Distorted Technology Adoption

Stephen Ayerst*

January 10, 2023

Abstract

To what extent do firm-level institutional distortions drive cross-country technology differences? I develop a quantitative model of technology adoption in which heterogeneous firms adopt technology depending on their underlying productivity and institutional environment. The institutional environment is represented by idiosyncratic firm-level wedges on revenues that distort firm decisions. The model is calibrated to match the US firm employment distribution and the average adoption length of new technologies. In less developed countries, measured institutional distortions target highly productive firms delaying the adoption of new technologies as these firms are (otherwise) early adopters. Quantitatively, measured institutional distortions explain around half of the observed cross-country variation in adoption lags. Additionally, distorted technology adoption over doubles the aggregate productivity loss compared with static misallocation alone.

Keywords: Productivity, Misallocation, Technology Diffusion.
JEL classification: O11, O14, 033, O41, O43.

---

*International Monetary Fund, 700 19th Street NW, Washington DC 20431. Email: stephen.b.ayerst@gmail.com. A previous version of this paper circulated under the title “Policy Distortions and Technology Adoption.” I thank Diego Restuccia for his advice and support. This paper has benefited from discussions with Rami Abou-Seido, Michelle Alexopoulos, Chaoran Chen, Murat Celik, Daniel Chippin, Diego Comin, Miguel Faria-e-Castro, Ashique Habib, Ridwan Karim, Peter Klenow, Soomin Lee, Nick Li, Gaelan MacKenzie, 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). The views expressed herein are those of the author and should not be attributed to the IMF, its Executive Board, or its management

1 Introduction

Despite the globalization of knowledge, large gaps in the adoption and use of technologies across countries persist.[1] Technologies used by firms in developing countries often lag several decades behind their counterparts in advanced economies, despite the large potential gains from adoption. What barriers prevent technologies from converging? Banerjee and Duflo (2005) note, for the case of India, that while many firms are aware of modern technologies, they lack sufficient scale to profitably adopt these technologies. The focus on the lack of scale coincides with a large literature, starting with Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), documenting that productive firms in developing countries often face challenges accumulating resources as institutional distortions prevent resources from being reallocated across producers.[2]

I examine the extent to which measured differences in institutional distortions act as a barrier to the adoption of new technologies in developing countries. Using a quantitative model of technology adoption by heterogeneous firms, I find that measured institutional distortions explain around half of the observed differences in cross-country adoption lags. Additionally, distorted technology use across firms over doubles the aggregate productivity loss compared to factor misallocation alone.

My first contribution is to develop a novel model of technology adoption by heterogeneous firms that face idiosyncratic institutional distortions. Following Hopenhayn (1992), firm heterogeneity arises from underlying differences in firm capabilities whereas institutional distortions are captured by an idiosyncratic wedge on firm revenues, as in Restuccia and Rogerson (2008). The idiosyncratic wedges act as a stand-in for institutional factors that cause a reallocation of resources across firms, such as firing costs, credit market constraints, or weak property rights (see, for example, Hopenhayn and Rogerson, 1993; Buera et al., 2011; Akcigit et al., 2021). Firms adopt costly technology, which determines the firm's productivity along with the firm's underlying capability. Building on Parente and Prescott (1994), the cost of adopting new technologies depends on the worldwide technology frontier and an aggregate barrier to adoption, which shifts the cost of adopting new technologies for all firms.

[1] See empirical evidence from Comin and Hobijn (2010) and Comin and Mestieri (2018).

[2] Others show that these same institutional distortions may be important for explaining other cross-country differences, such as firm life-cycles (Hsieh and Klenow, 2014) and average firm size (Bento and Restuccia, 2017). Other studies of technology find similar between-firm differences in technology use (Griliches, 1957; Suri, 2011; Chen, 2020).

The main tradeoff faced by firms is that newer, more productive technologies are also more expensive compared to older technologies. This leads to more capable firms and firms that benefit more from institutional distortions adopting technologies earlier. Consequently, the equilibrium features a non-degenerate distribution of technologies and employment in which larger firms (by employment) use newer technologies. The distribution of technologies implies a gradual adoption of new technologies that is used to discipline technology parameters in the quantitative analysis.

Cross-country adoption lags depend on both the difference in institutional distortions and the aggregate barrier to adoption. Institutional distortions delay the adoption of new technologies to the extent that they delay early adopters. This compresses the distribution of technologies used across firms leading to a shorter adoption period. Aggregate barriers affect the cost of adoption equally for all firms leading to a shift in the distribution of technologies used across firms. Given this, aggregate barriers are taken as a residual measure of technology use in the calibration.

The benchmark economy is calibrated to match United States' data. Following Hsieh and Klenow (2009), I discipline the distribution of firm-level productivities from the employment distribution of firms and estimates of institutional distortions. However, technology parameters are difficult to estimate in practice because detailed firm-level data on technology use is, generally, unavailable. My second contribution is to provide a novel solution to this problem by showing that aggregate time-series data on the adoption of new technologies can be mapped into the cross-sectional distribution of firm-level technologies. Technology parameters can then be disciplined using the adoption length of new technologies. Intuitively, the length of adoption informs the change in the relative costs and benefits of new technologies over time.

My third contribution is to use the model to quantify the extent that institutional distortions account for cross-country differences in technology adoption lags and productivity. In the data, I measure cross-country technology differences using technology adoption lags between the US and other countries. Technology adoption lags are used to calibrate aggregate barriers to adoption—the residual measure of technology in the model—and to provide a quantitative context for the results. I take technology adoption lags from Comin and Hobijn (2010)'s estimates of country-specific technology adoption lags constructed using a broad set of technologies (e.g., railways, cars, PCs). I follow Bento and Restuccia (2017) by measuring cross-country distortions directly using firm-level data for a range of countries. I parameterize cross-country differences using the relative correlation of measured institutional distortions to firm-level productivity, which tends to be higher in lower income countries, since this is the primary factor influencing the adoption lag of technologies in the model.

Moving the correlation of institutional distortions from the value in the US benchmark economy to the value in the lowest income countries delays technology adoption and decreases aggregate productivity. The counterfactual economy adopts technologies 16 years later than in the benchmark economy, around one-third (= 16/48) of the observed adoption lag in the data. While the impact is largest on early adopters, institutional distortions cause a sizable delay throughout the firm distribution. The median firm in the counterfactual economy uses technologies 11 years older than in the benchmark economy. The delays in technology adoption also magnifies the productivity cost of institutional distortions relative to factor misallocation alone. The benchmark economy is around 123% more productive than the counterfactual economy, explaining around 20% of the empirical productivity gap. Around two thirds of the productivity gap between the benchmark and counterfactual economies is explained by differences in technology adoption.

My paper builds on the literature examining drivers of cross-country technology differences. Parente and Prescott (1994) and Comin and Hobijn (2010) model technology differences as arising from country-specific aggregate barriers to technology adoption. I develop a novel quantitative framework in which institutional distortions are a micro-foundation for barriers to technology adoption. This channel is quantitatively important explaining around half of the variation in cross-country technology adoption across countries and around one-third of the gap in technology adoption lags between the lowest income countries and the US. Firm-level distortions provide a path for future work by connecting aggregate models of technology adoption to production heterogeneity and micro-level data.

My paper also builds on a series of papers that examine the dynamic effects of misallocation (Gabler and Poschke, 2013; Midrigan and Xu, 2014; Hsieh and Klenow, 2014; Bento and Restuccia, 2017; Guner et al., 2018). This literature finds that dynamic firm decisions amplify the effects of misallocation through the firm-level productivity distribution. I similarly find that institutional distortions compress the firm-level productivity distribution, resulting in lower aggregate productivity and magnifying the costs of misallocation. An upshot of the analysis is that this compression is stronger in more severely distorted economies implying that the magnification is increasing in the severity of distortions. Quantitatively, I find that the aggregate productivity cost for the lowest income countries is around 123% (around 20% of the observed gap in output-per-worker) with around two-thirds driven by distorted technology.

The remainder of this paper is organized as follows: Section 2 describes the model and its properties. Section 3 calibrates the model and discussed the fit with data. Section 4 examines the role of distortions for cross-country productivity and technology differences. Section 5 concludes.

2 Model

I develop a model of technology adoption decisions (as in Parente and Prescott, 1994) by heterogeneous firms that face institutional distortions (as in Restuccia and Rogerson, 2008). I show that the model implies a non-degenerate distribution of technologies used by firms and that this distribution implies an aggregate adoption curve for technology adoption. In the next sections, I use this model as a setting to quantify the cost of institutional distortions to cross-country aggregate productivity and technology adoption.

2.1 Economic Structure

Time is discrete and indexed by t. The economy is populated by a unit mass of infinitely lived households and an endogenous mass M_t of firms.[3] The economy is closed and represents an individual country in the quantitative analysis. Countries differ in their institutional environment, aggregate barriers to the adoption of new technologies, and the average productivity of firms. All countries share a common world technology frontier, which captures the global state of knowledge.

Technology and production. Technology is non-rivalrous and is modeled by an exogenous set of production techniques and the corresponding cost function. Technological progress is exogenous and described by the technology frontier $\bar{z}_t$, which grows at a rate g, such that $\bar{z} = \bar{z}_0 e^{gt}$ in period t. Firms j choose technology $z_{j,t}$ at the

[3] The mass of households is normalized to unity and aggregate values are interpreted as per capita.

beginning of each period t.[4] To adopt technology $z_{j,t}$ a firm j with current technology $z_{j,t-1}$ must hire x units of labour given by:

$$x(z_{j,t}, z_{j,t-1}) = \xi \frac{\pi}{\bar{z}_t} \int_{z_{j,t-1}}^{z_{j,t}} \left(\frac{\hat{z}}{\bar{z}_t}\right)^{\xi-1} \mathrm{d}\hat{z} = \pi \left[\left(\frac{z_{j,t}}{\bar{z}_t}\right)^{\xi} - \left(\frac{z_{j,t-1}}{\bar{z}_t}\right)^{\xi} \right], \tag{1}$$

where $\pi \geq 1$ is a country-specific parameter that captures aggregate barriers to adoption and $\xi > 1$ is the curvature of the cost function.[5] The aggregate barrier π acts as a stand-in for any residual factors that delay technology adoption (e.g. patent laws, trade barriers or human capital).

The functional form assumed for technology adoption costs follows Parente and Prescott (1994) (see also Comin and Hobijn, 2010). The cost function can be interpreted as a continuum of distinct technologies that are ordered by their net benefit to firms. For example, this could capture sequential technologies (e.g. adopting computers before email) or a pecking order of distinct unrelated technologies (e.g. adopting computers before just-in-time manufacturing practices). In this regard, firms do not need to repay for technologies they have already adopted and the cost of adopting newer technologies is independent from the firm's previous adoption. Technological progress through growth in $\bar{z}_t$ then both captures the introduction of new technologies as well as the decline in the cost of existing technologies. Appendix E shows that the quantitative results are identical when next period technology is chosen and the curvature of the cost function is on the change in technology $((z_{j,t+1} - z_{j,t})/\bar{z}_t)^{\tilde{\xi}}$ rather than the level of technology.

Firms differ in a component $s_{j,t}$ of productivity, which I refer to as the firm's ability. The value of $s_{j,t}$ could be thought of as capturing the ability of the manager of firm j. In addition to technology, firms choose to employ production labour $n_{j,t}$ to produce

[4] Changing the technology choice decision's timing does not significantly affect the results. The timing convention chosen is done to align the decisions made by the entering and incumbent firms.

[5] Setting the costs in terms of labour, as opposed to output, is primarily for consistency with the entry cost and does not substantially affect the results. Mechanically, labour costs have two affects on the economy. First, some fraction of the labour will be used for adoption and not used for production. This fraction is constant for all economies along the balanced growth path (BGP) equilibrium and so it does not affect the quantitative results. Second, the wage rate grows at rate g along the BGP causing technology adoption to be more expensive. This is identical to the common assumption in the technology adoption (e.g. Parente and Prescott, 1994) and innovation (e.g. Akcigit and Kerr, 2018) literatures that scale investment costs by productivity (in this case $\bar{z}_t$) in order to achieve stable growth.

output.[6] Output of firm j is given by:

$$y_{j,t} = (z_{j,t}(As_{j,t})^{1-\omega})^{1-\gamma} n_{j,t}^{\gamma}, \tag{2}$$

where $\gamma \in (0,1)$ determines the output share of labour and $\omega = 1/\xi$ is a scaling term for algebraic convenience. Since A and $s_{j,t}$ do not have natural units, scaling this term by $1-\omega$ does not affect the results. The term A is a productivity shifter that affects all firms equally and can be thought of as a residual measure of cross-country productivity differences that are not explicitly modeled.

Entry and exit. At the end of each period, firms face an exogenous probability $\lambda \in [0,1]$ of exiting. A mass $M_{E,t}$ of new firms enter at the beginning of the period by hiring $M_{E,t}c_E$ units of labour. After entering, firms draw an ability $s \in \mathcal{S}$ from distribution $\mu(s)$, which is constant throughout the firm's lifetime. Firms enter with zero technology, $z_{j,t-1} = 0$, and immediately adopt a technology, $z_{j,t}$, through investment.[7]

The mass of firms in the economy in the period t depends on the firms that do not exit in the previous period $(1-\lambda)M_t$ and the entry of new firms $M_{E,t}$. The mass of firms evolves according to:

$$M_{t+1} = (1-\lambda)M_t + M_{E,t}. \tag{3}$$

Households. Household preferences are given by:

$$U(\{C_t\}_{t=0}^{\infty}) = \sum_{t=0}^{\infty} \beta^t \log C_t. \tag{4}$$

Households supply one unit of labour inelastically as production workers ($N_{P,t} = M_t \int_j n_{j,t} dj$), in the establishment of new firms ($N_{E,t} = c_E M_{E,t}$) and in the adoption of new technologies ($N_{X,t} = \int_j x_{j,t} dj$).

[6] The model abstracts from capital as a production input. Extending the model to include capital is equivalent to setting $n = k^{\alpha}\ell^{1-\alpha}$, for capital k and labour ℓ, and does not change the main relationships between institutional distortions and technology adoption. Further, the inclusion of capital does not change the quantitative implications of the model if the results are interpreted as changes in total factor productivity, as opposed to output-per-worker.

[7] This assumption does not affect the results because, as I show below firms choose their optimal technology independently of $z_{j,t-1}$ in equilibrium. This implies that the setting the entry value of $z_{j,t-1}$ is quantitatively equivalent to changing the entry cost through c_E.

2.2 Market Structure and Institutional Distortions

Firms are perfectly competitive and produce a homogeneous good, which is taken as the numeraire. Institutional distortions are modeled as an idiosyncratic tax τ on firm revenues, as in Restuccia and Rogerson (2008), that act as a stand-in for institutions that reallocate factors of production across firms.[8] Some examples of distortions include hiring or firing constraints, credits constraints, or weak property rights (Hopenhayn and Rogerson, 1993; Buera et al., 2011; Akcigit et al., 2021). The government redistributes the tax revenues from the idiosyncratic taxes through a non-distortionary lump-sum transfer T_t to households.[9]

I follow Hopenhayn (2014) and define an institutional distortion as $\theta = (1-\tau)^{1/(1-\gamma)}$ to simplify algebra since this expression appears in many of the equilibrium results. A firm j with distortion θ_j earns profits

$$\theta_j^{1-\gamma} y_j - w n_j - w x_j, \tag{5}$$

where n_j and x_j are production and adoption employment and w is the wage rate.

Definition 1. *A system of institutional distortions (ψ, Θ) consists of a set of potential distortions Θ and a conditional probability distribution $\psi(\theta|s)$ that describes the probability a firm with ability s receives distortion θ.*

Definition 1 describes the institutional environment of the economy. Firms draw a permanent institutional distortion θ at entry after ability s is determined. The structure of distortions is flexible and can accommodate a range of institutional policies. While the distortion is modeled as depending on a firm's ability s, it is straightforward to map this relationship into how the distortions relate to the size of the firm as measured by n or y. Looking ahead, the interpretation of institutional distortions is such that a firm that receives a relative tax (below average θ) is allocated comparatively fewer resources while a firm that receives a relative subsidy (above average θ) is allocated comparatively more resources than in an undistorted economy. Misallo-

[8] The tax τ should not necessarily be thought of as a corporate tax in the model or calibration. This is because these taxes tend to be on profits and apply equally to firms. On its own, this implies that firms do not have an incentive to reduce their size to avoid taxation.

[9] The model results are not affected by the average level of $1-\tau$, such that there is always an equivalent system of distortions that would balance the government's budget with $T_t = 0$. Given this, I abstract from any financing or revenue concerns related to the idiosyncratic taxes by assuming a lump-sum transfer.

cation depends on the variation in institutional distortions θ rather than the average value of θ.[10]

2.3 Equilibrium

I focus on the balanced growth path equilibrium in which all variables and prices grow at constant rates. To simplify notation, I drop the t subscripts where it does not cause confusion and denote the prior period $(t-1)$ with subscript -1. I also refer to firms using by their type (s,θ) rather than j.

Solving the household's problem implies that standard Euler Equation given by $R = e^{g_C}/\beta - 1$, where g_C is the growth rate of consumption.

Firm's problem. The problem of a firm (s,θ) with previous technology z_{-1} is to choose labour n and technology z to maximize firm value, given by:

$$v_{(s,\theta)}(z_{-1}) = \max_{n,z} (z\theta(As)^{1-\omega})^{1-\gamma} n^{\gamma} - wn - wx(z, z_{-1}) + \frac{1-\lambda}{1+R} v_{(s',\theta')}(z), \qquad (6)$$

Proposition 1 characterizes the firm's value function and optimal choices of production labour and technology.[11] Distortions affect the economy through firm-level choices of both labour and technology.

Proposition 1. *Define the normalized value of technology as $\varphi = (z/\bar{z})^{\xi}$ and the normalized wage rate as $\bar{w} = w/\bar{z}^{1-\gamma}$. The value of a firm (s,θ) with prior technology z_{-1} is equal to $v_{(s,\theta)}(z_{-1}) = w\bar{v}_{(s,\theta)}(\varphi_{-1})$ where:*

$$\bar{v}_{(s,\theta)}(\varphi_{-1}) = A\frac{1-\omega}{1-D}\left[\frac{\omega^{\frac{\omega}{1-\omega}}(\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma))^{\frac{1}{1-\omega}}}{(\pi(1-De^{-\xi g}))^{\frac{\omega}{1-\omega}}}\right](\bar{w})^{-\frac{1}{1-\omega}\frac{1}{1-\gamma}}\theta^{\frac{1}{1-\omega}}s + \pi e^{-\xi g}\varphi_{-1}, \qquad (7)$$

where $D = \frac{1-\lambda}{1+R}e^{g_w}$ is the effective discount rate of firms and g_w is the growth rate of

[10]Policies that uniformly affect firm technology use would show up in aggregate barriers π and not the distortions θ. For example, the innovation literature shows that market power increases R&D spending by allowing firms to profit from their innovations. If market power was similarly used to aid adoption it would show up as a reduction in aggregate barriers π rather than a higher distortion θ.

[11]Proofs of all propositions are available in Appendix A.

wages. The firm's choices of labour and technology are characterized by:

$$n(s,\theta) = \left(\frac{\gamma}{\bar{w}}\right)^{\frac{1}{1-\gamma}} \theta\varphi(s,\theta)^{\omega}(As)^{1-\omega} \quad \textit{and} \quad \varphi(s,\theta) = \left(\frac{z(s,\theta)}{\bar{z}}\right)^{\xi} = \eta\theta^{\frac{1}{1-\omega}}s, \tag{8}$$

where $\eta = A\left[\frac{\omega(1-\gamma)\gamma^{\frac{\gamma}{1-\gamma}}}{\pi(1-De^{-\xi g})}\right]^{\frac{1}{1-\omega}} (\bar{w})^{-\frac{1}{1-\omega}\frac{1}{1-\gamma}}$.

The equilibrium value of a firm in (7) has two components. The first term describes the expected discounted stream of expected profits. The second term describes the value of the firm's current technology, described by technology proximity φ_{-1}.

The firm's technology proximity $\varphi(s,\theta)$ is a stationary measure of firm technology that scales the firm's technology $z(s,\theta)$ by the technology frontier $\bar{z}$. Proximity is a convenient measure because it does not depend on the technology frontier $\bar{z}$ and has a stationary distribution. The term η can be interpreted as a country-specific technology index that captures the relative suitability of the country's aggregate factors (e.g., productivity, wage, aggregate barriers) for technology adoption.

A newly created firm enters the economy with zero prior technology, $\varphi_{-1} = 0$ and a random type (s,θ). The normalized expected value of a new firm is given by:

$$\bar{v}_E = \int_{\mathcal{S}\times\Theta} \bar{v}_{(\hat{s},\hat{\theta})}(0)h(\hat{s},\hat{\theta})d\hat{s}d\hat{\theta} - c_E, \tag{9}$$

where $h(s,\theta) = \psi(\theta|s)\mu(s)$ is the joint distribution of ability and distortions. In equilibrium, the cost of entry must be at least as great as the expected value of a new firm ($\bar{v}_E \geq 0$) and equal ($\bar{v}_E = 0$) on the balanced growth path equilibrium where entry is strictly positive. Distortions affect the entry condition through both the expected value of entry and the wage rate, through general equilibrium effects.

Market clearing. Market clearing requires that both the output and labour markets clear. Output is only used as consumption by the household, such that the goods market clearing condition can be written as:

$$C = Y = M\int_{\mathcal{S}\times\Theta} (\bar{z}\varphi(\hat{s},\hat{\theta})^{\omega}(A\hat{s})^{1-\omega})^{1-\gamma} n(\hat{s},\hat{\theta})^{\gamma} h(\hat{s},\hat{\theta})d\hat{s}d\hat{\theta}. \tag{10}$$

Labor has three uses in the economy. First, labor can be employed by firms in the production of the final good, $N_P = M\int_{(s,\theta)} n(s,\theta)dsd\theta$. Second, labor can be employed by firms to invest in next technology, $N_X = M\int_{(s,\theta)} \pi\varphi(s,\theta) - (1-\lambda)M\int_{(s,\theta)} \pi e^{-g\xi}\varphi(s,\theta)$. The first term is total investment that would be needed for all firms to reach the value of technology $\varphi(s,\theta)$ in Proposition (1) while the second term subtracts the technology that incumbent firms already use from the previous period, since firms do not need to double pay for technology. Third, labor can be used in the creation of new firms, $N_E = \lambda c_E M$. Putting this together, the labor market clearing condition is:

$$1 = M\left[\int_{\mathcal{S}\times\Theta}\left[n(\hat{s},\hat{\theta}) + \pi(1-e^{-g\xi}(1-\lambda))\varphi(\hat{s},\hat{\theta})\right]h(\hat{s},\hat{\theta})d\hat{s}d\hat{\theta} + \lambda c_E\right], \tag{11}$$

where the left-hand side is the total supply of labour, which is normalized to one.

Equilibrium definition. Definition 2 introduces the equilibrium concept.

Definition 2. *Given the initial distribution of productivities $\mu(s)$ and a system of distortions (ψ,Θ), a balanced growth path equilibrium is the set of values:*

$$\{\bar{w}, R, C, n(s,\theta), \varphi(s,\theta), M, M_E, \bar{v}_{(s,\theta)}, \bar{v}_E, T, g_M, g_w, g_C\}$$

for all values of (s,θ), such that: (i) the household's Euler Equation is satisfied; (ii) firm value $\bar{v}_{(s,\theta)}(\varphi_{-1})$ is given by (7); (iii) the values $n(s,\theta)$ and $\varphi(s,\theta)$ are set as in (8); (iv) the entry value $\bar{v}_E$ is given by (9) and is equal to zero; (v) the mass of firms satisfies (3); (vi) the government budget is balanced; (vii) The goods market and the labor market clear, satisfying (10) and (11).

2.4 Equilibrium Characterization

Proposition 2 characterizes the solution to the balanced growth path equilibrium in Definition 2. For brevity, the proposition does not report the values of labor $n(s,\theta)$ or proximity $\varphi(s,\theta)$—the expressions follow from (8) and the wage rate in (13)—or the lump-sum government transfer T.

Proposition 2. *A balanced growth path equilibrium exists and is unique. The equi-*

librium is characterized by

$$C = Y = \zeta_Y M \left(\frac{\bar{z}A^{1-\omega}}{\pi^{\omega}}\right)^{1-\gamma} \left[\frac{\int_{\mathcal{S}\times\theta} \hat{\theta}^{\gamma+\frac{\omega}{1-\omega}} \hat{s} h(\hat{s},\hat{\theta})\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{\theta}}{\left(\int_{\mathcal{S}\times\Theta} \hat{\theta}^{\frac{1}{1-\omega}} \hat{s} h(\hat{s},\hat{\theta})\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{\theta}\right)^{\gamma+\omega(1-\gamma)}}\right], \tag{12}$$

$$w = \zeta_w \left(\frac{\bar{z}A^{1-\omega}}{\pi^{\omega}}\right)^{1-\gamma} \left[\int_{\mathcal{S}\times\Theta} \hat{\theta}^{\frac{1}{1-\omega}} \hat{s} h(\hat{s},\hat{\theta})\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{\theta}\right]^{(1-\gamma)(1-\omega)}, \tag{13}$$

where $\zeta_Y = \left(\frac{\gamma}{1-\gamma}\right)^{\gamma} \left(\frac{1-D}{1-\omega}\right)^{\gamma+\omega(1-\gamma)} \left(\frac{\omega}{(1-De^{-\xi g})}\right)^{\omega(1-\gamma)}$ *and* $\zeta_w = \left[\left(\frac{1-\omega}{1-D}\right)^{1-\omega} \left(\frac{\omega^{\omega}(\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma))}{(1-De^{-\xi g})^{\omega}}\right)\right]^{1-\gamma}$. *The growth rate of output and wages are given by* $g_w = g_Y = \ln Y_{t+1} - \ln Y_t = (1-\gamma)g$ *and the interest rate is* $R = e^{(1-\gamma)g}/\beta - 1$*. The mass of firms* M *is constant,* $g_M = 0$*, and given by*

$$M = \left[\left(\frac{1-D}{1-\omega}\right)\left[\frac{\gamma}{\omega(1-\gamma)} + \frac{1-e^{-\xi g}(1-\lambda)}{1-De^{-\xi g}}\right] + \lambda c_e\right]^{-1}. \tag{14}$$

The mass of entrants is equal to $M_E = \lambda M$*.*

The undistorted economy is a special case of Proposition 2 in which all firms draw the same distortion θ (e.g., $\theta = 1$). The equilibrium does not depend on the average level of distortions θ in that uniformly increasing θ has no affect on the allocations or aggregate output, due to offsetting general equilibrium adjustments to the wage rate w. It also follows from the proposition that distortions affect the economy only through the aggregate productivity level and not through the aggregate growth rate.

Growth in output and wages are both driven by growth in the technology frontier $\bar{z}$, which allows firms to employ better technologies, with the same proximity, at the same cost. Output and wages are also both decreasing in aggregate barriers π and increasing in both the productivity shifter A and the average ability s. Wages depend on the distortion-weighted ability, where distortions enter the expression with the same form as with firm-level labor demand (Proposition 1). This implies that wages will be higher when more productive firms have larger distortions and employ more workers (Recall that wages are increasing in the average value of distortions).

The impact of misallocation on output (in square brackets) depends on the the ratio of distortion-weighted average ability. In the undistorted economy, this expression simplifies to the average ability s to the exponent $(1-\omega)(1-\gamma)$, the same exponent on ability in the production function. The expressions provides an important insight into the interactions between distortions and technology adoption on aggregate productivity. The expression in square brackets nests the TFP impact of misallocation in the canonical model of misallocation (for example, Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009) for the case where $\omega \to 0$, which can be thought of as the no technology case. Technology adoption amplifies dispersion in distortions θ increasing the scope for misallocation.

Finally, the proposition shows that the mass of firms is stationary and does not depend on the technological frontier $\bar{z}$, the productivity shifter A, the aggregate barriers π, or the distribution of firm types (s, θ). This implies that changes in the system of distortions only affect the economy through reallocations across firms.

Aggregation. The previous results show that aggregate outcomes depend on the joint distribution of distortions and ability. Proposition 3 shows that despite the complexity of adding firm heterogeneity and institutional distortions, the model yields the same aggregate implications as a representative firm model.

Proposition 3. *Aggregate output is equivalent to a representative firm economy with mass M firms, a fixed labor cost λc_e, aggregate barrier $\pi^{rep} = \pi$, and representative firm (s^{rep}, θ^{rep}) with productivity*

$$s^{rep} = \left[\frac{\int_{\mathcal{S}\times\theta} \left[\hat{\theta}^{\gamma+\frac{\omega}{1-\omega}}\hat{s}\right] h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}}{\left(\int_{\mathcal{S}\times\Theta} \hat{\theta}^{\frac{1}{1-\omega}}\hat{s}h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}\right)^{\gamma+\omega(1-\gamma)}} \right]^{\frac{1}{1-\omega}\frac{1}{1-\gamma}}$$

and distortion $\theta^{rep} = 1$. Aggregate output can be written as:

$$Y = My(s^{rep}, \theta^{rep}) = M\left[\bar{z}\varphi(s^{rep}, 1)^{\omega}(As^{rep})^{1-\omega}\right]^{1-\gamma} n(s^{rep}, 1)^{\gamma}. \tag{15}$$

Proposition 3 shows that the model aggregates to a neoclassical growth model with technology adoption, nesting Parente and Prescott (1994), whereas the previous results show that the model nests a static misallocation model along the lines of Restuccia and Rogerson (2008). There are two main payoffs for combining these frameworks. First, as discussed in more detail in the calibration section, firm heterogeneity implies a gradual adoption of new technologies, which is used to identify technology parameters in the data. Second, the inclusion of firm heterogeneity allows for the impact of institutional distortions on technology adoption to be quantified.

Institutional distortions are a natural explanation for the slow adoption of new technologies in developing countries, but cannot be studied from an aggregate perspective. The quantitative section provides a bridge between measures of misallocation and observed adoption lags.

2.5 Adoption

I use the model to examine the implications for the adoption path of new technologies along the balanced growth path equilibrium. To start, I need to define a concept of adoption in the model. I consider the simplest definition in which a firm j that uses technology z_j also uses every technology invented prior to z_j (Definition 3). Since technologies are ordered by their productivity this implies that firm j uses all technologies z with $z_j \geq z$.

Definition 3. *A firm j with technology z_j has adopted a technology z if $z_j \geq z$.*

I write the definition in terms of the unscaled technology z (rather than scaled technology φ) since it highlights how new technologies become adopted. Over time, firms adopt more advanced (higher z) technologies as growth in the technology frontier $\bar{z}$ lowers the cost of adoption. The definition is equivalent to stating that a firm (s, θ) has adopted a technology with proximity φ if $\varphi(s, \theta) \geq \varphi$. An implication of the model is that the distribution of technologies across firms can be reframed in terms of the distribution of proximity $\varphi(s, \theta)$, which is stationary along the balanced growth path.[12] On its own, the stationarity of proximity is not immediately beneficial since both the firm's technology $z(s, \theta)$ and proximity $\varphi(s, \theta)$ would require detailed data on firm-level technology use to identify. However, I show next that the distribution of proximity $\varphi(s, \theta)$ implies a distribution of adoption lags, denoted $L(s, \theta)$, which can be mapped directly into observable data. This is a key property of the model and is used to identify the cost function of investment—the aggregate barriers π and the cost curvature ξ—in the calibration.

I define the optimal adoption lag $L(s, \theta)$ for firm (s, θ) as the number of periods following a technology's creation and its adoption by firm (s, θ). While adoption lags

[12]The cost structure assumed in (1) implies that firms immediately adopt technology up a specific value of φ regardless of their previous technology level. Appendix E considers an alternative cost function in which the curvature on the cost function is on the change in technology $(z_{j,t} - z_{j,t-1})$ rather than the level of technology. In this environment, I show that average scaled technology of firm types (s, θ) is constant over time. This leads to similar aggregate implications.

are stationary, firms eventually adopt all technologies as growth in the technology frontier $\bar{z}$ lowers the cost of existing technologies z by enough for firms to profitably adopt. The optimal adoption lag of a firm is given by:

$$L(s,\theta) = -\frac{1}{g\xi} \ln\left[\eta\theta^{\frac{1}{1-\omega}} s\right]. \tag{16}$$

The adoption lag is falling in the firm's proximity, $\varphi(s,\theta) = \eta\theta^{\frac{1}{1-\omega}} s$, and the rate at which the cost of technology declines, $g\xi$. The importance of (16) is that it allows the cross-section of firm technologies to be mapped to and from the distribution of adoption lags. This implies that the cross-sectional distribution of technologies used by firms can be recovered from aggregate time-series data on adoption. This mapping is further explored in Section 3. Proposition 4 summarizes the main comparative statics with the firm-level adoption lag and key variables.

Proposition 4. *Adoption lags do not depend on aggregate productivity and are increasing in aggregate barriers:*

$$\frac{\mathrm{d}L(s,\theta)}{\mathrm{d}\ln A} = 0; \tag{17}$$

$$\frac{\mathrm{d}L(s,\theta)}{\mathrm{d}\ln\pi} = -\frac{1}{g\xi}\frac{\mathrm{d}\ln\varphi(s,\theta)}{\mathrm{d}\ln\pi} = \frac{1}{g\xi} > 0. \tag{18}$$

Firms with higher ability (higher s) and institutional distortions (higher θ) adopt technologies earlier:

$$\frac{\partial L(s,\theta)}{\partial\ln\theta} = -\frac{1}{g\xi}\frac{\partial\ln n(s,\theta)}{\partial\ln\theta} = -\frac{1}{g\xi}\frac{1}{1-\omega} < 0; \tag{19}$$

$$\frac{\partial L(s,\theta)}{\partial\ln s} = -\frac{1}{g\xi}\frac{\partial\ln n(s,\theta)}{\partial\ln s} = -\frac{1}{g\xi} < 0. \tag{20}$$

For aggregate factors, adoption lags are not affected by the productivity shifter, A, and are increasing in aggregate barriers to adoption, π. Uniformly increasing firm productivity raises the profitability of firms leading firms to demand more labour, increasing the wage rate. The increase in the wage exactly offsets the increase in incentives to adopt, resulting in no change to adoption.[13] The increase in aggregate barriers, π, causes the overall cost of adoption to increase relative to the wage rate

[13]It is straightforward to show that uniform increases in the average institutional distortion θ similarly has no affect on adoption lags.

leading to firms investing less in technology adoption. Since aggregate barriers affect all firms equally, and there is no interaction with institutional distortions θ, the impact of aggregate barriers is to shift the adoption curve to earlier or later periods. Given this, I use the aggregate barriers π as a residual measure of technology across countries in the quantitative model.

Adoption lags are decreasing in the ability s and the institutional distortion θ of a firm. Increasing either s or θ makes firms more profitable incentivizing investment in the adoption of new technologies. The insight from Proposition 4 is that the impact of institutional distortions on the time-series adoption lag can also be understood through the impact of institutional distortions on the cross-sectional firm size distribution. The empirically relevant case are institutions that compress the size distribution by lowering employment at productive firms. These institutions will also tend to delay adoption by disincentivizing otherwise early adopters. For example, many policies notably limit firms from growing beyond a certain size in terms of employment, such as employment thresholds for taxes or competition from state-owned enterprises. These policy would then also delay the initial adoption of new technologies because they limit incentives for firms to adopt new technologies as firms are unable to fully realize the benefits of the adoption costs. The quantitative exercise measures how important this channel is for explaining cross-country technology differences.

3 Calibration

The benchmark economy is calibrated to replicate features of the United States' data on the distribution of firm sizes, distortions and technology adoption. I first match the distribution of firm types (s, θ) using estimates of institutional distortions and the firm size distribution. Given this distribution, the model implies an adoption curve, which can then be used to determine the technology cost curvature ξ. Section 4 quantifies the effect of changing institutional distortions and aggregate barriers on technology and aggregate productivity. Institutional distortions and technology barriers are disciplined using cross-country empirical evidence on misallocation and technology adoption lags.

3.1 Benchmark Economy

The benchmark economy calibration consists of setting values for the parameters $\{\gamma, g, \lambda, \beta, c_e, \xi, \omega\}$ and the distributions of distortions and firm ability. Average productivity A, technology barriers π and the technology frontier $\bar{z}$ are normalized as their level does not affect the results. Table 1 summarizes the calibrated parameters.

Table 1: Benchmark Calibration

Parameter		Value	Target
Discount Rate	β	0.98	Interest Rate
Labour Share	γ	0.67	Labour Share
Technology Growth	g	0.02	TFP Growth Rate
Exit Probability	λ	0.07	Establishment Exit Rate
Cost Function Elasticity	ξ	2.47	90-10 Lag
	ω	0.40	Implied by ξ
Firm Productivity	$\{s\}$	[1 , 267224]	Firm-Employment Distribution
Productivity Distribution	μ	-	Firm-Employment Distribution
Elasticity of Distortions	ρ_{BE}	0.09	Elasticity of TFPR w.r.t. TFPQ
Dispersion of Distortions	σ_{BE}	0.25	Standard Deviation of TFPR

Notes: The values of A, $\bar{z}$, π_{BE}, and c_E are normalized to one in the benchmark economy.

Preliminaries. I set the discount rate $\beta = 0.98$ to correspond to a yearly interest rate of around 4%. The growth rate of the technology frontier is set to $g = 7.2\%$ to a match an output-per-capita growth rate of 2.41%, which is the productivity growth rate reported by Aghion et al. (2019). I set the coefficient on labour to be $\gamma = 0.66$ to be consistent with the labour share of output. Along with the value of ω, the return on firm ability is $(1-\omega)(1-\gamma) = 0.22$, which falls in the range that is typically used in the literature (for example, Restuccia and Rogerson, 2008).

Distortions. Following Bento and Restuccia (2017) and Guner et al. (2018), the system of distortions is generated by:

$$\ln(1 - \tau(\theta)) = -\rho \ln s + \varepsilon, \tag{21}$$

where $\tau(\theta)$ is the implicit tax rate implied by the distortion θ, ρ is the elasticity of distortions with respect to ability s, and $\varepsilon \sim \mathcal{N}(0, \sigma)$ is a random component of the distortion that does not depend on the firm's ability s. Higher correlation ρ cause higher ability firms to be taxed more leading to a reallocation from high ability s to low ability s firms. Although this functional form abstracts from specific institutions, it is straightforward to see how they would map into (ρ, σ). For example, size-dependent employment laws or taxation would correspond to higher ρ.[14] The average value of distortions is not directly targeted since it does not affect the allocation of resources, aggregate productivity, or technology adoption.

Empirical studies generally focus on the relationship between total factor quantity productivity (TFPQ) and total factor revenue productivity (TFPR). TFPQ measures how efficiently firms transform inputs into output. In the model, TFPQ is given by:

$$\text{TFPQ}(s,\theta) = \Big(z(s,\theta)s^{1-\omega}\Big)^{1-\gamma}. \tag{22}$$

Measured TFPQ then depends on both firm ability s and the firm's distortion θ, through its affect on technology choice. In contrast, in models of static misallocation (e.g. Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009), TFPQ is fixed and so distortions do not affect TFPQ.

Total factor revenue productivity (TFPR) is commonly used as a measure of institutional distortions. TFPR measures how efficiently firms transform inputs into revenues. In the model, TFPR is given by

$$\text{TFPR}(s,\theta) = \frac{y(s,\theta)}{n(s,\theta)} \propto \frac{1}{1-\tau(\theta)}. \tag{23}$$

Dispersion in TFPR measures the extent that firms differ in their abilities to generate revenues, with higher dispersion indicating more misallocation.[15] I target institutional distortions (ρ, σ) to match moments on the joint distribution of TFPQ and TFPR.

The elasticity ρ is targeted to match the elasticity of TFPR with respect to TFPQ, denoted as ELAS. There are two challenges to relating the measured elasticity ELAS with the elasticity parameter ρ. First, technology implies that a firm's TFPQ and

[14] While the wedge $\tau(\theta)$ is modeled as depending on s, its relationship with other variables are directly implied through (21). For example, the elasticity of the wedge with respect to labour is equal to $\frac{-\rho(1-\omega)(1-\gamma)}{1-\rho(1-\omega)(1-\gamma)}$. In this regard, the elasticity can capture policies that are not directly related with the firm's ability s.

[15] This implicitly assumes that production labor n is reported in the data and not researchers x. The implications for TFPR are the same if the denominator is instead total labor $(n(s,\theta)+x(s,\theta,\varphi(s,\theta)))$ or the Cobb-Douglas composite $((\varphi^{\omega(1-\gamma)}n^{\gamma})^{1/(\gamma+\omega(1-\gamma))}$.

TFPR both depend on its institutional distortion creating a mechanical relationship between the two values. Second, TFPQ is a model-based measure implying that ELAS may differ depending on the model used to construct it. Proposition 5 summarizes a mapping that corrects for these two challenges.

Proposition 5. *Define $ELAS_i$ as the measured elasticity between TFPR and TFPQ in country i. The elasticity ρ_i in (21) that implies this relationship is equal to:*

$$\rho = \Gamma_1 \left[\frac{1}{ELAS} + \frac{\omega}{1-\omega} \Gamma_2 \right]^{-1} \tag{24}$$

where $\Gamma_1, \Gamma_2 > 0$ are parameters that depend on the empirical model.

Proposition 5 shows that the implied elasticity ρ is smaller than in static models without technology, where $\rho = \Gamma_1 \text{ELAS}$. Consequently, ignoring technology adoption implies that the severity of institutional distortions is overstated by measured elasticity. This result helps explain why specific institutional distortions are often unable to explain the full magnitude of indirect measures of misallocation (as noted by Restuccia and Rogerson, 2013). The proposition does not imply that the measured gains from reallocation are lower compared with static models, since the first-best static allocation must also be feasible with technology adoption.

I target an elasticity of $\text{ELAS}_{BE} = 0.13$, which is in the range of values reported by Bils et al. (2020) and implies $\rho_{BE} = 0.09$.[16] Given ρ_{BE}, the dispersion of distortions is set to $\sigma_{BE} = 0.25$ to match a standard deviation of log(TFPR) of 0.31, which is US value reported by Hsieh and Klenow (2009) after the Bils et al. (2020) correction for measurement error. I assume that the US is distorted as it provides a reasonable benchmark for the minimum level of distortions that could be realistically attained in other economies. While US moments may reflect institutional frictions, they may also reflect unavoidable factors (e.g. measurement error, random business shocks).

Finally, it is worthwhile to note that aggregate barriers π do not affect the measured relationship between TFPQ and TFPR since aggregate barriers affect all firms equally and only through TFPQ. This implies that country-specific factors that affect technology adoption but do not cause resource reallocation will not be captured in

[16]The correction terms Γ_1 and Γ_2 in Proposition 5 account for differences between the model used here and the empirical model used to construct measured elasticity ELAS. To compare with Hsieh and Klenow (2009) $\Gamma_1 = \nu/(\nu-1)^2 = 3/4$ and $\Gamma_2 = (\nu-1)/\nu = 2/3$ where $\nu = 3$ is the CES parameter.

the measures of institutional distortions.[17]

Firm size distribution. Given the calibrated institutional distortions, the distribution of productivities can be disciplined using the distribution of firms sizes. The relative employment of two firms (s, θ) and (s', θ') is given by:

$$\frac{n(s,\theta)}{n(s',\theta')} = \frac{\theta z(s,\theta) s^{1-\omega}}{\theta' z(s',\theta') s'^{1-\omega}} = \frac{\theta^{\frac{1}{1-\omega}} s}{\theta'^{\frac{1}{1-\omega}} s'}.$$

This implies a mapping between the firm size distribution (measured in employment) and the ability distribution $\mu(s)$ for firm ability $\mathcal{S} = \{s_1, s_2, ..., s_{1,000}\}$. I use the establishment size distribution from the Business Dynamic Statistics (BDS) database to construct the empirical distribution (see Figure 1a).[18]

Technology. The technology cost function has two components: the aggregate barrier π_{BE} that determines the level of the cost function and the curvature ξ of the cost function, which also determines ω. I normalize π_{BE} and interpret the values of aggregate barriers π in other economies relative to the US.

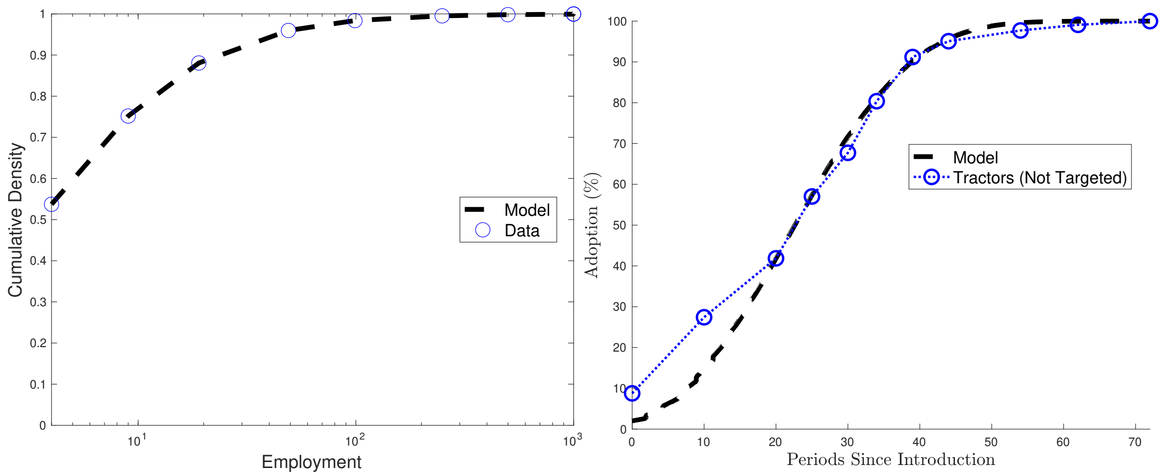
The ideal target for the curvature of the cost function ξ is the elasticity between a firm's measured productivity (TFPQ) and technology choice (measured by proximity φ): $\frac{\mathrm{d}\ln \mathrm{TFPQ}}{\mathrm{d}\ln\varphi} = \frac{1-\gamma}{\xi}$. However, the firm's choice of technology is not easily observed. Instead, the model suggests that the relative timing of adoption as an alternative measure. With this in mind, the curvature ξ is set to match the length of adoption in the United States measured by the 10-90 lag, which is the time between a technology being used in 10% of production and 90% of production. The assumption is that the main driving force of adoption of a long period is the relative cost and benefits of the technology. Proposition 6 characterizes the relationship between the 10-90 lag, the cost curvature ξ, and observables.

Proposition 6. *Let n_{10} and n_{90} be the employment of the marginal firms, ordered by output, that produce the bottom 10% and 90% of output. The number of periods for a*

[17] This holds even if the model is extended to allow for firm-specific technology barriers since barriers affect TFPQ and not TFPR.

[18] I set nodes such that the log distance between nodes is equal, which results in more nodes at lower productivities. Within nodes, productivities are set to be uniformly distributed. For the calibration of firm sizes, distortions are drawn from (21) with ρ equal to the value in Table 1 and $\sigma = 0$. This allows for a clean mapping from the data to the model and is equivalent to the assumption that σ captures measurement error.

Figure 1: Calibration of Productivity and Technology



(a) Firm Distribution by Employment (b) Calibrated Adoption Curve

Notes: Figure (a) plots the cumulative firm distribution by employment in the data and the calibrated model. Figure (b) plots the adoption curve implied by the calibrated model and the adoption curve of tractors in the data, which is not targeted in the calibration. Adoption (%) measures the fraction of total output produced by firms using technology introduced in period 0. Data on tractors measures the fraction of output produced by farms using tractors and is rescaled to 1 at the end of sample (see Chen (2020) for additional details).

technology to go from being used in 10% of production to 90% of production is:

$$\textit{10-90 Lag} = \frac{\ln n_{90}/n_{10}}{g_Y}\frac{1-\gamma}{\xi}. \tag{25}$$

The 10-90 lag in (25) is found by comparing the adoption lags of firms using $L_{(s,\theta)}$ in (16). The 10-90 lag is increasing in the size dispersion of firms (n_{90}/n_{10}) and decreasing in the growth rate g_Y and technology cost curvature ξ. Intuitively, increasing the dispersion of firm sizes implies larger dispersion of technology use across firms while increasing the growth rate or cost curvature implies that the cost of existing technologies decline more rapidly.

I target a 10-90 lag in the benchmark economy of 40 years in order to match the midpoints of lags for a wide range of technologies, rather than focusing on a specific technology. In a comprehensive study, Grubler (1991) finds an average 10-90 lag of 41 years for several hundred technologies. Other reports of 10-90 lags include: 54 years for steam locomotives and 25 years for diesel locomotives (Greenwood, 1999); 20 years for Blast Oxygen Furnaces (Oster, 1982); an average of 15 years for 21 consumer goods (Jovanovic and Lach, 1997). The periods of electrification and computerization reported by Jovanovic and Rousseau (2004) are also consistent with the adoption length. Appendix C summarizes the sensitivity of the results to the cost curvature ξ.

Along with the choice of the 10-90 lag, I set the growth rate g_Y and the labour coefficient γ as discussed above. I set $n_{90}/n_{10} = 1,169$ using the employment distribution implied by the BDS data (from Figure 1a).[19] Together, this implies a cost curvature of $\xi = 2.47$, which is close to values used in the literature. Parente and Prescott (1994) find a cost curvature of 1.8 using convergence in post-war Japanese growth. While not directly comparable, the curvature is also close to the R&D cost curvature of 2 typically used in the innovation literature (Akcigit and Kerr, 2018).

Figure 1b graphs the adoption curve in the benchmark economy. While the 10-90 lag is chosen to be consistent with a large range of technologies, I graph the adoption curve for tractors to provide a concrete comparison and external validation.[20] The figure shows that not only does the model fit the adoption length of tractors, the model replicates the S-shaped pattern of adoption, which is also a characteristic of other technologies (for example, hybrid corn in Griliches, 1957).

3.2 Cross-Country Moments

The quantitative analysis examines the impact of institutional distortions and aggregate barriers on cross-country technology and productivity differences. I calibrate economies i that differ from the benchmark economy in terms of the elasticity of distortions ρ_i and aggregate barriers π_i. Each economy i is disciplined to match moments of countries based on relative output-per-capita Y_i/Y_{US} in the data.[21] Figure 2 summarizes the model and data moments used to calibrate the parameters and Appendix D reports the corresponding values of ρ_i and π_i.

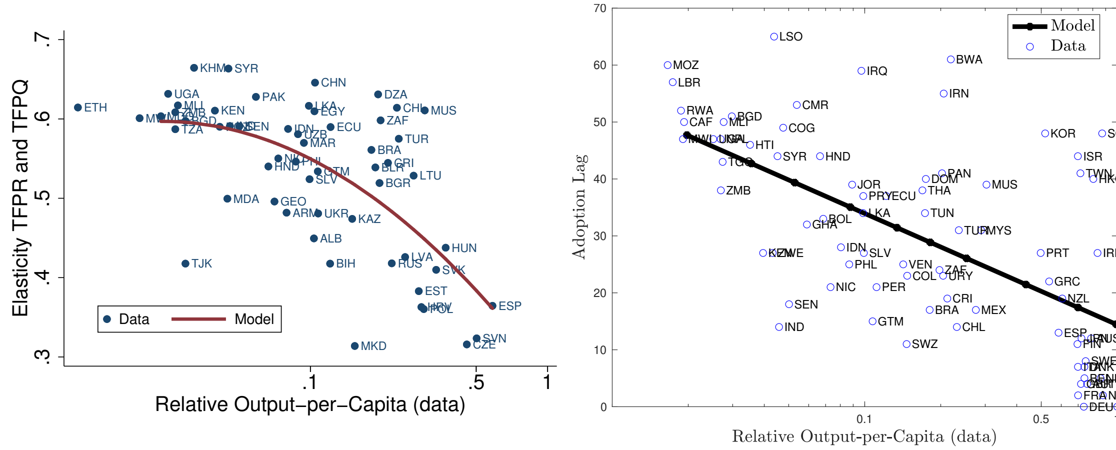
Institutional distortions. The elasticity ρ of distortions with respect to a firm's ability is the fundamental driver of cross-country differences in distortions. That is, distortions in less developed countries tend to be more targeted to productive or

[19] This assumes that the output distribution of firms is the same as the employment distribution, which will tend to hold in low distortion economies, but may not be exact. The calibrated model implies a 10-90 lag of 40 years in the benchmark economy showing this is a reasonable approximation.

[20] I thank Chaoran Chen for providing and helping with data on tractors used in Figure 1b. For additional details on the construction of the data see Chen (2020).

[21] Data for output-per-capita is from the Penn World Tables 9.1 for 2002.

Figure 2: Calibration of Productivity and Technology



(a) Elasticity of Inst. Distortions

(b) Calibrated Adoption Curve

Notes: Panel (a) plots the model implied elasticity of TFPR and TFPQ against empirical moments constructed using data from the World Enterprise Survey. Output-per-capita is from the Penn World Tables 9.1. Panel (b) plots the data and model implied technology adoption lags. The empirical value is the estimated country-level average adoption lag relative the US from Comin and Hobijn (2010). The model implied technology adoption lag is constructed as the number of periods, relative to the benchmark economy, it takes for the economy to produce 10% of output with a new technology.

large firms (see Hsieh and Klenow, 2009; Bento and Restuccia, 2017). I estimate the empirical elasticity of TFPR with respect to TFPQ using the World Bank Enterprise Survey following Bento and Restuccia (2017), where country estimates are reported in Figure 2a.[22] The elasticity for an economy i is given by:

$$\text{ELAS}_i = 0.275 - 0.171 \ln\left(\frac{Y_i}{Y_{US}}\right) - 0.023 \left[\ln \frac{Y_i}{Y_{US}}\right]^2. \tag{26}$$

Relative to a linear fit, the quadratic function implies lower values of elasticities for high income countries and does not substantially alter the estimates for low income countries. The intercept implies that the elasticity for an economy with output-per-capita equal to the US will be 0.275, which is larger than in the benchmark economy. Intuitively, this could reflect that wealthy European economies experience some de-

[22] All countries with fewer than 100 observations or with population less than 1 million are dropped from the sample. For comparability, the estimates of TFPR are constructed using the model structure in Bento and Restuccia (2017), which follows Hsieh and Klenow (2009). Equation (24) is used to construct the values of ρ_i from ELAS_i.

gree of size-dependent distortions beyond the US. The calibrated elasticities range from 0.275 to 0.596 in the most distorted economies, which is consistent with estimates found using higher quality micro data (see Appendix B). Figure 2a graphs the estimated elasticities for each economy along with the quadratic fit given in (26). The value of ρ_i is inferred using (24). The value of σ_i is held at the US level, which, along with the higher values of ρ_i, implies a higher value of the standard deviation of log(TFPR) and is consistent with empirical evidence (see Appendix B).

Aggregate barriers. A higher aggregate barrier π increases the cost of technology adoption to all firms resulting in delayed adoption (see Proposition 4). Aggregate barriers π are calibrated as a residual component of technology lags that cannot be explained by distortions. I define the adoption lag for economy i as the number of years between when economy i and the US benchmark economy produce 10% of output with a new technology.[23] I target estimates of adoptions lags from Comin and Hobijn (2010) who estimate the average adoption lags using a broad range of technologies (e.g., railways, cars, PCs) for each country.[24] Figure 2b reports the adoption lags for each country in the data and the implied adoption lags in the calibrated model.

Comin and Hobijn (2010) and Comin and Mestieri (2018) find that adoption lags have declined for newer technologies. For example, computers and the internet diffused more rapidly across countries than railways or steam ships. Targeting π_i to shorter adoption lags would imply smaller aggregate barriers π_i and a larger relative importance of institutional distortions for delaying technology adoption.

Discussion. Before moving to the quantitative results, I discuss the moments that are not included in the calibration and their potential impact on the model.

[23] The choice to measure adoption at 10% of output produced accounts for limitations in the data. Specifically, data may not be recorded in early periods of a technology's adoption leading to some bias in the adoption lags. It is unclear that this is an issue in Comin and Hobijn (2010) as they use a model to infer the adoption lags for countries. Using earlier points of adoption (e.g. first firm to use the new technology) would place a greater weight on higher ability s firms increasing the importance of institutional distortions in the quantitative analysis.

[24] Comin and Hobijn (2010) estimate adoption lags for country-technology pairs using a structural model and then decompose the adoption lag into a country- and technology-specific components. I use their estimates for the country-specific adoption lag as the target for aggregate barriers in the calibration. This allows for a country's technology lag to be captured more robustly than if a single technology were to be used.

- **Productivity** A: It would be straightforward to calibrate A_i to match residual differences in output-per-capita across countries. However, Proposition 4 shows that A_i does not affect technology adoption.

- **Ability Distribution** $\mu(s)$: I assume the cross-country ability distribution $\mu(s)$ is fixed. The average value of firm ability does not affect the results (by the same argument as A) and the distribution of $\mu(s)$ does not affect the measurement of any of the key model parameters, such as ρ. Additionally, the choice to hold $\mu(s)$ fixed does not imply that the firm size distribution (measured by either employment or output) or TFPQ distribution are fixed. The firm size distribution is compressed in more distorted economies because of the higher elasticity of distortions ρ_i, which is consistent with empirical evidence on the firm-size distribution (see, Adamopoulos and Restuccia, 2014; Poschke, 2018). The TFPQ distribution is also compressed in more distorted economies because of technology adoption. This is counterfactual to empirical evidence where TFPQ tends to be more dispersed in low income countries (see Appendix B). Targeting $\mu(s)$ to dispersion in TFPQ would increase the scope of institutional distortions in more distorted economies, making the fixed distribution a conservative assumption.

- **Entry Cost** c_E: The mass of firms M_i is constant across countries, which contradicts evidence showing that firms are smaller in lower income countries Bento and Restuccia (2017). This moment could be directly targeted through country-specific entry costs c_E. However, country-specific masses M_i would not affect the main results since there is no interaction with institutional distortions or aggregate barriers.

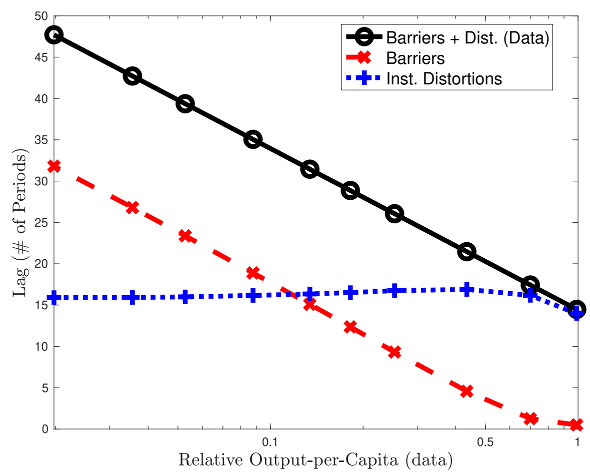
4 Quantitative Analysis

I use the calibrated model to quantify the extent that differences in institutional distortions and aggregate barriers explain cross-country technology adoption lags and aggregate productivity. I consider a counterfactual experiment in which either the correlation of institutional distortions ρ, aggregate barriers π, or both are adjusted in the benchmark economy to match the empirical evidence in Section 3.2. In this regard, the experiment can be thought of as adjusting the institutional structure of the US to match other countries. The gap between the benchmark and counterfactual economies then measure the relative importance of of institutional distortions and aggregate barriers. I use the US as the benchmark economy, with $(\rho_{BE}, \sigma_{BE}) > (0, 0)$, as opposed to the hypothetical undistorted economy, with $(\rho, \sigma) = (0, 0)$, to take seriously the issue that not all institutional distortions can be removed in practice. The results are divided between technology adoption and aggregate productivity.

4.1 Technology Adoption Lags

I focus primarily on the initial adoption lag as a measure of technology because it is directly comparable with the estimates in Figure 2b. Relative to the benchmark economy, the technology adoption lag can increase because either more correlated institutional distortions delay early adopters or aggregate barriers delay the adoption of all firms. Figure 3 summarizes the results. By construction, the counterfactual economy with both institutional distortions and aggregate barriers matches the adoption lag in the data (see Figure 2b) while the benchmark economy (not shown) has an adoption lag of zero.

Figure 3: Counterfactual Adoption Lag



Notes: The figure plots the relative adoption lag constructed as the number of periods between 10% of output being produced with a new technology in the counterfactual and benchmark economies. Barrier + Dist. refers to the economy with parameters (ρ_i, π_i). Inst. Distortions refers to the economy with parameters (ρ_i, π_{BE}). Barriers refers to the economy with parameters (ρ_{BE}, π_i). Other parameters are from Table 1.

The aggregate barrier line shows the benchmark economy with (ρ_{BE}, π_i) and the institutional distortion line shows the economy with (ρ_i, π_{BE}). The technology adoption lag of both channels (ρ_i, π_i) is equal to the sum the technology adoption lag in the economies with only aggregate barriers (ρ_{BE}, π_i) or only institutional distortions (ρ_i, π_{BE}) since there is no interaction between the two channels (Proposition 4). The contribution of the individual channels, ρ or π, to the adoption lag is then equal to the adoption lag in the economy with only that distortion.

Institutional distortions are important for cross-country technology adoption lags, but the profile is relatively flat across countries. The contribution of institutional distortions range from around 14 years, in high output-per-capita countries, to 17 years, in medium output-per-capita countries. The average contribution of institutional distortions to the gap in technology adoption lags is around 60% and ranges from one third in the lowest income countries to over 90% in the highest income countries.[25] Despite the large effect of institutional distortions in lower income countries, the overall contribution is smaller due to the larger base.

The delay in adoption from institutional distortions is driven by the higher correlation of distortions ρ in lower income countries disincentivizing firms from adopting new technologies because they lack sufficient scale or profitability to fully benefit from new technologies. Starting from the benchmark economy, small increases in ρ have relative large increases in the adoption lag because there is a high share of large early adopters. At higher values of ρ, the marginal impact from increasing ρ on adoption lags is small because there are fewer large early adopters and the firm distribution is more egalitarian. The flatness of the institutional distortion channel is then due to empirical elasticity of distortions implying that relatively high income countries are comparatively more distorted than the US. For example, (26) implies that economies with the same output-per-capita as the US would have around twice as high correlation of institutional distortions.

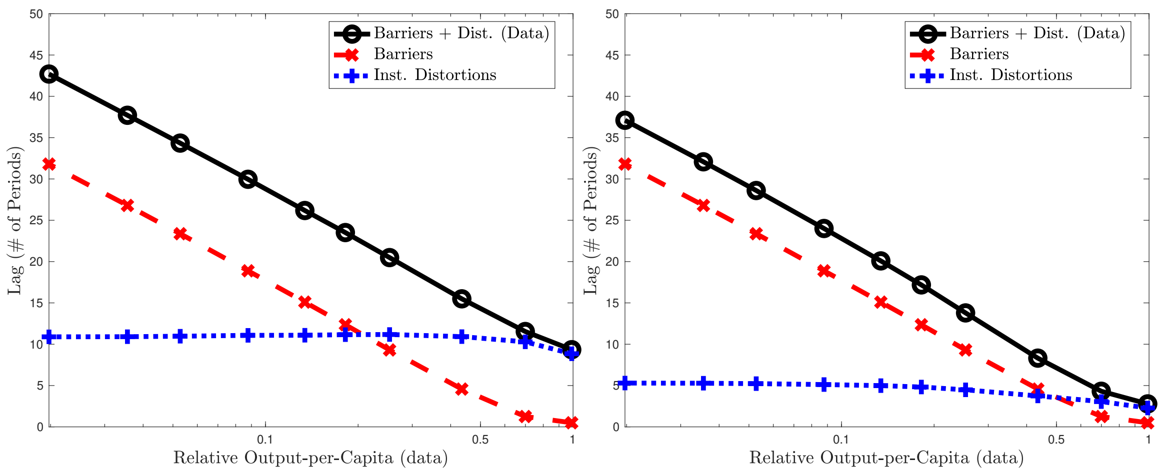
The main takeaway from the experiment is that institutional distortions are important across the distribution of countries. Institutional distortions in relatively high income economies provides a novel channel for of why European countries tend to lag behind American countries in the adoption of new technologies.[26] In low income countries, this channel helps explain why, despite the relative high productivity gains from using newer technologies, adoption tends to be low.

[25] Calculated as the adoption lag in the economy (ρ_i, π_{BE}) divided by the adoption lag in economy (ρ_i, π_i), which matches the average in the data by construction.

[26] See Comin and Hobijn (2010) for a broad set of evidence on technology adoption between the economies. Similarly Bloom et al. (2012) examine the relatively slow adoption of IT in Europe relative to the US.

Technology distribution. The model also has implications for the distributions of technology used across countries. I focus on two points, which I refer to as the 50th and 90th percentile adoption lags. The 50th (90th) percentile adoption lag is defined as the number of periods between 50% (90%) of output being produced with a new technology in the counterfactual and the benchmark economies. Figure 4 summarizes the adoption lags as well as the effects of aggregate barriers and institutional distortions.

Figure 4: Counterfactual Adoption Lag



(a) Adoption Lag: 50th Percentile (b) Adoption Lag: 90th Percentile

Notes: The figure plots the relative adoption lag constructed as the number of periods between 50% (Panel a) and 90% (Panel b) of output being produced with a new technology in the counterfactual and benchmark economies. Barriers + Dist. refers to the economy with parameters (ρ_i, π_i)*. Inst. Distortions refers to the economy with parameters* (ρ_i, π_{BE})*. Barriers refers to the economy with parameters* (ρ_{BE}, π_i)*. All other parameters are from Table 1.*

Comparing the adoption lags in Figure 4 with the previous figure shows that the adoption of new technologies is relatively compressed in lower income countries, matching evidence documented by Grubler (1991) that laggard countries tend to adopt technologies faster. For example, in the most distorted economy the 10-90 lag is around 10 years shorter than in the benchmark economy. The compression of the technology distribution is driven by more correlated distortions (higher ρ) disincentivizing high ability (high s) firms from adopting technologies. This similar explains why the overall impact of institutional distortions is decreasing at the 50th and 90th percentile adoption lags compared to the initial adoption lag. For low income countries, institutional distortions account for 11 years of the 50% adoption lag and 8 years of the 90th percentile adoption lag.

Given that low ability (low s) firms tend to benefit from the higher elasticity of distortions ρ_i in more distorted economies, the net effect on late adopters is ambiguous. All else equal, low ability s firms should adopt earlier in highly distorted economies because of beneficial distortions (higher θ) and lower wages, due to general equilibrium effects. Despite this, Figure 4b shows that institutional distortions still contribute to a positive gap for the 90th percentile adoption lag. This result can be understood through two effects of increasing ρ on the marginal adopter, i.e., the firm that produces the 90th percentile of output with the new technology. First, production is highly skewed in the benchmark economy implying that the marginal adopter is still relatively productive compared to the average firm, limiting the benefits that the firm receives from higher ρ. Second, increasing ρ causes output to be spread more equally across firms resulting in the marginal adopter in more distorted economies to also have lower ability s. Together, these effects lead to a positive contribution of institutional distortions to the 90th percentile adoption lag.

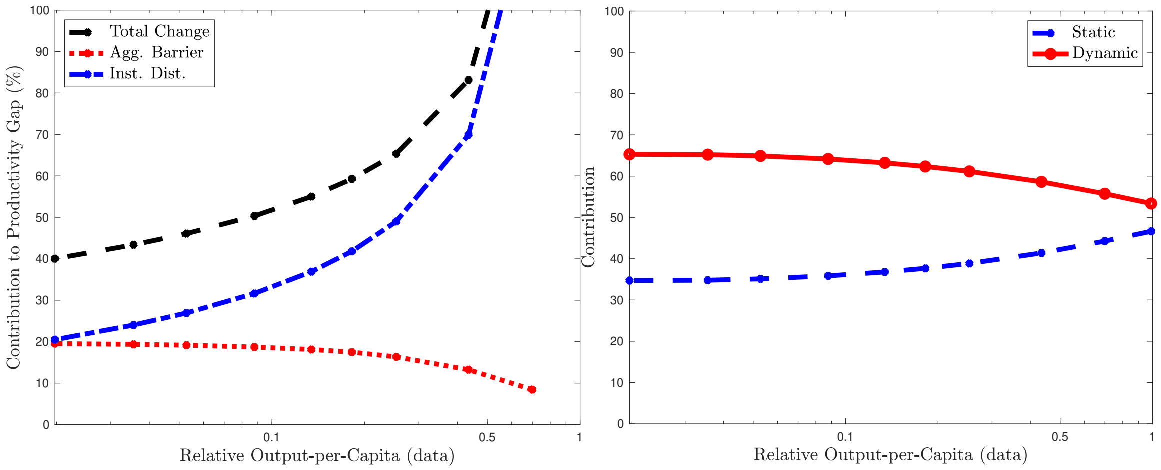
4.2 Aggregate Productivity

The previous results show that institutional distortions are an important driver of the observed differences in cross-country technology adoption lags. The second part of the experiment measures the productivity consequences of institutional distortions and aggregate barriers. Aggregate productivity is endogenous in the model and depends on both technology adoption and factor misallocation. Figure 5 reports the model implied contribution of institutional distortions and aggregate barriers to the cross-country productivity gap.[27]

The combination of institutional distortions and aggregate barriers have a substantial productivity cost with the relative productivity between the benchmark and counterfactual economies ranging from 43% in high income countries to around 380% in the lowest income countries. Both individual channels are also important throughout the distribution. That said, the impact of institutional distortions tends to dominate aggregate barriers because of its additional cost on productivity through the misallocation of labour.

[27]Results for countries with output-per-capita close to the US in the data are not reported because the small denominator causes the contribution to be very large.

Figure 5: Counterfactual Productivity



(a) Productivity Gap (b) Static and Dynamic Misallocation

Notes: Figure 5a reports the contribution of productivity gap in the model to the observed productivity gap in the data. The contribution of for economy i is calculated as $[\log(Y_i/L_i) - \log(Y_{BE}/L_{BE})]/[\log(Y_i^{data}/L_i^{data}) - \log(Y_{US}^{data}/Y_{US}^{data})]$. *Barrier + Dist refers to the economy with* (ρ_i, π_i). *Inst. Dist. refers to the economy with parameters* (ρ_i, π_{BE}). *Agg. Barriers refers to the economy with* (ρ_{BE}, π_i). *Results for the highest output-per-capita counterfactual economies are not reported because of the small base. Figure 5b reports the contribution of static and dynamic misallocation to the Inst. Dist. productivity gap in Figure 5a. Contribution is calculated change in productivity from static (dynamic) misallocation divided by the total change in productivity from misallocation and is based on (27).*

Aggregate barriers π act as a residual measure of technology differences in the calibration. In this regard, the results highlight the potential gain from removing barriers to technology adoption that are unrelated to institutional distortions. Aggregate barriers range from explaining a productivity gap of around 1% in the highest income countries to 115% in the lowest income countries, or around 20% of the observed gap in output-per-capita. The results indicate that, even after accounting for institutional distortions, aggregate technology barriers (i.e. those that affect firms equally) explain an important fraction of cross-country productivity differences. However, the results also indicate that policies designed to improve technology adoption through aggregate channels, without remedying misallocation, will not improve productivity by as much as suggested by technology adoption lags alone.

Institutional distortions account for a productivity gap of around 41% in high income countries and 123% in low income countries, or around 20% of the observed gap in output-per-capita. Institutional distortions are governed by the elasticity of distortions ρ such that a higher value of ρ causes more labour to be reallocated from high ability (high s) firms to low ability (low s) firms. This is the classic misallocation of resources. Similarly, there is also a "misallocation" of technology in the sense that increasing ρ causes high ability (high s) firms to invest less and low ability (low s) firms to invest more in technology.[28] I refer to the first channel as static misallocation and the second as dynamic misallocation.

The static and dynamic channels of misallocation are calculated as:

$$\frac{Y^{BE}}{Y^i} = \underbrace{\frac{\int_j \left(z(s_j,\theta_j^{BE})s_j^{1-\omega}\right)^{1-\gamma} n(s_j,\theta_j^{BE})^{\gamma}dj}{\int_j \left(z(s_j,\theta_j^{i})s_j^{1-\omega}\right)^{1-\gamma} n(s_j,\theta_j^{BE})^{\gamma}dj}}_{\text{Dynamic Misallocation}} \times \underbrace{\frac{\int_j \left(z(s_j,\theta_j^{i})s_j^{1-\omega}\right)^{1-\gamma} n(s_j,\theta_j^{BE})^{\gamma}dj}{\int_j \left(z(s_j,\theta_j^{i})s_j^{1-\omega}\right)^{1-\gamma} n(s_j,\theta_j^{i})^{\gamma}dj}}_{\text{Static Misallocation}}, \tag{27}$$

where the notation θ_j^i refers to the institutional distortion that firm j would receive if distortions where set given the parameterization in economy i. Specifically, this holds the noise component ε fixed in (21) and adjusts the elasticity parameter ρ. Since the effect of aggregate barriers π is orthogonal to institutional distortions, the comparison does not depend on its value.

Static and dynamic misallocation measure the productivity gaps between three economies: (1) an economy with elasticity ρ_i; (2) an economy with ρ_{BE} but where firm-level technology is restricted to its value in economy i, $z(s_j,\theta_j^i)$; and (3) the benchmark economy with ρ_{BE}. Economy (2) acts as an intermediate step in which only some reallocation occurs. The gap between economies (1) and (2) measures static misallocation following in the spirit of Hsieh and Klenow (2009), in which measured TFPQ is held fixed at the firm-level and resources are reallocated given the distribution of measured TFPQ. The gap between economies (2) and (3) measures dynamic misallocation where firms are free to adjust technology and to further adjust labour to the benchmark institutional distortions.

An implication is that economy (2) looks to be as efficient as the US benchmark from the perspective of the joint distribution of TFPQ and TFPR. The amplification channel when moving from economy (2) to (3) is through the change in the distribution of TFPQ. Figure 5b reports the contribution of each channel to the productivity

[28] Misallocation of technology is a misnomer because technology is non-rivalrous and so there is not the same sense of allocating technology between firms.

gap from institutional distortions in Figure 5a.

Comparing static and dynamic misallocation shows that the dynamic channel dominates, explaining around half of the gains in the top decile economy and around two-thirds in the bottom decile economy. In lower income countries dynamic misallocation becomes an increasingly important channel of productivity loss. The change in the relative importance is explained by the effect of institutional distortions on the productivity distribution. Higher distortions imply less dispersion in technology choice and consequently less dispersion in TFPQ. This limits static misallocation because there is less of a gap between low and high productivity firms limiting the scope of reallocation. This also expands the scope of dynamic misallocation because there is now a larger difference between the TFPQ distribution in the distorted and benchmark economies.

Overall, the results point to dynamic misallocation being an important amplifying channel for the productivity cost of misallocation. Additionally, the results suggest that in economies where the effects of misallocation is limited by the dispersion measured productivity, dynamic misallocation plays an even larger role.

5 Conclusion

Institutions that distort the incentives of potentially large and productive firms are prevalent in developing countries. In this paper, I examine the importance and consequences of these institutional distortions for delaying the adoption of new technologies and aggregate productivity. I start by developing a model of technology adoption that incorporates firm heterogeneity and institutional distortions. The model features a non-degenerate distribution of technologies used in equilibrium that implies a gradual adoption of new technologies, consistent with the empirical evidence. This feature of the model allows for aggregate time-series technology data to be mapped into key underlying parameters that govern the cost-benefit tradeoff of technology adoption. The calibrated model is used to quantify the overall importance of institutional distortions. Moving low income economies to US distortions would increase productivity by up to 123% and reduce the adoption lag of new technologies by up to 13 years.

The results have two implications that are important for future work. First, cross-country differences in technology use are inherently linked to firm-level decisions. In this regard, measures of aggregate barriers to adoption that neglect institutional distortions may overstate the importance of these channels for explaining cross-country productivity differences. This is because differences in technology adoption may be a symptom of other underlying inefficiencies in the economy rather than a cause. While I focus on institutional distortions as a source of firm-level inefficiencies, it is straightforward to extend the results to other mechanisms that might distort the distribution of firm-level productivities. Second, examining misallocation in a static setting may understate the importance of institutional distortions and may also distort the measurement of misallocation. The amplification channel of firm-level dynamics is not new in the literature (e.g. Gabler and Poschke, 2013; Hsieh and Klenow, 2014; Midrigan and Xu, 2014; Akcigit et al., 2021). The results here provide a comparable measure of the relative importance of this dynamic channel across countries. Additionally, the results show how the measurement of institutional distortions may be affected by the inclusion of technology (or other firm-dynamics). Given this, a path forward for future work would be to use firm-level panel data to examine the relationship between institutional distortions and the evolution of the distribution of productivities.

References

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2017). Misallocation, Selection, and Productivity: A Quantitative Analysis with Micro Data from China. *Working Paper*.

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

Aghion, P., Bergeaud, A., Boppart, T., Klenow, P. J., and Li, H. (2019). Missing Growth from Creative Destruction. *American Economic Review*.

Akcigit, U., Alp, H., and Peters, M. (2021). Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries. *American Economic Review*.

Akcigit, U. and Kerr, W. (2018). Growth through Heterogeneous Innovations. *Journal of Political Economy*, 126(4):1374–1443.

Ayerst, S., Brandt, L., and Restuccia, D. (2020). Market Constraints, Misallocation, and Productivity in Vietnam Agriculture. *Food Policy*.

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

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

Bento, P. and Restuccia, D. (2017). Misallocation, Establishment Size, and Productivity. *American Economic Journal: Macroeconomics*, pages 267–303.

Bils, M., Klenow, P. J., and Ruane, C. (2020). Misallocation or Mismeasurement? *Working Paper*.

Bloom, N., Sadun, R., and Van Reenen, J. (2012). Americans Do It Better: US Multinationals and the Productivity Miracle. *American Economic Review*.

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

Chen, C. (2020). Technology adoption, capital deepening, and international productivity differences. *Journal of Development Economics*.

Chen, C., Santaeulalia-Llopis, R., and Restuccia, D. (2019). The Effects of Land Markets on Resource Allocation and Agricultural Productivity. *Working Paper*.

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

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

Comin, D. and Mestieri, M. (2018). If technology has arrived everywhere, why has income diverged? *American Economic Journal: Macroeconomics*, 10:137–178.

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

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

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., and Ventura, G. (2018). Managers and Productivity Differences. *Review of Economic Dynamics*.

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

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

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

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

Jovanovic, B. and Rousseau, P. L. (2004). General Purpose Technologies. *Handbook of Economic Growth*.

Lanteri, A., Medina, P., and Tan, E. (2019). Capital-Reallocation Frictions and Trade Shocks. *Working Paper*.

Lucas, R. E. (1978). On the Size Distribution of Business Firms. *The Bell Journal of Economics*, 9:508–523.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71:1695–1725.

Midrigan, V. and 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. and Prescott, E. C. (1994). Barriers to Technology Adoption and Development. *Journal of Political Economy*, 102(2):298–321.

Poschke, M. (2018). The Firm Size Distribution Across Countries and Skill-Biased Change in Entrepreneurial Technology. *American Economic Journal: Macroeconomics*, 10:1–41.

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

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

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

Online Appendix (Not for Publication)

A Equilibrium Proof

Proof of Proposition 1 Following Parente and Prescott (1994), the investment cost function can be rewritten in terms of technology proximity as

$$\varphi_{j,t} = (1-\delta)\varphi_{j,t-1} + \frac{1}{\pi}x_{j,t}.$$

where technology drift is defined as $(1-\delta) = e^{-\xi g}$ such that the proximity of a fixed technology will decay by δ each period. The production function written in terms of technology proximity is

$$y_{j,t} = \left(\bar{z}_t (As_{j,t})^{1-\omega}\varphi_{j,t}^{\omega}\right)^{1-\gamma} n_{j,t}^{\gamma}. \tag{28}$$

To solve the value function, I use the guess and verify method with the guess:

$$v_{(s,\theta)}(\varphi_{-1}) = w\bar{v}_{(s,\theta)}(\varphi_{-1}) = w\left[K\theta^{\frac{1}{1-\omega}}s + \pi(1-\delta)\varphi_{-1}\right],$$

where K is a constant. The firm's problem can be written in terms of the normalized value as

$$\bar{v}_{(s,\theta)}(\varphi_{-1}) = \max_{n,\varphi} \left(\bar{z}\theta(As)^{1-\omega}\varphi^{\omega}\right)^{1-\gamma} n^{\gamma} - wn - w\pi(\varphi - (1-\delta)\varphi_{-1}) + D\bar{v}_{(s,\theta)}(\varphi).$$

The first-order conditions imply the solution to labor and technology proximity given in the main text

$$n(s,\theta) = \left(\frac{\gamma\bar{z}^{1-\gamma}}{w}\right)^{\frac{1}{1-\omega}\frac{1}{1-\gamma}} \eta^{\omega}\theta^{\frac{1}{1-\omega}}s, \qquad \textit{and} \qquad \varphi(s,\theta) = \eta\theta^{\frac{1}{1-\omega}}s,$$

where η is defined as in the main text. Substituting the expressions into the value function gives

$$\bar{v}_{(s,\theta)}(\varphi_{-1}) = A\left[\left(\frac{\bar{z}^{1-\gamma}}{w}\right)^{\frac{1}{1-\gamma}\frac{1}{1-\omega}}\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma)\left[\frac{\omega\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma)}{\pi(1-D(1-\delta))}\right]^{\frac{\omega}{1-\omega}}\right]\theta^{\frac{1}{1-\omega}}s$$
$$-\pi\left[\eta\theta^{\frac{1}{1-\omega}}-(1-\delta)\varphi_{-1}\right]+D\left[K\theta^{\frac{1}{1-\omega}}s+\pi(1-\delta)\eta\theta^{\frac{1}{1-\omega}}s\right].$$

The above expression solves

$$K = A\frac{1-\omega}{1-D}\left[\frac{\omega^{\frac{\omega}{1-\omega}}(\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma))^{\frac{1}{1-\omega}}}{(\pi(1-De^{-\xi g}))^{\frac{\omega}{1-\omega}}}\right](\bar{w})^{-\frac{1}{1-\omega}\frac{1}{1-\gamma}},$$

and confirms the initial guess.

On the balanced growth path, the technology frontier $\bar{z}_t$ grows at rate g and the wage rate w_t grow at rate $(1-\gamma)g$ implying that $\bar{v}_t(s,\theta)$ is constant.

Proof of Proposition 2 From the expressions in Proposition 1, the entry condition can be written as

$$c_e = A\frac{1-\omega}{1-D}\left[\frac{\omega^{\frac{\omega}{1-\omega}}(\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma))^{\frac{1}{1-\omega}}}{(\pi(1-De^{-\xi g}))^{\frac{\omega}{1-\omega}}}\right](\bar{w})^{-\frac{1}{1-\omega}\frac{1}{1-\gamma}}\int_{\mathcal{S}\times\theta}\left[\theta^{\frac{1}{1-\omega}}s\right]h(s,\theta)\,\mathrm{d}s\,\mathrm{d}\theta.$$

Rearranging gives

$$w = \left(\frac{\bar{z}A^{1-\omega}}{\pi^{\omega}}\right)^{1-\gamma}\left[\frac{1-\omega}{1-D}\left[\frac{\omega^{\frac{\omega}{1-\omega}}(\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma))^{\frac{1}{1-\omega}}}{(1-De^{-\xi g})^{\frac{\omega}{1-\omega}}}\right]\int_{\mathcal{S}\times\Theta}\hat{\theta}^{\frac{1}{1-\omega}}\hat{s}h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}\right]^{(1-\gamma)(1-\omega)}.$$

The expressions also shows that the growth rate of wages is equal to $(1-\gamma)g$

Output can then be solved by integrating over the production of all firms and using the expressions for labor and technology proximity form Proposition 1:

$$\begin{aligned}Y &= M\int_{\mathcal{S}\times\Theta}y(\hat{s},\hat{\theta})h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}\\ &= \zeta_Y M\left(\frac{\bar{z}A^{1-\omega}}{\pi^{\omega}}\right)^{1-\gamma}\left[\frac{\int_{\mathcal{S}\times\theta}\hat{\theta}^{\gamma+\frac{\omega}{1-\omega}}\hat{s}h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}}{\left(\int_{\mathcal{S}\times\Theta}\hat{\theta}^{\frac{1}{1-\omega}}\hat{s}h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}\right)^{\gamma+\omega(1-\gamma)}}\right]\end{aligned}$$

where $\zeta_Y = \frac{1}{1-\gamma}\left(\frac{1-D}{1-\omega}\right)^{\gamma+\omega(1-\gamma)}\left[\frac{\omega^{\frac{\omega}{1-\omega}}(\gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma))^{\frac{1}{1-\omega}}}{(1-De^{-\xi g})^{\frac{\omega}{1-\omega}}}\right]^{(1-\gamma)(1-\omega)}$. It also follows from the above expression that output will grow at a rate $g_Y = (1-\gamma)g + g_M$ where I show next that $g_M = 0$.

The labour clearing condition is used to solve for the equilibrium mass of firms. The labor market clearing condition is given by

$$1 = M\left[\int_{\mathcal{S}\times\Theta}\left[n(\hat{s},\hat{\theta}) + \pi(1-e^{-g\xi}(1-\lambda))\varphi(\hat{s},\hat{\theta})\right]h(\hat{s},\hat{\theta})d\hat{s}d\hat{\theta} + \lambda c_E\right].$$

Along with the expressions for labor and technology proximity, the above expression can be rearranged to show that

$$M = \left[\left(\frac{1-D}{1-\omega}\right)\left[\frac{\gamma}{\omega(1-\gamma)} + \frac{1-e^{-\xi g}(1-\lambda)}{1-De^{-\xi g}}\right] + \lambda c_e\right]^{-1}.$$

The above expression confirms that the mass of firms is constant along the balanced growth path, $g_M = 0$.

It follows from the expressions of wage, output and the mass of firms that the BGP values grow at constant rates if the distribution of firm types (s,θ) is stationary. Further, the wage expression implies that $\bar{z}_t^{1-\gamma}/w_t$ is constant if this condition holds. From the expressions for labour demand and technology proximity it follows that allocation across firms are stationary if the $\bar{z}^{1-\gamma}/w_t$ is constant and the distribution of types (s,θ) is stationary. Uniqueness and existence of a BGP equilibrium then follow from the stationarity of the distribution of ability s and distortions θ. More generally, with some modification of the expressions, the uniqueness and existence of the BGP equilibrium follows from s following an ergodic distribution.

Proof of Proposition 3 I show that aggregate output in the economy with fixed cost λc_e, a mass M of firms, aggregate barrier $\pi^{rep} = \pi$ and type (s^{rep}, θ^{rep}) equal to

$$s^{rep} = \left[\frac{\int_{\mathcal{S}\times\theta}\left[\hat{\theta}^{\gamma+\frac{\omega}{1-\omega}}\hat{s}\right]h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}}{\left(\int_{\mathcal{S}\times\Theta}\hat{\theta}^{\frac{1}{1-\omega}}\hat{s}h(\hat{s},\hat{\theta})\,\mathrm{d}\hat{s}\,\mathrm{d}\hat{\theta}\right)^{\gamma+\omega(1-\gamma)}}\right]^{\frac{1}{1-\omega}\frac{1}{1-\gamma}}$$

and $\theta^{rep} = 1$ is identical to the heterogeneous firm model described in the main text.[29]

[29]The mass of firms could be set to one in the representative firm economy with the appropriate adjustments to aggregate productivity (A in the baseline model) and aggregate barriers π. The

Output of a firm (s, θ) can be written as:

$$
\begin{aligned}
y(s,\theta) &= \left[\bar{z}_t \varphi(s,\theta)^{\omega} (As)^{1-\omega}\right]^{1-\gamma} n(s,\theta)^{\gamma} \\
&= \bar{z}\gamma^{\frac{\gamma}{1-\gamma}} \left(\frac{\bar{z}^{1-\gamma}}{w}\right)^{\frac{\gamma}{1-\gamma}} \eta^{\omega} A \theta^{\frac{\omega}{1-\omega}+\gamma} s.
\end{aligned}
$$

Through substitution of η and the wage rate w it follows that output of a firm (s, θ) can be written as

$$
y(s,\theta) = \zeta_Y \left(\frac{\bar{z}A^{1-\omega}}{\pi^{\omega}}\right)^{1-\gamma} \frac{\theta^{\gamma+\frac{\omega}{1-\omega}} s}{\left(\int_{\mathcal{S}\times\Theta} \hat{\theta}^{\frac{1}{1-\omega}} \hat{s} h(\hat{s},\hat{\theta})\, \mathrm{d}\hat{s}\, \mathrm{d}\hat{\theta}\right)^{\gamma+\omega(1-\gamma)}},
$$

and that

$$
y(s,1) = \zeta_Y \left(\frac{\bar{z}A^{1-\omega}}{\pi^{\omega}}\right)^{1-\gamma} s^{(1-\gamma)(1-\omega)},
$$

Then, it follows from the above expression that output is equivalent at the aggregate level to an economy with a representative firm with mass M, ability s^{rep}.

It is not generally the case that the wage rate in the representative firm model will be the same as in the heterogeneous firm model. However, it is always possible to scale the average value of distortions θ to ensure that the wage rates are equal in both economy, which does not affect aggregate output (or allocation or technology). Consequently, the representative firm model can be compared to the heterogeneous model with the same wage rate without loss of generality.

Finally, I show that the allocation of labour across production (N_P), adoption (N_X) and fixed costs (N_E) is the same in the two models. Note that fixed costs in the representative firm model take the place of entry costs in the heterogeneous firm model and use $M\lambda c_e$ units of labour in both cases. It then follows from the proof of Proposition 2 that the allocation of labour to production and adoption is the same in the two models.

Proof of Proposition 4 Follows from definition of $L(s, \theta)$.

intuition is the same.

Proof of Proposition 5 The inclusion of technology implies that a firm's TFPQ depends on its institutional distortion. The measured elasticity then needs to account for this relationship to be compared with the value of ρ in the calibration. I show the mapping for the decreasing returns to scale (DRS) model and the CES model used by Hsieh and Klenow (2009) and Bento and Restuccia (2017).

In both cases, let the empirical relationship be given by $\ln \text{TFPR} = \hat{\beta}_0 + \hat{\beta}_1 \ln \text{TFPQ}$.

DRS Case: It follows from the expression of TFPQ that

$$\begin{aligned}\ln \text{TFPQ}(s,\theta) &= \omega(1-\gamma)\ln \eta\theta^{\frac{1}{1-\omega}} s + (1-\omega)(1-\gamma)\ln s \\ &= (1-\gamma)\left[\omega \ln \eta + \frac{\omega}{1-\omega}\ln\theta + \ln s\right] \\ &= \omega(1-\gamma)\ln\eta - \frac{\omega}{1-\omega}\ln \text{TFPR}(s,\theta) + (1-\gamma)\ln s \\ &= \omega(1-\gamma)\ln\eta - \frac{\omega}{1-\omega}\left[\rho\ln s - \varepsilon\right] + (1-\gamma)\ln s\end{aligned}$$

where the last line follow from the definition of distortions in (21). Using the empirical relationship, it follows that the measured elasticity implies the relationship between $1-\tau$ and s is given by:

$$\ln(1-\tau) = -\hat{\beta}_0 - \hat{\beta}_1\left[\omega(1-\gamma)\ln\eta - \frac{\omega}{1-\omega}\left[\rho\ln s - \varepsilon\right] + (1-\gamma)\ln s\right],$$

which implies that:

$$\rho = \hat{\beta}_1\left[(1-\gamma) - \frac{\omega}{1-\omega}\rho\right],$$

The final expression for ρ is then given by:

$$\rho = (1-\gamma)\left[\frac{1}{\hat{\beta}_1} + \frac{\omega}{1-\omega}\right]^{-1}.$$

CES Case: Let ν denote the CES parameter in the model, the full model is outlined in Appendix F. The derivation of the relationship follows the same steps as in the DRS case, but the expression differ slightly to account for differences in variable relationships between the two models. TFPQ is given by:

$$\begin{aligned}\ln \text{TFPQ}(s,\tau) &= \frac{1}{\nu-1}\left[\omega \ln \eta^{\omega}(1-\tau)^{\nu\frac{\omega}{1-\omega}} s^{\omega} + (1-\omega)\ln s\right] \\ &= \frac{\omega}{\nu-1}\ln \eta^{\omega} + \left[\frac{1}{\nu-1} - \tilde{\rho}\frac{\omega}{1-\omega}\frac{\nu}{\nu-1}\right]\ln s\end{aligned}$$

Using the empirical relationship it follows then that the mapping between the elasticity $\tilde{\rho}$ and the measured elasticity is given by:

$$\tilde{\rho} = \frac{1}{\nu-1}\left[\frac{1}{\hat{\beta}_1} + \frac{\omega}{1-\omega}\frac{\nu}{\nu-1}\right]^{-1}.$$

Finally, an additionally adjustment is needed to convert from the elasticity in the CES model to the DRS model (discussed in Appendix F.3). The elasticity is then given by

$$\rho = \frac{\sigma}{\nu-1}\tilde{\rho} = \frac{\nu}{(\nu-1)^2}\left[\frac{1}{\text{ELAS}} + \frac{\omega}{1-\omega}\frac{\nu}{\nu-1}\right]^{-1}.$$

Consolidated Mapping: I write the relationship as:

$$\rho = \Gamma_1\left[\frac{1}{\text{ELAS}} + \frac{\omega}{1-\omega}\Gamma_2\right]^{-1}$$

where Γ_1 and Γ_2 are parameters that depend on the empirical model.

Alternative Approach An alternative approach would be to assume that the model is correct and find the value of ρ necessary to match the data as it is measured under the alternative data. In this case, TFPQ is incorrectly measured in the empirical analysis as $\widetilde{TFPQ}$, which is used in the measurement of elasticity implying that the empirical relationship is given by $\ln \text{TFPQ} = \hat{\beta}_0 + \hat{\beta}_1 \ln \widetilde{TFPQ}$. The relationship between $\widetilde{TFPQ}_j$ and the true TFPQ_j for firm j is given by $\widetilde{TFPQ}_j = (p_j y_j)^{\frac{\nu}{\nu-1}}/n_j = \text{TFPQ}^{\frac{1}{\nu-1}\frac{1}{1-\gamma}}$. Together, this implies that the elasticity of distortions is equal to:

$$\rho = (1-\gamma)\left[\frac{(1-\gamma)(\nu-1)}{\hat{\beta}_1} + \frac{\omega}{1-\omega}\right]^{-1}.$$

Proof of Proposition 6 The adoption lag of firm (s, θ) is given by

$$L(s,\theta) = -\frac{1}{g\xi} \ln \eta\theta^{\frac{1}{1-\omega}} s.$$

Using the equilibrium expressions, it follows that $\varphi(s,\theta) \propto n(s,\theta)$. It also follows that $g = \frac{g_Y}{1-\gamma}$. Comparing the adoption lag between firms j and j' can then be written as:

$$L(s_j, \theta_j) - L(s_{j'}, \theta_{j'}) = -\frac{1-\gamma}{g_Y \xi} \ln n_j / n_{j'}.$$

The 10-90 lag involves comparing firms based on their relative output. Defining n_{10} and n_{90} as the employment of the marginal firms that produce the 10th and 90th percentiles of output then implies that:

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

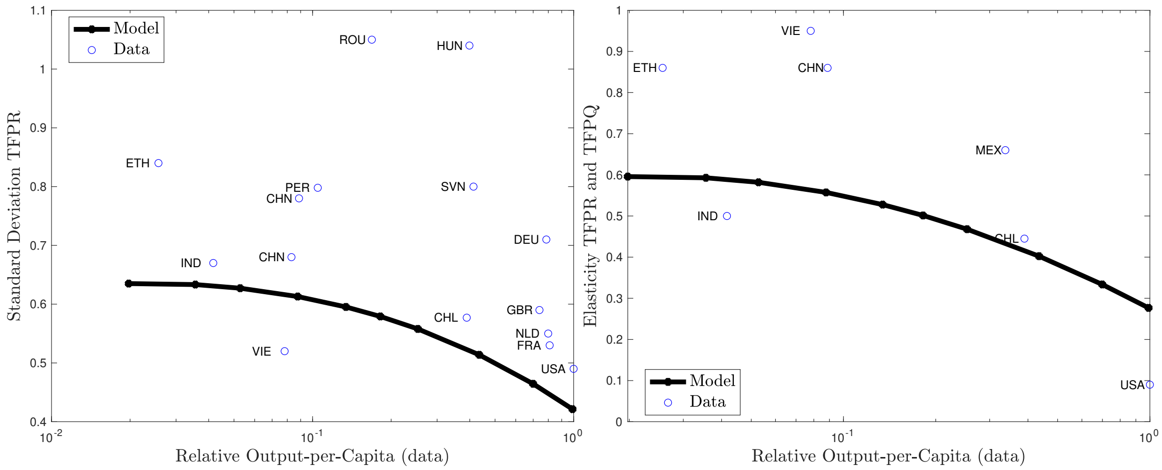
B Data Comparison

The data used to construct the elasticity of distortions ρ is from the World Bank Enterprise Survey. While this data is collected in order to be representative of firms within a country, it has a number of limitations. Additionally, the noise term of distortions σ is held at the US level for all counterfactual economies. To examine the reasonableness of these assumptions Figure 6 compares the values of moments from the calibrated model with those used in other misallocation studies.

Figure 6 shows that the standard deviation of TFPR and the elasticity of TFPR and TFPQ generated by the model are comparable with those found using high quality micro data. The values tend to be on the conservative side with many low income countries featuring higher elasticities relative to the calibration.

Table 2 reports the data used in the construction of Figure 6 and the corresponding sources. For the majority of the data in Table 2, the data is reported as in the cited paper. Many of the countries are examined in multiple papers and over multiple years. I use the source with either the most complete information (e.g. SD and Elas are listed) or with reported values within the range of other studies or years. As the values are not directly used in the calibration, this does not affect the results. For many of the empirical examinations of misallocation, the elasticity between TFPR

Figure 6: Calibrated Distortions



(a) Standard Deviation Log TFPR (b) Elasticity TFPR and TFPQ

Notes: Figure plots the model implied standard deviation of log TFPR and elasticity of TFPR and TFPQ against empirical moments for several sample countries.

and TFPQ is not reported and so the sample size is substantially smaller.

The table contains information from both studies on the manufacturing and agricultural sectors. Many of the agricultural studies focus on the most underdeveloped countries. While there is concern about comparison across industries, the agricultural sector employs the majority of labour in lower income countries implying that it is important for understanding how institutional distortions affect technology adoption in these economies.

C Sensitivity of Results

This Appendix Section discusses the sensitivity of the results with respect to the choice of technology cost curvature ξ. Technology costs depends on both the aggregate barriers to technology adoption π and the cost curvature parameter ξ. The aggregate barrier π does not interact with the effect of idiosyncratic policy distortions, and so it has no affect on the main results. On the other hand, the cost curvature ξ does have consequences for the effect of idiosyncratic distortions on aggregate productivity and technology adoption. I consider the sensitivity of the main results to the calibrated value of ξ. For brevity, I report only the results for the lowest income economies that correspond to the bottom decile of income-per-capita and focus only on variation in

Table 2: Data Sources

Country		year	Rel. Inc.	SD	Elas	Paper
China	CHN	2001	0.08	0.68		Hsieh and Klenow (2009)
India	IND	1991	0.04	0.67	0.5	Hsieh and Klenow (2009, 2014)
US	USA	1997	1.00	0.49	0.09	Hsieh and Klenow (2009, 2014)
Mexico	MEX	1998-2008	0.34		0.66	Hsieh and Klenow (2014)
Chile	CHL	1983	0.32	0.72	0.694	Chen and Irarrazabal (2014)
U.K.	GBR	1993-2002	0.74	0.59		Bartelsman et al. (2013)
Germany	DEU	1993-2002	0.79	0.71		Bartelsman et al. (2013)
France	FRA	1993-2002	0.81	0.53		Bartelsman et al. (2013)
Netherlands	NLD	1993-2002	0.80	0.55		Bartelsman et al. (2013)
Hungary	HUN	1993-2002	0.40	1.04		Bartelsman et al. (2013)
Romania	ROU	1993-2002	0.17	1.05		Bartelsman et al. (2013)
Slovenia	SVN	1993-2002	0.41	0.8		Bartelsman et al. (2013)
Peru	PER	2000-2014	0.10	0.798		Lanteri et al. (2019)
Vietnam	VIE	2006-2010	0.06	0.42	0.94	Ayerst et al. (2020)
China	CHN	1993-2002	0.09	0.78	0.86	Adamopoulos et al. (2017)
Ethiopia	ETH	2011-2016	0.03	0.84	0.86	Chen et al. (2019)

Notes: Rel. Income refers to real income per worker in the listed country relative to real income per worker in the United States for the same base year. All income and employment data is from the Penn World Tables 9.1. SD refers to the standard deviation of log TFPR. Elas refers to the elasticity between TFPR and TFPQ.

the institutional distortions through ρ_i (holding π_i fixed). To facilitate comparability, all parameters other than ξ are held at the baseline calibration values.

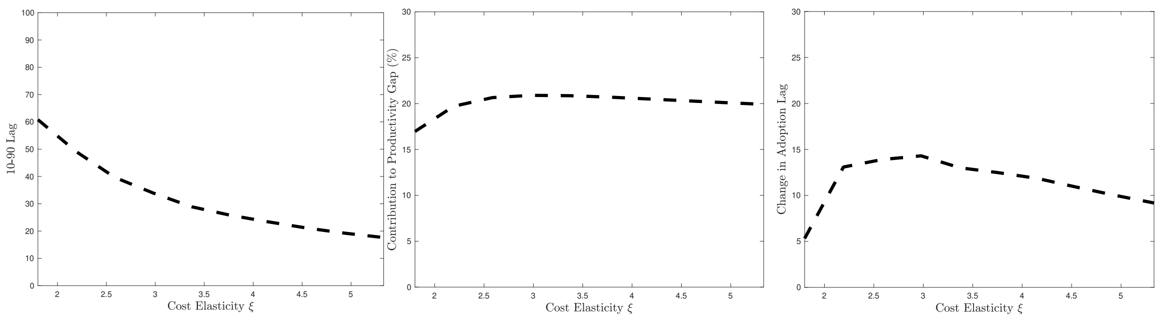
In the baseline calibration, the value of ξ is disciplined to a 10-90 lag of 40 years, which corresponds to an average value for a wide range of technologies in the literature. On the lower end of the reported values, Jovanovic and Lach (1997) find an average 10-90 lag of 15 years for 21 technologies. While this number is substantially smaller than considered in the baseline calibration, it is based on consumer-oriented technologies that are likely less relevant to focus on production in the model. In general, broader production technologies (e.g. tractors, electricity, computers) tend to have longer adoptions lengths, more in line with the baseline calibration. Despite this, I consider an upper bound on the curvature of $\xi = 5.3$, which implies a 10-90 adoption lag of around 17 years. On the other extreme, I set the lower bound on the curvature of $\xi = 1.8$ (or $\omega = 0.55$), which is similar to the calibrated value used by Parente and Prescott (1994).[30] The value of ξ implies a 10-90 lag of 61 years, which

[30]There is no direct mapping between my model and Parente and Prescott (1994) because of the different treatment of labour. I set $\omega = 0.55$ based on the elasticity between technology and

is substantially longer than the values reported for individual technologies.

Figure 7 reports the sensitivity of the main results under different values of ξ. Figure 7a reports the 10-90 lag in the benchmark economy under the alternate value of the technology cost curvature ξ. Figure 7b and 7c report the gains in productivity and decline in adoption lag attributable to moving from the most severely distorted economy to the benchmark economy level of distortions.

Figure 7: Sensitivity of Results to ξ



(a) 10-90 Lag (b) Productivity Contribution (c) Cross-Country Adopt Lag

Notes: Figure (a) plots the 10-90 lag for the Benchmark Economy under different values of the technology cost curvature ξ. Figure (b) plots the counterfactual gain in productivity from moving from the most distorted economy (lowest decile output-per-capita) to the benchmark economy. Figure (c) plots the change in adoptions lags from moving from the most distorted economy (lowest decile output-per-capita) to the benchmark economy.

The figure shows that the experiment for productivity and the adoption lag have hump shapes with respect to the technology cost elasticity ξ. The hump shapes for productivity and adoption are driven by two distinct but related effects. First, for productivity, both the benchmark and counterfactual economies are distorted. Increasing ξ monotonically affects productivity but to different extents in the benchmark and counterfactual experiments. The effects of misallocation tend to be more sensitive to changes in ξ for the benchmark economy than the counterfactual economy. Second, the value of ξ determines the point in which firm ranking (by employment or technology adoption) is reversed. At lower levels of ξ this reversal occurs at lower values of ρ. This implies that at low levels of ξ a further decrease in ξ causes early adopting low productivity firms to adopt technology earlier decreasing the adoption lag. Finally, note that while these are the mechanical drivers of the hump shape, the

productivity taking the response of labour as given.

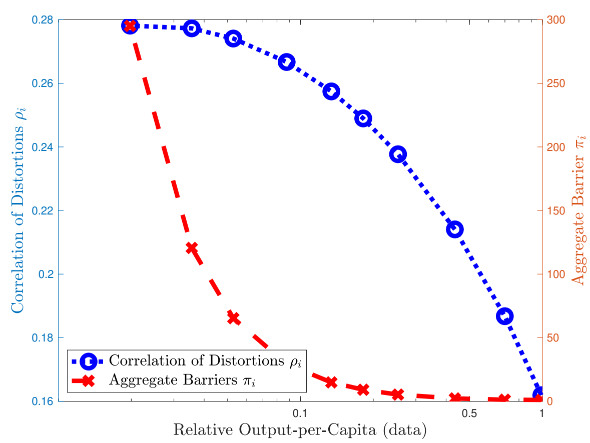
lack of recalibration is also an important component. Changing ξ has a direct affect on the magnitude of ρ and the consequences for productivity and adoption.

In terms of magnitude, the figures show that the results are generally robust to changes in the underlying value of ξ. The relative contribution of productivity is relatively robust throughout the examined values of ξ, with only a slight decline at lower values. While the cross-country adoption lag is more sensitive to ξ, it requires substantial changes in the value of ξ to significantly change the results. Further, decreasing the value of ξ results in counterfactually long 10-90 lags, which are not supported by data. The decline in the importance of institutional distortions for cross-country adoption lags is expected at higher values of ξ since this corresponds to the case where technology becomes diffused through the economy more rapidly. Consequently, there is less scope for institutional distortions because technologies used by firms are more similar, at least in terms of adoption lags.

D Additional Model Details

Figure 8 reports the calibrated values of ρ_i and π_i used in the cross-country analysis. The values are based on the data moments in Figure 2. Both the correlation of distortions ρ_i and the aggregate barriers to technology adoption π_i tend to be larger in lower income countries.

Figure 8: Calibrated Values of ρ and π



E Alternative Cost Function

In this Appendix section, I examine the sensitivity of the results to the functional form used for the cost function. While this functional form is commonly used in the technology adoption literature (e.g. Parente and Prescott, 1994; Comin and Hobijn, 2010), it leads to firm-level technology decisions that are independent from the firm's initial level of technology. In this section, I consider an alternative cost function in which the costs of new technologies become relatively cheaper if the firm uses better technologies. I then show that when the same parameters are used, the quantitative implications of the alternative cost function are the same as the baseline model.

E.1 Economic Environment

The economy is similar to the baseline model with the exception of the cost function. I define scaled ability $a = s^{1-\omega}$ to simplify some of the notation and refer to firms based on their scaled ability a and institutional distortion θ. To simplify exposition, I focus on the calibration case where firm types (a, θ) are fixed following entry. Throughout the section, I use tildes on parameters to denote parameters with the same interpretation as in the baseline model but with (potentially) different quantitative implications.

As in the baseline model, the state of technology is described by the technology frontier $\bar{z}_t$, which grows exogenously at rate $\tilde{g} \geq 0$. Improvements in the technology frontier reduces the costs to firms of adopting new technologies. The main difference with the cost function is now the curvature of the cost function is on the relative increase in technology, rather than on the level of technology. Specifically, the cost function is now given by

$$x(z_{t+1}, z_t) = \tilde{\pi} \left(\frac{z_{t+1} - z_t}{\bar{z}_t} \right)^{\tilde{\xi}}, \tag{29}$$

where $\tilde{\xi}$ and $\tilde{\pi}$ denote the curvature of the cost function and aggregate barriers to adoption, as in the baseline model. As in the baseline model, the adoption cost is paid in labour.

In the baseline model, the cost function is interpreted as capturing a menu of options that a firm faces when making investment decisions. That is, the cost represents the cost of purchasing, installing and learning to use a set of distinct technologies, In contrast, the alternative cost function in (29) can be thought of as capturing a specific type of dependence between technologies. In this interpretation, better technologies become less costly because the firm has accumulated some capacity or understanding of its current technologies (e.g., learning-by-doing).

The cost function now creates a dependence between the cost necessary to reach a particular technology level z_{t+1} and the firm's current technology z_t. Define the scaled technology as $\check{z}_t = z_t/\bar{z}_t$, where the interpretation of scaled technology is similar to proximity φ in the baseline model. The cost function in terms of the scaled technology is equal to

$$x(\check{z}_{t+1}, \check{z}_t) = \pi \left(e^{\tilde{g}} \check{z}_{t+1} - \check{z}_t\right)^{\tilde{\xi}}.$$

E.2 BGP Properties

I focus on the properties of the balanced growth path equilibrium for the analysis.

Firm's Problem Firm's are described by their ability a, institutional distortion θ, and current technology $\check{z}_t$. Firms choose labour equal to $n_{(a,\theta)}(\check{z}_t) = \left(\frac{\gamma \bar{z}_t^{1-\gamma}}{w_t}\right)^{\frac{1}{1-\gamma}} \check{z}_t \theta a$ in period t to maximize profit. The dynamic problem of the firm is to choose future scaled technology to maximize value:

$$v_{(a,\theta)}(\check{z}_t) = \max_{\check{z}_{t+1}} \tilde{\gamma} w_t \left(\frac{\bar{z}_t^{1-\gamma}}{w_t}\right)^{\frac{1}{1-\gamma}} \theta a \check{z}_t - w_t \tilde{\pi} \left(e^{\tilde{g}} \check{z}_{t+1} - \check{z}_t\right)^{\tilde{\xi}} + D v_{(a,\theta)}(\check{z}_{t+1}).$$

where $\tilde{\gamma} = \gamma^{\frac{\gamma}{1-\gamma}}(1-\gamma)$. Solving the firm's problem shows that the optimal adoption investment of a firm (a, θ) is equal to

$$x_{(a,\theta)} = \tilde{\pi} \left[\left(\frac{\bar{z}_t^{1-\gamma}}{w_t}\right)^{\frac{1}{1-\gamma}} \frac{\tilde{\gamma} D e^{-\gamma \tilde{g}} \theta a}{\tilde{\pi} \tilde{\xi} (1 - D e^{-\gamma \tilde{g}})} \right]^{\frac{\tilde{\xi}}{\tilde{\xi}-1}}, \tag{30}$$

where it follows that investment is increasing in the firm's type (a, θ) and falling in the aggregate barriers to technology adoption $\tilde{\pi}$. However, investment does not depend on the firm's current technology level z_t indicating that all firms that share a type (a, θ) choose the same investment level.

Substituting the expression in (30) into the firm's value function and solving yields:

$$\begin{aligned} v_{(a,\theta)}(\check{z}) = \; & w_t \left[\frac{\tilde{\gamma}}{1 - De^{-\gamma\tilde{g}}} \left(\frac{\bar{z}^{1-\gamma}}{w} \right)^{\frac{1}{1-\gamma}} \check{z} a \right] \\ & + w_t \left[\frac{1}{1 - De^{(1-\gamma)\tilde{g}}} \left[\frac{\tilde{\gamma} D e^{-\gamma\tilde{g}}}{\tilde{\xi}\tilde{\pi}(1 - De^{-\gamma\tilde{g}})} \left(\frac{\bar{z}^{1-\gamma}}{w} \right)^{\frac{1}{1-\gamma}} a \right]^{\frac{\tilde{\xi}}{\tilde{\xi}-1}} \pi(\tilde{\xi} - 1) \right], \end{aligned}$$

where I use that $\bar{z}_t^{1-\gamma}/w_t$ is constant on the BGP equilibrium in the derivation of the above expression. The value function is comprised of two terms which represent the current discounted value of the technology $\check{z}$ used by the firm and the discounted profits from technology improvements through adoption.

Technology Along the BGP, the value of $x_{(a,\theta)}$ is constant. Using the expression for x from the cost function, the technology value of a firm is equal to:

$$\begin{aligned} \check{z}_{(a,\theta)}(k) = \frac{z_{(a,\theta),t}(k)}{\bar{z}_t} &= \sum_{\hat{k}=1}^{k-1} \frac{\bar{z}_{t-\hat{k}}}{\bar{z}_t} \left(\frac{x_{(a,\theta)}}{\pi} \right)^{\frac{1}{\xi}} \\ &= \left[\sum_{\hat{k}=1}^{k-1} e^{-g\hat{k}} \right] \left(\frac{x_{(a,\theta)}}{\pi} \right)^{\frac{1}{\xi}} \approx k \left(\frac{x_{(a,\theta)}}{\tilde{\pi}} \right)^{\frac{1}{\xi}}, \end{aligned}$$

where k is the number of periods since the firm has entered the market.[31] Any two firms with the same type (a, θ) and age k will use the same scaled technology $\check{z}_{(a,\theta)}(k)$ regardless of the current period.

It follows from the entry and exit structure of the economy that the distribution of age k firms is $h_k(k) = \lambda(1-\lambda)^k$ with mean value $1/\lambda$. The average technology used by type (a, θ) firms is equal to

$$\bar{\check{z}}_{(a,\theta)} = \sum_{k=1}^{\infty} \check{z}_{(a,\theta)}(k) = \frac{1}{\lambda} \left(\frac{x_{(a,\theta)}}{\tilde{\pi}} \right)^{\frac{1}{\xi}},$$

where the expression follows from the average age of firms being $1/\lambda$, regardless of

[31] This model could be easily extended to include uncertainty in investment outcomes. The main difference would be that k would then be interpreted as the number of successful technology improvements and the distribution would be adjusted to reflect the possibility of failure.

type (s,θ). The relative average technology used between two firm types is equal to:

$$\frac{\bar{\check{z}}_{(a,\theta)}}{\bar{\check{z}}_{(a',\theta')}} = \left(\frac{\theta a}{\theta' a'}\right)^{\frac{1}{\tilde{\xi}-1}} . \tag{31}$$

Both the baseline model and the alternative cost function model highlight the same relationship between underlying productivity and average technology. The expression is not quantitatively comparable to the baseline model because ability a does not have natural units.

Adoption The definition of adoption is the same as in the baseline model. A new technology has scaled value $\check{z}_L = e^{-\tilde{g}L}$ after L periods. This implies that the optimal lag of a firm (a,θ) at age k is given by

$$L_{(a,\theta)}(k) = -\frac{1}{\tilde{g}} \ln k - \frac{1}{\tilde{g}\tilde{\xi}} \ln\left(\frac{x_{(a,\theta)}}{\tilde{\pi}}\right) . \tag{32}$$

Similarly to the baseline model, the firm's optimal adoption lag is declining in the firm's ability a and institutional distortion θ. The addition of k in the expression reflects that firms use more productive technologies over time indicating that older firms adopt newer technologies. The expression in (32) can be used to derive similar comparative statics as in Proposition 4.

BGP Characterization The equilibrium wage rate and output are given by:

$$w_t = \left(\frac{\tilde{\gamma} D e^{-\gamma\tilde{g}}}{\tilde{\xi}\tilde{\pi}(1 - De^{-\gamma\tilde{g}})} \left[\frac{\tilde{\pi}(\tilde{\xi}-1)}{c_e(1 - De^{(1-\gamma)\tilde{g}})}\right]^{\frac{\tilde{\xi}-1}{\tilde{\xi}}} \bar{z}_t \tilde{a}\right)^{1-\gamma},$$

$$Y_t = M\tilde{\pi}^{-\frac{1-\gamma}{\tilde{\xi}}} \left[\frac{\tilde{\gamma} D e^{-\gamma\tilde{g}}}{\tilde{\xi}(1 - De^{-\gamma\tilde{g}})}\right]^{-\gamma} \left[\frac{\tilde{\xi}-1}{c_e(1 - De^{(1-\gamma)\tilde{g}})}\right]^{\frac{-(1-\gamma)\tilde{\xi}-1}{\tilde{\xi}}} \tilde{\chi}\left(\bar{z}_t\tilde{a}\right)^{1-\gamma},$$

where $\tilde{a} = \left(\int_{\Theta\times\mathcal{A}} \left(\hat{\theta}\hat{a}\right)^{\frac{\tilde{\xi}}{\tilde{\xi}-1}} \mathrm{d}\hat{a}\,\mathrm{d}\hat{\theta}\right)^{\frac{\tilde{\xi}-1}{\tilde{\xi}}}$ and $\tilde{\chi} = \left(\int_{\Theta\times\mathcal{A}} \hat{\theta}^{\gamma+\frac{1}{\tilde{\xi}-1}} \hat{a}^{\frac{\tilde{\xi}}{\tilde{\xi}-1}} \mathrm{d}\hat{a}\,\mathrm{d}\hat{\theta}\right)/\tilde{a}^{\frac{\tilde{\xi}}{\tilde{\xi}-1}}$. The mass of firms is given by:

$$M = \left[\lambda c_e + \left[\frac{\tilde{\gamma} D e^{-\gamma\tilde{g}}}{\tilde{\xi}(1 - De^{-\gamma\tilde{g}})}\right]\left[\frac{\tilde{\xi}-1}{c_e(1 - De^{(1-\gamma)\tilde{g}})}\right]\left[\left(\frac{\tilde{\gamma} D e^{-\gamma\tilde{g}}}{\tilde{\xi}(1 - De^{-\gamma\tilde{g}})}\right) + \frac{\gamma^{\frac{1}{1-\gamma}}}{\lambda}\right]\right]^{-1},$$

which is constant along the BGP and does not depend on the distribution of institutional distortions or aggregate barriers $\tilde{\pi}$.

E.3 Comparison with Baseline Model

I conclude the section by discussing the parameterizations under which the main quantitative experiments would produce the same results in both the appendix and baseline models. It is not necessary for the appendix model to match the level of variables in the baseline model since the quantitative experiments examines relative outcomes, either over time or across economies.

Parameter Mapping I use the implied model moments to construct a mapping between parameter values in each model. The appendix model has three parameters $(\tilde{g}, \tilde{\xi}, \tilde{\pi})$, which correspond to the baseline model parameters (g, ξ, π). Throughout the section, I assume that the underlying distribution of abilities s (and implied value a) and institutional distortions θ are equal in the two models.

The comparisons of the growth rates and the aggregate barriers are straightforward. The technology growth rates are compared through their effect on output growth, which implies that $\tilde{g} = \frac{g_Y}{1-\gamma} = g$. The elasticity of the adoption lag $L_{(a,\theta)}(k)$ with respect to $\tilde{\pi}$ is equal to $1/g\xi$, which also implies that $\tilde{\pi} = \pi$.

In the baseline model, the choice of technology is closely related to the employment of the firm. In the appendix model, a similar intuition follows for the relationship between employment and investment, which can be written as $n_{(a,\theta)}(k) \propto kx_{(a,\theta)}$. Extending the intuition from the baseline model, the 10-90 Lag can be derived in the appendix model:

$$\text{10-90 Lag} = \frac{1-\gamma}{g_Y \tilde{\xi}} \ln n_{90}/n_{10} + \varepsilon,$$

where $\varepsilon = \frac{\tilde{\xi}-1}{\tilde{g}\tilde{\xi}} \ln k_{90}/k_{10}$ is a correction term that accounts for the relative ages of the 90th and 10th percentile firms. Since this is unobserved in the data, I assume that the correction term is equal to zero for the comparison of the models.[32] The above

[32]While the correction term can be positive or negative, it will tend to be positive as older firms are more likely to use better technologies. This will tend to imply a higher value of $\tilde{\xi}$ all else equal.

expression then implies that:

$$\tilde{\xi} = \xi.$$

Aggregate Outcomes The previous results show that setting the values of $(\tilde{g}, \tilde{\xi}, \tilde{\pi})$ equal to (g, ξ, π) imply similar moments in both models. I now show that the implications for aggregate output are the same in both models under this calibration.

The summary statistics $\tilde{a}$ and $\tilde{\chi}$ that capture the effects of average firm ability and institutional distortions on the aggregate economy. From the previous mapping it follows that:

$$\tilde{a} = \left(\int_{\mathcal{S}\times\Theta} \left(\hat{\theta} \hat{s}^{1-\omega} \right)^{\frac{1}{1-\omega}} d\hat{s} d\hat{\theta} \right)^{1-\omega} = \tilde{s}^{1-\omega}$$
$$\tilde{\chi} = \frac{\left(\int_{\mathcal{S}\times\Theta} \hat{\theta}^{\gamma+\frac{\omega}{1-\omega}} \left(\hat{s}^{1-\omega} \right)^{\frac{1}{1-\omega}} d\hat{a} d\hat{\theta} \right)}{\tilde{s}} = \chi.$$

The above expressions then imply that the relationship between the system of distortions and aggregate output is the same in both models. Additionally, it follows that the elasticity of aggregate output with respect to aggregate barriers π is equal to $\omega(1-\gamma)$ in both models.

Discussion The above results show that under the baseline parameters, the two models produce the same quantitative results. Under this interpretation, the baseline model can be thought of as capturing the average characteristics of firms that share the same ability s and institutional distortion θ.

There are two caveats with the comparison of the two models. The first caveat is that the distribution of abilities may not be the same in both models. Within type (s, θ) heterogeneity creates additional dispersion in employment that is not accounted for in the baseline model calibration. This has an ambiguous effect on the quantitative results as it is unclear how this affects the distribution of abilities s. Relatedly, the lag structure in the appendix model depends also on the ages of the marginal adopters. Given the relatively large employment bins used in the calibration of the baseline model, this additionally heterogeneity is unlikely to substantially alter the results.

The second caveat is that the relationship between TFPQ and TFPR implied by the abilities s and institutional distortions θ differs in the appendix model. This is because: (1) the inclusion of within type technology heterogeneity implies that TFPQ cannot be mapped one-to-one into firm type (s, θ); and (2) the inclusion of within type technology heterogeneity implies that the measured elasticity of TFPR with respect to TFPQ is biased towards zero. Both channels would imply that the elasticity ρ_i would need to be larger than in the calibrated model in order to match the data moments.[33] As larger ρ_i amplifies the quantitative results, the results in the baseline model can be viewed as a lower bound.

F CES Model

In this section, I outline the CES version of the model and discuss the properties that are necessary for the CES and baseline model, referred to as the decreasing returns to scale (or DRS) model, to be equivalent. This extends the analysis in Hsieh and Klenow (2009) who show that the static versions of the models have the same aggregate implications when the CES parameter is set to $\nu - 1 = \frac{1}{1-\gamma}$. For this section I assume that this relationship between parameters holds.

For the sake of brevity, I focus this section on the relevant properties of the model for the comparison with the DRS model. With this in mind, I focus primarily on the resulting allocations in the CES model.

F.1 Economic Environment

The model follows the baseline model with the exception of the final consumption goods, which is taken as the numeraire, and the intermediate production technology. The final consumption good aggregates intermediate goods according to:

$$Y_t = \left(\int_j y_j^{\frac{\nu-1}{\nu}} dj \right)^{\frac{\nu}{\nu-1}}$$

where $\nu > 0$ is the elasticity of substitution across goods and j denotes the good produced by firm j.

[33] Alternatively, the distribution of institutional distortions could be modeled as depending on TFPQ rather than ability s. This would create a disincentive for firm's to invest as they become more productive from higher technology z. This would add a within type technology distortion that would add to the productivity loss from institutional distortions. Similarly, this would also increase the importance of institutional distortions for explaining adoption lag as early adopters in the baseline model tend to be firms that have increased their technology through investment over several periods. As this investment would decline over time because of larger distortions, there would be fewer high technology (high $\check{z}_{j,t}$) firms that adopt early.

Each firm j produces a unique good j that is differentiated from each other good in the economy. The production of good j depends on firm j's technology z_j, ability s_j and labour input n_j:

$$y = \left(z_j s_j^{1-\omega}\right)^{\frac{1}{\nu-1}} n_j.$$

Scaling productivity by $\frac{1}{\nu-1}$ helps with the comparison of the two models, but otherwise does not affect the results. The cost of technology follows the baseline model.

F.2 Market Allocation

It follows from the expression for consumption that demand for good j is equal to $y_j = Y_t p_j^{-\nu}$. The firm's static problem is given by

$$\max_p (1-\tau) p y(p) - m y(p),$$

subject to demand and where the term $m = w_t/(z_t s^{1-\omega})^{1/(\nu-1)}$ is the marginal cost of producing good y. The solution to the problem implies that firms set prices equal to $p = \frac{\nu}{\nu-1}\frac{m}{1-\tau}$ and that firm revenue is equal to

$$r = py = Y_t \left(\frac{\nu}{\nu-1}\right)^{1-\nu} \left(\frac{w_t}{1-\tau}\right)^{1-\nu} (z_t s^{1-\omega}).$$

The firm's technology choice problem is then given by

$$v_{(s,\tau)}(z_{-1}) = \max_z \tilde{\nu} Y_t w_t^{1-\sigma} (1-\tau)^{\sigma} (z s^{1-\omega}) - w_t x_t + D v_{(s,\tau)}(z),$$

where $\tilde{\nu} = \nu^{-\nu}(\nu-1)^{1-\nu}$. Solving the problem shows that the choice of technology is given by:

$$\varphi(s,\tau) = \left(\frac{z_t(s,\tau)}{\bar{z}_t}\right)^{\xi} = \eta (1-\tau)^{\frac{\nu}{1-\omega}} s,$$

where $\eta = \left[\frac{\omega\tilde{\nu}}{\pi(1-De^{gw})}\frac{Y_t\bar{z}_t}{w_t^{\sigma}}\right]^{\frac{1}{1-\omega}}$. It follows that revenue and labour are given by

$$(1-\tau)r(s,\tau) = w_t\left(\frac{\nu}{\nu-1}\right)^{1-\nu}\left(\frac{\bar{z}_t Y_t}{w_t^{\nu}}\right)\eta^{\omega}(1-\tau)^{\frac{\nu}{1-\omega}}s$$
$$n(s,\tau) = \left(\frac{\nu}{\nu-1}\right)^{-\nu}\left(\frac{\bar{z}_t Y_t}{w_t^{\nu}}\right)\eta^{\omega}(1-\tau)^{\frac{\nu}{1-\omega}}s.$$

The expressions show a similar relationship between the allocations of revenues, labour and technology as in the baseline model. However, the elasticity of institutional distortions differs between the two models. This implies that if institutional distortions are set according to the same relationship with ability then the joint distribution of ability and resource allocations will differ between the two models.

F.3 Comparison with Baseline Model

Institutional distortions are set according to (21) in both the CES and DRS models. Given the solution in the previous section, the goal is to show the conditions under which the two models will produce the same joint distributions of productivity and resources. I show that this holds when the relationship between ability s and TFPQ is the same in both models.

In the CES model, measured TFPQ is equal to

$$\text{TFPQ} = \frac{(py)^{\frac{\nu}{\nu-1}}}{n} \propto (s^{1-\omega}\varphi^{\omega})^{\frac{1}{\nu-1}} \propto (1-\tau)^{\frac{\omega}{1-\omega}\frac{\nu}{\nu-1}}s^{\frac{1}{\nu-1}} = s^{\frac{1}{\nu-1}-\tilde{\rho}\left[\frac{\omega}{1-\omega}\frac{\nu}{\nu-1}\right]}.$$

In the DRS model, measured TFPQ is equal to:

$$\text{TFPQ} = \frac{py}{n^{\gamma}} \propto (s^{1-\omega}\varphi^{\omega})^{1-\gamma} \propto (1-\tau)^{\frac{\omega}{1-\omega}}s^{1-\gamma} = s^{1-\gamma-\rho\frac{\omega}{1-\omega}}.$$

It follows then that if ρ and $\tilde{\rho}$ are set appropriately, the implied distributions of s will be the same in the two models (up to a scalar constant). Specifically, this requires that the elasticities are set such that:

$$\rho = \frac{\nu}{\nu-1}\tilde{\rho}.$$

It is also straightforward to show that setting the elasticity as above will imply that the relationship between ability s (or equivalently TFPQ) and the other endogenous variables in the model (labour, technology and revenue). It also follows that when there is no technology choice ($\omega \to 0$) then only $1 - \gamma = \frac{1}{\nu - 1}$ is needed for the two models to be equivalent (as shown by Hsieh and Klenow, 2009).

A similar adjustment could be made for the standard deviation of TFPR, which is dictated by the parameter σ in the model. Following the same relationship as above, the implied relationship would be $\sigma = \frac{\nu}{\nu - 1}\tilde{\sigma}$. Given that TFPR is model independent (that is, $TFPR = py/n$) I use the standard deviation of TFPR in the data directly for the calibration.