An Intensive Exploration of Technology Diffusion

Diego Comin — Harvard University and NBER

Martí Mestieri — MIT

Preliminary Version

March 2010*

Abstract

We present a tractable model for the analysis of the relationship between economic growth and the intensive and extensive margins of technology adoption. At the aggregate level, our model is isomorphic to a neoclassical growth model. The microeconomic underpinnings of growth come from technology adoption of firms, both at the extensive and the intensive margin. We use a data set of 15 technologies and 166 countries to estimate the intensive and extensive margin of adoption using the structural equations derived from our model. We find that the variability across countries in the intensive margin is higher than in the extensive margin. The cross country variation in intensive margin of adoption accounts for around 40% of the variation in income per capita.

Keywords: Economic Growth, Technology Adoption, Cross-country studies.

JEL-code: E13, O14, O33, O41.

*We would like to thank Mar Reguant for comments and suggestions. Comin would like to thank the NSF (Grants # SES-0517910 and SBE-738101) for their financial assistance. The views expressed in this paper solely reflect those of the authors and not necessarily those of the National Bureau of Economic Research.

There is a growing consensus that differences in TFP play a major role in explaining the current differences in income per capita across countries. To the extent that productivity is embodied in the technologies used for production, this implies that cross country differences in TFP should translate to differences in technology usage. Cross-country differences in technology can take alternative, non-exclusive, forms. They can arise because of differences in *(i)* the range of technologies available for production, *(ii)* the number of producers that have adopted a given technology, and *(iii)* the number of units of the technology adopted by each adopter. The first margin is determined by the lag with which technologies are adopted in a country (i.e., the extensive margin), while the second and the third fall in the realm of the intensive margin of adoption. The goal of this paper is to document cross-country differences in the intensive margin of adoption and to explore its role in generating differences in technology usage, TFP and, ultimately, income per capita across countries.

Figure 1 reports the number of land line phone calls normalized by total output for the U.S., Japan, Paksitan and Malawi. One feature of this plot is that these are similar curves, but displaced vertically and horizontally. Intuitively, the horizontal shifts in the curves are associated with differences in the adoption lags, while vertical shifts reflect differences in the intensity of adoption. Note that the differences in the intensive margin can be sizable. For instance, the difference between the U.S. and Pakistan is on the order of two log points (while their log difference in income per capita in 2000 is around two and a half). Comin and Hobjin (2010) document cross country differences in the extensive margin (horizontal shifts in the figure). In this paper, we build a model and conduct an empirical analysis to document both the intensive and extensive margins of adoption. Thus, this work can be seen as complementing theirs along the intensive margin dimension.

There is a long list of factors that may affect the intensity of adoption of new technologies and, through this channel, aggregate productivity.[1] Our goal in this paper is not to assess how important a particular factor is in affecting the intensity with which new technologies are used. At this point, we are only interested in understanding how important are cross-country differences in

[1] One way to classify these factors is according to whether they affect the costs or the benefits for individual producers of adopting new technologies. Distortions in the product, capital and labor markets, may reduce a company's gains from adopting a new technology. On the cost side, differences in training, past adoption history, formal schooling, credit markets, and several other factors may generate cross-country differences in the costs faced by individual producers to adopt new technologies.

the intensity of adoption of new technologies to explain cross-country differences in productivity. To answer this more primitive question, it is not necessary to take a stand on the driver of the cross-country variation in the intensity of adoption.[2]

Given a range of technologies available for production, a relatively more intensive use of the new technologies leads to higher aggregate TFP because new technologies tend to embody a higher productivity. Yet, virtually nothing is known about the significance of the intensive margin of technology adoption for productivity. Clark' s (1988) classical study on spinning spindles documents large cross-country differences in the number of spindles that each worker operated circa 1900 and argues that this was a major contributor to differences in productivity. However, observing a more intensive adoption of a new technology in rich countries is not sufficient to establish the importance of the intensive margin for aggregate productivity, since this could be implied by reverse causality.[3]

Dealing with the effect of income on the observable measures of technology is a key challenge that any attempt to assess the importance of the intensive margin needs to confront. The approach we follow in this paper consists in developing a general equilibrium model of technology adoption that incorporates both an intensive and an extensive margin. Our model has two properties. First, at the aggregate level it is similar to the one-sector neoclassical growth model. Second, at the disaggregate level it has implications for the path of observable measures of technology adoption. In particular, it predicts the effect of income, adoption lags and the parameters that affect the intensity of adoption on our measures of technology adoption at any point in time. These two properties allow us to estimate for each country-technology pair the frictions that determine the intensive adoption margin. We then use the model to assess the effects of these frictions on aggregate productivity.

[2]For instance, cross-country differences in total factor productivity can be driven by differences in technology or by differences in the efficiency with which technologies are used. We do not try to disentangle these two in this paper. As will become apparent when we discuss the formulation of the model, we take a parsimonious approach when modelling the costs of adoption that can encompass the different examples described above.

[3]An example can clarify this point. We see more computers being adopted in Germany than in Kenya. This is in part the case because Germans are richer than Kenyans and they buy more computers (and more of everything). But there may be other reasons why Germans may buy more computers than Kenyans. In Kenya, it may be harder to obtain credit, it may be more costly to train workers to use computers, it may be harder to obtain the permits to start the business that will use the computers or there may be a higher risk that somebody steals the computers from the company.

As pointed out before, our model features the different margins that drive differences in technology adoption across countries. The first is the extensive margin, which determines the range of technologies used in a country. The second is the intensive margin, which can be decomposed into how many producers have adopted a given technology, and how many units of the technology each producer has adopted. All these three margins of decision are determined in equilibrium and are affected by a cost of bringing a technology to the country and the cost incurred by each individual producer to learn how to use the new technology. Moreover, we allow for a tax in capital to incorporate the possibility of differences in the efficiency with which technologies are used across countries.

A technology, in our model, is a group of production methods that is used to produce an intermediate good or service. Each production method is embodied in a differentiated capital good. A potential producer of a capital good decides whether to incur a fixed cost of adopting the new production method. If he does, he will be the monopolist supplying the capital good that embodies the specific production method. This decision determines whether or not a production method is used, which is the extensive margin of adoption.

Once the production method has been introduced, several factors determine how many units of the associated capital good are demanded, which reflect the intensive margin of adoption. These are the productivity it embodies, the price (and potential tax) of the capital good and the cost faced by individual producers to learn how to use it. Other things equal, these variables shift vertically the evolution of observable measures of technology adoption, such as the number of units of capital that embody a given technology or the output produced with this technology.

As in Comin and Hobijn (2009), an increase in the range of production methods used results in a gain from variety that boosts productivity. This variety effect leads to a non-linear trend in the embodied productivity level. Since adoption lags affect the range of production methods used, and thus the variety effect, adoption lags affect the curvature of the path of embodied productivity. Our model maps this curvature in embodied productivity into similar non-linearities in the evolution of observable measures of technology adoption. We use this curvature in the data to identify adoption lags.

The model delivers a system of structural equations for technology that can be estimated from the data. A key parameter in the estimating equations is the income elasticity of technology.

Our baseline model builds upon homothetic production functions (CES) and, as a result, imposes an income elasticity of one. We relax this restriction introducing a non-homethetic production function, thus not requiring any particular value for the income elasticity.

Our model delivers an expression for aggregate TFP that explicitly accounts for the intensive and extensive margins of adoption. We exploit this result to assess the differences in TFP that our estimated differences in the intensive margin generate. As our model has a clear mapping from TFP differences to income per capita, we can then explore how large are the differences in income per capita generated by differences in the intensive margin of adoption.

We use data for 15 technologies and 166 countries, as in Comin, Hobjin, and Rovito (2006). Our data cover major technologies related to transportation, telecommunication, IT, health care, steel production, and electricity. We obtain precise and plausible estimates of the adoption lags for two thirds of the 1278 technology-country pairs for which we have sufficient data.

Results There are three main findings that we want to emphasize from our exploration. First, the magnitude of cross-country differences in the intensive margin of adoption and adoption lags are large. On average, the logdifference of intensive margin of adoption with the U.S. is -.43. The average standard deviation of this measure is .72 and the average interquartile range is .79.[4] The analysis of variance reveals that 24% of this variation can be attributed to technology cross variation, a 51% can be attributed to cross-country variation and a 2%, to the covariance between the two. The remainder 23% remains unexplained. In contrast with the findings with adoption lags, we do not find a decrease over time in the differences of intensive margin of adoption across countries. They have remained more or less constant over time, and if anything, the dispersion has increased.

Second, we find that differences in intensive margin of adoption account for around 40% of the cross-country differences in income per capita.

Finally, we show that the empirical results obtained in the baseline model go through when we allow for non-homothetic production functions. For instance, the quantitative implications of the intensive margin on cross-country differences in income per capita are almost identical.

This paper is related to three strands of the literature. First, macroeconomic models of technology adoption (e.g. Parente and Prescott, 1994, and Basu and Weil, 1998) have tried to understand

[4]The interquantile range is defined as the difference between the 75th and the 25th percentile

the role of technology for TFP. However, these studies have used an abstract concept of technology that is hard to match with data. Second, the applied microeconomic technology diffusion literature (Griliches, 1957, Mansfield, 1961, Gort and Klepper, 1982, among others) focuses on the estimation of diffusion curves for a relatively small number of technologies and countries. These diffusion curves, however, are purely statistical descriptions which are not embedded in an aggregate model. Hence, it is difficult to use them to explore the aggregate implications of the empirical findings. Finally, Comin and Hobijn (2010) model only the extensive margin of adoption. Comin, Easterly and Gong (2010) are interested in exploring the persistence of technology over long periods of time but do not develop a model. As a result, they cannot estimate the extensive and intensive margin of adoption and cannot draw the implications of their estimates of the persistence of technology for aggregate productivity.

The paper is divided into four sections. Section 1 sets out a one-sector neoclassical growth featuring intensive and extensive margins of adoption. Section 2 describes the diffusion patterns of technology under the balanced growth path assumption, derives the structural equations that can be estimated from the data and explains how the margins of adoption are identified. Section 3 presents the results of the estimation and Section 4 concludes.

1 A one-sector growth model with extensive and intensive technology adoption

We next present a one-sector growth model with endogenous technology adoption at the extensive and intensive margins. The model maps the adoption margins into the time-path of observable measures of technology diffusion and illustrates how each adoption margin affects endogenous TFP differentials. In what follows, we omit the time subscript, t, where obvious.

1.1 Preferences

A measure one of households populates the economy. They inelastically supply one unit of labor every instant, at the real wage rate W, and derive the following utility from their consumption flow

$$U = \int_0^\infty e^{-\rho t} \ln(C_t) dt. \tag{1}$$

Here C_t denotes per capita consumption and ρ is the discount rate. We further assume that capital markets are perfectly competitive and that consumers can borrow and lend at the real rate $\widetilde{r}$.

1.2 Production

Technology:

Each instant, a new production method appears exogenously. We call these production methods, technology vintages or simply vintages. Production methods are capital embodied. The set of vintages available at time t is given by $\overline{V} = (-\infty, t]$. The productivity embodied in new vintages grows at a rate γ across vintages, such that

$$Z_v = Z_0 e^{\gamma v}. \tag{2}$$

Note that Z_v is constant over time. This characterizes the evolution of the world technology frontier. We shall choose the normalization parameter Z_0 such that vintage $\underline{v}$ has productivity $Z_{\underline{v}}$.[5]

A country does not necessarily use all the capital vintages that are available in the world because, as we discuss below, making them available for production is costly. The set of vintages actually used is given by $V = (-\infty, t - D]$. Here $D \geq 0$ denotes the *adoption lag*. That is, the amount of time between when the best technology in use in the country became available and when it was introduced in the country.

In order to map the model into the data, we introduce the concept of technology. A technology is a set of production methods used to produce closely related intermediates. In particular, we consider two technologies: an old one, denoted by o, and a new one, denoted by n. The old technology consists of the production methods introduced up till a fixed time $\underline{v}$, such that the set of vintages associated with the old technology is $V_o = (-\infty, \underline{v}]$. The new technology consists of the newest production methods, invented after $\underline{v}$, such that it covers $V_n = (\underline{v}, t - D]$.

Output:

The output associated to a technology τ, Y_τ, is given by:

$$Y_\tau = \left(\int_{V_\tau} Y_v^{\frac{1}{\mu}} dv \right)^{\mu}, \qquad \tau \in \{o, n\}, \tag{3}$$

[5] This implies that $Z_0 = Z_{\underline{v}} e^{-\gamma \underline{v}}$.

where Y_v denotes the intermediate output produced using technology vintage v. Final output Y, is produced competitively with the following production function:

$$Y = \left(\int_{-\infty}^{t-D} Y_v^{\frac{1}{\mu}} dv \right)^{\mu} = \left(\sum_{\tau \in \{o,n\}} Y_\tau^{\frac{1}{\mu}} \right)^{\mu}.$$

Once a technology vintage v is brought to the country, producers can find distinct ways to use it. Because each application developed solves a new problem, the larger the number of applications developed, N_v, the more efficient the production of intermediate service v is. In other words, there are efficiency gains from developing more applications.Each application yields a differentiated output, Y_{vi}. Differentiated oputputs are produced monopolistically. A competitive producer then aggregates these outputs in the form of intermediate v, Y_v, as follows:[6]

$$Y_v = N_v^{-(\mu-\mu')} \left(\int_0^{N_v} Y_{vi}^{\frac{1}{\mu}} di \right)^{\mu}, \text{ with } \mu > \mu' > 1. \tag{4}$$

Output Y_{vi} is produced by combining labor and capital, K_{vi}, that embodies production method, v, as follows:

$$Y_{vi} = Z_v L_{vi}^{1-\alpha} K_{vi}^{\alpha}, \tag{5}$$

Capital goods production and taxes:

Capital goods are produced by monopolistic competitors. Each of them holds the patent of the capital good used for a particular production method v. It takes one unit of final output to produce one unit of capital of any vintage. This production process is assumed to be fully reversible. For simplicity, we assume that there is no physical depreciation of capital. The capital goods suppliers rent out their capital goods at the rental rate $\phi_R R_v$. R_v is the price received by the capital goods producer, while the wedge ϕ_R captures a tax on the price of capital that the government rebates back to the consumers with a lump sum transfer. ϕ_R is constant across vintages and over time. Below we show that ϕ_R can capture a wide range of institutional distortions that affect the efficiency of the economy.

Technology adoption costs:

There are two types of adoption costs. The cost of bringing to the country the production method associated with a capital vintage, Γ_{vt}^e, and the cost incurred by each individual producer to find

[6]This specification is similar to Benassy (1996). The assumption that $\mu > \mu'$ ensures that the profits of a individual producer decline with N_v.

a distinct application of a production method that is already available, Γ^i_{vt}. We define the former as the extensive and the latter as the intensive adoption costs. Both of these are sunk costs. The extensive cost of adopting vintage v at time t is given by (6) while the intensive cost of adoption is given by (7).

$$\Gamma^e_{vt} = \frac{\alpha}{\epsilon}\overline{\Psi}\left(1+b_e\right)\left(\frac{Z_v}{Z_t}\right)^{\frac{1+\vartheta}{\mu-1}}\left(\frac{Z_t}{A_t}\right)^{\frac{1}{\mu-1}} Y_t, \text{ where } \vartheta > 0 \tag{6}$$

$$\Gamma^i_{vt} = \frac{\mu-1}{\mu}\overline{\Psi}\left(1+b_i\right)\left(\frac{Z_v}{Z_t}\right)^{\frac{1}{\mu-1}}\left(\frac{Z_t}{A_t}\right)^{\frac{1}{\mu-1}} Y_t. \tag{7}$$

In these expressions, A_t is the aggregate level of TFP to be defined below, b_e, b_i, and $\overline{\Psi}$ are constants. The parameters b_e and b_i reflect barriers to adoption for the agent that adapts the technology to the idiosyncrasies of the country or for individual producers that find a profitable use for the technology. $\overline{\Psi}$ is the steady state stock market capitalization to GDP ratio and is included for normalization purposes. The term (Z_v/Z_t) captures the idea that it is more costly to adopt technologies the higher is their productivity relative to the productivity of the frontier technology. The last two terms capture that the cost of adoption is increasing in the market size. We choose these formulations because, just like the adoption cost function in Parente and Prescott (1994), it yields the existence of an aggregate balanced growth path.[7]

1.3 Factor demands, output, and optimal adoption

The demand for the output produced with vintage v is:

$$Y_v = Y\left(P_v\right)^{-\frac{\mu}{\mu-1}}, \text{ where } P = \left(\int_{v\in V} P_v^{-\frac{1}{\mu-1}} dv\right)^{-(\mu-1)}. \tag{8}$$

We use the final good as the numeraire good throughout our analysis and normalize its price to $P = 1$. The demand faced by the ith producer of differentiated output associated to vintage v is:

$$Y_{vi} = Y_v\left(\frac{P_{vi}}{P_v}\right)^{-\frac{\mu}{\mu-1}} N_v^{-\frac{\mu-\mu'}{\mu-1}}, \text{ where } P_v = N_v^{\frac{\mu-\mu'}{\mu-1}}\left(\int_0^{N_v} P_{vi}^{-\frac{1}{\mu-1}} di\right)^{-(\mu-1)}. \tag{9}$$

Note that all producers of differentiated outputs associated to a given vintage face the same demand and have access to the same technology. As a result, they will charge the same price which

[7] It could of course be the case that the linearity in the adoption cost function is violated for some particular technology for some particular country, without necessarily violating balanced growth, but to the extent that we are documenting adoption lags across many technologies this is perhaps not so critical.

is given by a constant markup, μ, times the marginal cost of production:

$$P_{vi} = \mu \left[\frac{1}{Z_v} \left(\frac{\phi_R R_v}{\alpha} \right)^{\alpha} \left(\frac{W}{1-\alpha} \right)^{1-\alpha} \right], \text{for } i \in [0, N_v]$$

where R_v is the rental price of a unit of capital that embodies vintage v, ϕ_R is a tax on capital, and W is the wage rate. From (9), this implies that

$$P_v = N_v^{-(\mu'-1)} P_{vi}.$$

The revenue share of capital is α and labor exhaust the remaining revenue. This implies that the total demand faced by the producer of the capital good that embodies vinatge v is:

$$\begin{aligned} K_v &= \int_0^{Nv} K_{vi} di = \frac{\alpha P_v Y_v}{\phi_R R_v} \\ &= Y \left(\frac{Z_v}{\mu} \right)^{\frac{1}{\mu-1}} N_v^{\frac{\mu'-1}{\mu-1}} \left(\frac{(1-\alpha)}{W} \right)^{\frac{1-\alpha}{\mu-1}} \left(\frac{\alpha}{\phi_R R_v} \right)^{\epsilon}, \text{ where } \epsilon \equiv 1 + \frac{\alpha}{\mu - 1}. \end{aligned}$$

The supplier of each capital good takes as given the number of differentiated output producers but recognizes that the rental price he charges for the capital good, R_v, affects the price of the output associated with the capital good and, therefore, its demand, Y_v. The price elasticity of demand she faces, ϵ, is constant. As a result, the profit maximizing rental price equals a constant markup times the marginal production cost of a unit of capital, which we assume is equal to a unit of final output.

Because of the durability of capital and the reversibility of its production process, the per-period marginal production cost of capital is the user-cost of capital. Thus, the rental price that maximizes the profits accrued by the capital good producer is

$$R_v = R = \frac{\epsilon}{\epsilon - 1} \widetilde{r}, \tag{10}$$

where $\frac{\epsilon}{\epsilon-1}$ is the constant gross markup factor.

Aggregate representation:

Our model has the following aggregate representation of production:

$$Y = AK^{\alpha} L^{1-\alpha}, \text{ where } K \equiv \int_{-\infty}^{t} K_v dv, \ L \equiv \int_{-\infty}^{t} L_v dv \tag{11}$$

Aggregate TFP, A, can be expressed as

$$A = \left[\int_{-\infty}^{t-D} \left(N_v^{\mu'-1} Z_v \right)^{\frac{1}{\mu-1}} dv \right]^{\mu-1} \tag{12}$$

Optimal adoption:

The flow profits accrued by producers of differentiated outputs associated with vintage v are equal to

$$\pi_{vi} = \frac{\mu-1}{\mu} P_{vi} Y_{vi} = \frac{\mu-1}{\mu} Y \left(\frac{Z_v}{A} \right)^{\frac{1}{\mu-1}} N_v^{-\frac{\mu-\mu'}{\mu-1}}$$

The market value of each differentiated output producer equals the present discounted value of the flow profits. That is,

$$M_{vi,t} = \int_t^{\infty} e^{-\int_t^s \widetilde{r}_{s'} ds'} \pi_{vis} ds = \frac{\mu-1}{\mu} \left(\frac{Z_v}{Z_t} \right)^{\frac{1}{\mu-1}} \left(\frac{Z_t}{A_t} \right)^{\frac{1}{\mu-1}} N_v^{-\frac{\mu-\mu'}{\mu-1}} \Psi_t Y_t, \tag{13}$$

where

$$\Psi_t = \left(\frac{\mu-1}{\mu} + \frac{\alpha}{\epsilon} \right) \int_t^{\infty} e^{-\int_t^s \widetilde{r}_{s'} ds'} \left(\frac{A_t}{A_s} \right)^{\frac{1}{\mu-1}} \left(\frac{Y_s}{Y_t} \right) ds \tag{14}$$

is the stock market capitalization to GDP ratio.

Optimal adoption implies that, every instant, the value of becoming a user of a technology vintage v does not exceed the intensive cost of adoption. That is, for all vintages, v, that are adopted at time t

$$\Gamma_v^i \geq M_{vi}. \tag{15}$$

Thus, in equilibrium

$$N_v = \left(\frac{\Psi_t}{\overline{\Psi}_t (1+b_i)} \right)^{\frac{\mu-1}{\mu-\mu'}}. \tag{16}$$

Given N_v, the flow profits that the capital goods producer of vintage v earns are equal to

$$\pi_v = \frac{\alpha}{\epsilon \phi_R} P_v Y_v = \frac{\alpha}{\epsilon} N_v^{\frac{\mu'-1}{\mu-1}} \left(\frac{Z_v}{A} \right)^{\frac{1}{\mu-1}} Y \tag{17}$$

The market value of each capital goods supplier equals the present discounted value of the flow profits. That is,

$$M_{v,t} = \int_t^{\infty} e^{-\int_t^s \widetilde{r}_{s'} ds'} \pi_{vs} ds = \frac{\alpha}{\epsilon \phi_R} N_v^{\frac{\mu'-1}{\mu-1}} \left(\frac{Z_v}{Z_t} \right)^{\frac{1}{\mu-1}} \left(\frac{Z_t}{A_t} \right)^{\frac{1}{\mu-1}} \Psi_t Y_t. \tag{18}$$

Optimal adoption implies that, every instant, all the vintages for which the value of the firm that produces the capital good is at least as large as the adoption cost will be adopted. That is, for all vintages, v, that are adopted at time t

$$\Gamma_v^e \geq M_v \tag{19}$$

This holds with equality for the best vintage adopted if there is a positive adoption lag.

The adoption lag that results from this condition equals

$$D_v = \max\left\{\frac{\mu-1}{\gamma\vartheta}\left[\ln\left(1+b_e\right)+\ln\phi_R-\frac{\mu'-1}{\mu-1}\ln N_v+\left(\ln\overline{\Psi}-\ln\Psi\right)\right],0\right\}\equiv D$$

and is constant across vintages, v.[8] The lag with which new vintages are adopted is increasing in the adoption costs, b_e, and the tax wedge, ϕ_R, is decreasing in N_v and in the deviation of the stock market to output ratio from its steady state level. As shown in equation (16), the number of producers that develop distinct uses for technology vintage v, N_v, declines with the intensive cost of adoption, b_i, and increases with the deviation of the stock market to output ratio from the balance growth level.[9]

Conversely, there are several significant factors that do not influence the adoption decisions.First, given the specifications of the production function and the costs of adoption, the market size symmetrically affect the benefits and costs of adoption at both the intensive and extensive margins. Hence, variation in market size does not affect the timing of adoption, D, and the number of producers that use a new vintage, N_v. By the same token, the adoption margins are not affected by the productivity of technology at time zero, Z_0. Second, since on the balanced growth path $\Psi = \overline{\Psi}$, the steady-state adoption lags and number of producers do not depend on the stock market to output ratio.

These observations together with equation (12) help us understand what drives aggregate TFP in this model. Three factors can drive cross-country differences in TFP: The adoption lag, the number of producers that adopts each technology vintage, and the normalized productivity level of the initial vintage. Note that, Z_0 affects directly aggregate TFP but, as mentioned above, has no effect through D and N_v. The costs of adopting new technologies affect TFP because they influence the range of technologies available for production and how many different applications are

[8] In what follows we focus on the interior case where $\Gamma_{t\tau} \leq M_{t\tau}$.

[9] Note that ϕ_R does not affect the intensity of adoption as measured by the number of producers that adopt a new vintage. That is the case because, from the perspective of the potential producers of differentiated outputs, ϕ_R only affects aggregate demand. Aggregate demand, in turn, has a symmetric effect on the costs and benefits of adopting the vintage for the differentiated output producers. Instead, corporate income taxes (or expropriation risk) also affect the profit margin net of taxes. This assymetric effect would affect the number of producers that adopt the new vintage.

developed. Finally, the wedge tax (and other related frictions) only affect aggregate TFP through their effect on the adoption decisions.

2 Diffusion of the new technology

We define the equilibrium of this economy in Appendix B. In what follows, we focus on the balanced growth path of the economy.[10] Along the balanced growth path, adoption lags, D, are constant, the number of adopters that adopt each vintage once it is available in the country is constant and equal to N, and the economy grows at a constant rate equal to $\gamma/(1-\alpha)$. However, D and N can differ across countries.

So far, we have derived expressions for output and capital at the vintage level. However, because of the nature of available data, we are interested in the total demand for capital goods and the output produced with the production methods that make up the new technology $\tau = n$. We can express output produced with technology τ in the following Cobb-Douglas form

$$Y_\tau = A_\tau K_\tau^\alpha L_\tau^{1-\alpha}, \tag{20}$$

where

$$K_\tau \equiv \int_{v\in V_\tau} K_v dv,\ L_\tau \equiv \int_{v\in V_\tau} L_v dv, \tag{21}$$

$$\text{and } A_\tau = \left(\int_{V_\tau} \left(N_v^{\mu'-1} Z_v\right)^{\frac{1}{\mu-1}} dv\right)^{\mu-1} = N_\tau^{\mu'-1}\left(\int_{V_\tau} Z_v^{\frac{1}{\mu-1}} dv\right)^{\mu-1}. \tag{22}$$

The endogenous level of TFP for technology $\tau = n$ at time t can be expressed as

$$A_n = \left(\frac{\mu-1}{\gamma}\right)^{\mu-1} \underbrace{N_n^{\mu'-1}}_{\text{intensity of adoption}} Z_{\underline{v}} \underbrace{e^{\gamma(t-D-\underline{v})}}_{\text{embodiment effect}} \underbrace{\left[1-e^{-\frac{\gamma_\tau}{\mu-1}(t-D-\underline{v})}\right]^{\mu-1}}_{\text{variety effect}}. \tag{23}$$

The evolution of the new technology TFP is driven by the adoption margins. First, there are efficiency gains from the number of producers that adopt a given vintage. This affects the level of technology through the 'intensity of adoption' term in (23). The trend in TFP is driven by the economy-wide adoption of new, more productive, vintages. The adoption lag determines the best

[10] The transitional dynamics of the model are similar to the one described in the working paper version of Comin and Hobjin (2009).

vintage adopted and affects the level of TFP through the 'embodiment effect' term in (23). Finally, adoption lags drive the curvature of A_n. The marginal gain in productivity from adopting new vintages decreases as more vintages are adopted (variety effect). We shall use this result to identify the adoption lags in the data. These properties of the path for the level of TFP for technology τ affects the output and capital associated with the technology through its effect on the marginal cost of production of the technology-specific output measured by the price P_τ.

$$P_\tau = \frac{\mu}{A_\tau}\left(\frac{W}{1-\alpha}\right)^{1-\alpha}\left(\frac{\phi_R R}{\alpha}\right)^{\alpha}. \tag{24}$$

2.1 Empirical application

Our goal is to estimate the intensity of adoption and adoption lags for the different technology-country pairs in our data set. We extend the results above by allowing multiple sectors, each adopting a new technology.[11] We do that with a nested CES aggregator, where $\frac{\theta}{\theta-1}$ reflects the between sectors elasticity of demand and $\frac{\mu}{\mu-1}$ is, just as in the one-sector model, the within sector elasticity of demand. We allow $\frac{\theta}{\theta-1}$ to vary across sectors. Further, we allow the growth rate of embodied technological change, γ_τ, and the invention date, $\underline{v}_\tau$, to vary across technologies. We denote the technology measures for which we derive reduced form equations by $m_\tau \in \{y_\tau, k_\tau\}$. Small letters denote logarithms.

These modifications yield the following demand for technology τ output

$$y_\tau = y - \frac{\theta}{\theta-1}p_\tau. \tag{25}$$

Combining that with the intermediate goods price (24)

$$p_\tau = -\alpha\ln\alpha - a_\tau + (1-\alpha)(y-l) + \alpha r + \alpha\ln\phi_R, \tag{26}$$

we obtain the reduced form equation (27) for y_τ.

$$y_\tau = y + \frac{\theta}{\theta-1}\left[a_\tau - (1-\alpha)(y-l) - \alpha r + \alpha\ln\alpha - \alpha\ln\phi_R\right] \tag{27}$$

Similarly, we obtain the reduced form equation for k_τ by combining the log-linear capital demand equationwith (25) and (26),

$$k_\tau = y + \frac{1}{\theta-1}\left[a_\tau - (1-\alpha)(y-l) - \alpha r + \alpha\ln\alpha - \alpha\ln\phi_R\right] + \ln\alpha - r - \ln\phi_R. \tag{28}$$

[11] Comin and Hobijn (2008) derive the multi-sector version of a similar model in detail.

These expressions depend on the intensive margin and adoption lag D_τ through their effect on the productivity term, a_τ. Comin and Hobijn (2010) show that, to a first order approximation,

$$a_\tau \approx (\mu' - 1)n_\tau + z_{\underline{v}_\tau} + (\mu - 1)\ln(t - T_\tau) + \frac{\gamma_\tau}{2}(t - T_\tau), \tag{29}$$

where $T_\tau = \underline{v}_\tau + D_\tau$ is the time when the technology is adopted. In this approximation, the growth rate of embodied technological change, γ_τ, only affects the linear trend in a_τ.[12] Substituting this into (25) and (28) yields the following reduced form equation

$$m_\tau = \beta_1 + y + \beta_2 t + \beta_3\left((\mu - 1)\ln(t - T_\tau) - (1 - \alpha)(y - l)\right) + \varepsilon_\tau, \tag{30}$$

where ε_τ is the error term. The reduced form parameters are given by the β's. Note that, according to our theoretical model, the intercept β_1 is given by the following expression which depends on both the intensive and extensive margins of adoption,

$$\beta_1^y = \frac{\theta}{\theta - 1}\left[((\mu' - 1)n_\tau + z_{\underline{v}_\tau}) - \frac{\gamma}{2}T - \alpha(r + \ln\phi_R - \ln(\alpha))\right], \tag{31}$$

$$\beta_1^k = \frac{1}{\theta - 1}\left[((\mu' - 1)n_\tau + z_{\underline{v}_\tau}) - \frac{\gamma}{2}T - \alpha(r + \ln\phi_R - \ln(\alpha))\right] - r - \ln\phi_R. \tag{32}$$

2.2 Identification and estimation procedure

We use the reduced form equation in (30) to identify the adoption lags and the intensity of adoption. To this end, we assume that preference parameters (ρ) and technology parameters other than adoption costs (*i.e.* θ, μ, μ', γ α, and $z_{\underline{v}_\tau}$) are the same across countries. This implies that the equilibrium interest rate (r) is also the same across countries. The distortions that affect the efficiency of the economy (ϕ_R) and the adoption cost parameters (b_e and b_i) can vary across countries. As a result, the adoption lags (D_τ) and the number of producers that adopt a technology (n) can also differ across countries.

These assumptions impose some cross-country parameter restrictions. Since the intercept term, β_1, depends on n, D_τ and ϕ_R, it can vary across countries. The trend-parameter, β_2, just depends

[12] Intuitively, when there are very few vintages in V_τ the growth rate of the number of vintages, i.e. the growth rate of $t - T_\tau$, is very large and it is this growth rate that drives growth in a_τ through the variety effect. Only in the long-run, when the growth rate of the number of varieties tapers off, the growth rate of embodied productivity, γ_τ, becomes the predominant driving force over the variety effect.

on α and γ_τ, so it is assumed to be constant across countries.[13] β_3 only depends on the technology parameter, θ, and is therefore assumed to be constant across countries. We do not estimate μ and α. Instead, we calibrate $\mu = 1.3$, based on the estimates of the markup in manufacturing from Basu and Fernald (1997), and $\alpha = 0.3$ consistent with the post-war U.S. labor share.

The parameter β_1 is a technology-country specific constant. Therefore, it can be identified by a technology-country fixed effect. Once we have an estimate of β_1, we still need an estimate of the adoption lags to obtain an estimate of the intensity of adoption. We follow Comin and Hobijn (2010) and identify the adoption lags through the non-linear trend component in equation (30), which reflects the variety effect. Intuitively, after controlling for the observables such as GDP or labor productivity, only the adoption lag affects the curvature of m_τ. That implies that, *ceteris paribus*, if we see two countries one with a steeper diffusion curve than the other at a given point in time, this means that the former started adopting the technology later.

Given that there is no particular scale for intensive margin measures, we choose to report our estimates for each technology relative to the U.S. This has the additional advantage of removing any common cross-country component in β_1. To have comparable measures of intensive margin of adoption, we need to eliminate the differential effect of ϕ_R appearing on the technologies measured using capital. For each country, we regress the intercepts we obtain (relative to the U.S., say, $\beta_{1,\tau,c} - \beta_{1,\tau,US}$) on a constant and a dummy variable that takes value of 1 if the technology is measured in capital units. The coefficient on the technology fixed effect captures the differential effect of ϕ_R. Then, we can subtract it to obtain comparable measures of intensive margin of adoption.[14]

This identification strategy assumes that the underlying curvature of the diffusion curves is the same across countries. This would not be violated, for example, if the efficiency of an economy increased over time non-linearly inducing a similar pattern in the technology measure. Of course, a

[13] The output elasticity of capital is one minus the labor share in our model. Gollin (2002) provides evidence that the labor share is approximately constant across countries.

[14] An alternative approach could be to take advantage of the fact that we have two measures of railways that are measured as output (passengers and freight) and one that is measured as capital (rail lines). Under the plausible assumption that the average intercept of the output measures (passenger and freight measures) corresponds to the rail line measure used in the country, we can back out the additional effect of ϕ_R on the capital measures of technology. Then, we can subtract this additional effect from the other capital measures and can construct comparable measures of the intensive margin. The results we obtain using this procedure are similar to the ones reported in the main text.

priori, there is no reason why the distortions in the economy evolve to induce such a specific pattern of adoption rather than affecting the trend or adding noise to the evolution of technology. Nevertheless, Comin and Hobijn (2010) take seriously this hypothesis and test formally the identification assumption by allowing β_3 to vary across countries. Then, they see how often they can reject the null that the unrestricted and the restricted estimates are the same (i.e. $\hat{\beta}_3^u = \hat{\beta}_3$). They find that they cannot reject the null in two thirds of the technology-country pairs considered. Further, the estimated adoption lags in the restricted and unrestricted specifications are very highly correlated suggesting that, effectively, the deviations from a constant curvature pattern are not quantitatively important.

Because the adoption lag is a parameter that enters non-linearly in (30) for each country, estimating the system of equations for all countries together is practically not feasible. Instead, we take a two-step approach. We first estimate equation (30) using only data for the U.S. This provides us with estimates of the values of β_1 and D_τ for the U.S. as well as estimates of β_2 and β_3 that should hold for all countries. In the second step, we separately estimate β_1 and D_τ, using (30) and conditional on the estimates of β_2 and β_3 based on the U.S. data, for all the countries in the sample besides the U.S.

Besides practicalities, this two-step estimation method is preferable to a system estimation method for two other reasons. First, in a system estimation method, data problems for one country affect the estimates for all countries. Since we judge the U.S. data to be most reliable, we use them for the inference on the parameters that are constant across countries. Second, our model is based on a set of stark neoclassical assumptions. These assumptions are more applicable to the low frictional U.S. economic environment than to that of countries in which capital and product markets are substantially distorted. Thus, if we think that our reduced form equation is more likely to be misspecified for some countries other than the U.S., including them in the estimation of the joint parameters would affect the results for all countries.

We estimate all the equations using non-linear least squares. Since we estimate β_3 for the U.S., this means that the identifying assumption that we make is that the logarithm of per capita GDP in the U.S. is uncorrelated with the technology-specific error, ε_τ. However, because of the cross-country restrictions we impose on β_3 this risk of simultaneity bias is not a concern for all the other countries in our sample.

2.3 Non-homotheticities

One general concern in structural estimation exercises is model mis-specification. In the context of our model, the place where this concern probably is more relevant is in elasticity of technology with respect to income. For our model to have a balanced growth path, the production functions need to be homothetic. This implies that the income elasticity of technology measures is equal to one. In this subsection, we explore the implications for the estimated equations (30) if we replace the original production function (3) for a more general specification that allows for non-homotheticities.

Consider the following non-homothetic version of (3)

$$Y = \frac{1}{\bar{\theta}} \left(\sum_{\tau} \theta_{\tau} Y_{\tau}^{\frac{1}{\theta_{\tau}}} \right)^{\bar{\theta}},$$

where $\bar{\mu}$ is the long-run average of μ_{τ} over τ so that constant returns to scale are guaranteed (in the long-run),

$$Y_{\tau} = Y^{\frac{(\bar{\theta}-1)\theta_{\tau}}{\bar{\theta}(\theta_{\tau}-1)}} P_{\tau}^{\frac{-\theta_{\tau}}{\theta_{\tau}-1}}.$$

This yields the following reduced form equation

$$m_{\tau} = \beta_1 + \beta_y y + \beta_2 t + \beta_3 \left((\mu - 1) \ln (t - T_{\tau}) - (1 - \alpha) (y - l) \right) + \varepsilon_{\tau}, \tag{33}$$

which differs from (30) in that β_y is not retricted to be equal to 1. It will be greater than one if $\bar{\theta} > \theta_{\tau}$ and smaller otherwise.

We use an empirical strategy analogous to the previous section to estimate equation (33). This is, we estimate the technological parameters (β_2, β_3 and β_y) for the U.S. and then use this estimates for the other countries. The only difference is that now we estimate an additional parameter, the U.S. income elasticity of technology, β_y. The use of U.S. to estimate this parameter is attractive because we have data for U.S. income since 1820. Thus, we cover very different levels of income when we estimate β_y for the U.S.

To estimate the component β_y of equation (33) we estimate simultaneously the system of equations for all the U.S. technologies. The reason for proceeding in this fashion comes from the observation that income is highly correlated with the linear trend in (33). Thus, if the equations for the U.S. were to be run separately, the identification of the term β_y would come from the high frequency variation and not from the long-run variation of U.S. income. The elasticity that one

would obtain is conceptually wrong, as the non-homotheticities are meant to capture the effects of the long-run increase in income on the demand of technologies. To stress further this point we have filtered the income in two components: long run and short run frequencies, and we have estimated them separately. Indeed, the short-run elasticity is higher, capturing adjustment over the business cycle.

3 Results

We consider data for 167 countries and 15 major technologies, that span the period from 1820 through 2003. The technologies can be classified into 6 categories; (*i*) transportation technologies, consisting of steam- and motorships, passenger and freight railways, cars, trucks, and passenger and freight aviation; (*ii*) telecommunication, consisting of telegraphs, telephones, and cellphones; (*iii*) IT, consisting of PCs and internet users; (*iv*) medical, namely MRI scanners; (*v*) steel, namely tonnage produced using blast oxygen furnaces; (*vi*) electricity.

The technology measures are taken from the CHAT dataset. Real GDP and population data are from Maddison (2007). Appendix A contains a brief description of each of the 15 technology variables used and lists their invention dates.

For our estimation, we only consider country-technology combinations for which we have more than 10 annual observations. There are 1298 such pairs in our data. The third column of Table 1 lists, for each technology, the number of countries for which we have enough data.

We follow Comin and Hobjin (2010) and analyze only the technology-country pairs for which we have plausible and precise estimates of the adoption lags. These are estimates with an adoption date later than the invention year plus 10, and with small standard errors.[15,16,17] Two thirds of the the technology-country pairs are plausible and precise. In what follows, all our results are based on the sample of 807 plausible and precise estimates.[18]

[15] The 10 year cut off point is to allow for inference error.

[16] See Comin and Hobjin (2009) for a discussion of the reasons for obtaining implausible estimates.

[17] In particular, the cutoff that we use is that the standard error of the estimate of T_τ is bigger than $\sqrt{2003 - \underline{v}_\tau}$. This allows for longer confidence intervals for older technologies with potentially more imprecise data.

[18] Results that also include the imprecise estimates are very similar to the ones presented here.

Included in these plausible and precise estimates are 15 estimates of adoption lags for the U.S.

3.1 Estimated Intensive Margin

Table 1 presents the descriptive statistics of the estimates of the intensive margin of adoption relative to the U.S. Column 5 reports the cross-country average in the intensity of adoption by technology. Note that this statistic is negative for all the technologies but ships and freight railways. This means that for for all the technologies but these two, the U.S. intensity of adoption is higher than the average in our sample. Column 5 reports the cross-country standard deviation of the intensity of adoption by technology while column 10 reports the interquartile range.[19] There does not seem to exist evidence on convergence in the cross-country dispersion in the intensity of adoption for newer technologies. Table 2 reports the estimates from a simple regression of the mean and dispersion of the intensive margin on the year of invention. The coefficient on the average difference in intensive margin with the respect the U.S. is negative but insignificant, suggesting that the differences in intensive margin of adoption with respect the U.S. have not been reduced (much) over time. For the dispersion measures the coefficients are not significant, painting a similar picture of lack of convergence. These results are in contrast with the finding that the dispersion of adoption lags has declined monotonically over the last two centuries, as reported in in Comin and Hobijn (2010).

One explanation for the lack of trend in the dispersion of the intensity of adoption could the following: Some of the early technologies became dominated by superior technologies a long time ago. For early adopters, such as the U.S., the level of dominated technologies has been declining since a long while ago. As a result, the estimated intercept is lower than in late adopters, where these technologies have become dominated more recently. Alternatively, the persistence of cross-country dispersion in the intensity of adoption over time may be a real fact.

To disentangle these two hypothesis, we re-estimate our baseline regression for the old technologies using only data until 1939 where presumably none of the technologies was obsolete, yet. Table 4 compares the estimates of the average, standard deviation and inter-quartile range of the intensity of adoption for the countries for which we can precisely estimate the diffusion equation using data up to 1939.[20] The results vary a little by technology but, overall, there are no significant increases in the dispersion of the intensity of adoption when we restrict the sample to the pre-1939 period. This implies that the levels and the dispersion in the intensity of adoption for early technologies

[19] That is the difference between the adoption intensities in countries in the 75 and the 25 percentiles.

[20] For obvious reasons, we only report the estimates for technologies invented before 1900.

is not driven by the fact that early technologies have become dominated earlier in countries that adopted them first.

Variance decomposition

An additional aspect that we can explore is how much of the variation in the intensity of adoption is driven by country effects and how much by technology effects. Table 3 answers this question. Specifically, let $\Delta_{j\tau}$ be the relative measure of intensive margin of country j in technology τ. We can decompose $\Delta_{j\tau}$ as follows

$$\Delta_{j\tau} = \Delta_j + \Delta_\tau + u_{i\tau}, \tag{34}$$

where Δ_j is a country fixed effect, Δ_τ is a technology fixed effect and $u_{i\tau}$ is an error term. The first line of table 3 examines the contribution of the country fixed effects alone. That is, the R^2 when estimating (34) with only country fixed effects. Country-specific effects explain approximately 54% of the variation in the intensity of adoption. We calculate the contribution of technology-specific fixed effects in an analogous manner. We find that technology-specific effects explain 26% of the variation. The last row of Table 3 shows that country and technology fixed effects jointly explain about 77% of the variation in the intensity of adoption. Of this, 52% can be directly attributed to country effects, 24% can be directly attributed to technology effects, and the remaining 1% is due to the covariance between these effects that is the result of the unbalanced nature of the panel structure of our data. The drivers of the variance in the intensity of adoption differ quite a bit from the drivers of the variance for adoption lags. As shown in Comin and Hobijn (2010), the technology fixed effects account for 65% of the variance in the adoption lags. In contrast, country effects are the main component accounting for the variance of the intensity of adoption.

Non-homotheticities

As we have discussed in the identification section, we estimate the income elasticity of technology using U.S. data and then we use our results to estimate the reduced form equation (33) for the other countries. We obtain that short run elasticities are much higher than long run, as they reflect the demand of technology over the business cycle. The numbers that we obtain are 2.2 for the long-run elasticity and 6.6 for the short run elasticity. These elasticities are estimated very precisely. However, the additional flexibility that we give to the model comes at the cost of having imprecise estimates of the adoption lag for two U.S. technologies: Ships and Electricity. To estimate the differences in the intensive margin of adoption for these two technologies, we circumvent this

problem taking as a reference France rather than the U.S. Note that the dispersion of the intensive margin is not affected by the choice of the reference country, however the mean difference is affected. Thus the average mean effect across technologies is not directly comparable with the homothetic case.

The estimation results for which we obtain plausible and precise estimates is somewhat reduced with respect to the baseline exercise. We obtain plausible and precise estimates for 773 country-technology pairs. This represents 60% of our sample. The results on the adoption lags and the intensive margin are reported in tables 5 and 6, respectively.Adoption lags tend to be reduced slightly with respect the baseline model. The dispersion in our intensive margin measures remains quite similar, both for the standard deviation measures and the IQR.

Table 7 shows the correlation between the intensity of adoption estimates in the homothetic and non-homothetic cases. On average the correlations are high, approximately 76%. By technology, they range from 36% for freight railways to 95% for passanger aviation. Other than freight raylways, the correlation for all the technologies is above 50% and for 11 out of the 15 technologies it is above 70%.

Correlation with per capita income

Before using our model to conduct a development accounting exercise, it is revealing to explore the correlation between per capita income and the intensity of adoption for each technology. Table 8 reports these statistics for both the homothetic and non-homothetic estimates. The correlations are sizable. In the homothetic case the average correlation across technologies is 58% and in the non-homothetic case it is 54%. We find some variation across technologies. The correlation of adoption intensity with per capita income seems to be lower for the earlier technologies, specially for ships and railways. Contrary to the perception that information technologies may be closing the technological divide between rich and poor countries, we find that the intensity of adoption of these technologies (i.e. pcs, cellphones, internet) present quite high correlations with per capita income.

Table 8 also reports the correlation between the adoption lags and per capita income. As shown by Comin and Hobijn (2010) for the homothetic case, the correlation is also fairly high, approximately -46%. In the last column, we show that the there is also a significant correlation between the adoption lags in the non-homothetic case and per capita income though slightly lower

than in the homothetic case (-30% vs. -46%).

3.2 Development accounting

We next investigate how the estimated differences in the intensive margin of adoption translate into cross-country differences in per capita income. To answer this question, we have to approximate the aggregate effect of the differences in intensive margin of adoption. We draw from the the equilibrium results of our one-sector growth model. Aggregate production, Y, can be expressed as

$$Y = AK^{\alpha}L^{1-\alpha}, \quad \text{where } K \equiv \int_{-\infty}^{t} K_v dv,\ L \equiv \int_{-\infty}^{t} L_v dv. \tag{35}$$

Aggregate TFP, A, can be expressed as

$$A = \left[\sum_{\tau \in \{o,n\}} (A_\tau)^{\frac{1}{\mu-1}} \right]^{\mu-1} = \left(\frac{\mu-1}{\gamma} \right)^{\mu-1} N^{\mu'-1} Z_{\underline{v}} e^{\gamma(t-D-\underline{v})}. \tag{36}$$

The adoption lag affects aggregate TFP because a higher D reduces the productivity embodied in the best technology vintage available for production. The intensive margin of adoption, N, appears because of the variety effect. The higher the intensive margin of adoption, the higher the number of new varieties adopted per vintage, which results in higher productivity.

Substituting (36) into (35) and noting that $\phi_R KR = \alpha Y$ yields the following expression for labor productivity:

$$\frac{Y}{L} = A^{\frac{1}{1-\alpha}} \left(\frac{K}{Y} \right)^{\frac{\alpha}{1-\alpha}} = \left[\left(\frac{\mu-1}{\gamma} \right)^{\mu-1} N^{\mu'-1} Z_{\underline{v}} e^{\gamma(t-D_t-\underline{v})} \right]^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{\phi_R R} \right)^{\frac{\alpha}{1-\alpha}}. \tag{37}$$

If the only source of cross-country differentials in per capita income are differences in intensive margin of adoption of technology, then, in a balanced growth path, the log difference of country j's level of real GDP per capita with that of the U.S. is given by

$$(y_j - l) - (y_{US} - l) = \frac{1}{1-\alpha} \Delta_j, \tag{38}$$

where α is the capital share of the economy. We observe the left hand side of (38) in our data. To approximate the right hand side of equation (38), we assume that the country-specific relative measures of adoption lags we have estimated using our sample of technologies are representative of

the average intensive margin of adoption across all the technologies used in production. In other words, we assume that the average relative measure of our sample is approximately the average in the population.[21] This allows us to compute the contribution of the intensive margin of adoption on the aggregate TFP level and, ultimately, quantify cross country differences in TFP due to differences in the intensive margin of adoption.

Figure 2 plots the data for both sides of (38) for the countries in our dataset. The correlation between the two sides is 0.55. The income per capita data corresponds to year 2000 and comes from the Penn World Tables 6.2.[22] The thicker dashed line correspond to the regression line, while the light grey line is the 45°-line. The slope of the regression line is about 0.37, which can be interpreted as that our model and estimates explaining around 37% of the log per capita GDP differentials observed in the data.

Note that the there is an heteroskedastic pattern, as poor countries have more variance in the measure of the intensive margin. One possible explanation comes from the fact that we have less observations for poor countries than rich countries. As a result the approximation (39) may be less valid. We do not view this as an important concern because the lack of more data on technologies just makes our estimates noisier, but as can be seen in the figure, the tendency is similar to richer countries.[23]

Finally, we perform a development accounting exercise analogous to the previous section for the estimates obtained using a non-homothetic production function. Note, however, that the presence of non-homotheticities prevent the existence of a simple aggregate production function, as we had in our baseline model. Our assumption is that, as a first pass approximation, we can still use the aggregate production function results to assess whether the results are in the same ballpark. It turns out that the development accounting exercise yields very similar results to the baseline exercise, as it can be seen in figure 3 The correlation between the intensive margin and income per capita is .52, and the coefficient on the regression is .41. We view this result as highlighting the

[21] More formally, this means that

$$\frac{1}{N}\sum_{v=1}^{N}\frac{N_{v,i}^{\mu'-1}}{N_{v,US}^{\mu'-1}} = \mathbf{E}\left[\frac{N_{v,i}^{\mu'-1}}{N_{v,US}^{\mu'-1}}\right] \tag{39}$$

for all country i in our sample.

[22] Similar results obtain with data from Maddison.

[23] One possibility to overcome the problem could be to pool some similar developing countries together to obtain more precise estimates of the intensive margin.

robustness of the role of technology to explain differences in income per capita.

To sum up, this section suggests that differences in the intensive margin of adoption account for a substantial share of cross-country per capita income differences. They seem to account for a share of 40% of the variation in income per capita.

4 Conclusion

In this paper we have built and estimated a model of technology diffusion and growth that has two main characteristics. First, at the aggregate level, it is similar to the one-sector neoclassical growth model and has a well-defined balanced growth path. Second, at the disaggregate level, it has implications for the path of observable measures of technology adoption, such as the number of units of capital that embody a given technology or the output produced with this technology.

The main focus of our analysis is on the intensive margin of adoption. We can identify relative intensive margins. For each country, these are defined as the difference in the intensive margin of adoption for a particular technology relative to a baseline country. In our model, the intensive margin of adoption affects the aggregate TFP through two channels. First, there is a direct channel: the number of units adopted affects directly TFP because a higher number of units of technology adopted increase productivity. Second, there is an indirect effect through the adoption lag. It arises because low levels of adoption at the intensive margin delay the adoption of a technology by making adoption less profitable.

We estimate the intensive and extensive margins of adoption for 15 technologies and 161 countries over the period 1820-2003. We find that differences in the intensive margin of adoption are substantial. In fact, the magnitude of the variation across countries is higher than adoption lags.

We can use our model to assess how much of the variation in income per capita can be accrued to differences in the intensive margin of adoption. We find that around 40% of the variation can be attributed to differences in the intensive margin. Comin and Hobjin (2009) report that differences in the intensive margin account for at least 25% percent of the cross country variation. Taken together, these results imply that the role of technology is crucial to understand income per capita differences. In particular, the empirical estimates suggest that, through the lens of our model, two thirds of the cross-country income per capita differences can be explained by differences in

technology adoption.

A future avenue that we plan to explore is to understand the underpinnings of the variation in adoption costs across countries. This will involve taking a stand on the microfoundations for the cross-country variation in adoption costs, which our paper abstracts from.

References

[1] **Basu, Susanto, and John G. Fernald.** 1997. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy*, 105(2): 249-283.

[2] **Basu, Susanto, and David N. Weil.** 1998. "Appropriate Technology and Growth." *Quarterly Journal of Economics*, 113(4): 1025-1054.

[3] **Caselli, Francesco, and W. John Coleman.** 2001. "Cross-Country Technology Diffusion: The Case of Computers." *American Economic Review - AEA Papers and Proceedings*, 91(2): 328-335.

[4] **Clark, Gregory.** 1987. "Why Isn't the Whole World Developed? Lessons from the Cotton Mills." *Journal of Economic History*, 47(1): 141-173.

[5] **Comin, Diego, and Bart Hobijn.** 2004. "Cross-Country Technology Adoption: Making the Theories Face the Facts." *Journal of Monetary Economics*, 51(1): 39-83.

[6] **Comin, Diego, and Bart Hobijn.** 2008. "An Exploration of Technology Diffusion." HBS working paper 08-093.

[7] **Comin, Diego, Bart Hobijn, and Emilie Rovito.** 2006. "Five Facts You Need to Know About Technology Diffusion." NBER Working Paper #11928.

[8] **Comin, Diego, Bart Hobijn, and Emilie Rovito.** 2008. "A New Approach to Measuring Technology with an Application to the Shape of the Diffusion Curves." *Journal of Technology Transfer*, 33(2): 187-207.

[9] **Gollin, Douglas.** 2002. "Getting Income Shares Right." *Journal of Political Economy*, 110(2): 458-474.

[10] **Gort, Michael, and Steven Klepper.** 1982. "Time Paths in the Diffusion of Product Innovations." *Economic Journal*, 92: 630-653.

[11] **Griliches, Zvi.** 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica*, 25(4): 501-522.

[12] **Hall, Robert E., and Charles I. Jones.** 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114(1): 83-116.

[13] **Hsieh, Chang-Tai.** 2002. "What Explains the Industrial Revolution in East Asia? Evidence from the Factor Markets." *American Economic Review*, 92(3): 502-526.

[14] **Jerzmanowski, Michal.** 2007. "Total Factor Productivity Differences: Appropriate Technology vs. Efficiency.", *European Economic Review*, 51(8): 2080-2110.

[15] **Klenow, Peter J., and Andrés Rodríguez-Clare.** 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" *NBER Macroeconomics Annual,* 12: 73-103.

[16] **Maddison, Angus.** 2007. *Contours of the World Economy 1-2030 AD: Essays in Macro-Economic History.* Oxford: Oxford University Press.

[17] **Mansfield, Edwin.** 1961. "Technical Change and the Rate of Imitation." *Econometrica*, 29(4): 741-766.

[18] **Manuelli, Rodolfo, and Ananth Sheshadri.** 2003. "Frictionless Technology Diffusion: The Case of Tractors." NBER Working Paper # 9604.

[19] **Parente, Stephen L., and Edward C. Prescott.** 1994. "Barriers to Technology Adoption and Development." *Journal of Political Economy*, 102(2): 298-321.

[20] **Saxonhouse, Gary R., and Gavin Wright.** 2004. "Technological Evolution in Cotton Spinning, 1878-1933." in *The Fibre that Changed the World*, eds. Douglas A. Farnie and David J. Jeremy. Oxford: Oxford University Press.

[21] **Young, Alwyn.** 1992. "A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong and Singapore." *NBER Macroeconomics Annual*, 7: 13-54.

[22] **Young, Alwyn.** 1995. "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience." *Quarterly Journal of Economics*, 110(3): 641-680.

A Data

The data that we use are taken from two sources. Real GDP and population data are taken from Maddison (2007). The data on the technology measure are from the Cross-Country Historical Adoption of Technology (CHAT) data set, first described in Comin, Hobijn, and Rovito (2006). The fifteen particular technology measures, organized by broad category, that we consider are:

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

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

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

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

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

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

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

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

9. **Telephone:** Number of telegrams sent. *Invention year:* 1876; year of invention of telephone by Alexander Graham Bell.

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

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

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

13. **MRIs:** Number of magnetic resonance imaging (MRI) units in place. *Invention year:* 1977; first MRI-scanner built.

14. **Blast Oxygen Steel:** Crude steel production (in metric tons) in blast oxygen furnances (a process that replaced bessemer and OHF processes). *Invention year:* 1950; invention of Blast Oxygen Furnace.

15. **Electricity:** Gross output of electric energy (inclusive of electricity consumed in power stations) in KwHr. *Invention year:* 1882; first commercial powerstation on Pearl Street in New York City.

B Equilibrium and diffusion of the new technology

Let Γ_t denote the total adoption costs at instant t. Then

$$\begin{aligned}\Gamma_t &= \overline{\Psi}\left(1+b_e\right)\left(\frac{\gamma}{\mu-1}\right)e^{-\frac{\vartheta}{\mu-1}\gamma D}\left(\frac{Z_0 e^{\gamma\overline{v}}}{A_t}\right)^{\frac{1}{\mu-1}} Y_t\left(1-\dot{D}\right)\\ &+ \overline{\Psi}\left(1+b_i\right)\left(Z_0 A_t e^{\gamma t}\right)^{-\frac{1}{\mu-1}} Y_t \int_{-\infty}^{v_t} Z_v^{\frac{1}{\mu-1}}\dot{N}_v(t)dv. \qquad (40)\end{aligned}$$

where $\dot{D}$ denotes the time derivative of the adoption lags. Note that along the Balance Growth Path, the distribution over the vintages for which the measure of varieties adopted becomes degenerate around $\overline{v}_t$ and the aggregate costs become $\Gamma^i_{vt}N_v$.

The equilibrium path of the aggregate resource allocation in this economy can be defined in terms of the following nine equilibrium variables $\{C, K, I, \Gamma, Y, A, N, D, V\}$. Just like in the standard neoclassical growth model, the capital stock, K, is the only state variable. The eight equations that determine the equilibrium dynamics of this economy are given by

(i) The consumption Euler equation.

(ii) The aggregate resource constraint[24]

$$Y = C + I + \Gamma. \tag{41}$$

(iii) The capital accumulation equation

$$\dot{K} = -\delta K + I. \tag{42}$$

(iv) The production function, (11), taking into account that in equilibrium $L = 1$.

(v) The adoption cost functions (6) and (7).

(vi) The technology adoption equations, which determine the adoption lag (20) and the intensive margin of adoption (16).

(vi) The stock market to GDP ratio, (14).[25]

(vii) The aggregate TFP level, 12.

C Tables

[24] We assume that adoption costs are measured as part of final demand, such that Y can be interpreted as GDP.

[25] The dynamics of Ψ_t^e and Ψ_t^i are what are considered in the system of equilibrium equations. For example, the law of motion of for Ψ_t^e is (omitting superscripts and subscripts) $\frac{\dot{\Psi}}{\Psi} = \left\{\alpha \frac{\epsilon-1}{\epsilon} \frac{Y}{K} - \delta + \frac{1}{\mu-1} \frac{\dot{A}}{A} - \frac{\dot{Y}}{Y}\right\} - \frac{\alpha}{\epsilon} \frac{1}{\Psi}$.

Table 1: Estimated LogDifferences in Intensive Margins

Technology	Invention year	Number of Countries	Plausible and Precise	Log Intensive Margin (Δ_j)					
				Mean	sd	p5	p50	p95	IQR
Ships	1788	63	50	0.38	0.72	-0.74	0.22	1.82	0.86
Rail Passengers	1825	84	62	0.19	0.33	-0.38	0.19	0.77	0.38
Rail Freight	1825	89	42	-0.47	0.35	-0.92	-0.53	0.19	0.49
Telegraph	1835	69	46	-0.47	0.5	-1.36	-0.42	0.3	0.53
Telephone	1876	143	66	-0.78	0.69	-2.16	-0.67	0.23	0.9
Electricity	1882	138	97	-0.4	0.35	-1.02	-0.38	0.09	0.48
Trucks	1885	111	56	-0.76	0.69	-1.92	-0.83	0.35	0.75
Cars	1885	127	71	-0.79	0.83	-2.21	-0.72	0.44	1.03
Aviation Freight	1903	96	30	-0.57	0.83	-2.58	-0.37	0.38	0.72
Aviation Passengers	1903	99	51	-0.84	0.63	-2.1	-0.77	-0.02	0.72
Blast Oxygen Furnaces	1950	50	39	-0.79	0.89	-2.54	-0.36	0.21	1.13
PCs	1973	71	61	0.01	0.59	-0.87	-0.01	1.02	0.35
Cellphones	1973	87	72	-0.54	0.72	-1.94	-0.37	0.44	0.88
MRI	1977	12	12	-0.34	0.34	-1.07	-0.35	0.22	0.34
Internet	1983	59	52	-0.31	0.54	-1.15	-0.23	0.47	0.7
Total		1298	807	-0.43	0.72	-1.79	-0.35	0.57	0.79

IQR is defined as the difference between 75th percentile and 25th percentile.

Table 2: Mean, Std. Dev. and IQR regressed on Year of Invention

	Homothetic	Non-Homothetic
Mean	-0.17	-0.07
	(0.16)	(0.11)
Std.Dev.	0.05	-0.003
	(0.08)	(0.06)
IQR	0.01	0.01
	(0.01)	(0.01)

Table 3: Analysis of variance

	Model SS	**Country effect**	**Technology effect**	**Residual SS**	**Total SS**
Country effect alone	54%	54%		46%	100%
Technology effect	26%		26%	74%	100%
Joint effect	77%	24%	51%	23%	100%

Table 4: Comparison of estimates up to 1939 vs. whole sample (Capital Measures not corrected)

Technology Name	Adoption Year	Observations	mean 1939	sd 1939	iq 1939	mean	sd	iqr
Ships	1788	12	0.06	0.38	0.39	-0.11	0.63	0.67
Rail Freight	1825	12	-0.49	0.25	0.32	-0.57	0.31	0.45
Rail Passengers	1825	17	0.14	0.16	0.17	0.24	0.19	0.23
Telegraph	1835	20	-0.13	0.29	0.37	-0.54	0.54	0.62
Telephone	1876	9	-0.38	0.28	0.31	-0.35	0.28	0.35
Electricity	1882	18	-0.24	0.23	0.38	-0.20	0.18	0.26
Cars	1885	11	-0.99	0.56	0.63	-0.75	0.48	0.60
Trucks	1885	10	-0.96	0.63	1.00	-1.01	0.73	1.02
Total			-0.31	0.51	0.59	-0.37	0.56	0.69

Table 5: Quality of estimates and Adoption Lags with non-homotheticities

Technology	Invention year	Number of Countries	Plausible and Precise	Adoption Lags					
				Mean	sd	p5	median	p95	IQR
Ships	1788	63	31	101.61	59.72	-1.66	101.00	175.69	111.52
Rail Passengers	1825	84	58	82.28	26.30	30.75	85.87	122.05	33.55
Rail Freight	1825	89	40	61.78	31.51	14.28	67.47	113.54	54.43
Telegraph	1835	69	38	39.39	32.19	-4.62	27.14	107.06	49.41
Telephone	1876	143	65	45.61	31.94	2.87	41.66	101.19	52.29
Electricity	1882	138	51	35.17	25.90	-1.92	34.44	81.84	38.55
Trucks	1885	111	55	33.24	19.74	-1.88	28.83	66.63	27.16
Cars	1885	127	77	37.45	23.15	3.23	34.30	86.54	31.13
Aviation Freight	1903	96	36	37.58	14.89	17.86	38.06	64.43	21.60
Aviation Passengers	1903	99	53	27.98	13.95	10.71	24.27	52.96	14.74
Blast Oxygen Furnaces	1950	50	43	15.65	7.51	7.96	14.06	29.18	10.61
PCs	1973	71	69	13.98	2.94	7.74	14.11	18.82	2.93
Cellphones	1973	87	86	13.56	4.21	5.84	14.22	18.56	6.03
MRI	1977	12	12	3.02	2.38	0.16	2.60	7.72	4.01
Internet	1983	59	59	7.49	2.09	3.54	7.49	10.59	2.64
Total		1298	773	35.55	33.32	3.32	22.76	101.11	38.62

IQR is defined as the difference between the 75th percentile and 25th percentile.

Table 6: Estimated LogDifferences in Intensive Margins (Δ_j) with non-homotheticities

Technology	Invention year	Number of Countries	Plausible and Precise	Log Intensive Margin (Δ_j)					
				Mean	sd	p5	p50	p95	IQR
Ships	1788	63	31	0.61	0.69	-0.14	0.56	1.62	1.01
Rail Passengers	1825	84	58	0.38	0.42	-0.35	0.44	1.03	0.56
Rail Freight	1825	89	40	0.09	0.43	-0.63	0.04	0.76	0.49
Telegraph	1835	69	38	0.02	0.59	-0.99	0	1.22	0.47
Telephone	1876	143	65	0.15	0.62	-0.87	0.17	1.16	0.83
Electricity	1882	138	51	-0.03	0.63	-1.25	-0.03	0.86	0.9
Trucks	1885	111	55	0.01	0.64	-1.07	0.07	1.06	0.74
Cars	1885	127	77	0.09	0.79	-1.23	0.12	1.33	0.97
Aviation Freight	1903	96	36	-0.03	0.72	-1.51	0.15	0.92	0.81
Aviation Passengers	1903	99	53	-0.42	0.76	-1.87	-0.39	0.55	0.97
Blast Oxygen Furnaces	1950	50	43	0.03	0.5	-0.78	0.1	0.67	0.74
PCs	1973	71	69	0.55	0.58	-0.4	0.56	1.38	0.64
Cellphones	1973	87	86	-0.02	0.74	-1.53	0.08	0.95	1.02
MRI	1977	12	12	0.08	0.33	-0.54	0.08	0.47	0.53
Internet	1983	59	59	0.23	0.62	-0.87	0.28	1.27	0.69
Total		1298	773	0.11	0.68	-1.11	0.16	1.1	0.78

IQR is defined as the difference between the 75th percentile and 25th percentile.

Table 7: Comparison of Homothetic and Non-Homothetic Intensive Margins

Technology Name	**Observations**	**R^2**
Ships	30	0.89
Rail Passengers	53	0.67
Rail Freight	38	0.36
Telegraph	34	0.84
Telephone	59	0.79
Electricity	45	0.71
Trucks	48	0.71
Cars	58	0.72
Aviation Freight	29	0.89
Aviation Passengers	51	0.95
Blast Oxygen Furnaces	38	0.87
PCs	55	0.69
Cellphones	63	0.83
MRI	12	0.53
Internet	50	0.86
Total	663	0.76

R^2 from regressing the non homothetic intensive margin on the homothetic.

Table 8: Correlation of log difference in Intensive Margins and Adoption Lags with log income per capita in 2000

	Invention	Homothetic			Non Homothetic		
Technology	**year**	Observations	Intensive Margin	Adoption Lag	Observations	Intensive Margin	Adoption Lag
Ships	1788	50	0.18	-0.48	31	0.31	-0.47
Rail Passengers	1825	62	0.38	-0.62	56	0.33	-0.49
Rail Freight	1825	42	0.16	-0.67	40	0.46	-0.30
Telegraph	1835	46	0.46	-0.49	37	0.49	-0.28
Telephone	1876	66	0.79	-0.49	64	0.53	-0.37
Electricity	1882	97	0.76	-0.57	50	0.68	-0.27
Trucks	1885	56	0.50	-0.36	50	0.38	-0.25
Cars	1885	71	0.46	-0.33	66	0.48	-0.04
Aviation Freight	1903	30	0.85	-0.09	36	0.78	0.20
Aviation Passengers	1903	51	0.80	-0.13	53	0.79	0.23
Blast Oxygen Furnaces	1950	39	0.90	-0.33	41	0.80	-0.35
PCs	1973	61	0.56	-0.32	61	0.37	-0.38
Cellphones	1973	72	0.66	-0.59	69	0.65	-0.49
MRI	1977	12	0.66	-0.39	12	0.23	-0.44
Internet	1983	52	0.50	-0.74	55	0.55	-0.82
Total	—	807	0.58	-0.46	721	0.54	-0.30

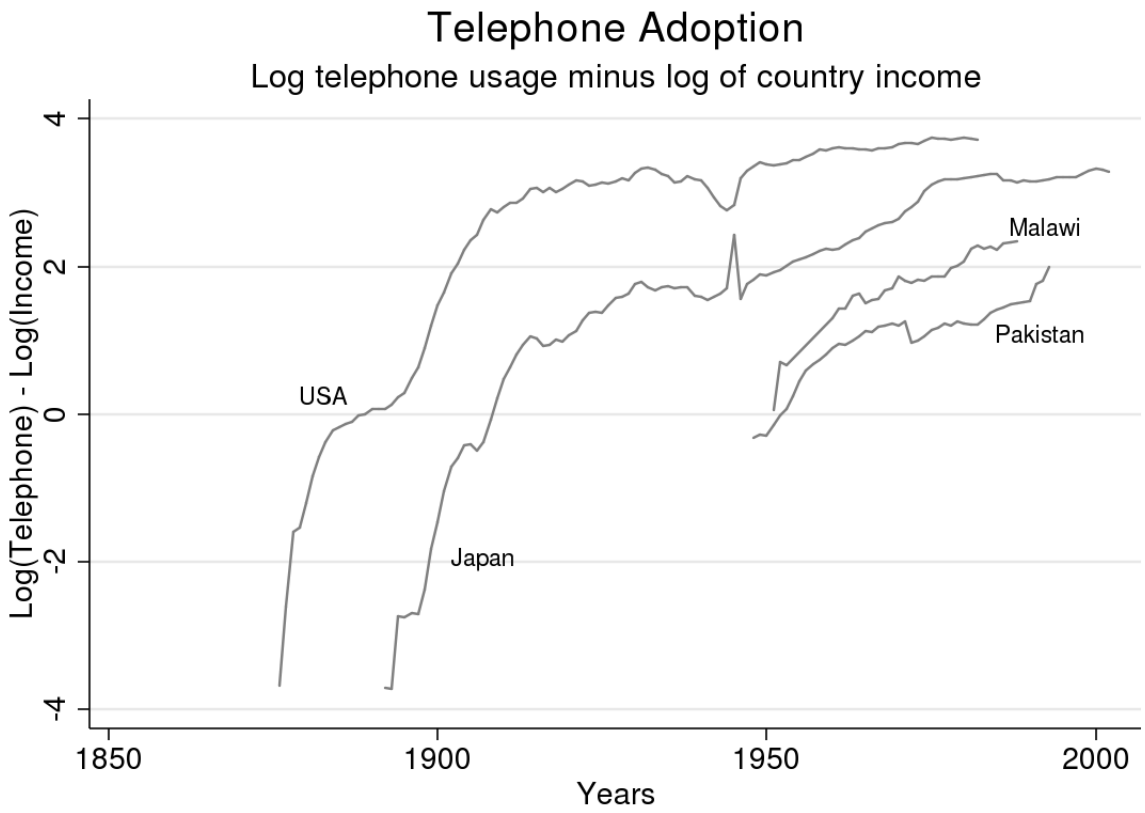

Figure 1: Differences in telephone adoption subtracting own country income for four different countries.

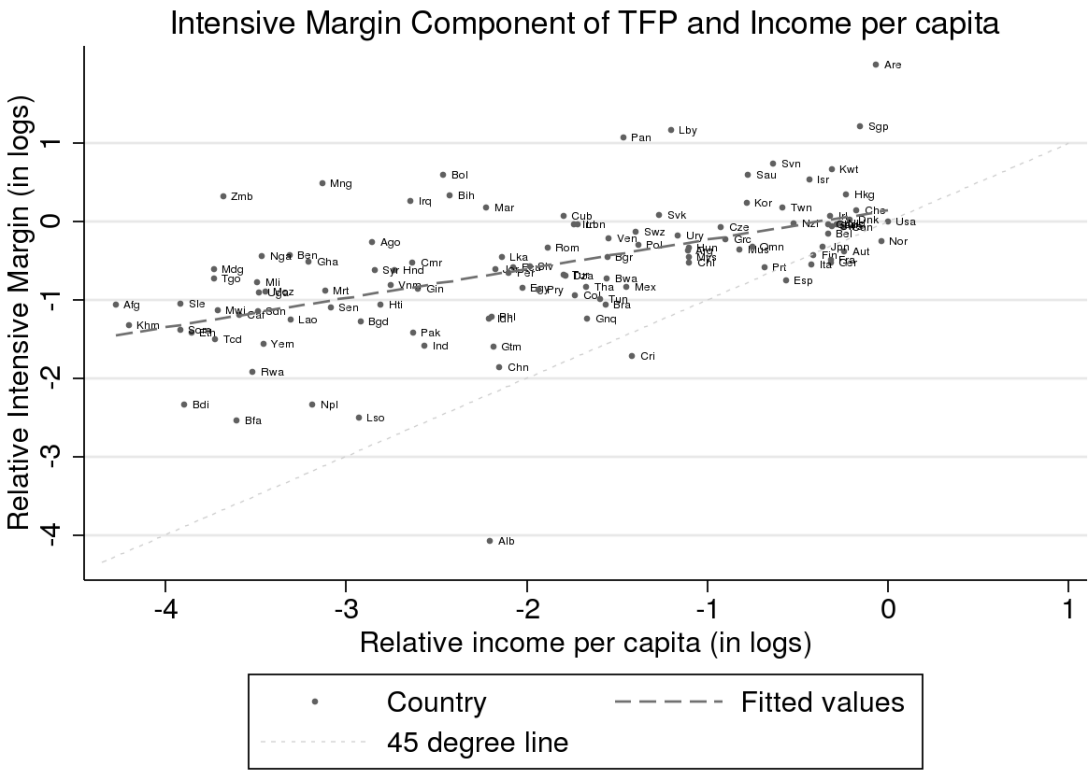

Figure 2: Intensive Margin component of TFP and differences in income per capita.

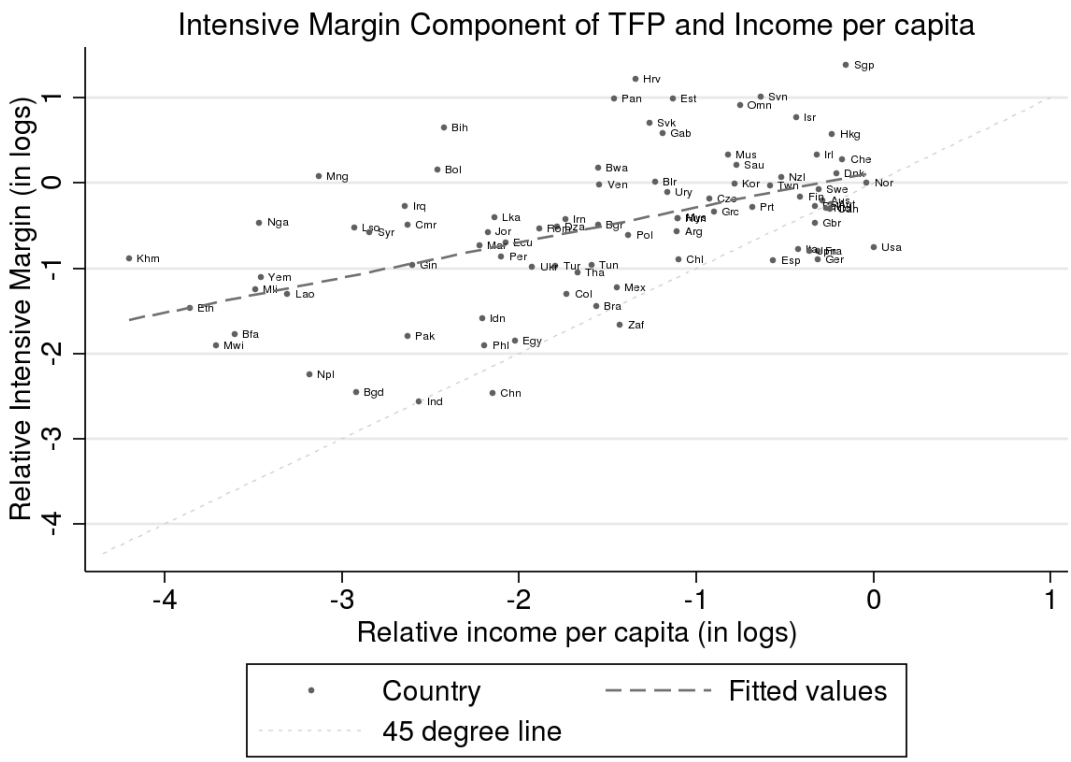

Figure 3: Intensive Margin component of TFP and differences in income per capita with non homotheticities.