall	Within Tech.
Unrestricted Curvature	0.91	0.78	0.80	0.78
		[0.68, 0.86]		[0.74, 0.91]
Quintiles	0.96	0.91	0.99	0.99
		[0.85, 0.96]		[0.99, 0.99]
No Country × Time FE	0.94	0.95	0.99	0.99
		[0.90, 0.99]		[0.99, 1.00]
Non-homotheticities	0.95	0.91	0.87	0.90
		[0.83, 0.98]		[0.86, 0.93]
Estimated μ	0.86	0.82	0.57	0.89
		[0.67, 0.87]		[0.82, 0.93]
Obsolescence	0.94	0.93	0.97	0.98
		[0.65, 0.95]		[0.95, 0.99]
No correction intensity	-	-	0.99	1.00
				[0.99, 1.00]

Note: Overall refers to the correlation of all estimates in the baseline and in the alternative specification. Within Tech. reports the median correlation of the estimates within technologies. The 25th and 75th percentiles of the correlation within technologies are reported in brackets.

Eliminating Aggregate Country Trends across Technologies Our results are robust to omitting the aggregate country trends across technologies, β^c_{t0} (See row 3 in Table 4 and row 4 in Table 5). The fact that the adoption estimates hardly change when excluding β^c_{t0} in the estimation implies that controlling for total income y^c_t in the diffusion curve is sufficient to partial out aggregate country-specific trends. That is, exogenous aggregate TFP plays a minor role in the evolution of the two margins of adoption.

Non-homotheticities We investigate the robustness of our estimates and the dynamics of adoption margins once we allow for non-homotheticities in the demand for technology. Non-homotheticities alter our baseline estimating equation (28) by introducing an income elasticity in the demand for technology, $\beta_{\tau y}$, potentially different from one,

$$y^c_{\tau t} = \beta^c_{t0} + \beta^c_{\tau 1} + \beta_{\tau y} y^c_t + \beta_{\tau 2} t + \beta_{\tau 3}\left((\mu - 1)\ln(t - D^c_\tau - \tau) - (1-\alpha)(y^c_t - l^c_t)\right) + \varepsilon^c_{\tau t}. \quad (36)$$

A practical difficulty in estimating (36) is the colinearity of the time-trend and log-income, y^c_t. To overcome this problem, we group technologies by their invention date in four groups and estimate a common income elasticity for each group (indexed by T).[49] As in the baseline, we

[49] The four groups are pre-1850, 1850-1900, 1900-1950, and post-1950. We have implemented a similar approach grouping the technologies according to the sector rather than the invention date, obtaining similar results.

Table 5: Time Trend Coefficient Across Alternative Specifications

Dependent Variable:	Log(Lag)		Log(Intensity)
	Western (1)	Rest World (2)	Rest World (3)
Baseline	-0.0076	-0.0106	-0.0050
	(0.0009)	(0.0005)	(0.0005)
Unrestricted Curvature	-0.0059	-0.0101	-0.0065
	(0.0006)	(0.0005)	(0.0007)
Quintiles	-0.0084	-0.0109	-0.0054
	(0.0008)	(0.0004)	(0.0005)
No Country $\times$ Time FE	-0.0080	-0.0111	-0.0054
	(0.0006)	(0.0003)	(0.0005)
Non-homotheticities	-0.0069	-0.0104	-0.0044
	(0.0009)	(0.0006)	(0.0005)
Estimated μ	-0.0084	-0.0112	-0.0089
	(0.0010)	(0.0005)	(0.0013)
Obsolescence	-0.0069	-0.0111	-0.0041
	(0.0009)	(0.0005)	(0.0005)
No correction intensity	-	-	-0.0037
			(0.0005)

Note: This table reports the coefficient ω on the time trend resulting from regressing the log of adoption lag and the intensity of use parameter on $\rho + \omega(\text{Invention Year}_\tau - 1820) + \varepsilon^c_\tau$ for the different country groupings. Robust standard errors in parentheses. Each technology observation is weighted so that each technology carries equal weight.

estimate equation (36) in two-steps. In the first step, we jointly estimate the income elasticity, β_{Ty}, along with $\beta_{\tau 2}$, and $\beta_{\tau 3}$ from the diffusion curves of the U.S., UK, and France for each of the four technology groupings. Effectively, this method identifies the income elasticity of technology out of the time series variation in the baseline countries in income and technology.[50] 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 approach.

The estimates of the income elasticity for the technologies invented in the four periods, β_{Ty}, range from 1.58 for $T =$ (pre-1850), to 1.99 for $T =$ (1850-1900).[51] The estimates of the slopes of the Engel curves do not vary much across technology groups and they do not have a clear trend.

Once we have obtained the estimates for the income elasticity, we proceed as in the baseline estimation, but instead of imposing an income elasticity of one as our theoretical model suggests, we use the estimated income elasticity. Therefore, in the second step, we

[50] See Comin *et al.* (2015) for evidence on the similarity across countries in the slope of Engel curves and a theory that would generate heterogeneous Engel curves analogous to the ones estimated here.

[51] Table C.2 in the Appendix reports the estimates.

estimate β^c_{t0}, $\beta^c_{\tau 1}$ and D^c_τ for each country-technology pair from the equation

$$y^c_{\tau t} = \beta^c_{t0} + \beta^c_{\tau 1} + \hat{\beta}_{Ty} y^c_t + \hat{\beta}_{\tau 2} t + \hat{\beta}_{\tau 3} \left((\mu - 1) \ln(t - D^c_\tau - \tau) - (1 - \alpha)(y^c_t - l^c_t) \right)) + \varepsilon^c_{\tau t}, \quad (37)$$

where $\hat{\beta}_{\tau 2}$, $\hat{\beta}_{\tau 3}$ and $\hat{\beta}_{Ty}$ are the values of $\beta_{\tau 2}$, $\beta_{\tau 3}$ and β_{Ty} estimated for the U.S., U.K. and France in the first step.

The estimates of the two margins that we obtain are highly correlated with our baseline estimates (See row four in Table 4). Moreover, both the convergence of adoption lags and the divergence of the intensity of use are statistically and economically robust to allowing for non-homotheticities. (See row 5 in Table 5.)

Estimated μ In our baseline estimation, we calibrate the elasticity of substitution between vintages, $\mu/(\mu - 1)$, because it is difficult to separately identify $\beta_{\tau 3}$ and μ in the baseline diffusion equation (28). However, it is possible to identify them simultaneously in the exact structural equation that results from substituting expression (17) for z_τ in (18) rather than its log-linear approximation.

In the fifth row of Table 4 we compare the adoption margins obtained using this alternative approach with our baseline estimates. Both sets of estimates are similar. Furthermore, the evolution of the adoption margins across countries quantitatively resembles very much that of our baseline estimates. If anything, the divergence pattern in the intensity of use parameter seems stronger. (See row 6 in Table 5.)

Obsolescence Some technologies eventually become dominated by others. This is for example the case of the telegraph which was rendered obsolete by the telephone. The obsolescence of technology may affect the shape of the diffusion curves (especially in the long run) and therefore the estimates of our adoption margins. Since our theory just concerns the diffusion process (and is silent about the phase out process) it does not provide any guide on how obsolescence impacts our technology measures. As a robustness check, we re-estimate equation (28) over a time sample where obsolescence dynamics are unlikely to be relevant. For each technology, we censor the sample at the point where the leading country starts to experience a decline in the per capita adoption level.[52] This affects the estimation period in six of the twenty-five technologies in our sample. The estimates and evolution of both adoption margins are robust to controlling for the potential obsolescence of technologies. (See row 5 of Table 4 and row 7 in Table 5.)

[52] More precisely, we censor all observations that are 90% below the peak level of technology usage for the leading country (after the peak has been attained). This is to allow for some fluctuations in the level of adoption.

Measurement of the intensity of use parameter In our baseline estimation, we identify the intensity of use parameter by removing from the intercept $\beta_{\tau 1}$ the productivity gains arising from the timing of adoption. This effect operates through the adoption lag, which we subtract according to (33),

$$\ln \hat{I}_{\tau}^{c} = \frac{\beta_{\tau 1}^{c} - \beta_{\tau 1}^{\text{Western}}}{\beta_{\tau 3}} + \frac{\gamma}{2}(D_{\tau}^{c} - D_{\tau}^{\text{Western}}). \tag{33}$$

In our theory, correcting for differences in adoption lags in the intercept has a clear interpretation. For the intensity of use parameters to be comparable across countries, they should be computed as if adoption had started at the same time in all countries. This way, any remaining vertical differences in the resulting corrected diffusion curves can be attributed to differences in the intensity of use parameter of technology.

However, in the light of the convergence in adoption lags, one might worry that the divergence in the intensity of use parameter that we find is a mechanical result inherited from the trends in the estimated adoption lags. To address these concerns, we compute an alternative measure of the intensity of use parameter under the extreme assumption that the productivity growth of a technology is zero, $\gamma = 0$. Note that, in this case, there is no correction in the intensity of use parameter arising from differences in the timing of adoption in (33).

The last row in Table 4 shows the correlation between these estimates of the intensity of use parameter and our baseline estimates. The overall correlation between these two estimates is .99. The divergence in the intensity of use parameter is also robust to this alternative computation. In particular, the time trend on the intensity of use parameter for non-Western countries is $-.37\%$, compared to $-.50\%$ in our baseline estimate. Thus, even in the extreme case of no growth in productivity embodied in new technology vintages, we find significant divergence in the intensity of use parameter, albeit somewhat smaller than in our baseline estimates.

5 Income Dynamics

In this section, we evaluate quantitatively the implications of the cross-country patterns of technology diffusion for the evolution of the world income distribution. We focus on two questions: *(i)* the model's ability to generate pre-industrial income differences, and *(ii)* the model's account of the Great Divergence.

To explore these questions, we feed the estimated adoption trends (equations 34 and 35) into the aggregate representation of our model (equation 24). In particular, equation (24) in Proposition 2 shows how aggregate TFP A depends on sectoral adoption patterns through their effect on technology-level TFPs, A_{τ}. We then use the result in equation (23) that output is proportional to $A^{1/(1-\alpha)}$ to compute the effect of TFP growth on output growth.[53]

It is important to note that the estimation of the adoption margins is based on micro-level data on technology diffusion and that the estimation does not impose any constraint that links these margins to aggregate productivity dynamics. Therefore, there is no a priori reason to expect that the model's world income distribution and its evolution resemble their data counterparts.

5.1 Calibration

To simulate the evolution of productivity growth we need to calibrate four common parameters. γ and σ define the evolution of the technology frontier; α, is the share of intermediate goods, and θ determines the elasticity of substitution between sectors, $1/(\theta-1)$.

Prior to the Industrial Revolution, the technology frontier grew at $(1-\alpha)\cdot 0.2\%$.[54] This is the growth rate of Western Europe from 1500 to 1800 according to Maddison (2004, 2010). We model the Industrial Revolution as a one time increase in the growth rate of the frontier.[55] We study the robustness to modelling the Industrial Revolution as more gradual changes in the rate of arrival of new technologies in the robustness section 5.3. We set the date of the Industrial Revolution to year $T = 1765$ (year in which James Watt developed his steam engine). After 1765, the frontier growth rate, $\sigma+\gamma$, jumps to $(1-\alpha)\cdot 2\%$ per year. This ensures that Modern growth along the balanced growth path is 2%. As previously discussed, we set $1-\alpha = .7$. In our baseline simulation, we split evenly the sources of growth in the frontier between γ and σ.[56] Finally, we calibrate the elasticity of substitution in the final good production function, $1/(\theta-1)$, using the average estimates of $\beta_{\tau 3}$, which implies a value of $\theta = 1.28$.

We assume that the adoption margins evolve continuously between 1765 and 1983 as described by equations (34) and (35). To reiterate the discussion in Section 4.1, this assumption implies that we treat the cross-country evolution of technology adoption margins we have uncovered for our sample of technologies as representative of the universe of new technologies invented over the last two centuries. We also note that all technologies enter symmetrically into the production of the final good as specified in the theory section, equation (9). Prior to 1765 and after 1983, the adoption margins remain constant.[57] Constant adoption margins at

[53]Thus, when simulating output growth, we are factoring in the general equilibrium effect that higher TFP levels generate higher levels of intermediates. As such, our simulations on productivity growth should be interpreted as labor productivity growth that is driven by technology adoption.

[54]Alternatively, we can set the beginning of the Industrial Revolution at 1779, year of invention of the first technology in our sample (the mule spindle), without any significant change to our findings,.

[55]According to Mokyr (1990) and Crafts (1997), this captures well the Industrial Revolution.

[56]From our reading of the literature, it is unclear what fraction of frontier growth comes from increases in productivity due to new technologies (σ) and new vintages (γ). In Section 4.4 we conducted robustness checks on this division.

[57]We select 1983 as our end date because this is the last invention year in our data (the internet). For the

the end and the beginning of the sample ensure that the economies transition between the two balanced growth paths (described by equation 25). Note that this assumption imposes that all economies should converge to the same long-run growth rates going forward.[58] Thus, the only difference across countries in our simulations from 1765 onward comes through differences in the transition path between the pre-Modern and Modern balanced growth paths. Since all technology parameters are the same across countries, these differences in the transition come entirely from the differences in the estimated adoption margins that we feed into the model.

5.2 Cross-country evolution of income growth

Initial productivity differences We start by studying the model's ability to account for the initial differences in productivity between Western and non-Western countries. Our estimates from Table 3 imply that the difference between the average adoption lag in the sample of Western countries and the rest of the world is 42.5 years in 1820. The average gap in the (log) intensity of use parameter is 0.6, which implies that technologies in non-Western countries are 55% as productive as in Western countries. 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 (25) implies that Western countries had an income per capita 2.5 higher than the rest of the world.[59] This is roughly in line with the results from Maddison (2004, 2010), who reports a two-fold income gap between the West and the rest.

Great Divergence Next we study the evolution of productivity growth for the 1765-2000 period (see Figure 4).[60] Despite that the acceleration in the growth rate of the frontier starts in 1765, the growth rate of the Western countries remains at its pre-industrial level until around 1830 due to the long initial adoption lags. At this point, Western economies benefit from the higher productivity embodied in industrial technologies and experience a gradual acceleration in productivity growth. By 1900, Western economies have reached the steady state growth rate of 2%.[61] Growth in the non-Western economy does not take off until approximately 1870. After that, it accelerates at a slower pace than the average Western economy. By year 2000,

adoption lag, and given that it tends towards zero at the end of the sample, similar results would emerge if we allowed for a continuing trend. For the intensity of use parameter, and given our finding of divergence, allowing for a continuing trend would exacerbate the differences between Western and non-Western countries.

[58] Our baseline simulations indicate that by year 2100 the growth rates would have approximately converged.

[59] That is, $\exp(.2\% \cdot 42.5 + .6/(1-\alpha)) = 2.5$.

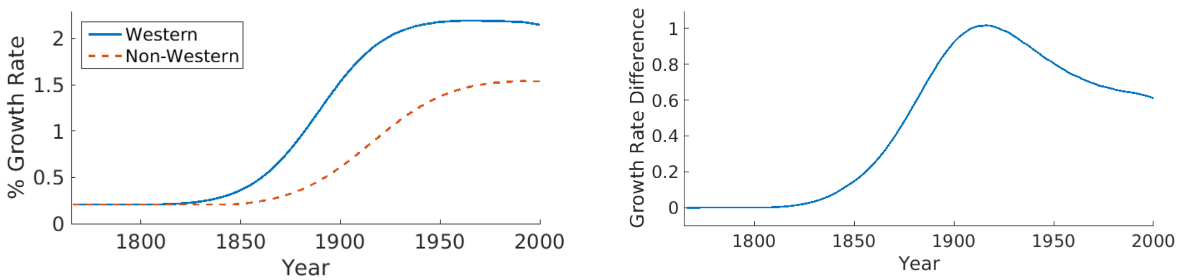
[60] In the working paper version (Comin and Mestieri, 2013), we characterize analytically the transition after a permanent increase in the growth rate of the technology frontier and after changes in the adoption margins. We show that the path of the growth rate during the transition is S-shaped. We also provide approximate expressions for the half-life of the system during the transition. We show that the half-life depends on both adoption margins, and that it is an order of magnitude larger than in conventional calibrations of the neoclassical model.

[61] The transitional dynamics implied by the model are very protracted: the half-life of the growth rate is a hundred years. This is an order of magnitude greater than the neoclassical growth model with a conventional calibration, e.g., Barro and Sala-i-Martin (2003).

Figure 4: World Income Dynamics, 1820-2000

(a) Output Growth: Western vs. non-Western

(b) Difference in Growth Rates



Note: Growth rates of Western countries and the rest of the World obtained imputing the estimated evolution of the adoption margins using the baseline calibration. Western countries are defined as in Maddison (2004, 2010).

it is still 1.5%. Figure 4a shows the gap in productivity growth between Western and non-Western economies. The gap in growth rates between Western and non-Western economies is considerable. The average annual growth rate is 0.6 percentage points lower in non-Western countries over a period of 180 years. The gap peaks around 1913 at 1%. From then on, the gap declines monotonically until reaching 0.6% by 2000.

Table 6 reports the average growth and growth gaps of our simulation and in the Maddison (2010) data. The evolution of productivity growth in both country groups traces quite well the data, though the model slightly underpredicts growth during the nineteenth century.

The sustained cross-country difference in growth rates generated by the model accumulate into a 2.95-fold income gap between the Western countries and the rest of the world from 1820 to 2000. Maddison (2010) reports an actual income widening by a factor of 3.9. Hence, differences in the technology adoption patterns account for 75.6% of the Great Income Divergence between Western and non-Western countries over the last two centuries. When compounding the increase in the gap with initial income differences in 1820, it follows that differences in technology imply an income gap between Western and non-Western countries of 7.38 ($2.5 \cdot 2.95$) by 2000, which is of the same order as the the 7.2-fold gap reported by Maddison (2004, 2010).[62]

Mechanisms at Work We next decompose the three mechanisms at work in our simulation: the acceleration of the technological frontier and the evolution of the two adoption

[62]The simulation does also well in replicating the time series income evolution of each country group separately. For Western countries, Maddison (2010) 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 13.3-fold increase over the same period, and 16% of this increase is generated prior to 1913. For non-Western countries, Maddison (2010) reports an almost 5-fold increase, with around 37% of the increase being generated prior to 1913. Our simulation generates a 4.5-fold increase in the 1820-2000 period with 31% of this increase occurring in pre-1913.

Table 6: Growth rates of GDP per capita

		Time Period		
		1820-2000	1820-1913	1913-2000
Simulation	Western Countries	1.44%	.80%	2.13%
	Rest of the World	.84%	.37%	1.34%
	Difference West-Rest	.60%	.43%	.79%
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 (2010). We use 1913 instead of 1900 to divide the sample because there are more country observations in Maddison (2010). 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 of 145.

margins. We start by simulating the dynamics of our model after a common acceleration of the technology frontier for both countries, keeping constant the adoption margins at their initial levels. Figure 6 shows that these initial conditions are an important source of cross-country income divergence. Longer adoption lags in the non-Western country imply a delay of around 80 years to start benefiting from the productivity gains of the Industrial Revolution. As a result, the income gap increases by a factor of 2.1 by year 2000.

To assess the role of adoption lags in cross-country growth dynamics, we simulate the evolution of our two economies, but keeping the intensity of use parameters at pre-industrial levels (i.e., no divergence). Figure 7a presents the results from this simulation. It shows that cross-country differences in adoption lags are an important driver of income divergence during the nineteenth century. By 1900, the income gap between Western and non-Western countries has grown by a 16%, as opposed to 32% in the full-blown simulation. Thus, the initial differences and subsequent evolution in adoption lags accounts for half of the income gap generated in the nineteenth century, while the other half is accounted by the intensity of use. In the twentieth century, the faster reduction in adoption lags in non-Western countries produces a higher growth rate in non-Western countries. In fact, had the intensity of use parameters remained constant at the pre-industrial levels, the relative income between Western and non-Western countries in 2000 would almost coincide with the relative levels in 1820.

To analyze the role of the intensity of use, we simulate the evolution of the two economies as in our baseline model, but keeping the adoption lags constant at their pre-industrial levels. Figure 7b presents the dynamics of income growth in each country. In this simulation, the growth acceleration in both Western and non-Western countries starts much later than in the baseline (Figure 4). This is a consequence of omitting the productivity gains from a reduction in adoption lags. As shown in the bottom panel of Figure 7b, the growth gap between the two groups of countries is smaller than in the baseline during the nineteenth century but greater during the twentieth century. As a result, the income gap between Western countries and the rest of the world accumulated between 1820 and 2000 is 3.5.

To sum up, the main findings from our simulations are:

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 2.95 over the last 200 years. This represents 75% of the actual increase in the income gap observed in the data.

2. Our model generates very protracted transitional dynamics. This is due both to the effect of adoption lags and to the fact that it takes time for new technologies to become a significant share of aggregate output. This latter effect is affected by the intensity of use.

3. Large cross-country differences in adoption lags account for half of the income divergence during the nineteenth century between Western countries and the rest of the world. However, convergence in adoption lags by itself is a strong force towards convergence in the twentieth century. Absent a divergence in the intensity of use parameter, Western and non-Western income per capita would not have diverged in the last two-hundred years.

4. The Great Divergence continued during the twentieth century because of the divergence in intensity of use of technologies between Western countries and the rest of the world.

Implications for the World Income Distribution We now explore the dynamics of productivity using a finer classification of countries by income quintiles (rather than classifying countries in only two groups).[63] Figure 5 presents the income distribution relative to the first quintile generated by the model and the empirical distribution in the data for 1820, 1913 and 2000.

Figure 5a shows that for 1820 our simulation underpredicts the relative income at the second, fourth and fifth quintiles relative to the first quintile. This deviation is largest for

[63] At each of these years, we classify countries by income quintiles. For each quintile and date, we estimate the evolution of the adoption margins as in the baseline equations (34) and (35). We conduct the estimation of the evolution using information over the whole sample. (The results reported are robust to restricting the time period to prior to 1913 for the years 1820 and 1913). This gives a total of 15 evolutions for each of the two adoption margins. For each year and income quintile, we feed the estimated technology dynamics and simulate the income growth between 1765 and the year of consideration.

Table 7: Annual Growth rates of GDP per capita by regions

	Simulation		Maddison	
	1820-1913	1913-2000	1820-1913	1913-2000
USA and Canada	.73%	1.84%	1.63%	1.90%
Western Europe	.64%	1.93%	1.29%	2.16%
Africa	.33%	1.06%	.36%	.90%
Asia	.36%	1.38%	.49%	1.70%
Latin America	.40%	1.28%	.59%	1.50%

Note: Annual growth rates of GDP per capita by regions. Simulation results and growth rates from Maddison (2010). For the missing income per capita values in 1820 and 1913, we impute the minimal within group income reported in that year.

the second quintile, where our model generates a relative income of .9 compared to 1.3 in the data. For the third quintile, we slightly overpredict the relative income, 1.7 compared to 1.5 in the data. For 1913, Figure 5b, the model is on mark for the relative income of the second and fifth quintile. It somewhat overpredicts the relative income of the third (2.55 vs. 1.8) and fourth (3.6 vs. 2.9) quintiles. For 2000, Figure 5c, the model is on mark for the second and third quintiles, and slightly underpredicts the relative income for the fourth (8.6 vs. 9.9) and overpredicts the fifth (30.7 vs. 27.5). We conclude from this exercise that the output growth processes across the world income distribution that result from the quintile diffusion patterns are remarkably similar to those observed over the last two centuries.

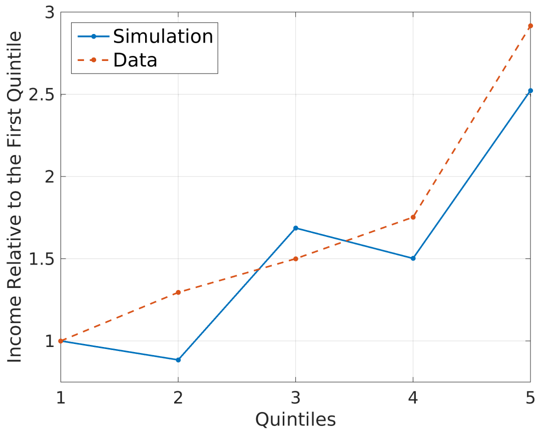
The Geography of Growth We further explore the implied growth paths by continent. For each continent, we estimate the trends in the adoption lags and intensity of use parameters. Then, we simulate the evolution of productivity growth since the industrial revolution. Table 7 shows that technology dynamics in each continent have induced income dynamics that resemble those observed in the data. As in the baseline case, the model tends to slightly underpredict the average growth rates during the nineteenth century. However, it is able to capture very well the relative growth rates across regions in the world. 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 0.93.

5.3 Robustness

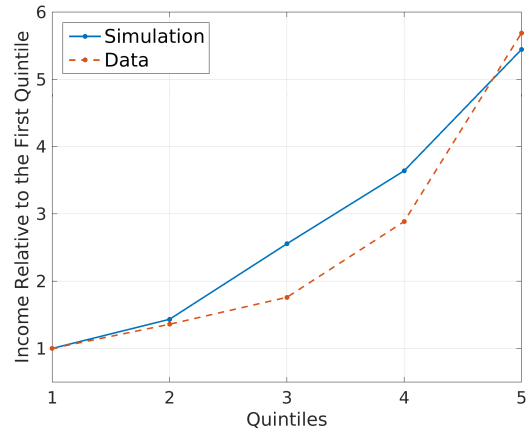
We conduct four robustness checks to our analysis: *(i)* using the technology dynamics when we allow for non-homotheticities in production, *(ii)* alternative calibrations of γ and σ, *(iii)* interpreting the divergence of the intensity of use parameter as a structural break instead of

Figure 5: Evolution of the World Income Distribution

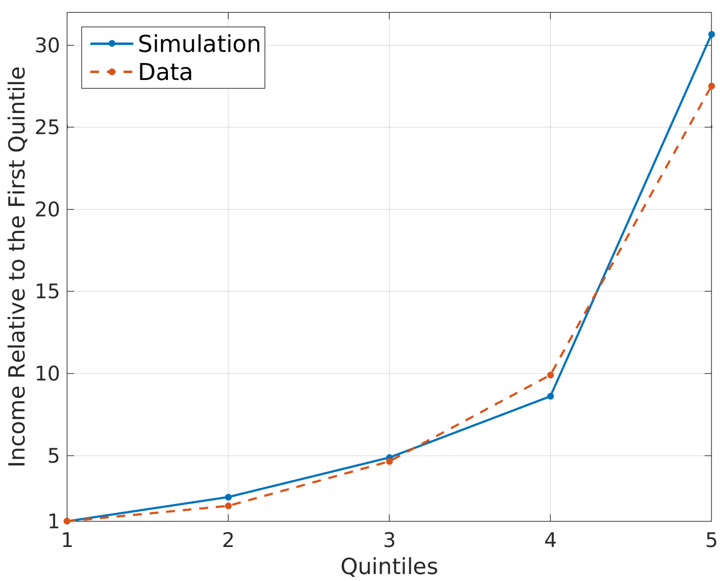
(a) World Income Distribution in 1820



(b) World Income Distribution in 1913



(c) World Income Distribution in 2000



Note: Data refers to the world income distribution computed from Maddison (2010) for the corresponding year. Simulation obtained feeding the diffusion pattern for adoption lags and the intensity of use parameter of the corresponding quintile to the baseline calibration described in Section 5. Quintile groups are computed for each of the three years separately.

a smooth process and *(iv)* a slow acceleration of the technology frontier.

Non-homotheticities in production As shown in Section 4.4, the trends we have identified for the evolution of adoption are robust to the presence of non-homotheticities in the demand for technology. When feeding in these alternative trends in adoption into our model the productivity gap between Western and non-Western countries increases by a factor of 2.62 over the period 1820-2000, which represents 67% of the Great Divergence. Moreover, the growth patterns are very similar to the baseline exercise (see Figure C.1 in the online appendix).

Calibration of γ and σ Our baseline results assumed that the productivity growth after the Industrial Revolution was equally shared between the productivity growth of new technologies (σ) and of new vintages (γ). 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 C.2a in the online appendix depicts the dynamics of productivity growth when balanced growth comes only from the development of better vintages (i.e., $\sigma = 0$), while C.2b 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 (2.50 when growth comes from γ, 3.29 when it comes from σ, vs. 2.95 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.

Trend vs. structural break in the intensity of use In our estimation, we have assumed that the evolution of the intensity of use parameter follows a linear trend. We analyze an alternative specification in which the intensity of use parameter has a structural break, rather than following a linear function.[64] Figure C.3a in the online appendix shows that the key features of the cross-country evolution of productivity growth are robust to this modeling choice. The magnitude of the increase in the income gap between Western and non-Western countries

[64]The the linear model provides a better statistical fit as measured by the R^2. The R^2 is 0.11 in the baseline exercise, and 0.02 with a structural break in 1913.

is similar (2.63 vs. 2.95 in the benchmark), and the predicted timing of the divergence is very similar to the baseline formulation.

Slow acceleration of the technology frontier In our baseline exercise, we characterize the transition to modern growth with a one-time acceleration of the technology frontier in 1765. This stark assumption allows us to cleanly illustrate that protracted transitional dynamics arise only from the dynamics of technology adoption. However, there is some contention among historians about whether the industrial revolution induced a one time change or a smoother transition in the growth rate of the world technology frontier (see e.g., Broadberry *et al.* (2015)). We investigate how sensitive our results are to the one-time acceleration assumption. We re-do our baseline calibration assuming that the technology frontier slowly accelerated from .2% to 2% over a hundred years. In particular, we assume a linear trend in the growth rate starting in 1765. Figure C.3b in the online appendix shows the evolution of income of Western vs. non-Western countries in this case. The bottom line is that the transition looks similar to our baseline exercise. The slow transition of the frontier makes the transition to modern growth slower for all countries, but specially for non-Western countries because it takes them longer to adopt the newer, more productive technologies relative to the baseline simulation. This exacerbates the accumulated income per capita differences over the 1820-2000 period which represent an increase in the income gap by a factor of 3.1 vs. 2.95 in the baseline simulation.

6 Conclusions

What accounts for cross-country differences in income growth over the long term? Often economists answer this question by attributing cross-country differences in income growth to TFP growth. In this paper, we have developed a methodology that permits us to explore further this fundamental question. A critical intermediate step in our approach is to analyze what drives differences in the diffusion curves across countries and over time. In particular, we have separately identified the role on diffusion of economy-wide factors such as TFP, and two distinct margins of adoption that are specific to each technology and country: the adoption lag and the intensity of use.

We have documented two distinct trends of these adoption margins over the last two centuries using a panel that covers twenty-five technologies and 139 countries. Adoption lags have converged across countries, while the intensity of use has diverged. We have evaluated the importance of the cross-country evolution of technology diffusion for productivity growth by feeding the estimated evolution of the technology adoption margins in the aggregate representation of our model economy. Our simulations have shown that differences in technology diffusion patterns account for a major part of the evolution of the world income distribution over the last two centuries. In particular, differences in the evolution of adoption margins in Western and non-Western countries account for around 75% of the income per capita divergence observed between 1820 and 2000.

These findings have important implications for the study of the long-run drivers of income differences and the Great Divergence in particular. First, our findings show that the majority of the divergence in productivity can be accounted for by technology diffusion. Second, the critical dimension of technology diffusion to understand the divergence in productivity is the intensity of use of technology. Third, economy-wide factors such as exogenous TFP not only played a minor role in explaining cross-country productivity dynamics. They also played a minor role in explaining differences in the evolution of technology diffusion across countries.

Our findings prompt new questions that we plan to investigate in future work. Probably, the most significant question is: what are the drivers of the cross-country divergence in the intensity of use? We have divided our tentative answers in two classes of explanations. The first class is based on intrinsic differences between the technologies invented early in our sample and those invented later on. We have superficially explored the relevance of this explanation for the technologies in our sample. We do not find much promise for the dimensions we have considered. For example, the magnitude of capital required to adopt a technology appears to be uncorrelated with the invention date.[65] We reach similar conclusions when exploring the correlation between the invention date and the tradability of the goods and services produced with technologies, and whether the technology is a general-purpose technology.

The second class of explanations relates the evolution of the intensity of use with the divergence of some relevant attribute across countries. The difficulty of building an argument along these lines is that some of the obvious candidates are either stable over time (e.g., geography) or have converged over the last 200 years (e.g., the quality of institutions). Despite the apparent difficulty of uncovering fundamental drivers of technology diffusion, we consider that successful theories of current cross-country differences in productivity will have to properly account for the divergence in the intensity of use of technology. These explorations shall complement our analysis towards a fuller understanding of cross-country income dynamics.

[65] In the first half of the nineteenth century we have technologies that require little capital (e.g., spindles, and mail) and others that require significant capital deployment by adopters (e.g., railways, telegraph). The split is also relatively even for technologies invented in the last fifty years in our sample. Cellphone services and blast-oxygen furnaces are quite capital intensive, while transplants, PCs and internet are less capital intensive. Conversely, among the technologies with the largest gap between Western and non-Western countries in the intensity of use, there are both capital intensive (e.g., blast oxygen furnaces), and less capital intensive technologies (e.g., liver and heart transplants).

References

Acemoglu, D., Johnson, S. and Robinson, J. (2004). *Institutions as the Fundamental Cause of Long-Run Growth*. NBER Working Papers 10481, National Bureau of Economic Research, Inc.

Aghion, P. and Howitt, P. (1992). A Model of Growth through Creative Destruction. *Econometrica*, **60** (2), 323–51.

Allen, R. C. (2012). Technology and the Great Divergence: Global economic development since 1820 . *Explorations in Economic History*, **49** (1), 1 – 16.

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.

Benassy, J.-P. (1996). Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters*, **52** (1), 41 – 47.

Broadberry, S., Campbell, B. M. S., Klein, A., Overton, M. and van Leeuwen, B. (2015). *British Economic Growth, 12701870*. No. 9781107070783 in Cambridge Books, Cambridge University Press.

Clark, G. (1987). Why isn't the whole world developed: Lessons from the cotton mills. *The Journal of Economic History*, **47** (1), 141–173.

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

— and — (2009b). Lobbies and Technology Diffusion. *The Review of Economics and Statistics*, **91** (2), 229–244.

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

— and Mestieri, M. (2010). *An Intensive Exploration of Technology Adoption*. Tech. rep., NBER.

— and — (2013). *If Technology Has Arrived Everywhere, Why has Income Diverged?* NBER Working Papers 19010, National Bureau of Economic Research, Inc.

— and Nanda, R. (2014). *Financial Development and Technology Diffusion*. CEPR Discussion Papers 10251, C.E.P.R. Discussion Papers.

Comin, D. A., Lashkari, D. and Mestieri, M. (2015). *Structural Change with Long-run Income and Price Effects*. Working Paper 21595, National Bureau of Economic Research.

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.

Easterly, W. and Levine, R. (2001). It is 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.

Gancia, G. A., Müller, A. and Zilibotti, F. (2013). *Advances in Economics and Econometrics: Theory and Applications*, Tenth World Congress of the Econometric Society (Econometric Society Monographs), chap. Structural Development Accounting.

Jerzmanowski, M. (2007). Total factor productivity differences: Appropriate technology vs. efficiency. *European Economic Review*, **51** (8), 2080 – 2110.

Jones, C. I. and Williams, J. C. (2000). Too Much of a Good Thing? The Economics of Investment in R&D. *Journal of Economic Growth*, **5** (1), 65–85.

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.

Klette, T. J. and Kortum, S. (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy*, **112** (5), 986–1018.

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*.

— (2010). *Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD*. Tech. rep., University of Groningen.

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

— (2000). Innovation and Its Enemies: The Economic and Political Roots of Technological Inertia. In *A Not-so-dismal Science*, Oxford University Press, pp. 61–91.

— (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.

Olson, M. (1982). *The Rise and Decline of Nations: Economic Growth, Stagflation, and Social Rigidities*. Yale University Press.

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.

Parente, S. L. and Prescott, E. C. (1994). Barriers to technology adoption and development. *Journal of Political Economy*, **102** (2), 298–321.

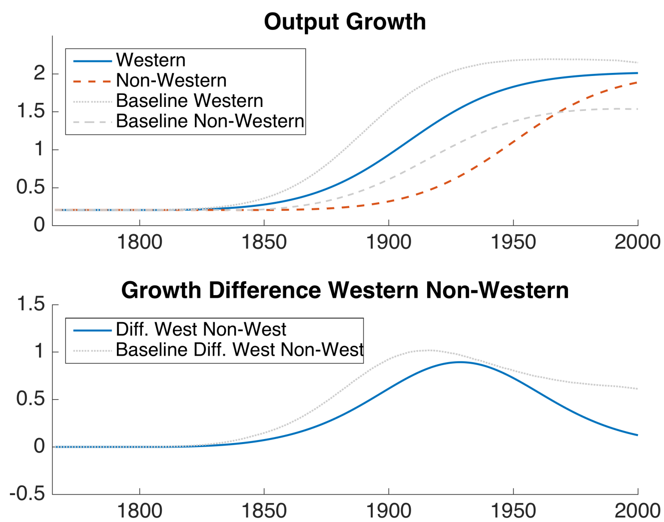
Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, **11** (4), 707 – 720.

Rios-Rull, J.-V. and Krusell, P. (1999). On the Size of U.S. Government: Political Economy in the Neoclassical Growth Model. *American Economic Review*, **89** (5), 1156–1181.

Appendix

A Additional Figures

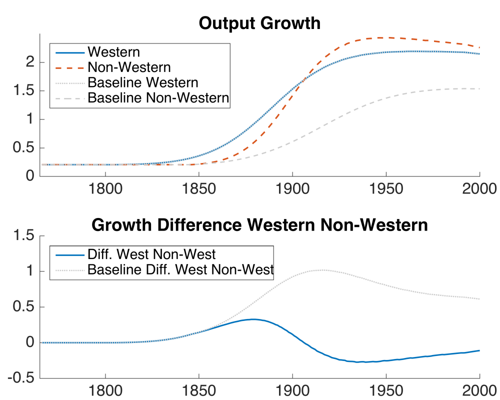
Figure 6: Acceleration of the World Technology Frontier



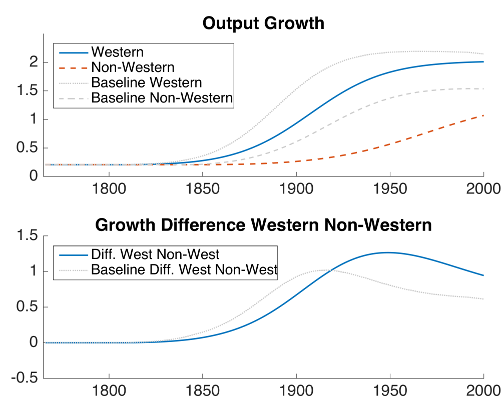
Growth of Western and non-Western countries with *only* an acceleration of the technology frontier. Both margins of adoption are held constant.

Figure 7: Role played by the different margins of adoption

(a) Dynamics Generated by Adoption Lags



(b) Dynamics Generated by the Intensity Of Use Parameter



B Proof of Propositions

Proof of Proposition 1: The demands for labor and intermediate good at the vintage level are given by:

$$\alpha \frac{P_{\tau,v} Y_{\tau,v}}{X_{\tau,v}} = (1+\zeta_{\tau x}) R_x \tag{B.1}$$

$$(1-\alpha) \frac{P_{\tau,v} Y_{\tau,v}}{L_{\tau,v}} = (1+\zeta_L) W \tag{B.2}$$

where R_x, and W are the prices of intermediate good and labor, and $\zeta_{\tau x}$ and ζ_L are the tax rates on intermediate goods and labor. The price of vintage v output is

$$P_{\tau,v} = \frac{\left(\frac{(1+\zeta_{\tau x})R_x}{\alpha}\right)^{\alpha} \left(\frac{(1+\zeta_L)w}{1-\alpha}\right)^{1-\alpha}}{e^{\chi_t} Z(\tau, v)}. \tag{B.3}$$

The demand for the producer of output associated with technology vintage τ, v is

$$Y_{\tau,v} = \int_0^{N_\tau} Y^i_{\tau,v} di = N_\tau Y^i_\tau \left(P^i_\tau / P_{\tau,v}\right)^{-\mu/(\mu-1)}, \tag{B.4}$$

where the second equality takes advantage of the symmetric demands across producers of differentiated output producers.

The unit cost of a technology τ differentiated output, P^i_τ, is

$$P^i_\tau = \left(\int_\tau^{t-D_\tau} P_{\tau,v}^{-1/(\mu-1)} dv\right)^{-(\mu-1)} \tag{B.5}$$

$$= \frac{\left(\frac{(1+\zeta_{\tau x})R_x}{\alpha}\right)^{\alpha} \left(\frac{(1+\zeta_L)w}{1-\alpha}\right)^{1-\alpha}}{e^{\chi_t + (\sigma+\gamma)\tau}} \left(\frac{\mu-1}{\gamma}\right)^{-(\mu-1)} \left(e^{\frac{\gamma(t-D_\tau-\tau)}{\mu-1}} - 1\right)^{-(\mu-1)} \tag{B.6}$$

Where the second equality follows from substituting (B.3) and integrating.

Given the symmetry in the pricing of differentiated outputs, the unit price of technology τ final output, P_τ, is

$$P_\tau = P^i_\tau N_\tau^{-(\psi-1)} \tag{B.7}$$

Combining, equations (B.1) through (B.7), we can derive the following expressions for the demands for labor and intermediate goods associated with the production of technology τ

output

$$L_\tau = \int_\tau^{t-D_\tau} L_{\tau,v} dv = \int_\tau^{t-D_\tau} \frac{(1-\alpha)P_{\tau,v}Y_{\tau,v}}{(1+\zeta_L)W} dv = \frac{(1-\alpha)P_\tau Y_\tau}{(1+\zeta_L)W}, \tag{B.8}$$

$$X_\tau = \int_\tau^{t-D_\tau} X_{\tau,v} dv = \int_\tau^{t-D_\tau} \frac{\alpha P_{\tau,v}Y_{\tau,v}}{(1+\zeta_{\tau x})R_x} dv = \frac{\alpha P_\tau Y_\tau}{(1+\zeta_{\tau x})R_x}. \tag{B.9}$$

Given these demands, sectoral output can be represented by

$$Y_\tau = A_\tau X_\tau^\alpha L_\tau^{1-\alpha} \tag{B.10}$$

where the sectoral TFP level is equal to

$$A_\tau = e^{\chi_t} N_\tau^{(\psi-1)} \left(\int_\tau^{t-D_\tau} Z(\tau,v)^{1/(\mu-1)} dv \right)^{\mu-1} \tag{B.11}$$

$$= \left(\frac{\mu-1}{\gamma} \right)^{\mu-1} \overbrace{e^{\chi_t}}^{\text{Exog. TFP}} \underbrace{N_\tau^{(\psi-1)}}_{\text{Intensity}} \underbrace{e^{\sigma\tau+\gamma(t-D\tau)}}_{\text{Embodiment}} \overbrace{\left(1 - e^{\frac{-\gamma}{\mu-1}(t-D_\tau-\tau)}\right)^{\mu-1}}^{\text{Variety}}. \tag{B.12}$$

This proofs claims (i) and (iii) in proposition 1. Claim (ii) follows from the demand for technology τ output, Y_τ. Cost minimization in the production of Y, implies that

$$Y_\tau = Y P_\tau^{-\theta/(\theta-1)} \tag{B.13}$$

where we have imposed that the price of aggregate output is the numeraire. This proves claim (ii). □

It follows from (B.13) and (8) that

$$Y_\tau^i = Y_\tau N_\tau^{-\psi} \tag{B.14}$$

Proof of Proposition 2: Aggregating across technologies, τ, we can define X and L as follows

$$\begin{aligned} X &= \int_{-\infty}^{t-D_\tau} X_\tau d\tau = \int_{-\infty}^{t-D_\tau} \frac{\alpha P_\tau Y_\tau}{(1+\zeta_x)R_x} d\tau \\ &= \frac{\alpha}{(1+\zeta_x)R_x} \int_{-\infty}^{t-D_\tau} P_\tau Y_\tau d\tau = \frac{\alpha Y}{(1+\zeta_x)R_x}, \end{aligned} \tag{B.15}$$

$$\begin{aligned} L &= \int_{-\infty}^{t-D_\tau} L_\tau d\tau = \int_{-\infty}^{t-D_\tau} \frac{(1-\alpha)P_\tau Y_\tau}{(1+\zeta_L)W} d\tau \\ &= \frac{(1-\alpha)}{(1+\zeta_L)W} \int_{-\infty}^{t-D_\tau} P_\tau Y_\tau d\tau = \frac{(1-\alpha)Y}{(1+\zeta_L)W}, \end{aligned} \tag{B.16}$$

where the last equality follows from the demand (B.13), and the fact that the price of aggregate output is the numeraire.

It follows from (B.15) and (B.16) that:

$$X^\alpha L^{1-\alpha} = \frac{Y}{\bar{c}} \tag{B.17}$$

where

$$\bar{c} \equiv \left(\frac{(1+\zeta_L)W}{1-\alpha}\right)^{1-\alpha} \left(\frac{(1+\zeta_x)R_x}{\alpha}\right)^{\alpha} \tag{B.18}$$

The price of aggregate output

$$\begin{aligned} P &= 1 = \left(\int_{-\infty}^{t-D_\tau} P_\tau^{-1/(\theta-1)} d\tau\right)^{-(\theta-1)} \\ &= \bar{c}\left(\int_{-\infty}^{t-D_\tau} A_\tau^{1/(\theta-1)} d\tau\right)^{-(\theta-1)} \\ &= \frac{\bar{c}}{A} \end{aligned} \tag{B.19}$$

where

$$A = \left(\int_{\tau}^{t-D_\tau} A_\tau^{1/(\theta-1)} d\tau\right)^{\theta-1} \tag{B.20}$$

Substituting (B.19) into (B.17) yields the expression for the aggregate production function in the proposition:

$$Y = AX^\alpha L^{1-\alpha} \tag{B.21}$$

□