University of Toronto
Department of Economics

Working Paper 571

Idiosyncratic Distortions and Technology Adoption

By Stephen Ayerst

December 18, 2016

Idiosyncratic Distortions and Technology Adoption

Stephen Ayerst*

December 12, 2016

Abstract

Empirical evidence indicates that resource misallocation has a substantial negative effect on aggregate productivity in developing countries. I show that the same underlying institutions that create misallocation are also important for explaining cross-country technology differences. I study a model of heterogeneous firms that choose both labour and technology inputs. Distortions are modeled as idiosyncratic wedges on firm revenues and delay adoption by disincentivizing firms from investing in newer technologies. At the aggregate level, distortions targeting high productivity firms delay the initial adoption of new technologies. In the calibrated model, distortions account for a large portion of the observed cross-country technology differences. Moving from the distortions of the bottom decile economy to the United States' level explains just under half of the observed adoption lag and increases productivity by 89%. Over half of the productivity increase is from firms adjusting technology.

Keywords: Productivity, Misallocation, Technology Diffusion.
JEL classification: 011, 014, 033, 041, 043.

*Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada M5S 3G7. Email: stephen.ayerst@mail.utoronto.ca. I thank Diego Restuccia for his advice and encouragement. This paper has benefited from discussions with Rami Abou-Seido, Michelle Alexopoulos, Chaoran Chen, Daniel Chippin, Diego Comin, Miguel Faria-e-Castro, Ashique Habib, Ridwan Karim, Soomin Lee, Nick Li, Michael Peters, Baxter Robinson, Xiaodong Zhu as well as participants of ECO 4060, the Macro Brown Bag workshop at the University of Toronto, the CEA conference (U. Ottawa) and EconCon 2016 (Princeton).

1 Introduction

Cross-country differences in income are largely attributed to differences in productivity. A recent literature, starting with Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), shows that the misallocation of resources in developing countries is an important source of cross-country variation. The view of this literature is that institutions distort firm-level incentives differently, leading to a misallocation of resources.[1] As resources are reallocated away from their optimal use, the aggregate productivity of the economy falls. In this paper, I argue that these distortions are also important for explaining cross-countries differences in technology usage. Specifically, I show that distortions act as a barrier to the adoption of new technologies explaining up to 43% (19 of 44 years) of the observed adoption lags in developing countries.

I study a model of firm heterogeneity in which firms independently choose their technology based on their underlying productivity and institutional setting. The basic structure of the model follows Hopenhayn (1992) in which firms differ in an underlying component of their productivity. The model differs by allowing firms to choose a technology input, which partially endogenizes productivity. Following Parente and Prescott (1994), technology is modeled as an exogenous set of productivities that are available to firms at a convex cost. Importantly, firms face a tradeoff between more productive modern technologies and less costly antiquated technologies. Firms with high underlying productivity benefit more from technologies due to complementarity between a firm's underlying productivity and its technology. Consequently, high productivity firms adopt technologies earlier and the equilibrium features a distribution of technologies used within a country.

In the model, new technologies are introduced each period, which expands the set of technologies available to firms. Existing technologies become cheaper over time, which in turn drives the adoption of technologies. Coupled with the distribution of technologies, this

[1] See Restuccia and Rogerson (2013) for a review of the misallocation literature. Some specific examples of institutions include: size-dependent tax policies; financial constraints on small and young firms; state-owned enterprises; and restrictions on land markets.

implies an equilibrium distribution of adoption lags. The model structure is motivated by three observations from the literature. First, the adoption of new technologies (within a country) is gradual with more productive firms adopting earlier than less productive firms. Second, the cost of existing technologies are falling over time. Third, developing countries are not typically the source of technological innovations. Rather, technologies tend to originate from developed countries before being adopted later by developing countries.

The institutional setting is captured by firm-specific (idiosyncratic) distortions, which are modeled as a wedge on firm revenue (as in Restuccia & Rogerson, 2008). The wedges act as a stand-in for the types of institutional frictions that distort the relative incentives of firms leading to a misallocation of resources. These institutional frictions are more prevalent in developing countries. For example, Hsieh and Klenow (2009) document numerous institutional frictions in the manufacturing sector for China and India, such as preferential treatment for State-Owned Enterprises, or the licensing and size restriction of plants. Other examples include credit constraints that prevent firms from reaching their optimal size (see Buera, Kaboski, & Shin, 2011) or labour market rigidities resulting from firing costs (see Hopenhayn & Rogerson, 1993). For the agricultural sector, Adamopoulos and Restuccia (2014) document a number of frictions that prevent farms from acquiring more land, such as explicit size constraints or subsidies targeted at small farms.

For intuition on the key mechanisms, consider a distortion that taxes a firm's revenues. The firm responds similarly to this distortion as it would to a fall in productivity. This is because the distorted firm is less profitable than if it were undistorted and, as a result, the firm has less incentive to hire workers or invest in modern technologies. Depending on the distribution of distortions, the distorted economy may then lag in the adoption of new technologies relative to the undistorted economy. This is the case if distortions tend to target high productivity firms (correlated distortions in the literature). However, if distortions tend to reallocate resources randomly, then the adoption period becomes prolonged and the adoption lag decreases. This is due to some (ex ante) high productivity firms becoming more profitable and adopting earlier.

Having established the relationship between distortions and the adoption pattern, I use the model to assess the quantitative importance of distortions to cross-country productivity and technology differences. I discipline the benchmark economy to United States' data on the establishment size distribution and the adoption of new technologies. I then vary parameters related to distortions in the benchmark economy to target moments found in empirical studies of distortions in developing countries. Distortions are calibrated to match empirical moments from Hsieh and Klenow (2009, 2014). Additionally, an aggregate barrier is calibrated to match the cross-country adoption lags found by Comin and Hobijn (2010) using data on a robust sample of technologies and countries. The aggregate barrier acts as a residual measure of technology differences between countries and provides a benchmark to compare the magnitude of distortions. The quantitative results are summarized in two counterfactual experiments.

The first experiment examines the gap in productivity between the benchmark economy and counterfactual economies. The gap in productivity is decomposed into three components: (1) changing technology through the aggregate barrier; (2) reallocating static factors of production (labour) *holding firm-level technologies fixed*; and (3) changing the distribution of technologies across firms. The second and third channel represent a static and dynamic channel of misallocation. The model predicts that moving from the distortions and aggregate barriers of the bottom decile economy to the United States benchmark would increase productivity by 212%. This gain in productivity is composed of an 67% gain from aggregate barriers and an 89% gain from distortions ($213\% = (1+67\%)(1+89\%)-1$). Both the static and dynamic channels of misallocation contribute substantially to the change in productivity from distortions. While the static channel is widely discussed in empirical studies of misallocation, the dynamic channel is not. This suggests a larger gain from removing distortions than is typically found.

Another takeaway from the experiment involves the relative size of the gap attributable to aggregate barriers and idiosyncratic distortions. The relative magnitude is heterogeneous across countries with developed countries benefiting less from changing distortions and more from changing aggregate barriers. For example, the calibrated distortions in the top decile economies are similar to the United States, but rich countries still tend to lag the United States in technology adoption. This implies in a productivity gap of 35% from aggregate barriers, but close to 0% from idiosyncratic distortions. Relative to the bottom decile economy discussed previously, this shows the change in the relative importance of the two channels. The results suggest that technology differences are relatively more important for developed countries while idiosyncratic distortions are relatively more important for developing countries.

The second experiment quantifies the effect of distortions on cross-country differences in technology. Specifically, I measure the extent to which adoption lags increase when distortions in the benchmark economy are increased to the level found in developing countries. This experiment offers three major takeaways. First, distortions can act an important barrier to the adoption of new technologies in developing countries. Increasing distortions in the benchmark economy increases the initial adoption lag by up to 19 years. This is just under half of the observed 44 year difference between the bottom decile of countries and the United States. Second, while distortions are important for most countries, they are much less important than aggregate barriers for developed countries. Third, the effect of distortions is heterogeneous based on when the adoption lag is measured. For the bottom decile of countries, distortions explain 34% (13 of 38 years) of the median adoption lag - measured as when 50% of output is produced with a new technology - compared to half for the initial adoption lag. The third takeaway highlights the importance of correlated distortions, which tend to delay the initial adoption of new technologies, but have less of an effect on later adoption.

Taken together, the quantitative exercises suggest that technology differences may be symptomatic of underlying institutional issues.

Related Literature: My paper is related to two broad literatures: one on cross-country differences in technology; and the second on the misallocation of resources in developing countries. In the first literature, my paper builds on the work of Parente and Prescott (1994) and Comin and Hobijn (2010). Both papers model technology differences between countries as a result of aggregate barriers to adoption. My model is similar at the aggregate level but differs by explicitly allowing firm heterogeneity and distortions to play a role. Further, work by Comin and Mestieri (2014) shows that the intensive margin of technology adoption is an important component of aggregate productivity. Firm heterogeneity provides an intuitive explanation for the intensive margin and distortions provide an explanation of why it may differ across countries. I also show that the contribution of technology to aggregate productivity depends not only on the level of technology, but also on the distribution of technology across firms.

In the second literature, my paper is related to a series of papers that focus on the effect of distortions on the long-run distribution of productivities (Gabler & Poschke, 2013; Bento & Restuccia, 2016; Hsieh & Klenow, 2014; Atkeson & Burstein, 2014). This literature finds that incorporating dynamic decisions by firms amplifies the effects of distortions by changing the distribution of productivities. I contribute to this literature by examining the effect of distortions on technology adoption decisions. Focusing on technology allows the model to map into observable data on technology adoption lags, which provides a benchmark for the importance of distortions.[2] However, in order to maintain the simplicity of the model, I abstract from many of the dynamic considerations that are important in this literature.[3]

The remainder of this paper is organized as follows. Section 2 provides an overview of the facts on technology adoption. Section 3 presents and characterizes the model. Section

[2] Bento and Restuccia (2016) take a similar approach by examining the relative effect of distortions on R&D expenditures. This provides them with a similar benchmark of how well their model is able to match an observable measure of productivity improvement.

[3] An implication is that the model does not fully incorporate dynamic effects of technology choices (for example, spillovers to other firms). The dynamic aspect is likely important for fully explaining how technology differences impact cross-country differences in productivity.

4 provides a calibration of the model to United States' data and discusses the quantitative effect of distortions on the adoption pattern. Section 5 continues the quantitative analysis by considering two experiment that quantify the important of distortions to cross-country differences in productivity and technology. Section 6 concludes.

2 Evidence on Technology Adoption

In this section, I present three observations on technology adoption found in the literature. These observations serve two purposes: they provide foundation for the non-standard features of the model (presented in Section 3); and they help discipline the calibration (Section 4).

Observation 1. *The adoption of new technologies is gradual.*

The first observation is that technology adoption is a relatively gradual process, taking up to several decades for major innovations. This observation is related to the 10-90 lag, which measures the number of years it takes for the technology to move from 10% to 90% of output produced using the technology. In a comprehensive sample of 265 innovations, Grubler (1991) finds that the 10-90 lag ranges from several years to hundreds of years. On average, he finds a 10-90 lag of 41 years in his sample. Jovanovic and Lach (1997) examine a sample of 21 goods and find a 10-90 lag of 15 years. Other studies on individual technologies report 10-90 lags of 54 years for steam locomotives; 25 years for diesel locomotives (Greenwood, 1999); around 20 years for Blast Oxygen Furnaces (Oster, 1982); and around 40 years for tractors (Manuelli & Seshadri, 2014).

The observation is also related to the widely reported S-shaped pattern of adoption.[4] The S-shaped pattern is characterized by a relatively small number of early and late adopters with the majority of adoption occurring in an intermediate transitory period. The shape is

[4]For example, see work by Griliches (1957) on hybrid corn; David (1966) on the mechanical reaper; Oster (1982) on blast oxygen furnaces; Levin, Levin, and Meisel (1987) on optical scanners; or Manuelli and Seshadri (2014) on tractors.

explained by heterogeneity in the marginal benefits of the new technology to adopter.[5] Early adopters are those with the highest benefit, while laggards wait until later improvements make the technology profitable before adopting.[6]

Observation 2. *The cost of existing technologies falls over time.*

The decline in prices has been examined in connection with the adoption of new technologies. Specifically, papers have examined this relationship for tractors (Manuelli & Seshadri, 2014); equipment (Greenwood, Hercowitz, & Krusell, 1997); computers (Yorukoglu, 1998); and a number of other technologies (Gort and Klepper (1982)provide evidence on 23 innovations).

The fall in cost may allow developing countries to catch up through periods of relatively cheap growth.[7] Parente and Prescott (1994) show that the number of years for countries to double their income has fallen systematically over time.[8] Consistent with this view, Comin and Hobijn (2010) find that the periods of rapid growth for a number of South-East Asian countries (Hong Kong, South Korea, Taiwan and Singapore) coincided with these countries reducing their technology adoption lags.

Observation 3. *Technology lags are declining with income per capita and are persistent over time and across technologies.*

The third observation differs from the first two as it relates to adoption at the country-level rather than at the firm-level. The observation relates to Comin and Hobijn (2004)'s finding that technology follows a "trickle-down" pattern in which technologies originate in developed countries and then are progressively adopted by less developed countries.

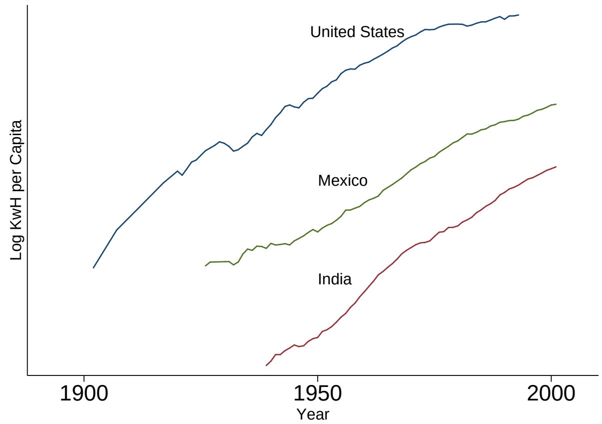
[5] Another popular explanation for the S-shaped adoption pattern is that firms have imperfect access to information on new technologies. This is consistent with evidence on the adoption of production techniques of soccer balls in Pakistan (Atkin et al., 2014); managerial practices in India (Bloom et al., 2013); the adoption of pineapples in Ghana (Conley & Udry, 2008); among others.

[6] Griliches (1957) finds that hybrid corn was adopted in more profitable areas earlier. In the development context, Suri (2011) finds a similar patterns for hybrid crops in Kenya.

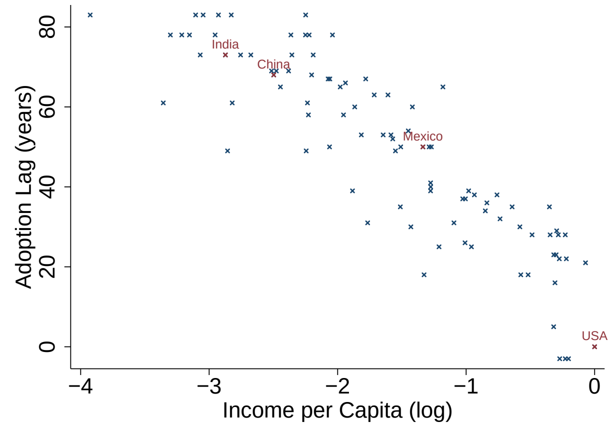
[7] See Gerschenkron (1965) or Abramovitz (1986) for early discussions and evidence of this hypothesis.

[8] Stokey (2014) extends this analysis to include more countries and discusses the cases of growth miracles and disasters in particular.

Figure 1: Adoption of Electricity



(a) Electricity Usage



(b) Electricity Lags (1990)

Notes: Data on electricity usage and population is from the Cross-country Historical Adoption of Technology (CHAT) dataset (Comin & Hobijn, 2009). Adoption Lag (Figure 1b) is defined as the number of years since KwH per capita was the same in the United States.

Figure 1 shows two trends for the case of electricity that are illustrative of this observation.[9] Figure 1a highlights the trends found in the study of the persistence of technology usages by Comin, Hobijn, and Rovito (2008). Specifically, they show that the rankings of countries in terms of technology usage tend to persist over time. Further, a country's ranking in a technology tends to be closely related to its ranking in other technologies as well as its ranking in output per capita. This leads to the second part of this observation (Figure 1b), that there is a strong negative relationship between a country's adoption lag and their income per capita (Comin & Hobijn, 2004, 2010). The takeaway is that differences in adoption lags should be thought of as relating to fundamental features of a country, rather than being thought of as idiosyncratic to the technology.

[9]I note that electricity is chosen for brevity and because it is not an industry-specific technology. I note that the trends hold more generally across technologies.

3 Model

I consider a model that incorporates technology choice along the lines of Parente and Prescott (1994) into a standard model of firm heterogeneity (for example, Hopenhayn, 1992). For ease of exposition, I begin the section by discussing the model without distortions. Distortions are then introduced and the distorted equilibrium is defined and characterized. The section ends with a discussion of key variables and adoption.

3.1 Economy without Distortions

Time is discrete and indexed by t. The economy is populated by a unit mass of infinitely lived households and an endogenous mass M of firms.[10] Households supply labour inelastically and are firm owners. Firms differ in an exogenous component of productivity and choose labour and technology inputs. Technology is non-rivalrous and is modeled by an exogenous set of productivities and a cost function.

Technology and Production: The set of available technologies is taken to be exogenous and is given by $[0, \bar{z}_t]$ in period t, where $\bar{z}_t$ is the technology frontier. The exogeneity of technologies follows evidence in Section 2 showing that developing countries lag significantly in the adoption of technologies. The technology frontier grows at a rate g, such that in period t the technology frontier is

$$\bar{z}_t = \bar{z}_0 e^{gt}.$$

A firm with technology z_t produces

$$y_t = z_t \tilde{y}(z_t), \tag{1}$$

[10]The mass of households is normalized for notational convenience. Aggregate variables are interpreted as per capita values.

where $\tilde{y}$ is a stand-in for the firm's static choices, which is discussed later. Firms choose technology z_t at the beginning of each period t.[11] Following Parente and Prescott (1994), the labour cost x_t of adopting z_t for a firm with z_{t-1} is given by

$$x_t = \pi \left[\left(\frac{z_t}{\bar{z}_t} \right)^{\xi} - \left(\frac{z_{t-1}}{\bar{z}_t} \right)^{\xi} \right], \tag{2}$$

where $\pi \geq 1$ is a parameter that captures aggregate barriers to adoption and $\xi > 1$ is the curvature of the cost function.[12] Aggregate barriers π are the same for all firms within a country but may differ between countries. The barrier acts as a stand-in for any other differences between countries that may lead to systematic differences in technology usage.[13] Otherwise, firms are unconstrained in their ability to purchase technology from the frontier.[14]

Technology is modeled to capture two channels in which technology innovation occurs. First, new technologies are introduced in the economy expanding the set of technologies. This is captured in the model by the technology frontier $\bar{z}_t$ growing over time. Second, innovations on existing technology make them less costly to adopt. This is captured in the model by the price of existing technologies falling over time. The cost function is also consistent with the observations in Section 2. The convexity of the cost function is necessary for some firms to adopt below the technology frontier as in Observation 1, while the falling cost over time is consistent with Observation 2. Finally, Observation 3 suggests that developing countries tend to be importers of innovations and so the technology frontier is considered exogenous to the economy.[15]

[11] Changing the technology choice to the period prior does not significantly affect the results. This timing aligns the decisions made by the entering and incumbent firms.

[12] For example, Comin and Hobijn (2010) and Stokey (2014) assume similar cost functions.

[13] For example, patent protection laws, trade barriers, or human capital differences may all results in cross-country differences in technology that are not firm-specific. The quantitative exercises are agnostic on the causes of these barriers π.

[14] Specifically, I assume that there are no informational barriers that prevent firms from adopting otherwise profitable technologies. This is consistent with the evidence presented in Section 2.

[15] A concern with modeling the cost function in a reduced form is that there may be an important interaction with idiosyncratic distortions. Comin and Hobijn (2004) note that technological innovations and improvements tend to occur in developed countries. I take this as evidence that at least for initial adoption, any reduction in the cost occurs externally to the country. A key channel abstracted from in this paper is the role of spillovers (for example learning-by-doing) in the adoption of new technologies.

I define proximity to the technology frontier as $\varphi_t = \left(\frac{z_t}{\bar{z}_t}\right)^{\xi} \in [0,1]$. A technology can be written as $z_t = \bar{z}_t \varphi_t^{\frac{1}{\xi}}$, which decomposes a technology's productivity into the frontier productivity and the technology's proximity to the frontier. It is convenient to rewrite the firm's problem in terms of proximity φ as the transformed problem is invariant over time. The technology cost (2) is rewritten as a law of motion for φ as

$$\varphi_t = \frac{1}{\pi} x_t + (1-\delta)\varphi_{t-1}, \tag{3}$$

where $(1-\delta) = e^{-g\xi}$ is the rate at which a technology's proximity drifts away from the technology frontier between periods. I refer to $(1-\delta)$ as the technological drift. In what follows, I drop the time t subscript and use -1 to denote the prior period.

Firm Heterogeneity: Firms are perfectly competitive and produce a homogeneous good. The static component of production is given by $\tilde{y} = (\kappa a)^{(1-\omega)(1-\gamma)} n^{\gamma}$, where n is labour; a is an firm-specific component of productivity; and κ is a country-specific component of productivity.[16] Intuitively, the term a is thought of as capturing relative differences in managerial talent between firms. The term κ is thought of as a residual measure of productivity that captures factors not explicitly modeled. In this spirit, I assume that the average value of a (denoted by $\bar{a}$) is constant across countries so that a only captures between firm differences.

The production function (1) can be written in terms of φ as

$$y(n, \varphi; a) = \bar{z}((\kappa a)^{1-\omega} \varphi^{\omega})^{1-\gamma} n^{\gamma}, \tag{4}$$

where $\gamma \in (0,1)$ and $\omega = \frac{1}{\xi}\frac{1}{1-\gamma} \in (0,1)$ relate to a firm's span-of-control. The term ω then measures the decreasing returns on proximity φ that corresponds to a given curvature ξ on the cost function of technology z. As $\omega(1-\gamma)+\gamma \in (0,1)$ firms have decreasing returns

[16] The firm productivity κa is thought of as a rescaling of the true firm productivity such that the exponent implies a labour and technology demand function that is linear in κa.

to scale in their chosen inputs and so they will have an optimal size that depends on their productivity a. For now, firms only differ in productivity. Later, productivity and distortions will determine a firm's optimal size.

The value of a firm is given by

$$\begin{aligned} v_a(\varphi_{-1}) &= \max_{n,\varphi,x} \bar{z}(\varphi^{\omega}(\kappa a)^{1-\omega})^{1-\gamma} n^{\gamma} - wn - wx + D v_a(\varphi) \\ \text{s.t. } \varphi &= \frac{1}{\pi} x + (1-\delta)\varphi_{-1}. \end{aligned} \tag{5}$$

The discount rate $D = \frac{1-\lambda}{1+R} \in (0,1)$ depends on the rate at which firms discount future profits R and the exogenous probability of exit $\lambda \in (0,1)$.

I solve problem (5) in two steps: first, I solve for labour demand conditional on the level of technology; and second, I solve for the proximity choice given the labour demand. Given proximity $\varphi(a)$, the firm's static problem is

$$\max_{n} \bar{z}(\varphi(a)^{\omega} a^{1-\omega})^{1-\gamma} n^{\gamma} - wn. \tag{6}$$

The solution to (6) is given by $n(a) = \bar{z}^{\frac{1}{1-\gamma}} \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} \varphi(a)^{\omega}(\kappa a)^{1-\omega}$. Given the conditional labour demand, the firm's dynamic problem is

$$\begin{aligned} v_a(\varphi_{-1}) &= \max_{\varphi,x} B(\varphi^{\omega} a^{1-\omega}) - wx + D v_a(\varphi), \\ \text{s.t. } \varphi &= \frac{1}{\pi} x + (1-\delta)\varphi_{-1}. \end{aligned} \tag{7}$$

The term B is a ratio of the firm's variable profits to the idiosyncratic components of the firm's productivity. Then, B is the component of the marginal benefit of technology that is common to all firms in the economy. The term $a^{1-\omega}$ generates heterogeneity in the marginal benefit of technology between firms. The solution to (7) is

$$\varphi(a) = \max\left\{\eta a, 1\right\}, \tag{8}$$

where $\eta = \kappa \left[\bar{z}^{\frac{1}{1-\gamma}} \gamma^{\frac{\gamma}{1-\gamma}} w^{-\frac{1}{1-\gamma}} \frac{\omega(1-\gamma)}{\pi(1-D(1+g_w)(1-\delta))} \right]^{\frac{1}{1-\omega}}$ and g_w is the growth of the wage rate w. I assume that the distribution of a is such that no firm operates at the technology frontier, or equivalently that $\varphi(a) \leq 1$ for all a.

The term η is a measure of the level of technology in an economy. To see this, consider that the average technology level in an economy is $\eta\bar{a}$ and recall that $\bar{a}$ is assumed to be constant across countries. Variation in the level of technology between countries is then driven by differences in η with a higher η indicating that each firm is closer to the technology frontier. The level of technology η is increasing in the country-specific productivity level κ, decreasing in the wage rate w, and decreasing in the aggregate barrier π.

Entry and Exit: At the end of the period, each firm faces an exogenous probability λ of exiting.

New firms enter at the beginning of a period by paying a labour cost c_e. Entrants draw an underlying productivity a from the distribution $f(a)$ on $\mathcal{A}$ after paying the entry cost. Firms enter with zero prior technology, which is equivalent to setting prior proximity $\varphi_{-1} = 0$. This assumption does not affect the results as firms choose their optimal proximity (8) independently of their prior proximity φ_{-1}.[17] This implies that entrants immediately invest to their optimal proximity to the technology frontier, as opposed to transitioning slowly. The choice of setting $\varphi_{-1} = 0$ is then equivalent to setting $\varphi_{-1} > 0$ and raising the cost of entry c_e.

Households: There is a unit mass of households in the economy that act as firm owners and consume their income. Households supply one unit of labour each period. Aggregate labour is supplied to firms (N_w), to the establishment of new firms (N_E) and to investment in new technologies (N_X).[18]

[17]A potentially important source of variation is that firm in richer countries may enter the economy with relatively higher levels of technology. By the arguments mentioned, this is equivalent to having lower cost of entry in more developed countries. This channel would then to magnify the impact of technology differences between countries. I choose to abstract from this channel because the focus of the current paper is not to examine how technology affects the cost of entry.

[18]I assume that the cost of entry and investment are both paid in units of labour. Alternatively, it could

3.2 Distorted Economy

Distortions are modeled as an idiosyncratic tax τ on firm revenues as in Restuccia and Rogerson (2008). The government redistributes the tax revenues from these idiosyncratic taxes through a non-distortionary lump-sum transfer T_t to households. This tax acts as a stand-in for institutions that reallocate factors of production between firms. Some examples include: credit market constraints (Buera et al., 2011; Midrigan & Xu, 2014); size-dependent policies (Guner, Ventura, & Xu, 2008); and agricultural frictions that restrict the transfer of land, such as inheritance rules, explicit size constraints and tenancy restrictions (Adamopoulos & Restuccia, 2014). The model nests a static misallocation model, which corresponds to the case when $\omega \to 0$.

Rather than work directly with τ, a distortion is defined as $\theta = (1-\tau)^{\frac{1}{1-\gamma}}$. A system of distortions is the set of possible values of the distortions Θ and the joint probability distribution g of the distortions and firm idiosyncratic productivities. For simplicity, I assume that firms draw θ upon entry (after paying the cost of entry) and that θ is constant over a firm's life.[19] A distortion $\theta < 1$ corresponds to resources being reallocated from a firm, while a distortion $\theta > 1$ corresponds to resources being reallocated to a firm.

The distorted firm's problem is

$$v_{(a,\theta)}(\varphi_{-1}) = \max_{n,\varphi,x} \bar{z}(\theta\varphi^{\omega}a^{1-\omega})^{1-\gamma}n^{\gamma} - wn - wx - Dv_{(a,\theta)}(\varphi), \tag{9}$$
$$\text{s.t. } \varphi = \frac{1}{\pi}x + (1-\delta)\varphi_{-1}.$$

be assumed that these costs are paid in units of output. This would result in the cost of entry falling over time and consequently, the mass of firms being a function of productivity. Evidence presented by Bento and Restuccia (2016) and Bollard, Klenow, and Li (2016) show that the mass of firms is not positively related to productivity. However, qualitatively and quantitatively, the results do not depend much on the choice of specification. Hence, I choose the specification that is most consistent with the empirical evidence.

[19] This abstracts from two important features that have been emphasized in the literature. The first is that distortions may have an asymmetric affect on different inputs. In this case, labour and technology may be differently affected by distortions. The second is that distortions may affect the dynamic decisions of firms. This could occur if the distortion a firm received was determined by an observable level of inputs such as labour or output.

The resulting labour demand function and choice of proximity to the technology frontier are

$$n(a,\theta) = \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}} \left[\bar{z}^{\frac{1}{1-\gamma}} \theta \varphi(a,\theta)^{\omega} a^{1-\omega}\right] \quad \text{and} \quad \varphi(a,\theta) = \max\left\{\eta \theta^{\frac{1}{1-\omega}} a, 1\right\}, \tag{10}$$

where η is defined as in the undistorted economy. As is apparent in (10), distortions affect the economy through firm-level choices of labour and technology.

The free-entry condition, given by

$$wc_e = \int_{\mathcal{A}\times\theta} v_{(a,\theta)}(0) g(a,\theta)\, \mathrm{d}a\, \mathrm{d}\theta. \tag{11}$$

The free-entry condition equates the expected value of an entrant (right-hand side) to the cost of entry (left-hand side). Distortions affect the free-entry condition through the expected value of an entrant (right-hand side) and through the wage rate (left-hand side).

3.3 Equilibrium

Definition 1. *Given an initial distribution of φ_0 and a system of distortions Θ and $g : \mathcal{A} \times \Theta \to [0,1]$, a competitive equilibrium consists of the prices w_t; the labour function $n_t : \mathcal{A} \times \Theta \to \mathbb{R}_+$; the proximity function $\varphi_t : \mathcal{A} \times \Theta \to [0,1]$; the investment function $x_t : \mathcal{A} \times \Theta \to \mathbb{R}_+$; the mass of firms M_t; and lump-sum transfers T_t; such that*

1. *Taking prices as given, households maximize utility in each period t;*
2. *Taking prices and previous technology φ_{t-1} as given, firms choose $n_t(a,\theta)$, $x_t(a,\theta)$ and $\varphi_t(a,\theta)$ to solve problem (9) in each period t;*
3. *Potential firms enter if optimal, such that (11) holds in each period t;*
4. *T_t balances the government's budget, $T_t = \int_{\mathcal{A}\times\Theta} \tau(\theta) g(a,\theta) da d\theta$, in each period t;*
5. *Markets clear in each period t*

- *Goods Market:* $Y_t = Consumption_t$;

- *Labour Market:* $1 = N_{wt} + N_{Et} + N_{Xt}$;

The effects of distortions on the aggregate economy are summarized by two sufficient statistics. The first sufficient statistic is an average of the underlying productivities with weights $\theta^{\frac{1}{1-\omega}}$. The statistic is given by

$$\tilde{a} = \int_{\mathcal{A}\times\Theta} \left[\theta^{\frac{1}{1-\omega}} a\right] g(a,\theta)\, \mathrm{d}a\, \mathrm{d}\theta. \tag{12}$$

Notice that $\frac{n(a,\theta)}{n(a)} = \frac{\varphi(a,\theta)}{\varphi(a)} = \theta^{\frac{1}{1-\omega}}$, such that the weights $\theta^{\frac{1}{1-\omega}}$ correspond to the ratio of input demand in the distorted and undistorted economies. The term $\tilde{a}$ is an index of firm productivity and can be interpreted similarly to an effective average productivity of the firms in the economy.[20]

The second sufficient statistic describes the allocative efficiency of the economy. The statistic is given by

$$\chi = \frac{\int_{\mathcal{A}\times\theta} \left[\theta^{\gamma+\frac{\omega}{1-\omega}} a\right] g(a,\theta)\, \mathrm{d}a\, \mathrm{d}\theta}{\int_{\mathcal{A}\times\Theta} \left[\theta^{\frac{1}{1-\omega}} a\right] g(a,\theta)\, \mathrm{d}a\, \mathrm{d}\theta}. \tag{13}$$

The term χ is a measure of how close the allocation of labour and technology is to the efficient allocation. In the undistorted economy the numerator and denominator of (13) are equal, $\chi = 1$, indicating that resources are used optimally.

Proposition 1 characterizes the equilibrium.[21]

Proposition 1. *An equilibrium exists and is unique. The equilibrium is characterized by*

$$\ln Y = \zeta_Y + \ln \bar{z} - \omega(1-\gamma)\ln \pi + (1-\omega)(1-\gamma)\ln \kappa\tilde{a} + \ln \chi, \tag{14}$$

$$\ln w = \zeta_w + \ln \bar{z} - \omega(1-\gamma)\ln \pi + (1-\omega)(1-\gamma)\ln \kappa\tilde{a}, \tag{15}$$

[20] This is a similar interpretation as the one suggested by Hopenhayn (2014). Specifically, the relevant effect of distortions is captured by how the distortions reallocate resources across firms.

[21] A proof is available in Appendix A.

where ζ_i for $i \in \{Y, w\}$ is a collection of constants. The growth rate of output is constant and given by

$$g_Y = \ln Y_{t+1} - \ln Y_t = g \tag{16}$$

and is equal to the growth rate of wages g_w. The mass of firms M_t is a collection of constants and does not grow over time.

The undistorted economy is a special case of Proposition 1 in which distortion have a degenerate distribution where all firms draw $\theta = 1$. The equilibrium growth rate in (16) does not depend on distortions, such that the effect of distortions is only on the level of output. The output (14) and wage rate (15) are increasing in the technology frontier $\bar{z}$, the country-level productivity κ, and decreasing in aggregate barriers π. The wage rate is increasing in the effective average firm productivity $\tilde{a}$ and output is increasing in the allocative efficiency of the economy χ.

Corollary 1 states that a firm's proximity to the technology frontier is constant over time. The corollary is important because it allows a firm's adoption decision to be characterized solely by $\varphi(a, \theta)$.

Corollary 1. *In equilibrium, a firm's choice of the proximity to the technology frontier is constant,*

$$\varphi_t(a, \theta) = \varphi_s(a, \theta),$$

for all periods t and s.

3.4 Discussion of Equilibrium

This section provides a discussion of key equilibrium outcomes at the aggregate- and firm-level. These results are important for the calibration and quantitative exercises presented in

Section 5.

Aggregate Economy: Aggregate output in (14) can be written in terms of a representative firm $(\chi\tilde{a}, 1)$ with mass M:

$$Y = My(\chi\tilde{a}, 1) = M\bar{z}\left[(\kappa\chi\tilde{a})^{(1-\omega)}\varphi(\chi\tilde{a}, 1)^{\omega}\right]^{1-\gamma} n(\chi\tilde{a}, 1)^{\gamma}, \tag{17}$$

The terms $n(\chi\tilde{a}, 1)$ and $\varphi(\chi\tilde{a}, 1)$ can be interpreted as indices that capture the contribution of labour and technology to aggregate productivity. The technology index can be written as: $\varphi(\chi\tilde{a}, 1) = \eta\chi\tilde{a}$. The technology index captures both the level of technology through $\eta\tilde{a}$ and how efficiently technology is distributed across firms through χ. This implies that even if countries have the same level of technology (η is the same) the contribution of technology to productivity may differ because of how it is distributed (χ is different). The takeaway is that both the level and the *distribution* of technology are important for aggregate outcomes.

At the aggregate level, the model is similar to Parente and Prescott (1994) or Comin and Hobijn (2010). The benefit of adding firm heterogeneity is that it allows for distortions to affect firm choices. Focusing on distortions in the disaggregate economy provides two main insights. First, idiosyncratic distortions can have a substantial affect on aggregate productivity. This is well established in the existing literature.[22] Second, idiosyncratic distortions act as a barrier to the adoption of new technologies. Further, the importance of this channel is quantifiable using existing evidence in the misallocation literature. This is the main contribution of this paper.

Another benefit of focusing on firm-level decisions is that it provides a natural setting to study the gradual aggregate adoption of technologies. Aggregate adoption is determined by both the number of firms that use the technology and how intensely the technology is used. Comin and Mestieri (2014) show that while many countries superficially adopt similar technologies, they differ significantly in how intensively technologies are used. The term

[22]For example, Restuccia and Rogerson (2008) or Hsieh and Klenow (2009).

$\varphi(\chi\tilde{a}, 1)$ provides insight on how the efficiency of technology use (measured by χ) helps explain cross-country differences in technology intensity.

Another useful decomposition of aggregate output is to consider the loss to the economy of moving from an undistorted economy to distorted economy. This loss is given by

$$\ln \text{Loss} = (1-\omega)(1-\gamma)\ln\frac{\tilde{a}}{\bar{a}} + \ln\chi \tag{18}$$

Recall that the productivity index $\tilde{a}$ is equal to the average productivity $\bar{a}$ in the undistorted economy and that the allocative efficiency χ is equal to 1 in the undistorted economy. The first term in (18) is the difference in effective productivity between the distorted and undistorted economy. The second term in (18) is the difference in allocative efficiency between the distorted and undistorted economy.[23]

It is important to note that this loss does not depend on either the level of aggregate barriers π or the country-specific productivity κ. The intuition for this is that these parameters affect all firms equally and so they do not influence the relative incentives of firms. These parameters can then be normalized in the calibration of the benchmark economy.

Firm-Level Variables: A firm's labour allocation and output are given by

$$n(a,\theta) = \frac{1}{M}\frac{\theta^{\frac{1}{1-\omega}}a}{\mathbb{E}[\theta^{\frac{1}{1-\omega}}a]}, \qquad \text{and} \qquad y(a,\theta) = \frac{Y}{M}\frac{\theta^{\frac{\omega}{1-\omega}+\gamma}a}{\mathbb{E}[\theta^{\frac{\omega}{1-\omega}+\gamma}a]}. \tag{19}$$

Recall that the undistorted economy is the special case where $\theta = 1$ for all firms. Then by (19), in the undistorted economy, a firm's share of total labour is equal to its share of total output. However, in the distorted economy $\theta \neq 1$ for some firms and so it is no longer the case that a firm's share of total labour is equal to its share of total output.

Combining $n(a,\theta)$ and $y(a,\theta)$ provides a measure of the revenue productivity of firms

[23] It is straightforward to show that any proportional increase to all distortions θ would have no effect on aggregate productivity. While not immediately apparent in (18), this follows from also considering the definitions of χ and $\tilde{a}$.

TFPR, given by

$$\text{TFPR}(a,\theta) = \frac{y(a,\theta)}{n(a,\theta)} \propto \frac{1}{1-\tau(\theta)}, \tag{20}$$

where $\tau(\theta)$ is the tax rate that corresponds to the distortion θ. TFPR is useful because it is often reported in firm-level studies of distortions and so (20) provides a mapping to the distortions reported in these studies.[24] Since the variance of $\ln \text{TFPR}$ is often reported, I use that statistic to calibrate distortions in Section 5. Formally, the variance is given by

$$\text{Var}(\ln \text{TFPR}) = \text{Var}(\ln(1-\tau(\theta))). \tag{21}$$

The variance of firm-level productivities is driven solely by the tax rate, $\tau(\theta)$. Importantly, (21) does not depend on the parameter ω implying that technology does not affect the variance in distortions inferred from variance in $\ln \text{TFPR}$. Put another way, the dispersion of distortions is not larger in a model with technology than in a model without.

3.5 Adoption

Finally, I discuss adoption of new technologies along the balanced growth path. Corollary 1 states that firms maintain their proximity to the technology frontier over time. Consequently, the adoption pattern in the economy is driven by the falling cost of new technologies over time. I begin by discussing firm-level adoption and deriving a firm's optimal adoption lag. Specifically, the number of periods it takes for a firm to use a technology after the technology is invented. I then discuss how the optimal adoption lag varies with regard to key country-specific parameters. The section ends with a discussion of aggregate adoption.

Firm-Level Adoption: When a new technology is introduced it has proximity $\varphi = 1$.

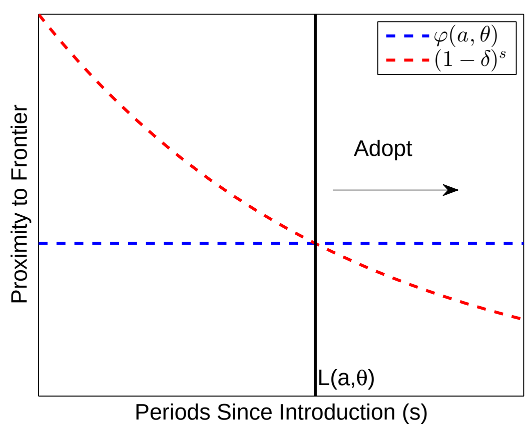
[24] Many of the firm-level studies focus use models with CES preferences and monopolistic competition. Hsieh and Klenow (2009) show the mapping of $\text{TFPR}(a,\theta)$ between these models and the class of span-of-control models used here.

After s periods the technology has proximity $(1-\delta)^s$. A firm has adopted the technology if the firm uses any technology with proximity greater than the technology. Formally, firm (a, θ) adopts the technology in any period s such that

$$\varphi(a, \theta) \geq (1-\delta)^s. \tag{22}$$

The form of (22) is a consequence of Corollary 1. Since firms maintain their proximity to the technology frontier over time, the adoption lag depends only on the drift of the technology frontier, which is the right-hand side of (22). Figure 2 provides a graphical representation.

Figure 2: Adoption Decision



I define the optimal adoption lag $L(a, \theta)$ for firm (a, θ) as the number of periods between when a technology is introduced and when it becomes viable for the firm to adopt.[25] Solving for the period in which (22) holds with equality gives

$$L(a, \theta) = \max\left\{-\frac{1}{g\xi} \ln\left[\eta\theta^{\frac{1}{1-\omega}} a\right], 0\right\}. \tag{23}$$

Intuitively, $L(a, \theta)$ is the length of time for a firm (a, θ) to adopt a new technology after it is introduced. The adoption lag in (23) must be bounded by 0 as the firm can only adopt

[25]For ease of exposition, I define the adoption lag using the smooth function $L(a, \theta)$. However, since time is discrete in this model, the adoption lag is a step function. The true adoption lag then corresponds to the smallest integer larger than $L(a, \theta)$.

technologies that have been invented. Otherwise, the adoption lag is falling in the firm's proximity $\varphi(a,\theta) = \eta\theta^{\frac{1}{1-\omega}}a$ and the technological drift, $(1-\delta) = e^{-g\xi}$.

It is apparent in (23) that a firm's proximity maps directly into the adoption lag. The adoption pattern of an economy conveys the same information as the distribution of technologies. In the quantitative exercises, I consider how distortions affect technology by showing the effect of distortions on the adoption pattern. This is convenient as it is comparable to the existing literature and it allows the adoption lags to be compared with empirical evidence.

Relation with Parameters: Before introducing aggregate adoption, it is useful to build intuition by examining the relationship between the firm-level adoption lag $L(a,\theta)$ and key parameters. In particular, the country-level productivity κ; the aggregate barrier π; and the distribution of distortions θ. These relationships provide a foundation for the numerical exercises in Sections 4 and 5.

First, the relationship between $L(a,\theta)$ and κ is given by

$$\frac{\mathrm{d}L(a,\theta)}{\mathrm{d}\ln\kappa} = 0. \tag{24}$$

The firm-level adoption lag does not vary with changes in the country-level productivity. Intuitively, increasing κ raises the profitability of firms and so firms demand more labour and more firms enter the market. The increase in labour demand through these channels increases the wage rate by exactly enough to offset the increase in the incentive of firms to invest in technology resulting in no net change in the adoption lag. Consequently, κ also does not have an effect on the aggregate adoption pattern. This highlights the importance of general equilibrium effects on adoption. Based on (24), we might expect that, as the quantitative results in Section 4 confirm, any change in distortions that affect all firms equally will be offset by a corresponding change in the wage rate.

Second, the relationship between $L(a,\theta)$ and π is given by

$$\frac{\mathrm{d}L(a,\theta)}{\mathrm{d}\ln\pi} = -\frac{1}{g\xi}\frac{\mathrm{d}\ln\varphi(a,\theta)}{\mathrm{d}\ln\pi} = \frac{1}{g\xi} > 0. \tag{25}$$

Perhaps unsurprisingly, a firm's adoption lag is increasing in the aggregate barrier π.[26]

Third, the relationship between $L(a,\theta)$ and θ is given by

$$\frac{\partial L(a,\theta)}{\partial\ln\theta} = -\frac{1}{g\xi}\frac{\partial\ln n(a,\theta)}{\partial\ln\theta} = -\frac{1}{g\xi}\frac{1}{1-\omega} < 0. \tag{26}$$

A firm's adoption lag is decreasing in the its distortion θ. The partial elasticity of the adoption lag with respect to the distortion is proportional to the firm's elasticity with respect to the distortion. This is similar to the result in (19) showing that a firm's proximity is proportional to its labour input and provides a nice intuition for the aggregate effects of distortions on the adoption pattern. Specifically, the effects of distortions on the adoption pattern can be related to the effects of distortions on the size distribution. If firms become more similar in terms of size, the length of adoption will shorten while if they become less similar, the length of adoption will be prolonged. On the other hand, if distortions tend to reallocate resources away from high productivity (ex-ante large) firms, the initial adoption of a technology will be delayed.

Together (25) and (26) show that there is no interactions between π and θ on adoption lag. This implies that the lag caused by distortions is separate from the lag caused by aggregate barriers. In the calibration, this allows π to be used as a residual measure of adoption for a given system of distortions.

Aggregate Adoption: With these relationships in mind, I turn to the aggregate adop-

[26]Additionally, (25) is used to compare the relative lag between two countries: $\frac{\pi_i}{\pi_j} = \exp\{g\xi \text{ Rel. Lag i-j}\}$, where Rel. Lag i-j is the number of periods between when a firm (a,θ) would adopt a technology in country j relative to country i. This provides a convenient target in the calibration as π can be varied to match adoption lags between two countries.

tion pattern. To consider the adoption pattern, it is necessary to define a statistic that characterizes the adoption pattern in the economy.

A possible statistic to measure adoption is the fraction of firms that have adopted the new technology. This statistic is given by the mass of firms with $\varphi(a, \theta) \geq (1-\delta)^s$ for each period s. The statistic only measures technology adoption through the extensive margin. However, this statistic has some issues that are pointed out by Comin and Mestieri (2014). Notably, only measuring adoption in terms of the extensive margin misses some key variation between countries.

My preferred statistic is the fraction of output produced with a technology of age s, which is given by

$$H(s) = \int_{\varphi(a,\theta) \geq (1-\delta)^s} \frac{y(a, \theta)}{Y/M} g(a, \theta) \, \mathrm{d}a \, \mathrm{d}\theta. \tag{27}$$

The statistic H incorporates both the intensive and extensive margin of adoption.[27] The extensive margin is the number of firms that have adopted the technology. The intensive margin is the average production of firms that have adopted the technology, such that firms that produce more are given more weight. In the quantitative exercises that follow, I report adoption based on (27).

4 Calibration and the Effect of Distortions

In this section, I calibrate the benchmark economy (BE) to the United States and examine the effect of changing distortions on the pattern of adoption. In particular, I vary distortions along three dimensions that relate to increasing the dispersion of distortions, increasing the correlation of distortions and increasing the average level of distortions. The consensus view in the literature is that the first two dimensions are important determinants of aggregate

[27] A third measure is the fraction of resources allocated to firms that have adopted the new technology. In a richer model that includes capital, this measure could be useful as it relates directly to measures of adoption that focus on the accumulation of technology-specific capital.

productivity. On the other hand, the average level of distortions is considered as a test of whether a similar effect of distortions could be captured in a representative firm framework. I find that the average level of distortions does not affect the adoption pattern.

4.1 Calibration

The calibration includes choosing values for the nine parameters $\{\gamma, R, g, \lambda, \bar{z}, c_e, \xi, \pi_{BE}, \kappa_{BE}\}$; the distribution of idiosyncratic distortions f; and the joint distribution of distortions and idiosyncratic productivities g. The calibration is summarized in Table 1.

Table 1: Calibration Values and Targets

Parameter	Value	Target
γ	2/3	Standard
g	2.00%	US Growth Rate $g_Y = 2\%$
λ	0.07	Standard
R	0.04	Standard
$\bar{z}$	1.00	Normalization
c_e	1.00	Normalization
κ_{BE}	1.00	Normalization
a_{min}	1.00	Normalization
a_{max}	15,865	US Firm Size Distribution
π_{BE}	200	5% Adoption in Period 0 (Normalization)
ξ	8.74	40 Year 10-90 Lag
ω	0.34	Implied by ξ

The parameters $\{\gamma, R, g, \lambda\}$ are standard and correspond to a period being one year. The growth rate of the technology frontier g is chosen such that output growth g_Y is 2%. The exogenous exit rate λ is 7%; the discount rate R is 4%; and the labour share γ is 2/3. The parameters $\{\bar{z}, c_e\}$ are normalized to unity.

A firm in the model corresponds to an establishment in the data. I choose f to match the employee distribution of establishments from the Business Dynamics Statistics (BDS) for 2000. I approximate f by a linear piecewise function with support $[a_{min}, a_{max}]$ such that the CDF of firm sizes matches the CDF of establishment sizes reported by the BDS.[28] The

[28] The BDS reports the fraction of establishments with employment less than or equal to specific levels.

lower support a_{min} is normalized to unity and the upper support a_{max} is chosen such that the average (across draws of distortions) firm with underlying productivity a_{max} employs $1,000$ times as many workers as the average firm with underlying productivity a_{min}.[29]

Cost Function: The cost function has two components: the aggregate barrier π_{BE} that determines the level of the cost and the curvature ξ, which determines ω. The choice of π_{BE} is not critical as π does not change the effect of distortions on aggregate productivity or on the adoption lag, as shown in (14) and (25).[30] With this in mind, I set π_{BE} to normalize adoption in the benchmark economy in period 0 to be 5%. The periods t in the model are then interpreted as the number of years since the United States began using the technology.

I set ξ to target the length of adoption measured by the 10-90 lag (discussed in Section 2). The 10-90 lag measures the number of periods it takes for a technology to move from being used in 10% of production to 90% of production. Denote n_{10} and n_{90} as the size of the marginal firms that adopt the technology in the periods where the technology is used in 10% and 90% of production. The 10-90 lag is found by comparing the adoption lags in (23) of these two firms:

$$\text{10-90 Lag} = \frac{\ln n_{90}/n_{10}}{g_Y}\frac{1}{\xi}. \tag{28}$$

The 10-90 lag in (28) is related to the ratio n_{90}/n_{10}, the output growth rate g_Y, the labour coefficient γ, and the curvature of the cost function ξ. I set ξ to target the 10-90 lag given the choices of n_{90}, n_{10}, g_Y and γ discussed previously.[31]

I target a 10-90 lag in the benchmark economy of 40 years. This number roughly corre-

For example, the fraction of establishments with less than or equal to four employees. I set f such that the fraction of firms at each employment level matches the numbers reported by the BDS.

[29]The largest establishment size in the BDS is 1,000 workers.

[30]Implicit in this statement is the assumption that firms are not constrained in their choice of technology. Otherwise, the choice of π would determine number of constrained firms and could change the effect of distortions. The assumption that no firms are constrained follows evidence that many technologies were invented years before they were ever put to use by firms. (Comin & Hobijn, 2010) discuss several examples.

[31]An alternative interpretation of ξ is the elasticity between the technology choice (φ) and a firm's pro-

sponds to the 10-90 lag of tractors in the United States (see Manuelli and Seshadri, 2014). This choice falls in the range of 10-90 lags discussed in Section 2. For ease of language, I refer to the 10% and 90% levels of adoption as initial and final adoption for the remainder of the paper.

Parente and Prescott (1994) set the parameter ω in the range $[0.50, 0.55]$. I find that these values of ω imply a 10-90 lag of 54 years, which is substantially longer than what is found in the literature. Appendix B discusses the sensitivity of the results with respect to the choice of ξ and the implied 10-90 lag.

Distortions: Following a firms draw of its idiosyncratic productivity a it draws a distortions according to

$$\ln(1-\tau) = \mu - \rho \ln a + \epsilon \tag{29}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$. The functional form assumed in (29) is similar to distortions in Bento and Restuccia (2016) or Guner, Parkhomenko, and Ventura (2015).

The system of distortions (29) has three parameters μ, σ, and ρ. Increasing the level of distortions μ increases the expected distortions firms receive when they enter the economy. Increasing the dispersion of the distortions σ increases the variance of the distortions for a given value of a. Increasing the elasticity of the distortions ρ reallocates more resources from high a firms to low a firms.

I set $(\sigma_{BE}, \rho_{BE}) = (0.466, 0.081)$ based on Hsieh and Klenow (2009) and set μ_{BE} such that $\tilde{a} = \bar{a}$. I do not assume that the United States is undistorted (as is common in the literature) so that the benchmark economy is comparable with other countries in the cross-

ductivity ($TFPQ$). Formally, the relationship is

$$\log(TFPQ) = \log \bar{z} + \frac{1}{\xi}\log(\varphi) + (1-\gamma)(1-\omega)\log a.$$

where the elasticity of a firm's technology choice with respect to their measured productivity is $\frac{d \log \varphi}{d \log TFPQ} = \xi$. A higher value of ξ then corresponds to a higher elasticity between a firm's technology and its productivity.

country analysis (Section 5).

4.2 Adoption Pattern

The remainder of this section shows the comparative statics exercise of varying distortions along the three dimensions (μ, σ, ρ) on the adoption pattern. The results of the exercise show how the distortions found in other studies may affect the adoption pattern. Although I abstract from the specific sources of distortions, it is not difficult to consider how specific institutions would map into (μ, σ, ρ). The results also provide a foundation for the cross-country analysis in Section 5 by showing how each dimension of the distortions affect the adoption pattern. The results show sensitivity of the analysis in Section 5 to changes in (μ, σ, ρ).

Table 2 summarizes the results of the exercise. Table 2 is partitioned into three sections that correspond to the number of years to reach initial, median and final levels of adoption, defined as 10%, 50% and 90% of total output produced with a new technology.

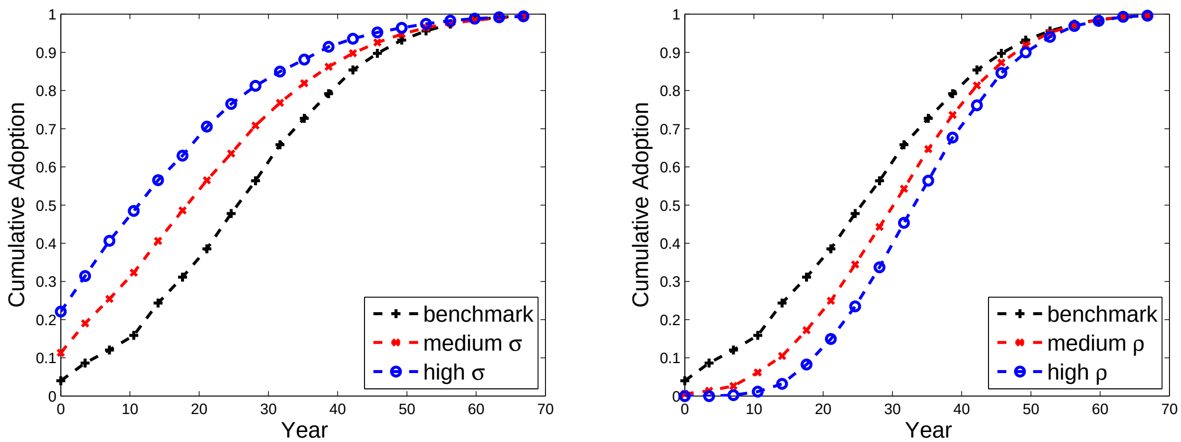
Table 2: Effect of Distortions on Adoption Lags

	Initial		Median		Final	
ρ / σ	0.25	0.50	0.25	0.50	0.25	0.50
0.00	3.4	0.0	17.6	14.3	41.5	40.2
0.15	22.6	13.3	34.4	29.6	45.6	47.1
0.30	31.5	21.3	39.8	34.4	48.3	50.5

Notes: Values are the number of years to reach initial adoption (defined as 10% of total output produced with new technology), median adoption (50%) and final adoption (90%). Distortions are given by (29) with variation in parameters (σ, ρ).

The first parameter is the level of distortions μ, which has no affect on the adoption pattern. Because μ does not affect the adoption pattern, I do not list the results in Table 2 for changes in μ. Intuitively, increasing μ results in two effects on the adoption of technologies, which exactly offset. The first effect is that the increase in the distortions makes all firms less profitable and, consequently, lowers investment in technology. The second effect is that firms demand less labour causing the wage rate to fall. This makes technology relatively

Figure 3: Adoption Pattern



(a) Dispersion (σ) (b) Elasticity (ρ)

Notes: Cumulative adoption is the fraction of output produced with the new technology. Figure 3a shows the effect of increasing the dispersion of distortions σ relative to the benchmark economy. Figure 3b shows the effect of increasing the elasticity of distortions with respect to firm productivity ρ relative to the benchmark economy.

cheaper increasing firm investment in technology. This result reiterates why the effect of distortions on the adoption pattern could not be studied in a representative firm framework.

The second parameter is the dispersion of distortions σ. Increasing σ prolongs the adoption period and results in both earlier initial adoption and later final adoption. Figure 3a shows graphically the effect of changing σ on the adoption pattern. The intuition is that increasing σ also increases the dispersion of firm sizes meaning there are more very large and very small firms. The large firms adopt earlier decreasing the initial adoption lag while the small firms adopt later increasing the final adoption lag.[32]

The third parameter is the elasticity of distortions ρ. Increasing ρ shortens the adoption period and delays initial adoption. Figure 3b shows graphically the effect of changing ρ on the adoption pattern. For intuition, consider how changing distortions affects the firm's technology decision in (10) when the wage rate is fixed and $\sigma = 0$. Firms with high underlying

[32]An implication of this result is that if technology spillovers are large enough, increasing σ may be beneficial. The intuition for this result is that since large firms adopt technologies sooner, the spillovers of having new technologies sooner may outweigh the negative affects of misallocation. I leave the study of this interaction to future research.

productivity (high a firms) now draw a smaller θ causing them to adopt technologies that are further away from the technology frontier. Low a firms now draw a higher value of θ causing them to adopt technologies that are closer to the technology frontier. Overall, the increase in ρ narrows the distribution of technologies used by firms, which shortens the adoption period. The wage rate tends to fall in response to an increase in ρ, which results in all firms adopting technologies closer to the technology frontier. However, this does not offset the effect of drawing a smaller θ on the high a firms' adoption choice.

5 Cross-Country Differences

In this section, I examine the role of distortions in cross-country productivity and technology differences. I consider two sets of counterfactual experiments in which distortions and aggregate barriers are varied to match features of developing countries. The first exercise decomposes the total gap in productivity from changing the average level of technology, the allocation of labour across firms, and the distribution of technology across firms. The second exercise measures the importance of idiosyncratic distortions as a barrier to adoption by comparing the adoption pattern under different levels of distortions.

5.1 Cross-Country Calibration

In the counterfactual experiments, I consider only changes to the parameters $\{\kappa, \pi, \mu, \sigma, \rho\}$ holding the distribution of other parameters, such as firm-level productivities $f(a)$, fixed. Other parameter values are taken from the benchmark economy (BE) in Section 4. In this regard, the experiment is thought of as changing the institutions in the United States to resemble those of other countries. This is similar in spirit to the quantitative exercise in Restuccia and Rogerson (2008) where distortions are varied holding the distribution of productivities constant. It differs from the quantitative exercise in Hsieh and Klenow (2009) in which the distribution of productivities and distortions vary across countries.

I construct sets of parameters to match targets corresponding to representative economies for countries in each decile of the income (per capita) distribution. Each of the counterfactual economies is indexed by i corresponding to the decile is represents.

I set the country-specific component of productivity κ such that $\kappa_i = \kappa_{BE}$ for all economies i. This follows the results that κ does not change the effect of distortions on aggregate productivity or the adoption pattern, as is shown by (14) and (24).[33] For similar reasons, I set $\mu_i = \mu_{BE}$. The numerical results in Section 4.2 show that the choice of the average level of distortions μ does not affect the adoption pattern. Similarly, the expression (14) can be used to show that changing μ does not affect aggregate output.

This leaves three parameters to be calibrated. I target $\{\sigma_i, \rho_i\}$ based on the distortions found by Hsieh and Klenow (2009, 2014) and π_i based on the adoption lags found by Comin and Hobijn (2010). These targets are listed in Table 3.

Table 3: Cross-Country Targets

Decile	Y_{US}/Y_i	Lag_i	$ELAS_i$	SD_i
10	0.95	14.3	0.122	0.486
9	1.22	16.4	0.162	0.503
8	1.55	18.4	0.200	0.519
7	2.20	21.4	0.256	0.542
6	3.06	24.2	0.309	0.564
5	4.11	26.7	0.356	0.584
4	5.55	29.3	0.404	0.604
3	8.69	33.1	0.476	0.634
2	15.42	37.9	0.568	0.673
1	34.35	44.7	0.696	0.726

Notes: Y_{US}/Y_i is the ratio of labour productivity in the data. Lag the number of years it takes economy i to adopt a new technology after it has been adopted in the United States. SD is the standard deviation of log total factor revenue productivity. ELAS is the elasticity between $1-\tau$ and idiosyncratic productivity.

I set the dispersion σ_i and elasticity ρ_i of distortions to jointly target two moments: the standard deviation of $\ln TFPR$ (denoted by SD) and the elasticity of distortions with respect to firm productivity (denoted by $ELAS$). Mechanically, this is done by setting

[33] It is straightforward to set κ_i to target the residual productivity between what is captured by $(\pi_i, \mu_i, \sigma_i, \rho_i)$ and what is observed in the data. This does not change the results.

ρ_i to match $ELAS_i$ and then adjusting σ_i until the SD_i implied by the model matches Table 3. Interestingly, the value of ρ_i inferred from $ELAS_i$ is smaller than in a model without technology.[34] This is because firms adjust their productivity in response to the distortion causing the measured elasticity to be larger than than the actual elasticity between distortions and a firm's underlying productivity. Combined with the results showing that the variance of TFPR is identical to a model without technology, this suggests that the distortions are more moderate in a model with technology. Specifically, distortions need to be less targeted on productive firms in order to achieve the same observable outcomes.

I discipline SD_i and $ELAS_i$ with the moments reported by Hsieh and Klenow (2009, 2014) for India and the United States. The targets are based on the values for India in 1994 ($ELAS_{India} = 0.50$ and $SD_{India} = 0.67$) and the United States in 1997 ($ELAS_{US} = 0.13$ and $SD_{US} = 0.49$). To construct the targets, I assume that SD_i and $ELAS_i$ are linearly related to log income per capita and infer the values based on the average income per capita in each decile, reported in Table 3.[35]

I focus on the United States and India because the value of SD_i and $ELAS_i$ reported by Hsieh and Klenow (2009, 2014) are comparable for the two countries. To put the values of SD_i in perspective, I contrast the values reported by other micro-studies with those in Table 3. Hsieh and Klenow (2009) report values of 0.63 to 0.74 for China between 1998 and 2001, which is consistent with the SD_i for the bottom two deciles. Bartelsman, Haltiwanger, and Scarpetta (2013) report a higher value of 0.58 for the United States between 1993 and 2001. Over a similar time period they report values similar to the United States (0.53 to 0.71) for

[34] The parameter ρ_i is calculated directly from the value of $ELAS_i$ using the relationship

$$\rho_i = (1-\gamma)(k-1)\left[\frac{1}{ELAS_i} + \frac{\omega}{1-\omega}\right]^{-1}.$$

where k is the CES term used by Hsieh and Klenow (2009) ($k = 3$). Recall that the limiting case when $\omega \to 0$ corresponds to a static misallocation model. Then, it follows from the above expression that $\frac{\partial \rho}{\partial \omega} < 0$ implying that the inclusion of technology lowers ρ. Since firms adjust technology in response to distortions, distortions need to be less elastic with respect to the idiosyncratic component of productivity to match the observed elasticity in the data.

[35] The relationships used are $SD_i = 0.49 + 0.067 \log Y_{US}/Y_i$ and $ELAS_i = 0.13 + 0.160 \log Y_{US}/Y_i$.

developed European economies (deciles 8 to 10) and significantly larger values (0.80 to 1.05) for transitioning European economies (deciles 4 to 6). Overall, these trends are consistent with the targets in Table 3.

The comparison of $ELAS_i$ is not as straightforward as it is not a commonly reported statistic in micro-studies.[36] Bento and Restuccia (2016) provide estimates of $ELAS_i$ from establishment-level data on a large set of 134 countries. They find a substantial negative trend in $ELAS_i$, ranging from 0.09 for the United States to around 0.70 for the most severely distorted countries. This matches the targets in Table 3 closely.

The aggregate barrier π_i is chosen after (σ_i, ρ_i) such that the initial adoption lag (10% of output) matches the country component of lags reported by Comin and Hobijn (2010). I estimate the lags based on a linear relationship between log income per capita and their reported adoption lags (see Figure 6 in Comin & Hobijn, 2010).[37] The adoption lags also serve as a benchmark to compare the magnitude of the reduction in the adoption lag in the second counterfactual experiment.

Table 4 summarizes the parameter values for each economy i. The calibration is done both with and without distortions to highlight the effect of distortions on aggregate barriers π_i. The partial calibration is similar to a Parente and Prescott (1994) type model with all technology differences attributed to differences in π.

For the lower deciles, the aggregate barriers in the partial calibration are substantially larger than in the full calibration. The difference between the calibrations shows that the decrease in aggregate barriers is heterogeneous across countries. For developing countries, distortions reduce the necessary aggregate barriers by a significant amount while for developed countries the effect is relatively small. Overall, the table shows that distortions are an important driver of technology variation and help explain a substantial portion of the

[36] Chen and Irarrazabal (2014) report values of $ELAS_i$ for Chile to be 0.91 in 1983 and 0.88 in 1996. Their reported values of SD_i are 0.97 and 0.86. For the agricultural sector, Adamopoulos, Brandt, Restuccia, and Leight (2016) report a value of $ELAS_i$ of 0.88 for China. I take the targets in Table 3 as a more conservative estimate of the values of $ELAS_i$.

[37] The regression that the adoption lags are based on is $L_i = 14.73(3.05) - 8.48(1.88) \ln Y_i/Y_{US} + \varepsilon$, where the standard errors are given in parenthesis.

Table 4: Cross-Country Parameters

	Partial Calibration (Without Distortions)	Full Calibration (With Distortions)		
Decile	π_i/π_{BE}	π_i/π_{BE}	ρ_i	σ_i
10	12	14.3	0.076	0.47
9	18	11.1	0.099	0.47
8	25	9.1	0.121	0.47
7	42	7.9	0.151	0.47
6	69	7.9	0.177	0.46
5	107	9.3	0.200	0.45
4	167	8.3	0.222	0.44
3	325	17.6	0.254	0.43
2	760	32.8	0.292	0.40
1	2492	87.9	0.340	0.36

Notes: Targets are given in Table 3. π is chosen to match the initial adoption lag, denoted by Lag_i. ρ is chosen to match the elasticity of distortion with respect to idiosyncratic productivity, denoted by $ELAS_i$. σ is chosen to match the standard deviation of total factor revenue productivity, denoted by SD_i.

residual barriers π. This compliments other studies (such as Parente and Prescott (1994)) by exploring factors that cause technology barriers.

Before moving to the counterfactual experiments, it is useful to discuss the trends of ρ and σ in relation to the institutional differences between developed and developing countries. Despite the decrease in SD_i with income, σ_i is increasing in income. On the other hand, the elasticity ρ_i is decreasing in income which is expected as it maps directly from $ELAS_i$. Other than distortions being more severe in developing countries, Table 4 is indicative of more directed reallocation from high productivity (high a) firms to low productivity (low a) firms in poorer countries. This is consistent with cross-country evidence showing that large firms tend to be more constrained in developing countries.[38] The calibration may also be thought of in terms of more specific institutions that tend to reallocate resources away from more productive firms.[39]

[38]See, for example Hsieh and Olken (2014) who find that large manufacturing firms tend to be more constrained in India, Indonesia and Mexico. Similarly, Adamopoulos and Restuccia (2014) show that the average farm size in developing economies is substantially smaller than developed countries.

[39]Banerjee and Duflo (2005) survey the micro evidence on specific institutions that may lead to misallocation. Restuccia and Rogerson (2013) survey the macro literature on specific institutions that lead to

5.2 Counterfactual Productivities

The first exercise measures the aggregate and relative importance of distortions and aggregate barriers to productivity. The exercise involves measuring the gap in productivity between the benchmark economy and the economies under the counterfactual aggregate barriers π and idiosyncratic distortions (σ, ρ). This gap is interpreted as the difference in productivities attributable to institutional differences between the United States and the representative economies.[40]

Table 5 summarizes the results of the experiment. Column 1 reports the total gap in productivity from changing both aggregate barriers and idiosyncratic distortions. Columns 2 and 3 decompose the total gap into changes attributable to changes in the aggregate barriers and idiosyncratic distortions.

Column 2 summarizes the gap in productivity from changing the aggregate barrier π. Increasing π results in all firms in the economy using less productive technologies. This can clearly be seen by considering the technology index $\varphi(\chi\tilde{a}, 1) = \eta\chi\tilde{a}$. Recall that η is a common measure of productivity to all firms and that it is decreasing in π ($\frac{\partial\eta}{\partial\pi} < 0$). In this regard, π acts as a residual measure of technology and Column 2 captures the productivity gap from all other possible channels that would lead to technology differences.

Column 3 summarizes the gap in productivity from changing distortions (σ, ρ). The change in productivity results from labour and technology being misallocated relative the benchmark economy. From the perspective of the technology index, this gap is captured by a gap in the allocative efficiency term χ between the benchmark and counterfactual economies. The values reported in column 3 are similar in magnitude to the potential gains found in micro-studies of misallocation (for example, Hsieh & Klenow, 2009; Bartelsman et

misallocation (what they refer to as the "Direct Approach").

[40] It would be more accurate to list the results as the loss in productivity that occurs from applying the counterfactual distortions and barriers to the benchmark economy. However, it is common in the literature to list this value as a productivity gain. Hence, I adopt this convention for the sake of comparability and for ease of exposition in the decomposition listed in Table 6.

Table 5: Productivity Counterfactuals

	Productivity Gap $((Y_{BE}/Y_i - 1) \times 100)$		
Decile	(1) Total	(2) Aggregate Barriers	(3) Idiosyncratic Distortions
10	33	35	-1
9	39	32	5
8	45	28	12
7	56	27	23
6	72	27	35
5	85	28	45
4	96	27	54
3	132	39	67
2	163	49	79
1	213	67	89

Notes: Productivity gap refers to the difference in productivity between the benchmark economy and the counterfactual economy: $(Y_{BE}/Y_i - 1) \times 100$. *Column 1 is the total gap from changing both aggregate barriers and idiosyncratic distortions* $(\pi, \rho, \sigma) = (\pi_i, \rho_i, \sigma_i)$. *Column 2 is the gap from changing only aggregate barriers:* $(\pi, \sigma, \rho) = (\pi_i, \sigma_{BE}, \rho_{BE})$. *Column 3 is the gap from changing only the aggregate barriers:* $(\pi, \sigma, \rho) = (\pi_{BE}, \sigma_i, \rho_i)$.

al., 2013; Chen & Irarrazabal, 2014).[41] The gap is substantial and shows the importance of idiosyncratic distortions, with the gap being as high as 89% in the bottom decile economy.

Comparing column 2 to column 3 shows that the importance of distortions relative to aggregate barriers varies across the income distribution. In richer countries the role of aggregate barriers is more important distortions. On the other hand, both channels are important in poorer countries, but the role of distortions is larger than aggregate barriers. To give a comparison, consider the amount of the total gap in column 1 that is accounted for by changing only distortions (Column 3).[42] For the top decile economy this value is negative and for the ninth decile economy removing distortions accounts for only 15% of the total gain ($15\% = \frac{\log(1+5\%)}{\log(1+39\%)}$). This value increases to 55% for the sixth decile economy

[41] A caveat with this comparison is that the distribution of underlying productivities is fixed and the same as the benchmark economy for the results in Table 5. The underlying productivities distribution of firms in the countries studied in the cited papers are unlikely to be identical to the United States.

[42] As is apparent in the expression for aggregate output given in (14), there is no interaction between the level of aggregate barriers and the effect of distortions. Hence, the gap is the just the multiplication of the two separate effects. For example, $213\% = (1 + 67\%)(1 + 89\%) - 1$ for the bottom decile economy.

($55\% = \frac{\log(1+35\%)}{\log(1+72\%)}$) and is relatively stable for the lower decile economies, with a value of 56% for the bottom decile economy.

Decomposition: I now consider a further decomposition of the productivity gap from idiosyncratic distortions into a static and a dynamic channel. The static channel captures the productivity gap from only reallocating labour when firm-level technologies are held constant. In this regard, the static channel is similar to the gains reported in empirical studies of misallocation that consider the reallocation of factors (labour, capital) across firms with fixed TFPQs. The dynamic channel captures the additional productivity gap from allowing firms to also adjust their technology in response to distortions. This channel allows for TFPQ to change in response to distortions and is more closely related to the recent literature on dynamic misallocation.

The results from the decomposition are listed in Table 6. For the decomposition, I take the counterfactual economies generated above as data and consider the gains in productivity from moving to the benchmark in two steps. The productivity gain from changing idiosyncratic distortions listed in Table 6 corresponds to Column 2 of Table 5 and is the total gain that is considered in this decomposition. The first step (Static Channel) involves changing distortions to the benchmark economy and holding firm-level technology fixed. Firms are only able to adjust labour in response to the new distortions. The second step (Dynamic Channel) involves allowing firms to adjust both technology and to further adjust labour to the optimal level, given the benchmark distortions. For both the static and dynamic channels, I list the gain in productivity from that step as well as the contribution of that step to the overall productivity gain. The contribution is a normalized measure of the two channels that captures their relative importance.

Comparing the columns in Table 6 shows that in general the two channels are both important components of productivity. For the poorest economies, the contribution of the dynamic misallocation channel is substantial and larger than the static channel. This suggests larger

Table 6: Decomposition of Idiosyncratic Distortions

Decile	Idiosyncratic Distortions	Static Channel		Dynamic Channel	
	% Gain	% Gain	Contribution	% Gain	Contribution
10	-1	-1	56	0	44
9	5	3	56	2	44
8	12	7	56	5	44
7	23	12	55	10	45
6	35	17	53	15	47
5	45	21	51	20	49
4	54	24	49	25	51
3	67	26	46	32	54
2	79	28	42	40	58
1	89	26	37	50	63

Notes: Contribution is calculated as Contribution of Channel X $= \frac{\log(1+\textit{\% Change from Channel X})}{\log(1+\textit{\% Total Change})}$. *The static misallocation channel is the gain from reallocating labour in the distorted economy, holding firm-level technology (equivalently, TFPQ) fixed. The dynamic misallocation channel is the gain from allowing firms to re-optimize both labour and technology.*

gains from alleviating misallocation than what is reported in empirical studies.

The takeaway from Tables 5 and 6 is that the interaction between idiosyncratic distortions and technology is an important consideration for aggregate productivity. This complements the calibration (Table 4) by showing that not only can idiosyncratic distortions account for a large portion of the aggregate barriers π, but they are also an important channel for aggregate productivity. The results also show that policies that target the level of technology without accounting for the distribution of technology may not be as effective as suggested by cross-country differences in technology. On the other hand, policies that alleviate misallocation may be additionally beneficial because they improve the contribution of technology to productivity.

5.3 Counterfactual Adoption Time

The second exercise quantifies the extent to which distortions explain cross-country technology differences. Rather than focus on differences in technology level, the exercise focuses on differences in the adoption lags of new technologies. The benefit of focusing on adoption lags is that it allows the prediction to be compared against empirical estimates of adoption lags and different lags contain information on the distribution of technology across firms.

The exercise consists of three steps. First, a representative economy with distortions $(\pi_i, \sigma_i, \rho_i)$ is simulated and the adoption lags are recorded at the initial (10% of output produced with the technology), median (50%) and final (90%) levels (Panel A in Table 7). The years are recorded relative to the benchmark economy, such that the first entry in Table 7 indicates that the top decile economy initially adopts technologies 14 years after the benchmark economy. Recall that the aggregate barriers π are chosen such that the initial adoption lag in Table 7 matches the adoption lags found by Comin and Hobijn (2010). Second, distortions in the economy (σ_i, ρ_i) are changed to the benchmark economy distortions (σ_{BE}, ρ_{BE}) and the new adoption lags are recorded. Third, the difference between the two adoption lags are calculated and recorded in Panel B of Table 7. This change is the hypothetical reduction that would occur in these countries from moving to the same institutions as the United States.

Qualitatively, the results in Table 7 are explained by the trends discussed in the initial examination of the effect of distortions on the adoption pattern (Section 4.2). The large reduction in the initial adoption lag in Panel B relative to Panel A is caused by a combination of decreasing ρ and increasing σ when moving to the BE distortions (see Table 4). Recall that both a decrease in ρ and an increase in σ are expected to lower the initial adoption lag. On the other hand, the reduction in the final adoption lag is much smaller than the reduction in the initial adoption lag. This is a consequence of the increase in σ and the decrease in ρ having opposing affects on the final adoption lag (see Figure 3).

Distortions account for about 13 to 14 years of the initial adoption lag in the median economy. This ranges from having no effect on the top decile economy to decreasing the adoption lag by almost half ($43\% \approx 19/44$) in the bottom decile economy. It is also reassuring to note that the reduction in the initial adoption lag is large across the lower half of the

Table 7: Counterfactual Adoption Pattern

	(A) Counterfactual Adoption Lag (Years)			(B) Change in Distortions (Change in Years)		
Decile	Initial	Median	Final	Initial	Median	Final
10	14	14	15	1	1	0
9	16	15	14	-3	-2	0
8	18	16	14	-6	-4	-1
7	21	18	15	-10	-7	-3
6	24	20	16	-13	-9	-4
5	26	22	17	-14	-10	-4
4	29	21	17	-17	-9	-5
3	32	28	22	-16	-12	-6
2	37	32	26	-18	-13	-6
1	44	38	32	-19	-13	-7

Notes: Adoption pattern is summarized by the initial (10% of output produced with the technology), median (50%) and final (90%) adoption lags. Panel A records the number of years the counterfactual economy with parameters $(\pi_i, \sigma_i, \rho_i)$ *lags the benchmark economy in the adoption of new technologies. Panel B summarizes the reduction in adoption lags that occurs when idiosyncratic distortions are the same as in the benchmark economy:* $(\pi, \sigma, \rho) = (\pi_i, \sigma_{BE}, \rho_{BE})$.

income distribution. This provides some evidence on the robustness of the numbers in the bottom decile economy to changes in σ or ρ.

Distortions account for a smaller fraction of the final adoption lag when compared to the initial adoption lag. In the bottom decile economy for example, distortions account for only around 22% (22% $\approx$ 7/32) of the final adoption lag compared to 43% for the initial adoption lag. A caveat regarding the predictions for the median and final adoption lags is that technology spillovers are likely important determinant of these lags. The concern is that the adoption experience of early adopters influences the adoption decisions of late adopters. For example, learning-by-doing spillovers would imply that late adopter benefit from the experiences of early adopters with the technologies. However, this caveat does not apply to the initial adoption lag as early adopters do not benefit from these types of spillovers. With this in mind, I take the median and final adoption lags as being informative about qualitative changes in the adoption pattern, but not magnitudes.

Together the results of this section show that technology differences between countries may be symptoms of underlying institutional problems. Table 7 shows that there are substantial technology differences between countries and that distortions account for a large portion of these differences. Tables 4 and 5 show similar patterns. This implies that policies that target technological catch-up but neglect misallocation can be less effective than would be expected from just examining the technology gap. This is because distortions act as a wedge in technology efficiency that prevents distorted economies from catching up in terms of productivity. The conclusion is that the distribution of technology is an important factor to consider in the study of technology differences.

6 Conclusion

The interactions between institutional distortions and technology is an important channel for explaining cross-country differences in productivity and technology. Institutions that are characterized by a systematic reallocation of resources from high productivity firms to low productivity firms significantly delay the adoption of modern technologies. These are the specific types of institutions that are common in developing countries. In the cross-country analysis, distortions account for as much as half of the observed differences in adoption lags between countries. This suggests that improving institutions associated with misallocation would also improve technologies used in developing countries. Further, the presence of endogenous productivity choice through technology suggests that the productivity gains from reallocating resources are larger than previously thought.

My analysis has abstracted from some of the dynamics of technology adoption. Notably, a richer model including technology spillovers may magnify the effect of distortions by emphasizing the importance of early adoption. Specifically, distortions that delay initial adoption also delay learning from early adopters and make later adoption more costly. Another path would be to use firm-level data that includes technology use to get a fuller picture of the interaction between technology and the allocation of resources. I leave an exploration of these paths for future research.

References

Abramovitz, M. (1986). Catching Up, Forging Ahead, and Falling Behind. *Journal of Economic History*, *46*(2), 385–406.

Adamopoulos, T., Brandt, L., Restuccia, D., & Leight, J. (2016). Misallocation , Selection and Productivity : A Quantitative Analysis with Panel Data from China. *Working Paper*, 1–51.

Adamopoulos, T., & Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, *104*(6), 1667–1697.

Atkeson, A., & Burstein, A. (2014). Aggregate Implications of Innovation Policy. *Working Paper*, 1–59.

Atkin, D., Chaudhry, A., Chaudhry, S., Khandelwal, A., & Verhoogen, E. (2014). Organizational Barriers to Technology Adoption : Evidence from Soccer-Ball Producers in Pakistan. *Working Paper*.

Banerjee, A. V., & Duflo, E. (2005). Growth Theory Through the Lens of Development Economics. In *Handbook of economic growth.*

Bartelsman, B. E., Haltiwanger, J., & Scarpetta, S. (2013). Cross-Country Differences in Productivity : The Role of Allocation and Selection. *American Economic Review*, *103*(1), 305–334.

Bento, P., & Restuccia, D. (2016). Misallocation , Establishment Size , and Productivity. *University of Toronto Working Paper*.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, *128*(1), 1–51.

Bollard, A., Klenow, P. J., & Li, H. (2016). Entry Costs Rise with Development. *Working Paper*, 1–46.

Buera, F. J., Kaboski, J. P., & Shin, Y. (2011). Finance and development: A tale of two sectors. *American Economic Review*, *101*(5), 1964–2002.

Chen, K., & Irarrazabal, A. (2014). The role of allocative efficiency in a decade of recovery. *Review of Economic Dynamics*.

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

Comin, D., & Hobijn, B. (2009). The CHAT Dataset. *NBER Working Paper Series No. 15319*.

Comin, D., & Hobijn, B. (2010). An Exploration of Technology Diffusion. *American Economic Review*, *100*(December), 2031–2059.

Comin, D., Hobijn, B., & Rovito, E. (2008). Technology usage lags. *Journal of Economic Growth*, *13*(November), 237–256.

Comin, D., & Mestieri, M. (2014). If Technology has Arrived Everywhere , why Has Income Diverged ? *Working Paper*.

Conley, T. G., & Udry, C. R. (2008). Learning About a New Technology : Pineapple in Ghana. *American Economic Review*.

David, P. (1966). The mechanization of reaping in the Ante-Bellum Midwest. In *H. rosovsky, ed., industrialization in two systems (harvard university press, cambridge, ma)* (pp. 3–39).

Gabler, A., & Poschke, M. (2013). Experimentation by firms , distortions , and aggregate productivity. *Review of Economic Dynamics*, *16*(1), 26–38.

Gerschenkron, A. (1965). *Economic Backwardness in Historical Perspective: A Book of Essays*.

Gort, M., & Klepper, S. (1982). Time paths in the diffusion of product innovations. *Economic Journal*, *92*(367), 630–653.

Greenwood, J. (1999). The Third Industrial Revolution : Technology , Productivity , and Income Inequality. *Economic Review*, 2–12.

Greenwood, J., Hercowitz, Z., & Krusell, P. (1997). Long-Run Implications of Investment-Specific Technological Change. *American Economic Review*, *87*(3), 342–362.

Griliches, Z. (1957). Hybrid Corn: An Exploration in the Economics of Technological Change. *Econometrica*, *25*(4), 501–522.

Grubler, A. (1991). Diffusion: Long Term Patterns and Discontinuities. *Technological Forecasting and Social Change*, *39*, 159–180.

Guner, N., Parkhomenko, A., & Ventura, G. (2015). Managers and Productivity Differences. *Working Paper*(August).

Guner, N., Ventura, G., & Xu, Y. (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, *11*(4), 721–744.

Hopenhayn, H. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica*, *60*(5), 1127–1150.

Hopenhayn, H. (2014). On The Measure of Distortions. *Working Paper*.

Hopenhayn, H., & Rogerson, R. (1993). Job Turnover and Policy Evaluation : A General Equilibrium Analysis. *Journal of Political Economy*, *101*(5), 915–938.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, *124*(4), 1403–1448.

Hsieh, C.-T., & Klenow, P. J. (2014). The Life Cycle of Plants in India and Mexico. *The Quarterly Journal of Economics*, *129*(3), 1035–1084.

Hsieh, C.-T., & Olken, B. A. (2014). The Missing "Missing Middle". *Journal of Economic Perspectives*, *28*(3), 89–108.

Jovanovic, B., & Lach, S. (1997). Product Innovation and the Business Cycle. *International Economic Review*, *38*(1), 3–22.

Levin, S. G., Levin, S. L., & Meisel, J. B. (1987). A Dynamic Analysis of the Adoption of a New Technology: The Case of Optical Scanners. *Review of Economics and Statistics*, *69*, 12–17.

Manuelli, R. E., & Seshadri, A. (2014). Frictionless Technology Diffusion : The Case of Tractors. *American Economic Review*.

Midrigan, V., & Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, *104*(2), 422–458.

Oster, S. (1982). The Diffusion of Innovation among Steel Firms : The Basic Oxygen Furnace. *The Bell Journal of Economics*, *13*(1), 45–56.

Parente, S. L., & Prescott, E. C. (1994). Barriers to Technology Adoption and Development. *Journal of Political Economy*, *102*(2), 298–321.

Restuccia, D., & Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics*, *11*, 707–720.

Restuccia, D., & Rogerson, R. (2013). Misallocation and productivity. *Review of Economic Dynamics*, *16*(1), 1–10.

Stokey, N. L. (2014). Catching Up and Falling Behind. *Journal of Economic Growth*, *20*, 1–36.

Suri, T. (2011). Selection and Comparative Advantage in Technology Adoption. *Econometrica*, *79*(1), 159–209.

Yorukoglu, M. (1998). The Information Technology Productivity Paradox. *Review of Economic Dynamics*, *1*, 551–592.

A Equilibrium Proof

Proof of Proposition 1: Note that from the definition in the main text, the following expressions hold:

$$\varphi_t(a,\theta) = \zeta_\varphi \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \theta^{\frac{1}{1-\omega}} a, \quad n_t(a,\theta) = \zeta_n \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \theta^{\frac{1}{1-\omega}} a,$$
$$y_t(a,\theta) = w_t \zeta_y \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\gamma}\frac{1}{1-\omega}} \theta^{\frac{1}{1-\omega}} a.$$

Define $\tilde{v}_t = \mathbb{E} v_{(a,\theta),t}(0)$ and note that this implies that the free-entry condition can be written as $w_t c_e = \tilde{v}_t$. By definition $\tilde{v}_t$ can be written as

$$\tilde{v}_t = \int_{\mathcal{A}\times\Theta} \Big(y_t(a,\theta) - w_t \pi \varphi_t(a,\theta) - w_t n_t(a,\theta) + D v_{(a,\theta),t+1}(\varphi_t(a,\theta))\Big)\, g(a,\theta)\,\mathrm{d}a\,\mathrm{d}\theta \qquad \text{(A)}$$

Integrating over the expected next period value of a firm can be written as

$$\int_{\mathcal{A}\times\Theta} v_{(a,\theta),t+1}(\varphi_t(a,\theta)) g(a,\theta)\,\mathrm{d}a\,\mathrm{d}\theta = \tilde{v}_{t+1} + w_{t+1}\pi(1-\delta)\zeta_\varphi \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \tilde{a} \qquad \text{(B)}$$

Intuitively, the right-hand side of (B) is equal to the expected value of an entrant plus the expected present value of past investment in technology by the firm. This is a result of the distribution of (a,θ) being fixed. Substituting (B) into (A) and using the fact that $\tilde{v}_t = w_t c_e$ in all periods t implies that

$$\begin{aligned} w_t c_e =& w_t \zeta_y \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\gamma}\frac{1}{1-\omega}} \tilde{a} - w_t \zeta_n \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \tilde{a} - w_t \pi \zeta_\varphi \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \tilde{a} \\ &+ D(1+g_w) w_t \left[c_e + \pi(1-\delta)\zeta_\varphi \left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \tilde{a}\right] \end{aligned}$$

The above expression immediately implies that $g_w = g$ and can be rearranged to solve for the equilibrium wage rate:

$$\ln w = \zeta_w + \ln \bar{z} - \omega(1-\gamma)\ln \pi + (1-\omega)(1-\gamma)\ln \tilde{a}$$

where $\zeta_w = (1-\omega)(1-\gamma)\ln\left\{\frac{1-\omega}{1-D}[\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma)]^{\frac{1}{1-\omega}}\left[\frac{\omega}{1-D(1+g_v)(1-\delta)}\right]^{\frac{\omega}{1-\omega}}\right\}$. The growth rate of the wage rate is $g_w = \ln w_{t+1} - \ln w_t = \ln \bar{z}_{t+1} - \ln \bar{z}_t = g$.

Labour Clearing: The labour clearing condition is used to solve for the equilibrium mass of firms,

$$1 = N_w + N_X + N_E$$
$$\ln M = -\ln\left[\exp(\zeta_1) + (1+\lambda(1-\delta))\exp(\zeta_2) + \lambda c_e\right].$$

where ζ_1 and ζ_2 are collections of constants. The mass of firms does not grow ($g_M = 0$).

Aggregate Output and Growth: Aggregate output is given by

$$Y = M\int y(a,\theta)$$
$$\ln Y = \zeta_Y + \ln \bar{z} - \omega(1-\gamma)\ln \pi + (1-\omega)(1-\gamma)\ln \kappa\tilde{a} + \ln \chi.$$

The growth rate is given by $g_Y = \ln Y_{t+1} - \ln Y_t = \ln \bar{z}_{t+1} - \ln \bar{z}_t = g$.

Corollary: From the proof of proposition 1, $\varphi_t(a,\theta) = \zeta_\varphi\left(\frac{\bar{z}_t}{w_t}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}}\theta^{\frac{1}{1-\omega}}a$ and $g_w = g$ implying that $\varphi_t(a,\theta) = \varphi_s(a,\theta)$ for any periods t and s.

B Sensitivity Analysis

In this section, I consider the sensitivity of the results with respect to different values for the curvature on the cost function ξ. By definition ω is related to the value of ξ by $\omega = \frac{1}{\xi}\frac{1}{1-\gamma}$. Additionally, recall that the 10-90 lag is given by 10-90 Lag $= \frac{\ln n_{90}/n_{10}}{g_Y}\frac{1}{\xi}$. The choice of ξ is equivalent to the choice of ω and the choice of the 10-90 lag. I use these two values for the alternative calibrations. The first alternative is to set $\omega = 0.55$ which corresponds to the upper end of values suggested by Parente and Prescott (1994). The second alternative is to set 10-90 Lag = 15 which corresponds to the average value in the sample of technologies used by Jovanovic and Lach (1997). These targets imply values of ξ equal to 5.5 and 25, which is roughly half and double the value used in the main text. In addition, I report the results when $\xi \in \{7.87, 9.61\}$ which are roughly 10% below and above the value of ξ used in the main text.

Table 8 summarizes the main predictions of the model under the alternative parameterizations. The targets are in bold. For brevity, I only show the results for the bottom decile economy that correspond to changes in distortions. Columns 4 and 5 correspond to Table 5 and 7.

Table 8: Results under Alternative Parameterizations

			Counterfactual		
ξ (1)	ω (2)	10-90 Lag (3)	Productivity Gap (4)	Adoption Lag (5)	Notes
5.5	**0.55**	54	127%	-21	Parente and Prescott (1994)
7.87	0.38	47	92%	-21	10% below benchmark ξ
8.74	0.34	**40**	89%	-19	Benchmark Calibration
9.61	0.31	36	85%	-18	10% above benchmark ξ
20.4	0.15	**15**	72%	-10	Jovanovic and Lach (1997)

Notes: Targets are in bold, otherwise the calibration is the same as in the main text. ξ is the curvature on the technology cost function. ω is the decreasing returns to a firm's proximity to the technology frontier. 10-90 Lag is the number of years it takes between 10% of output being produced with a technology to 90% of output. (4) is the predicted gap in productivity between the benchmark economy and the bottom decile distortions $(Y_{BE}/Y_1 - 1) \times 100$. (5) is the change in the initial adoption lag (10% of output) from moving the benchmark economy to the bottom decile distortions.

Table 8 shows that under $\omega = 0.55$ the 10-90 Lag is 54 years, which is 14 years longer than the target in the main text. Additionally, this value is substantially larger than the values recorded in the literature. The results under this parameterization imply a larger role for distortions. Both the productivity gap attributed to distortions and the reduction in the adoption lag are larger. Qualitatively, the results are very similar to the main text.

Under the Jovanovic and Lach (1997) alternative parameterization, the results are consistent with the main text. The increase in productivity (column 4) is smaller by 17 percentage points compared to the value in Table 5. On the other hand, the reduction in the initial adoption lag falls by around half compared to the -19 years in Table 7. Comin and Hobijn (2010) note that many modern technologies (for example, computers) have shorter adoption lags than older technologies. If these technologies also have shorter 10-90 lags, then the results in Table 8 are not necessarily suggestive of distortions being less important.