Big Push in Distorted Economies[*]

Francisco Buera

Washington University in St. Louis

Hugo Hopenhayn

UCLA

Yongseok Shin

Washington University in St. Louis

Nicholas Trachter

Federal Reserve Bank of Richmond

July 2, 2021

Abstract

Why don't poor countries adopt more productive technologies? Is there a role for policies that coordinate technology adoption? To answer these questions, we develop a quantitative model that features complementarity in firms' technology adoption decisions: The gains from adoption are larger when more firms adopt. When this complementarity is strong, multiple equilibria and hence coordination failures are possible. More important, even without equilibrium multiplicity, the model elements responsible for the complementarity can substantially amplify the effect of distortions and policies. In what we call the Big Push region, the impact of idiosyncratic distortions is over three times larger than in models without such complementarity, without coordination failures playing a role.

1 Introduction

Many countries have industrialized and grown rapidly by adopting modern technologies. Why don't poor countries adopt more productive technologies? What policies can effectively promote technology adoption? The standard view emphasizes the role of distortions or barriers to technology adoption (e.g., Parente and Prescott, 1999; Hsieh and Klenow, 2014; Cole et al., 2016; Bento and Restuccia, 2017). According to this view, eliminating the distortions is the obvious policy response. An alternative view emphasizes the role of complementarity and coordination failures: Firms in poor countries use unproductive technologies because other firms use unproductive technologies, even though the gains from adoption increase with the number of adopters so that they will all benefit from a coordinated decision to adopt. According to this view, policies can have large aggregate effects by coordinating firms' decisions. This view has a long tradition in policy circles (e.g., Rosenstein-Rodan, 1943; Hirschman, 1958) and is supported by more recent theoretical works (e.g., Murphy et al., 1989; Matsuyama, 1995; Ciccone, 2002). However, there have been few quantitative analyses, if any, of the coordination failure view of economic development. Our paper bridges these two paradigms in a quantitative framework.

This paper makes two contributions. First, our theoretical analysis goes beyond determining the existence of multiple equilibria to show that, even if they do not exist, complementarity in technology adoption can amplify the effect of distortions and policies. Second, our quantitative analysis based on aggregate and micro-level data examines the empirical relevance of multiple equilibria and coordination failures and, more important, shows the economic significance of the amplification channel in the absence of multiple equilibria. In what we call the Big Push region, the impact of idiosyncratic distortions is over three times larger than in models without such complementarity. This amplification in our model helps account for the vast income gap between poor and rich countries without coordination failures playing a role.

In our model, firms are ex-ante heterogeneous, produce differentiated goods, are subject to idiosyncratic distortions, and are connected to one another through input-output linkages. Firms first choose whether or not to pay a fixed cost and enter the market. Active firms can operate a traditional technology or, upon paying adoption costs, a more productive modern technology.

We first theoretically analyze the conditions under which the complementarity in

firms' technology adoption decisions can amplify the effect of policies and distortions and even support multiple equilibria. When more firms adopt the productive technology, the aggregate price falls, which has the direct effect of reducing the price and the profit of the marginal adopter. However, the lower aggregate price also means a larger demand for the marginal firm's output, a lower cost of the intermediate input for production, and less expensive goods portion of the adoption costs, raising the marginal adopter's gains. If the latter forces dominate the direct competition effect, we have complementarity.

More heterogeneity in firm productivity implies that the density of the firms near the adoption margin is smaller, implying a weaker feedback effect from the marginal firm's adoption to the aggregate economy and hence weaker complementarity.[1] On the other hand, idiosyncratic distortions correlated with firm productivity (i.e., more productive firms being distorted more) compress the effective heterogeneity across firms, which strengthens complementarity.

For simple versions of our model, we derive the conditions for amplification and equilibrium multiplicity. What contributes to amplification are, consistent with the discussions above: a small elasticity of substitution across differentiated goods, adoption costs in goods rather than labor, a high intermediate input intensity of the modern technology, small heterogeneity in firm productivity, and a close correlation between idiosyncratic distortions and firm productivity. Multiplicity is an extreme form of amplification, and indeed the conditions for the existence of multiple equilibria are stronger than those for amplification, a result corroborated by our quantitative model that exhibits potent amplification even without multiple equilibria.

Next, we use aggregate and micro-level data from the US and India and conduct a quantitative analysis. The US is an undistorted benchmark, and India is a large developing country for which relevant micro-level data is available. The full version of our model has several layers, but it is tractable enough that most of the parameters can be identified transparently from the data. In spite of the potential presence of multiplicity, under the assumption that the data comes from an equilibrium where adopters and non-adopters coexist, the values of the key model parameters are obtained in closed form from the establishment size distribution.

The quantitative analysis has three main results. First, although our calibration

---

[1]This may explain why multiplicity results in this literature invariably come from models with homogeneous firms. In contrast, firm heterogeneity is essential for our quantitative strategy.

targets the moments from the establishment size distribution but not the income level of either country, the model generates nearly as large an income gap between the US and India—a factor of seven—as in the data, due to the higher adoption costs and degrees of idiosyncratic distortions in India, whose impact is amplified by the complementarity. Second, the US is in the unique equilibrium region but India is in the multiplicity region. However, India is found to be in the good (or high adoption/output) equilibrium, and hence coordination failures do not explain why India is poorer than the US.

The final result is that the effect of idiosyncratic distortions on aggregate productivity and GDP can be highly non-linear. When the complementarity in firms' adoption decisions is strong enough, multiple equilibria appear above a threshold degree of idiosyncratic distortions, but eventually only the bad equilibrium survives with even more distortions. This implies that, for a heavily distorted economy in a unique bad equilibrium, a reform that reduces the distortions just enough to place it in the multiplicity region gives it a chance to coordinate to the good equilibrium featuring more adoption and vastly higher GDP. More important, even when the model has a unique equilibrium for any degree of distortions, the model elements responsible for the complementarity still amplify the impact of distortions. In an empirically plausible range of distortions, a small reduction in distortions can disproportionately increase adoption and aggregate productivity by several times more than in the standard models without complementarity (e.g., Hsieh and Klenow, 2009; Hopenhayn, 2014; Restuccia and Rogerson, 2017). These positive non-linear effects with or without multiplicity are the Big Push in distorted economies, and we call this range of distortions the Big Push region.

Our quantitative results offer two broader implications. First, the powerful amplification of the impact of distortions through complementarity in our model can help account for the huge income differences across countries. Second, the existence of the Big Push region, where a small reduction in distortions can unleash massive improvements, can be an explanation of why some distortion-reducing reforms are more successful than others.

Related literature   The idea that underdevelopment can result from coordination failures goes back to Rosenstein-Rodan (1943). It has been formalized by Murphy et al. (1989) in a model with aggregate demand spillovers and by Ciccone (2002) in

a model with differences in intermediate input intensities across technologies.[2]

Some empirical works have applied the idea of multiple equilibria and coordination failures to historical contexts (Davis and Weinstein, 2002, 2008; Redding et al., 2011; Kline and Moretti, 2014; Lane, 2019; Crouzet et al., 2020). The evidence so far is mixed, suggesting that the possibility of multiple equilibria depends on the details of the economic environment, a theme we emphasize.

Such advances in the theoretical and empirical literature have not been actively followed by quantitative work with few exceptions. Valentinyi et al. (2000), although a theoretical work, makes the important point that multiplicity is overstated in representative agent models. Using a heterogeneous agent version of the two-sector model of Matsuyama (1991), in which the economies of scale that are external to individual producers cause multiplicity, they show that sufficient heterogeneity restores a unique equilibrium. Graham and Temple (2006) study a representative agent version of a similar two-sector model and find that a quarter of the world's economies are stuck in a low output equilibrium. Caucutt and Kumar (2008) numerically explore a model in the theoretical literature.[3]

Relative to these papers, our contribution is to quantitatively analyze a richer, more granular model of coordination failures, bringing together elements emphasized in the theoretical literature and disciplining the analysis with micro-level data. More important, we find that, even in the absence of multiplicity, these model elements amplify the impact of idiosyncratic distortions and policies.

Our model builds on widely-used models of heterogeneous firms, including those of Hopenhayn (1992) and Melitz (2003). We extend the standard model to incorporate discrete technology adoption choices.[4] Our modeling choice is partly motivated by the evidence in Holmes and Stevens (2014), who show wide variations in the size of plants,

---

[2]Krugman (1992) and Matsuyama (1995) review the earlier papers on this topic and the more recent theoretical contributions. Additional examples include Okuno-Fujiwara (1988), Rodríguez-Clare (1996) and Rodrik (1996), which analyze open-economy models of coordination failures.

[3]Owens et al. (2018) study a quantitative urban model in which residential externalities cause multiple equilibria at the neighborhood level. Another related literature explores the role of coordination failures in accounting for the Great Recession (Kaplan and Menzio, 2016; Schaal and Taschereau-Dumouchel, 2019) in the tradition of Cooper and John (1988), but this literature abstracts from micro-level heterogeneity.

[4]Yeaple (2005) and Bustos (2011) also consider firms' technology choice decisions, but they either ignore the possibility of multiplicity or make assumptions that happen to guarantee uniqueness. The small- vs. large-scale sector choice in Buera et al. (2011) can also be thought of as a technology choice, but that model also has a unique equilibrium.

even within narrowly-defined industries. In our model, small firms producing with the traditional technology coexist with large firms operating the productive modern technology, with the technology choice driven by and magnifying the underlying heterogeneity in firm-level productivity.

Another important element of our model is the input-output linkages in the form of round-about production as in Jones (2011), which helps make firms' adoption decisions complementary in our model and amplifies the effect of distortions in general.

Finally, following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), we introduce idiosyncratic distortions, which stand in for various types of frictions, including barriers to technology adoption as modeled by Parente and Prescott (1999) and Cole et al. (2016). The interaction between distortions and technology adoption in our model is related to the impact of distortions on productivity-enhancing investment in Bento and Restuccia (2017) and Bhattacharya et al. (2013). Our emphasis is the amplification of the effect of distortions through the complementarity in firms' adoption decisions, which results in highly non-linear effects of distortions with or without multiple equilibria. Relative to the distortion literature, our model is unique in its ability to generate large income differences across countries with moderate degrees of idiosyncratic distortions.

2   Setup

The economy is populated by a mass $L = 1$ of workers and measure one of potential firms, each of them producing a differentiated good j.[5] Workers supply their labor inelastically and use their labor income to consume a final good. The differentiated goods produced by firms are combined to produce an intermediate aggregate,

$$X = \left[\int y_j^{\frac{\eta-1}{\eta}} \, dj \right]^{\frac{\eta}{\eta-1}} , \ \eta > 1 \ ,$$

where η is the elasticity of substitution, governing how complementary are differentiated goods in the production of the intermediate aggregate. The intermediate aggregate can be transformed with a linear technology to produce the final consumption good and the intermediate input to be used by firms.

Firms are heterogeneous in their productivity z, drawn from a cumulative

---

[5]Thus, j indexes a differentiated good or a particular firm.

distribution $F(z)$, with density $f(z) = \partial F(z)/\partial z$. Based on their productivity, firms choose to be active or inactive. An active firm with idiosyncratic productivity z must incur κ_e units of labor to enter the market and operate. An active firm produces using technology $i \in \{t, m\}$, labor l, and intermediate aggregate x, according to:

$$y = z\frac{A_i}{\nu_i^{\nu_i}(1-\nu_i)^{1-\nu_i}}l^{1-\nu_i}x^{\nu_i}, \ \nu_i \in [0,1],$$

where ν_i is the intermediate input elasticity. Technology m, the "modern" technology, is more productive and intermediate input intensive than technology t, the "traditional" technology. Specifically, we assume that $A_m/\left[\nu_m^{\nu_m}(1-\nu_m)^{1-\nu_m}\right] > A_t/\left[\nu_t^{\nu_t}(1-\nu_t)^{1-\nu_t}\right]$ and $\nu_m \geq \nu_t$.

The modern technology requires κ_a units of an adoption good. The adoption good is produced by a competitive fringe in the adoption good sector by combining labor and the intermediate aggregate using a Cobb-Douglas production function. We denote by P_a the price of the adoption good.

Finally, firms are subject to idiosyncratic gross output distortions given by $\tau z^{-\xi}$, where $\xi \in [0, 1]$ is the elasticity of distortions with respect to productivity and τ is a budget-balancing scale parameter. With $\xi > 0$, low productivity firms are subsidized and high productivity firms are taxed, which is the empirically relevant range in the misallocation literature as we show in Section 5.

As is standard, the demand for differentiated good j and the price index of the intermediate aggregate are

$$y_j = \left(\frac{P}{p_j}\right)^{\eta} X \ \text{ and } \ P = \left[\int p_j^{1-\eta}dj\right]^{\frac{1}{1-\eta}}.$$

We assume the labor market is perfectly competitive with equilibrium wage w.

Problem of Intermediate Input Producers An active firm with productivity z producing with technology i earns operating profits $\pi_i^o(z)$. The firm chooses the price p for its differentiated good, and the amount of labor l and intermediate input m required for production. The problem is:

$$\pi_i^o(z) = \max_{p,l,x} \ \tau z^{-\xi}p\left(\frac{P}{p}\right)^{\eta}X - wl - Px \tag{1}$$

$$\text{s.t. } z\frac{A_i}{\nu_i^{\nu_i}(1-\nu_i)^{1-\nu_i}}l^{1-\nu_i}x^{\nu_i} \geq y = \left(\frac{P}{p}\right)^{\eta}X. \tag{2}$$

From the first order conditions of this problem we obtain expressions for the optimal price and input demands:

$$p_i(z) = \frac{\eta}{\eta-1} \frac{w^{1-\nu_i} P^{\nu_i}}{A_i \tau} \frac{1}{z^{1-\xi}}, \tag{3}$$

$$l_i(z) = \left(\frac{\eta-1}{\eta}\right)^{\eta} (1-\nu_i) \tau^{\eta} \left(\frac{P}{w}\right)^{(1-\nu_i)\eta+\nu_i} X A_i^{\eta-1} z^{\eta(1-\xi)-1}, \tag{4}$$

$$x_i(z) = \frac{\nu_i}{1-\nu_i} \frac{w}{P} l_i(z). \tag{5}$$

From these expressions, we obtain an expression for the maximized operating profit,

$$\pi_i^o(z) = \frac{1}{\eta} \left(\frac{\eta-1}{\eta} \frac{1}{w^{1-\nu_i}}\right)^{\eta-1} \tau^{\eta} P^{\eta(1-\nu_i)+\nu_i} X A_i^{\eta-1} z^{\eta(1-\xi)-1},$$

which is increasing in z provided $\eta - 1/(1-\xi) > 0$.

It is useful to analyze the effect of distortions ξ on a firm's decisions in two steps. First, the firm acts as if having a lower productivity—lowered by a factor $z^{-\xi}$, as evidenced in equation (3). This effect also translates into an equivalent drop in output (equation 2). Thus, this effect is captured by defining effective productivity as $z^{1-\xi}$. Second, as seen in equations (4) and (5), distortions imply a lower elasticity of factor demands and profits with respect to z than in the undistorted case, i.e. $\eta - 1/(1-\xi) \leq \eta - 1$. To see this, notice that we can rewrite the productivity term in equation (4) as $z^{(1-\xi)(\eta-1/(1-\xi))}$; the first term in parentheses corresponds to the productivity reduction effect, while the second one corresponds to the lower demand elasticity resulting from the wedge.

Finally, the profits of a firm with productivity z, given the optimal entry and adoption decisions, are

$$\pi(z) = \max_{\text{inactive,active}} \{0, \max\{\pi_t^o(z), \pi_m^o(z) - P_a \kappa_a\} - w\kappa_e\}. \tag{6}$$

The entry and adoption decisions are characterized by thresholds z_e and z_a, where $z_e \leq z_a$. That is, a firm with productivity z will be active if and only if $z \geq z_e$, and will adopt the modern technology if and only if $z \geq z_a$. This is because, under the assumption that $\eta - 1/(1-\xi) > 0$, the operating profit $\pi_i^o(z)$ is increasing in productivity z for $i \in \{t, m\}$.

Problem of Adoption Good Producers A representative competitive firm producing the adoption good takes the adoption good price P_a as given and solves

$$\max_{L_a, X_a} \frac{1}{\gamma^\gamma (1 - \gamma)^{1-\gamma}} P_a L_a^{1-\gamma} X_a^\gamma - w L_a - P X_a \ , \tag{7}$$

where γ is the intermediate aggregate share in the production of the adoption good.

Equilibrium We consider symmetric equilibria where all firms of a given productivity make the same decision.

Definition 1. *A symmetric equilibrium consists of entry and adoption decisions by differentiated goods producers, factor demands by producers of differentiated and adoption goods, and relative factor prices w/P and $P_a/P = (w/P)^{1-\gamma}$ such that (i) firms maximize profits and (ii) the markets for labor and intermediate aggregate clear:*

$$\int_{z_e}^{z_a} l_t(z)\, dF(z) + \int_{z_a}^{\infty} l_m(z)\, dF(z) + (1 - F(z_e))\, \kappa_e$$

$$+ (1 - F(z_a))(1 - \gamma)\left(\frac{P}{w}\right)^\gamma \kappa_a = L \ , \tag{8}$$

$$C + \int_{z_e}^{z_a} x_t(z)\, dF(z) + \int_{z_a}^{\infty} x_m(z) dF(z) + (1 - F(z_a))\gamma \left(\frac{w}{P}\right)^{1-\gamma} \kappa_a = X \ . \tag{9}$$

Equation (8) is the labor market clearing condition, and the four terms in the left-hand side are the labor used for traditional technology production, modern technology production, entry costs, and adoption good production. Equation (9) states that the intermediate aggregate used for consumption, traditional technology production, modern technology production, and adoption good production must add up to the quantity produced on the right-hand side.

Manipulating the equilibrium conditions, we can characterize the equilibrium by three equations in three variables, z_e, z_a and P/w, which are shown in the appendix.

3 Understanding Amplification and Multiplicity

A firm's incentives to adopt the modern technology can be affected by other firms' adoption decisions. We show that the strength of this complementarity has important implications for the amplification of the impact of distortions and policies and also for the possibility of multiple equilibria and coordination failures.

Before examining the role that different parameters play in our model, we first consider the complementarity at a more abstract level for the case of a policy that subsidizes technology adoption.

Let $D(z; a, s)$ denote the net gains from adoption for a firm with productivity z, when the mass of adopters in the economy is a and the adoption subsidy is s:

$$D(z; a, s) \equiv \pi_m^o(z; a, s) - \pi_t^o(z; a, s) - (1 - s) P(a) \kappa_a , \qquad (10)$$

where $\pi_i^o(z; a, s)$ is the operating profit of a firm with productivity z using technology i, and where we made explicit the dependence on the aggregate variables a and s. The aggregate price index when the mass of adopters is a is $P(a)$. We are assuming that the adoption costs are in units of goods only ($\gamma = 1$ in our model).

A firm adopts the modern technology if $D \geq 0$ and operates the traditional technology otherwise. Consistent with our model, we assume that D is increasing in z, so the adoption decision rule is characterized by a productivity threshold z_a: that is, a firm adopts the modern technology if and only if $z \geq z_a$, with the corresponding mass of adopters being $a = 1 - F(z_a)$, where z_a satisfies

$$D(z_a; a, s) = 0 . \qquad (11)$$

We now consider a small change in subsidy, ds. We define the *direct* effect as the impact on z_a holding fixed the value of a at the initial equilibrium level:

$$\left[\frac{dz_a}{ds}\right]^{direct} = -\frac{D_s(z_a; a, s)}{D_z(z_a; a, s)}. \qquad (12)$$

The *total* effect of this policy includes the equilibrium response of a and is:

$$\left[\frac{dz_a}{ds}\right]^{total} = -\frac{D_s(z_a; a, s)}{D_z(z_a; a, s) - D_a(z_a; a, s) f(z_a)} .$$

The ratio between the total and the direct effect is

$$\frac{[dz_a/ds]^{total}}{[dz_a/ds]^{direct}} = \frac{1}{1 - \frac{D_a(z_a; a, s) f(z_a)}{D_z(z_a; a, s)}} , \qquad (13)$$

which can be interpreted as a multiplier measuring the degree of the feedback effect. When this ratio is greater than 1, the feedback is positive, an indication of complementary in adoption decisions. The intensity of complementarities depends

on the feedback ratio:

$$r\left(z_a, a, s\right) \equiv \underbrace{D_a\left(z_a; a, s\right)}_{\text{incentive}} \underbrace{\frac{f\left(z_a\right)}{D_z\left(z_a; a, s\right)}}_{\text{feedback}} \ . \tag{14}$$

The feedback ratio can be interpreted as the product of two terms that determine the adoption complementarities. The incentive term is the effect of a change in the mass of adopters on the incentives (net gains) for adoption $D_a\left(z_a; a, s\right)$. The feedback term shows how the changes in the net gains from adoption feed back into a change in the adoption threshold, $f\left(z_a\right)/D_z\left(z_a; a, s\right)$. The expression $1/D_z\left(z_a; a, s\right)$ measures how the productivity of the marginal adopter relates to changes in the incentives to adopt, and $f(z_a)$ is the density of firms at the adoption threshold. A positive feedback ratio $r(z_a, a, s)$ implies a multiplier exceeding one and hence amplification. In addition, the closer the feedback ratio is to one, the higher the multiplier will be.

The feedback ratio also plays an important role in the analysis of equilibrium determination and the possibility of multiplicity. An equilibrium for a given s consists of a threshold z_a and $a = 1 - F\left(z_a\right)$ satisfying equation (11). In examining the equilibrium, it is useful to consider the following mapping:

$$T\left(a\right) = \left\{a'|a' = 1 - F\left(z_a\right) \text{ and } D\left(z_a, a, s\right) = 0\right\} \ .$$

This function is analogous to a best response map, giving the fraction of adopters $T(a)$ when firms behave optimally in response to the equilibrium prices that prevail with an adoption rate a. An equilibrium is given by a fixed point of this mapping $a = T\left(a\right)$. The slope of this mapping is:

$$T'\left(a\right) = \frac{D_a\left(z_a; a, s\right) f\left(z_a\right)}{D_z\left(z_a; a, s\right)},$$

which is precisely the feedback ratio $r(z_a, a, s)$. The equilibrium is unique if $T'\left(a\right) < 1$ for all a. The possibility of multiple equilibria arises if there is a range where $T'\left(a\right) > 1$ followed by a range where $T'\left(a\right) < 1$ as shown in the left panel of Figure 1. In the right panel, we zoom in on the right-most equilibria of the left panel. We show the effect of a change in the subsidy rate s, breaking it into the direct and the total effects. Note that the condition for multiplicity is stronger than the condition for amplification, $r(z_a, a, s) > 0$.

In the remainder of this section we illustrate the role of various model elements
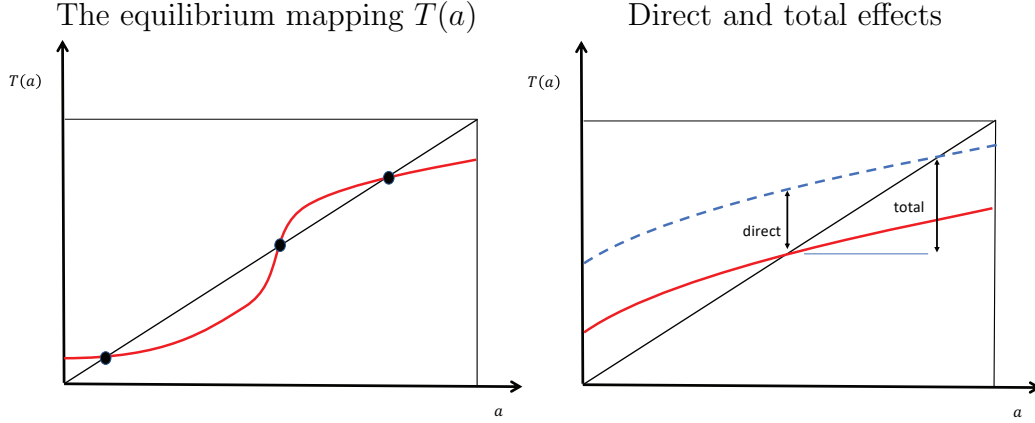
The equilibrium mapping $T(a)$ Direct and total effects

Figure 1: Equilibrium Mapping and Amplification: An Example

in generating amplification and multiplicity using simple examples.

3.1 Simple Model I: Heterogeneity

Consider the case with no intermediate inputs in the production of differentiated goods, $\nu_i = 0$, and an exogenous population of active firms, i.e., no entry margin. Adoption costs are in units of goods only ($\gamma = 1$). We further assume that the differentiated goods producers' productivity z follows a Pareto distribution, $z \in [1, \infty) \sim 1 - z^{-\zeta}$. There is no idiosyncratic distortion.

In this case, the net gains from adoption presented in equation (10) are

$$D(z; a, s) = \frac{1}{\eta} \left(\frac{\eta - 1}{\eta} \right)^{\eta - 1} P(a)^\eta X(a) \left(A_m^{\eta-1} - A_t^{\eta-1} \right) z^{\eta-1} - (1 - s)P(a)\kappa_a \ ,$$

and the feedback ratio presented in equation (14) becomes

$$r\left(z_a, a\right) = \left[(1 - \eta)\epsilon_{Pa} + \epsilon_{Xa}\right] \frac{\zeta}{\eta - 1} \ , \tag{15}$$

where $\epsilon_{Pa} = -(\partial P/\partial a)(a/P)$ and $\epsilon_{Xa} = (\partial X/\partial a)(a/X)$ are the elasticities of the price index and the aggregate demand with respect to the mass of adopters. We dropped the dependence of the feedback ratio on s, as the subsidy rate does not show up in this particular formulation.

As in the general case, the feedback ratio is the product of the incentive and the feedback terms. The incentive term measures the impact of the mass of adopters on the profitability of the marginal adopter. It is a function of the adoption elasticities of

12

the price index and the aggregate demand with respect to the adoption rate, ϵ_{Pa} and ϵ_{Xa}, and the demand elasticity, η. More adoption affects a firm's profit through two channels. First, it shifts the demand in proportion to the change in the intermediate aggregate X. This is captured by the elasticity ϵ_{Xa}. The second channel is through the impact on the price index. This is a negative competitive effect, whose strength is governed by the demand elasticity η. In addition, the decline in the price index has a direct effect on the net gains from adoption by lowering the cost of adoption. The net effect of the price index on the feedback ratio is captured by the term $(1 - \eta)\epsilon_{Pa}$ in equation (15).

The feedback term in equation (15) captures the response of the mass of adopters to the increase in the net gains from adoption for the marginal adopter, and depends on the demand elasticity η and the elasticity of the density of the productivity distribution to the mass of adopters ζ. In the feedback term, $\eta - 1$ measures the elasticity of the net gains from adoption with respect to productivity z, and its inverse measures the response of the adoption threshold to the changes in the net gains from adoption. The impact of changes in this threshold on the mass of adopters depends on the density of firms at this point, captured by the parameter ζ.

To further understand the feedback ratio, we now analyze the adoption elasticities of the price index ϵ_{Pa} and the aggregate demand ϵ_{Xa}:

$$\epsilon_{Pa} = \epsilon_{Xa} = \left(\frac{1}{\eta - 1} - \frac{1}{\zeta} \right) \left(1 - \frac{A_t^{\eta-1}}{A_m^{\eta-1}} \right) \mathcal{M}(a) \ ,$$

where $\mathcal{M}(a) = A_m^{\eta-1} a^{1-\frac{\eta-1}{\zeta}} / \left[A_t^{\eta-1} \left(1 - a^{1-\frac{\eta-1}{\zeta}} \right) + A_m^{\eta-1} a^{1-\frac{\eta-1}{\zeta}} \right] \in (0,1)$ is the modern sector's share of total value added and employment. The demand elasticity η has a negative effect on the adoption elasticities ϵ_{Pa} and ϵ_{Xa}. A higher elasticity of substitution η implies that the additional contribution of the marginal, less productive, adopter will be smaller relative to the contribution of the infra marginal, and more productive, adopters. On the other hand, a more elastic density of the productivity distribution (higher ζ) implies that the productivity of the marginal adopter declines less as the mass of adopters rises. This naturally results in larger elasticities with respect to the mass of adopters.[6]

---

[6]In making these arguments, we are holding fixed the values of $A_t^{\eta-1}/A_m^{\eta-1}$ and $\mathcal{M}(a)$, which also depend on η and ζ. One interpretation is that we are implicitly adjusting A_t/A_m and the aggregate adoption rate a by changing the adoption cost κ_a.

Using the expressions for the adoption elasticities, we can express the feedback ratio in terms of the model parameters:

$$r(z_a, a) = \left(\frac{2-\eta}{\eta-1}\right)\left(\frac{\zeta+1-\eta}{\eta-1}\right)\left(1 - \frac{A_t^{\eta-1}}{A_m^{\eta-1}}\right)\mathcal{M}(a) \ . \tag{16}$$

For the effect of policies to be amplified, the multiplier in equation (13) must be greater than one. This requires a positive feedback ratio. In this case, the feedback ratio is positive if and only if

$$2 - \eta > 0 \ , \tag{17}$$

since the last three terms in the right-hand side of equation (16) are strictly positive.[7] When η is low, the negative competition effect is small, and thus the positive effect on the gains from adoption of a larger mass of adopters dominates, generating a positive feedback ratio and hence amplification.[8] While the existence of amplification depends solely on η, the degree of amplification does depend on ζ, the ratio A_m/A_t, and $\mathcal{M}(a)$. Intuitively, a higher density of marginal firms (high ζ) and a larger difference in technology A_m/A_t give larger amplification.

Finally, we can obtain a necessary condition for multiplicity, i.e., the condition $T'(a) > 1$. Using the fact that the last two terms in the right-hand side of equation (16) are strictly less than 1, a necessary condition for multiplicity is

$$2 - \eta > \frac{1}{1+\zeta} \ . \tag{18}$$

This implies that the necessary condition for multiplicity is stricter than the condition for amplification—that is, amplification is possible even in the absence of multiple equilibria. Multiplicity requires that complementarity is strong enough to countervail the negative effect of heterogeneity on the density of marginal adopters, as well as that on the elasticities ϵ_{Pa} and ϵ_{Xa}. This may explain why most multiplicity results in the literature come from models with homogeneous producers ($\zeta \to \infty$).

---

[7]For aggregates to be finite, $\zeta + 1 - \eta$ must be positive.

[8]With limit pricing, as in Murphy et al. (1989), the negative competition effect is inoperative. As a result, amplification and multiplicity do not require restrictions on η.

3.2 Simple Model II: Idiosyncratic Distortions

We now introduce idiosyncratic output distortions (controlled by ξ) into the simple example with heterogeneity above, and explore how they affect complementarity and amplification. We retain all other simplifying assumptions: no entry margin, $\nu_i = 0$, and $\gamma = 1$. As we show, distortions have two effects. First, distortions compresses the distribution of the effective productivity of firms. In particular, it is as if the tail parameter of the productivity distribution were $\zeta/(1-\xi)$ instead of ζ. Second, distortions create a wedge between the value of output produced by firms and their profitability and hence their factor demand, reducing the elasticity of profits and labor demand with respect to firms' productivity.

With distortions, the feedback ratio is now

$$r\left(z_a, a\right) = [(1-\eta)\epsilon_{Pa} + \epsilon_{Xa}] \frac{\zeta/(1-\xi)}{\eta - 1/(1-\xi)} . \tag{19}$$

Comparing the feedback ratio with the one with no distortions in equation (15), we see that the feedback term is adjusted by the distortion parameter ξ. This adjustment can be understood as a more concentrated firm productivity distribution, i.e., one with a tail parameter $\zeta/(1-\xi)$, and a lower elasticity of the net gains from adoption with respect to firms' productivity. Both adjustments make the feedback term larger, holding fixed the adoption elasticities ϵ_{Pa} and ϵ_{Xa}.

We can also derive expressions for the two adoption elasticities:

$$\epsilon_{Pa} = \left(\frac{1}{\eta - 1} - \frac{1-\xi}{\zeta}\right)\left(1 - \frac{A_t^{\eta-1}}{A_t^{\eta-1}}\right)\mathcal{M}^v(a) ,$$

$$\epsilon_{Xa} = \eta\varepsilon_{Pa} - (\eta - 1)\frac{1 - \frac{1-\xi}{\zeta}\left(\eta - \frac{1}{1-\xi}\right)}{1 - \frac{1-\xi}{\zeta}(\eta - 1)}\frac{\mathcal{M}^l(a)}{\mathcal{M}^v(a)}\varepsilon_{Pa} ,$$

where $\mathcal{M}^v(a)$ and $\mathcal{M}^l(a)$ are the modern firms' shares of total value added and employment, respectively.[9]

The effect of distortions on the elasticity ϵ_{P_a} is analogous to what we obtained

---

[9]In particular, $\mathcal{M}^l(a) = \mathcal{M}^v(a^{1+\xi/\zeta})$ and

$$\mathcal{M}^v(a) = \frac{A_t^{\eta-1}a^{1-\frac{(1-\xi)(\eta-1)}{\zeta}}}{A_t^{\eta-1}\left(1 - a^{1-\frac{(1-\xi)(\eta-1)}{\zeta}}\right) + A_m^{\eta-1}a^{1-\frac{(1-\xi)(\eta-1)}{\zeta}}} .$$

in the model with no distortions, adjusted for the compression of the effective productivity distribution with the tail parameter $\zeta/(1-\xi)$. Accordingly, with more distortions, and for a given level of $\mathcal{M}^v(a)$, the elasticity of the price index with respect to the mass of adopters is higher.

The effect of distortions on the elasticity of the aggregate demand ϵ_{Xa} involves two effects. The first effect is the same as the one for ϵ_{Pa}, since the aggregate demand is a function of the price index. This implies that the compression of the productivity distribution increases both elasticities. The second is the effect that distortions have on the aggregate demand, creating a wedge between firms' input demand and revenues. Here, holding fixed the modern firms' shares and ϵ_{Pa}, an increase in distortions reduces ϵ_{Xa}. However, one can see that the net effect of ξ on ϵ_{Xa} is positive: The larger the degree of distortions, the more responsive the aggregate demand is to an increase in the mass of adopters.

The relative effect of distortions on the two elasticities, ϵ_{Pa} and ϵ_{Xa}, depends crucially on the mass of adopters, a, through its effect on the modern firms' share of value added and employment, $\mathcal{M}^v(a)$ and $\mathcal{M}^l(a)$.

The modern firms' share of employment is smaller than that of value added, i.e., $\mathcal{M}^l(a) \leq \mathcal{M}^v(a)$, with strict inequality when $a < 1$. This is because distortions introduce a wedge between the value of the output produced by the modern firms and their profitability and hence factor demands. This effect is clearest in the limit with an arbitrarily small mass of modern firms, where the employment share of the modern firms is arbitrarily small relative to their value added share, i.e., $\lim_{a\to 0} \mathcal{M}^l(a)/\mathcal{M}^v(a) \to 0$, as the modern firms are infinitely productive and hence face arbitrarily large distortions. Here the elasticity of the price index is smaller than that of the aggregate demand, $\epsilon_{Pa} < \epsilon_{Xa}$. At the other extreme where the mass of adopters is close to one, we have $\mathcal{M}^l(a)/\mathcal{M}^v(a) \approx 1$, and the elasticity of the price index is larger than that of the aggregate demand, $\epsilon_{Pa} > \epsilon_{Xa}$.

Using the expressions for the adoption elasticities, we rewrite the feedback ratio:

$$r(z_a, a) = \left[\frac{1 - (\eta - 1)\frac{\mathcal{M}^l(a)}{\mathcal{M}^v(a)}}{\eta - 1} - \frac{\xi}{\zeta + 1 - \eta}\left(\eta \frac{\mathcal{M}^l(a)}{\mathcal{M}^v(a)} - 1\right) \right]$$
$$\times \left(\frac{\zeta + 1 - \eta}{\eta(1 - \xi) - 1}\right)\left(1 - \frac{A_t^{\eta - 1}}{A_m^{\eta - 1}}\right)\mathcal{M}^v(a). \tag{20}$$

Again, a multiplier exceeding one (i.e., amplification) requires $r(z_a, a) > 0$, which

16

is equivalent to

$$1 - (\eta - 1)\frac{\mathcal{M}^l(a)}{\mathcal{M}^v(a)} > \frac{\xi(\eta - 1)}{\zeta + 1 - \eta}\left(\eta\frac{\mathcal{M}^l(a)}{\mathcal{M}^v(a)} - 1\right) \ . \tag{21}$$

This condition is automatically satisfied when the mass of adopters is small, $a \approx 0$, because then $\mathcal{M}^l(a)/\mathcal{M}^v(a) \approx 0$ and $\zeta + 1 - \eta$ is assumed to be positive for aggregates to be finite. Thus, with idiosyncratic distortions, an economy with high adoption costs (i.e., a small mass of adopters) will feature a multiplier greater than one for adoption subsidies. To the contrary, distortions tend to make amplification less likely when the mass of adopters is large, $a \approx 1$, since condition (21) is stronger than condition (17) when $\mathcal{M}^l(a)/\mathcal{M}^v(a) \approx 1$.

How do distortions affect the necessary condition for multiple equilibria? As before, because multiplicity requires $T'(a) > 1$, the condition for multiplicity is stronger than that for amplification. From equation (20), holding fixed $\mathcal{M}^v(a)$, more distortions relax the condition for multiplicity. In the limiting case of $\lim_{a \to 1} \mathcal{M}^l(a)/\mathcal{M}^v(a) = 1$, the necessary condition for multiplicity is

$$2 - \eta > \frac{1}{1 + \frac{\zeta}{1-\xi}} \ ,$$

which is weaker than condition (18) for $\xi > 0$. How ξ reduces the denominator of the right-hand side of equation (19) is the dominant force. When $a < 1$, we have $\mathcal{M}^l(a)/\mathcal{M}^v(a) < 1$, which further relaxes the necessary condition for multiplicity.

3.3   Discussion of the Role of Additional Elements

In this section we briefly describe two additional elements of the model. We first discuss the role of intermediate input in production. The use of intermediate input strengthens the role of complementarities in adoption, making both amplification and multiple equilibria more likely. We then describe the role of the entry margin, and discuss its relevance for the size of firms in the data.

Intermediate Input in Production. One important difference between our benchmark quantitative model and the simple models explored in Sections 3.1 and 3.2 is the use of differentiated goods. While in the simple models differentiated goods are used only as intermediate input for the production of the final good (to be

used for consumption and adoption costs), in the quantitative model an aggregated quantity of the differentiated goods is also used as an intermediate input in the production of the differentiated goods themselves (i.e., round-about production). The importance of intermediate input in production is governed by the intermediate input intensity in the production of the adoption good, γ, and the intermediate input intensity of the traditional and the modern technologies in the production of the differentiated goods, ν_t and ν_m, respectively.

In the simple examples, we set the intensity of the intermediate aggregate in the production of the adoption good to be equal to one, i.e. $\gamma = 1$. If $\gamma = 0$ in the simple examples, the effect of policies would be dampened by a negative equilibrium feedback, and neither amplification nor multiplicity would arise. A higher γ leads to stronger complementarity: More adoption reduces the adoption costs by more when a larger fraction of the costs is in units of goods rather than labor. The same logic applies to the intensity of the intermediate aggregate in the production of the differentiated goods themselves.

Consider first the case where the intermediate input intensity in the production of the differentiated goods is the same for both the traditional and the modern technologies: $\nu_t = \nu_m = \nu$. The use of the intermediate aggregate as an input in the production of the differentiated goods strengthens complementarities and the equilibrium feedback for two reasons. First, more adoption now reduces the price of firms' intermediate input and hence their production costs, partly offsetting the negative competition effect on firms' profits, which in turn relaxes the restrictions on the elasticity of substitution across differentiated goods η. For instance, when $\nu = 1$ in the simple examples, the necessary condition for amplification and multiplicity is always satisfied for any $\gamma \in [0, 1]$, independently of η. Second, as in the simple examples, the strength of the net effect of more adoption on the gains from adoption depends on η. When the intermediate aggregate is itself an input in the production of differentiated goods, the strength of the net effect also depends on ν, through a standard intermediate input multiplier $1/(1 - \nu)$, as in Jones (2011).

When the modern technology uses intermediate input more intensively, that is, $\nu_m > \nu_t$ as in our quantitative model, there is an additional positive effect of more adoption on firms' incentive to adopt the modern technology. More adoption reduces the price of the intermediate input relative to labor, raising the relative profitability of the technology that uses the intermediate input more intensively. In this case,

amplification and multiplicity can arise even when the production of the adoption goods requires only labor ($\gamma = 0$), which shares some similarity with the result in Ciccone (2002).

In summary, the goods share of the adoption costs and the intermediate input intensity in the differentiated goods production both contribute to amplification and multiplicity, but neither is absolutely necessary for such outcomes.

Entry Margin. Relative to the simple models explored above, our quantitative model also has the firm entry margin: A firm must pay a fixed labor cost $\kappa_e > 0$ to become active. If firms' productivity follows a Pareto distribution, the entry margin is summarized by the productivity of the marginal entrant z_e, which is given by

$$
z_e = \max \left\{ 1, \left[1 + \frac{(\eta - 1)\zeta}{\zeta - (\eta - 1)} \left[1 + \left(\left(\frac{A_m}{A_t} \right)^{\eta - 1} - 1 \right) a^{1 - \frac{\eta - 1}{\zeta}} \right] \right]^{\frac{1}{\zeta}} \kappa_e^{\frac{1}{\zeta}} \right\}. \tag{22}
$$

The productivity of the marginal entrant z_e and the average size of firms z_e^ζ are an increasing function of the fraction of adopters a, the cost of entry κ_e, and the ratio of the productivity shifter of the modern to tradition technology, A_m/A_t.

4 Identification of Model Parameters

Section 3 shows that several features of the model can generate multiple equilibria. In general, parameter identification is not granted when a model features multiple equilibria (see, for example, Jovanovic, 1989), because the mapping from the data to the model parameters may not be unique. In this section, we construct an argument that allows us to uniquely identify the parameters of the model despite potential multiplicity. The key assumption for our identification strategy is that traditional and modern firms coexist in the data. However, the strategy does not presuppose multiplicity. Rather, once the parameters are uniquely identified from the data, we check whether or not the model has any other equilibrium for those parameter values.

To keep things as simple as possible, we provide our constructive argument for the case where the intermediate input elasticity is the same for the modern and the traditional technology, i.e., $\nu_t = \nu_m = \nu$, and the adoption good production only uses intermediate input ($\gamma = 1$). In addition, following the common practice in the literature, we maintain the assumption in Section 3 that firm-level productivity

distribution F is Pareto with a tail parameter ζ. We set the elasticity of substitution across differentiated goods η from outside of the model, using estimates common in the literature on demand or profit share estimation. We also assume that there is no distortion $\xi = 0$. However, our identification strategy also holds without these parameter restrictions.

Our goal here is to identify the following six parameters: the technology parameters A_t and A_m, the entry and adoption costs κ_e and κ_a, the parameter of the productivity distribution ζ, and the intermediate input elasticity ν. We normalize the productivity of the modern technology to one, $A_m = 1$. This leaves us with five parameters. For identification, we rely on the size distribution of establishments in the data $G(l)$, where size is defined as the number of employees, as well as the intermediate input share in the data.

We begin by showing that, for a given η, we can identify the intermediate input elasticity ν directly from the intermediate input share in the data:

$$\nu = \frac{\eta}{\eta - 1} \times \text{intermediate input share} . \tag{23}$$

The other four parameters are identified from the establishment size distribution in the data. In particular, we rely on the implications of the theory for the relationship between the log of employment ($\log l$) and the log of the fraction of establishments with size larger than l as illustrated in Figure 2. We hereafter refer to this relationship simply as the log-log relationship. Our identification strategy relies on the three thresholds in the figure, which must exist under the assumption that both traditional and modern technologies are used in the economy: (i) the size of the smallest entrant $\underline{l}_t \equiv l_t(z_e)$, (ii) the size of the largest establishment using the traditional technology $\bar{l}_t \equiv l_t(z_a)$, and (iii) the size of the smallest establishment operating the modern technology $\underline{l}_m \equiv l_m(z_a)$.

From the slope in the right tail of the log-log relationship, i.e., $l > \underline{l}_m$, we identify the tail parameter of the productivity distribution ζ for a given η:

$$\frac{\zeta}{\eta} = -\frac{d \log(1 - G(l))}{d \log l} \quad \Rightarrow \quad \zeta = -\eta \frac{d \log(1 - G(l))}{d \log l} . \tag{24}$$

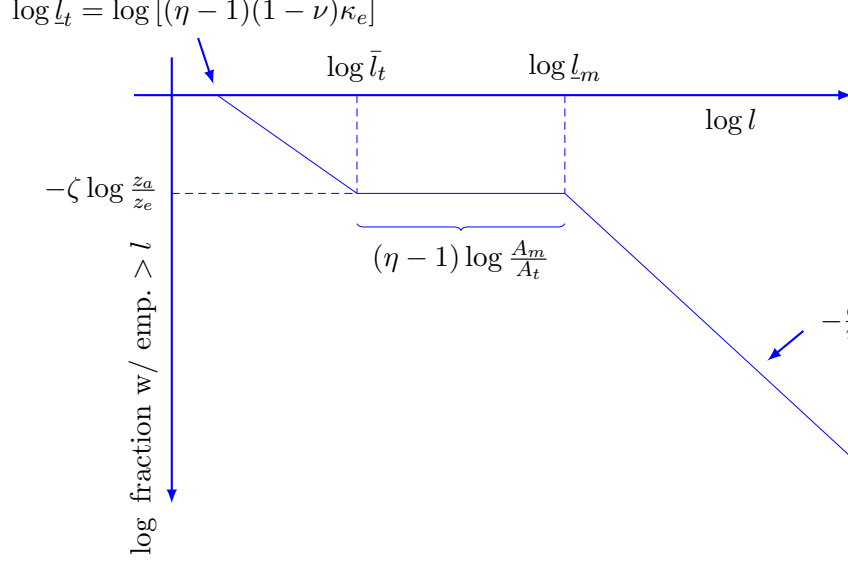Given the value of the intermediate input elasticity ν, the size of the smallest

Figure 2: Identification from the Establishment Size Distribution

establishment is a simple function of the entry cost, pinning down κ_e:

$$\underline{l}_t = [(\eta - 1)(1 - \nu)\kappa_e] \;\Rightarrow\; \kappa_e = \frac{\underline{l}_t}{(\eta - 1)(1 - \nu)}. \tag{25}$$

The theory implies that there should be a gap in the size distribution of establishments, if both the modern and the traditional technologies are operated in the economy.[10] In particular, there should be no establishment larger than the largest establishment using the traditional technology $\bar{l}_t$ but smaller than the smallest establishment operating the modern technology $\underline{l}_m$; i.e., $G(l) = G(\underline{l}_m) = G(\bar{l}_t)$ for $l \in [\bar{l}_t, \underline{l}_m]$. The difference between these two employment levels is a function of the relative productivity of the two technologies, A_m/A_t, which, given the knowledge of η and the normalization of $A_m = 1$, identifies A_t:

$$\log \underline{l}_m - \log \bar{l}_t = (\eta - 1)\log\left(\frac{A_m}{A_t}\right) \;\Rightarrow\; A_t = \left[\frac{\bar{l}_t}{\underline{l}_m}\right]^{\frac{1}{\eta - 1}}.$$

Finally, to identify the adoption cost κ_a we use the equilibrium condition relating the ratio of the two thresholds to the ratio of the adoption and the entry costs:

$$\kappa_e \left(\frac{z_a}{z_e}\right)^{\eta - 1} = \frac{\frac{P}{w}}{\frac{A_m^{\eta-1}}{A_t^{\eta-1}} - 1}\kappa_a \;\Rightarrow\; \kappa_a = \frac{G\left(\bar{l}_t\right)^{-\frac{\eta-1}{\zeta}}\left(\frac{A_m^{\eta-1}}{A_t^{\eta-1}} - 1\right)}{\frac{P}{w}}\kappa_e\;.$$

[10]This is akin to the concept of missing middle in Tybout (2000).

21

5 Calibration

We use data from the US and India to calibrate the parameters of the model. We think of the US as an economy that is distortion-free ($\xi = 0$) and therefore informative about the tail parameter of the firm productivity distribution ζ, among others. India is a large developing economy with relatively good data availability and is informative about the degree of idiosyncratic distortions ξ in particular. This choice is consistent with the evidence on the relationship between productivity and idiosyncratic distortions in the US and India in Hsieh and Klenow (2014, p.1059).

For our benchmark model without the simplifying assumptions of Section 3 or 4, the following 11 parameters need to be calibrated: the elasticity of substitution among differentiated goods η; the share of intermediate input in the production of the adoption good γ; the intermediate input elasticity of the modern and the traditional technology ν_m and ν_t; the productivity levels of the modern and the traditional technology A_m and A_t; the Pareto tail parameter of the firm productivity distribution ζ; the entry and the adoption costs κ_e and κ_a; and finally the degree of idiosyncratic distortions ξ and the budget-balancing scale parameter τ.

Six of the eleven parameters are assumed to be the same for the US and India: η, γ, ν_m, ν_t, A_m and ζ. Four of these six are fixed outside of the model. We maintain the normalization of $A_m = 1$. We set $\eta = 3$, which is considered to be on the lower side, as discussed in Hsieh and Klenow (2009).[11] As in Section 4, we set $\gamma = 1$, i.e., the adoption good production uses intermediate goods but not labor. In addition, we set $\nu_t = 0$, so that labor is the only input of the traditional technology. This last assumption maximizes the difference between the two technologies' intermediate input elasticity, ν_m and ν_t. Our choices of η, γ and ν_t make amplification and multiplicity more probable as explained in Section 3, but the conclusions from our quantitative analysis do not rest on these assumptions. In Section 6.3, we show the result with $\gamma = 0$—i.e., labor is the only input of the adoption good production—and with $\nu_t = \nu_m$. Appendix B reports a sensitivity analysis with different values of η and ν_t.

One of the remaining common parameters, ν_m, is then calibrated to match the intermediate input share in the US data, yielding $\nu_m = 0.70$.[12] The other, the

---

[11]In our simple example of Section 3.1, the necessary condition for amplification and multiplicity was $\eta < 2$. However, as explained in Section 3.3, with intermediate input in the production of the differentiated goods, the restriction on η is relaxed.

[12]The intermediate input share in the US in 2007 was 0.46, calculated from the BEA input-

tail parameter of the productivity distribution, is calibrated to match the tail of the establishment size distribution for the US in the Census Bureau's 2007 Business Dynamics Statistics (BDS), giving $\zeta = 2.42$ under the assumption of no distortion ($\xi = 0$ and $\tau = 1$).

We now describe how we calibrate the five parameters that differ between the US and India: two distortion parameters (ξ, τ) and three technology parameters $(A_t, \kappa_e, \kappa_a)$. Following the procedure in Section 4, we calibrate these country-specific parameters from their establishment size distribution only. In particular, we do not use any information on the level of income or productivity in either country. Since there is a priori no tight relationship between a country's establishment size distribution and its income level, it is an open question how the model-predicted income gap between the US and India will measure up to the data.

Since we assume that the US has no distortions, the US calibration has $\xi = 0$ and $\tau = 1$. For India, given the common tail parameter of the productivity distribution ζ calibrated to the US data, we calibrate its degree of distortions ξ to match the tail of the establishment size distribution, utilizing

$$\frac{\zeta}{\eta\,(1 - \xi) - 1} = -\frac{\partial \log(1 - G(l))}{\partial \log l} \ , \tag{26}$$

which is the distortion-augmented version of equation (24).

The calibration of the other three country-specific parameters, A_t, κ_e, and κ_a, closely follows the procedure in Section 4: They are chosen to match many features of the empirical establishment size distribution (the log-log relationship) for each country. The empirical moments are chosen to capture the non-linearities of the log-log relationship in Figure 2. For the US we use eight points from the empirical log-log relationship, and for India 26 points.

Figure 3 is the log-log plots of the establishment size distribution from the calibrated model for the US and India, together with the empirical log-log relationship. The US data comes from the 2007 BDS. To produce the figure for India, we use the size distribution of establishments for the entire Indian economy constructed by Buera et al. (2020), who combine data from the Fifth Economic Census in 2005 and the 2003 Survey of Land and Livestock Holdings carried out in

---

output tables. Because we assume that the traditional firms use no intermediate input ($\nu_t = 0$), the intermediate input share of the modern firms has to be 0.47 in order for the share in the entire economy to be 0.46. Multiplying 0.47 by $\frac{\eta}{\eta - 1}$ as in (23), we obtain $\nu_m = 0.70$.
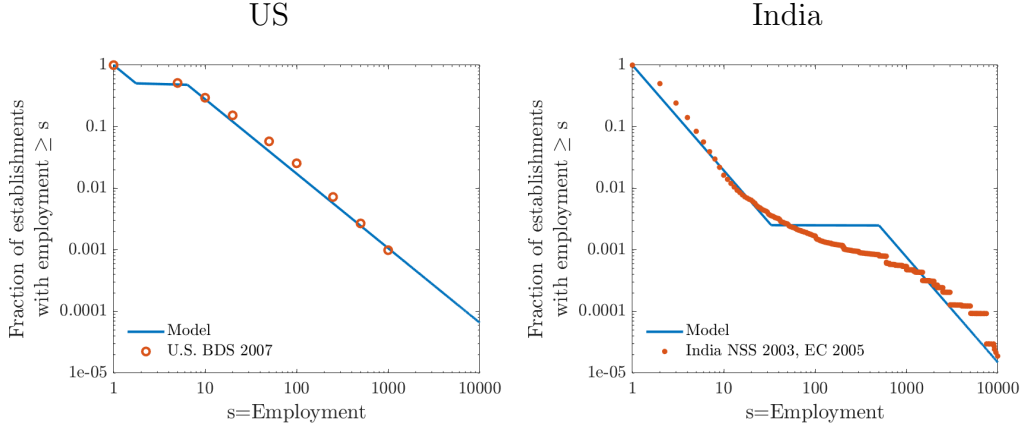
US　　　　　　　　　　　　　India

Note: The source of the US data is the 2007 Business Dynamics Statistics. For the Indian data, we combine the 2003 National Sample Survey and the 2005 Economic Census. See Buera et al. (2020) for detail.

Figure 3: Establishment Size Distribution: Model and Data

the 59th round of the National Sample Survey (NSS). The Census has comprehensive information for all entrepreneurial units, excluding agriculture. The NSS provides information on employment by productive units in agriculture. In order to obtain an accurate establishment size distribution for the entire Indian economy, it is crucial to account for its agricultural sector, which accounted for 57 percent of the total employment in 2004.[13]

As shown in the left panel, the model calibrated to the US (solid line) generates a flat region that is the size gap between firms using the traditional technology and those using the modern technology, in order to match the concavity of the log-log relationship in the data (circles) for small establishments. The vertical location of the flat region shows that roughly half of all establishments use the modern technology. The Indian calibration in the right panel also shows both the traditional and the modern technology in use in equilibrium, separated by a flat region (solid line). However, in India a much larger fraction of firms uses the traditional technology, as evidenced by the fact that less than 1 percent (0.25 percent, to be precise) of establishments are to the right of the flat region in the calibrated economy. The calibrated model captures the conspicuously flat region over intermediate establishment sizes in the Indian data (dots): a missing middle.[14]

---

[13]By comparison, the agricultural employment share was only 1.4 percent in 2007 in the US.

[14]This does not contradict Hsieh and Olken (2014), who do not find a missing middle in Indian manufacturing. Buera et al. (2020) find a missing middle when studying the broad cross-section of

24

Parameter	US	India
Elasticity of substitution, η		3
Intermediate aggregate share in adoption good production, γ		1
Productivity distribution Pareto tail parameter, ζ		2.42
Modern technology productivity, A_m		1
Modern technology intermediate input elasticity ν_m		0.70
Traditional technology intermediate input elasticity, ν_t		0
Entry cost, κ_e	0.50	0.50
Traditional technology, A_t	0.43	0.07
Adoption cost, κ_a	16	272
Degree of distortions, ξ	0	0.19
Distortion scale parameter, τ	1	2.14

Table I: Calibrated Parameters

In Table I we report the calibrated parameters. Some remarks are in order. First, the US and India have the same entry cost κ_e, which is identified from the size of the smallest establishment using (25). Because the smallest establishment is of the same size in both countries (one employee) and we assume that η and ν_t are the same for both countries, so is κ_e. However, this does not mean that the entry rate of firms is the same in the two countries. In fact, as shown in Table II, the fraction of firms that enter in India is three times that in the US. Second, the traditional technology parameter A_t for the US is six times that for India. Since both countries have the same productivity level of the modern technology A_m by assumption, the technology gap between the modern and the traditional technology is six times as high in India. Third, the cost of adoption is 17 times higher in India. The cost of adoption in India must be higher in order to rationalize the minuscule fraction of firms adopting the modern technology in spite of the enormous productivity gains from doing so.[15] When measured in units of labor ($P\kappa_a$, since $\gamma = 1$), the cost of adoption in India is over 80 times higher, because the (endogenous) price of the intermediate aggregate is five times higher in India.

Finally, the Indian calibration exhibits significant idiosyncratic distortions, as given by $\xi = 0.19$. The tail of the establishment size distribution in India is thinner and hence the log-log relationship in the right tail is steeper. Equation (26) pins down

---

establishments, and within narrow sectors, e.g., the agricultural sector.

[15]The high adoption cost can be viewed as standing in for other inhibitors of technology adoption that are not explicitly modeled in our theory, such as the shortage of skilled labor necessary for using the modern technology, financial constraints, and bureaucratic or anti-competitive barriers to adoption.

|	US	India
Gross domestic product per capita	4.34	0.66
Consumption per capita	3.92	0.54
Average establishment size	19.0	5.7
Fraction of entrants	0.05	0.17
Fraction of entrants that adopt A_m	0.50	0.003
Employment share of A_m firms	0.96	0.58
Value added share of A_m firms	0.98	0.81

Table II: Statistics from the Calibrated Economy

ξ.[16] The implied firm-level revenue taxes and subsidies are not extreme. As we show in Figure 12 in the appendix, the firm at the top 10^{-4} percentile of the active firm productivity distribution is taxed at about 40 percent, while the maximum subsidy is 80 percent for the least productive firm with $z = z_e$.

At face value, some of the calibrated parameters in Table I seem contradictory to the fraction of entrants reported in Table II: The US exhibits less firm entry than India in spite of the substantially lower cost of adopting the modern technology and the significantly better traditional technology. However, as explained in Section 3, when more firms adopt the modern technology, they crowd out less productive firms through the negative competition effect as well as through higher wages. That is, the general equilibrium effects from more firms adopting the modern technology are responsible for the lower entry rate in the US. In spite of the lower entry rate in the US, the significantly higher rates of modern technology adoption means that the US GDP per capita is nearly seven times that of India.

This last result is a success for the model. Even though the calibration is based on the difference in the establishment size distribution between the US and India and does not use any information on the income or productivity gap between the two, the model generates a huge income gap. In the Penn World Tables the GDP per worker of India is 6 percent of the US level in 2005, while in the model it is 15 percent. That is, the model accounts for 73 percent of the US-India income gap.[17]

---

[16]A back-of-the-envelope calculation of the Indian ξ from Hsieh and Klenow (2014, p.1059) gives 0.4. One possible interpretation is that our model elements amplify the impact of distortions so that a lower degree of idiosyncratic distortions can match the Indian data.

[17]We use the output-side real GDP at chained PPPs and the number of persons engaged from the Penn World Tables 9.0. The income gap in the data is a factor of 16.7 and in our model 6.7. Since $16.7 \approx 6.7 \times 2.5$, the model explains $6.7/(6.7 + 2.5) \times 100 \approx 73$ percent of the actual income gap.
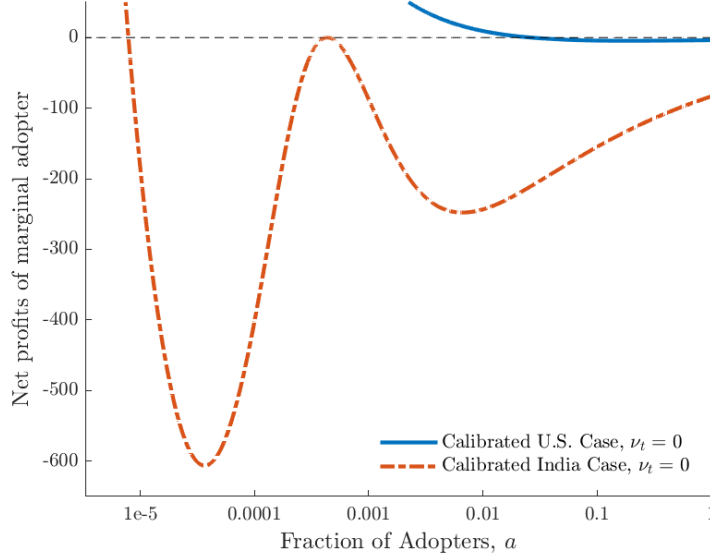
Note: The figure shows, for the US and India, the gains for the marginal adopter from operating the modern technology, net of adoption costs.

Figure 4: Net Gain of the Marginal Adopter, US and India

6   Quantitative Exploration

In this section we explore quantitatively the role of the various model elements—mechanisms and parameter values—in amplifying the impact of distortions, generating coordination failures, and ultimately accounting for the US-India income gap. We also discuss what a Big Push in distorted economies is.

6.1   Multiple Equilibria and the US-India Income Difference

The calibrated US economy has a unique equilibrium but the Indian economy features multiple equilibria. We show this by examining the net gains from adopting the modern technology for the marginal adopter with productivity z_a,

$$\pi_m^o(z_a) - \pi_t^o(z_a) - P\kappa_a \ ,$$

which must be 0 in an equilibrium. Figure 4 shows this object for the US and India. As we vary z_a, the price of the intermediate aggregate and the wage adjust to clear markets.

For the US (solid line), the net gains are monotonically decreasing in the fraction

	US w/ Indian Parameters	India w/ US Parameters
Benchmark	1.0	0.14
Adoption cost, κ_a	0.37	0.71
Degree of distortions, ξ	0.41	0.34
Traditional technology, A_t	1.03	0.19

Table III: Explaining Consumption Difference

of adopters (and hence increasing in the productivity of the marginal adopter z_a) and intersect the zero line once. This intersection is the unique, stable equilibrium. For India (dashed line), the net gains cross the zero line three times, twice from above and once from below. The leftmost intersection is the stable, low adoption or "bad" equilibrium, and the rightmost one is the stable, high adoption or "good" equilibrium. The one in the middle is unstable. Our calibration selects the good equilibrium to match the Indian data. That is, despite equilibrium multiplicity, coordination failures do not explain the observed income difference between the US and India. As we show in Section 6.2, if Indian firms were to fail to coordinate and get trapped in the bad equilibrium, India's GDP would further shrink by a factor of 4. That is, the US-India income gap could have been a factor of 28 rather than 7 in Table II.

If coordination failures do not account for the income difference between the US and India, then what does? The two countries in our model have different productivity of the traditional technology A_t, adoption costs κ_a, and distortions ξ, all identified only from their establishment size distribution.[18] In Table III, we calculate the contribution of each of these elements to the US-India gap in per-capita consumption. To do so, we compute the hypothetical aggregate consumption of the US by replacing one of the parameters with its value in the Indian calibration, holding all others constant. This result is in the first column, where we replace κ_a, ξ, and A_t one by one. In the second column, we do the reverse: Starting from the Indian calibration, we replace one of the parameters with its value in the US calibration. All per-capita consumption is normalized by the US level in the benchmark calibration.

The first row shows that the model generates a factor of 7 difference between the US and Indian consumption ($=1/0.14$).[19] Starting from the US calibration, we see that the adoption cost difference has the largest impact: Giving the US

---

[18]The entry cost is also country-specific, but the calibrated κ_e's for the US and India coincide. As ξ changes, τ adjusts to balance the budget.

[19]The gap in consumption is slightly larger than the gap in output, because entry costs and adoption costs are a larger fraction of the output in India than in the US.

the high adoption cost of India shrinks the US consumption by a factor of 2.7 (=1/0.37). The role of idiosyncratic distortions is of a similar magnitude: Introducing idiosyncratic distortions of Indian proportions ($\xi = 0.19$) to the US economy reduces the consumption by a factor of 2.5 (=1/0.41). The last row of the first column shows that, if we replace the traditional technology productivity A_t of the US with the lower value from India, the US consumption actually rises modestly. This is because the very low A_t leads to more adoption of the modern technology.[20]

The same set of counterfactual exercises for India in the second column leads to similar conclusions, although now adoption costs play a more important role. Giving India the much lower adoption cost of the US while holding all other parameters constant results in a five-fold increase in consumption, which is much larger than the factor of 2.7 in the first column: The rise in adoption caused by the lower adoption costs represents a larger increase in productivity when the traditional technology is less productive as in India. Eliminating idiosyncratic distortions in India raises consumption by a factor of 2.4 (=0.34/0.14), which is nearly identical to the result in the first column (2.5). Finally, replacing the traditional technology A_t with the higher US value has a modest positive effect on Indian consumption. The higher A_t nearly doubles the number of firms but further discourages the adoption of the modern technology.

To summarize, the model nearly replicates the large income gap between the US and India in the data, without directly targeting the income or productivity level of either country. Coordination failures turn out not to be part of the story despite the existence of multiple equilibria, and adoption costs and distortions explain most of the income gap.

6.2   Multiplicity, Amplification and Distortions

In this section we further explore the role of adoption costs κ_a, relative technology productivity A_m/A_t, and idiosyncratic distortions ξ. We first identify the set of these three parameter values that generates multiple equilibria, holding fixed the other parameters as calibrated. Second, we show how per-capita income and the average size of firms change with the idiosyncratic distortion ξ. This exercise showcases the potentially huge effect of distortions with or without equilibrium multiplicity.

---

[20]In this case, the number of entrants is nearly halved, but all the entrants adopt the modern technology.
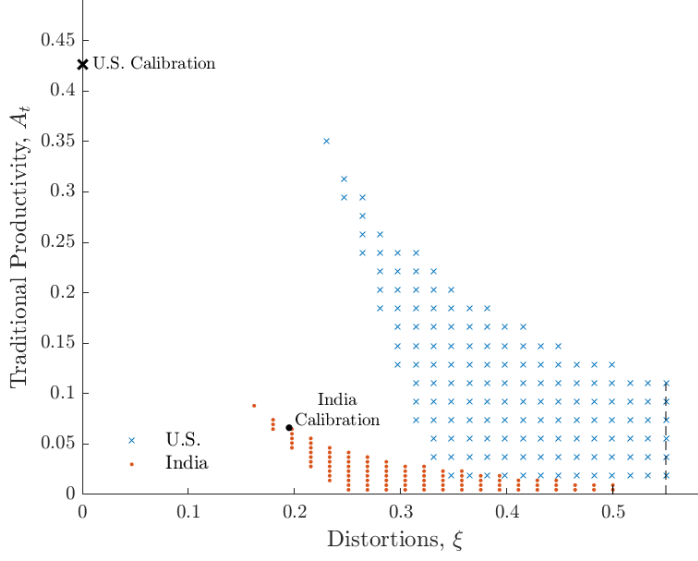
Note: The figure shows, for the US (cross) and India (dot), the combination of ξ and A_t for which multiple equilibria exist. The larger cross and dot correspond to the calibrated US and India respectively.

Figure 5: Region of Multiple Equilibria, (ξ, A_t) Space

Region of Multiplicity Figure 5 shows, for a low adoption cost economy (the US, cross) and a high adoption cost economy (India, dot), the combination of the distortion parameter ξ and the traditional productivity A_t that generates multiple equilibria. (Recall that A_m for both is normalized to 1.) We hold all other parameters fixed at their respective calibrated values, except that we adjust τ so that the budget balances. The larger cross and dot represent the calibrated US and Indian economies, respectively. We see that India is in the region of multiplicity but the US is not.

Multiple equilibria arise for economies with both high degrees of distortions and unproductive traditional technology, toward the lower right corner of the figure. One interesting result is that, holding fixed the productivity of the traditional technology A_t, as we increase idiosyncratic distortions ξ (moving horizontally), we enter and then exit the region of multiplicity. To the right of the region, the only equilibrium is the one with nearly no adoption. Similarly, holding fixed the degree of idiosyncratic distortions, as we lower the productivity of the traditional technology A_t (moving downward), we enter and then exit the region of multiplicity, although it is hard to see this for the high adoption cost case (India, dot). The unique equilibrium with A_t close to 0 has a small number of entrants, nearly all of whom adopt the

US **India**
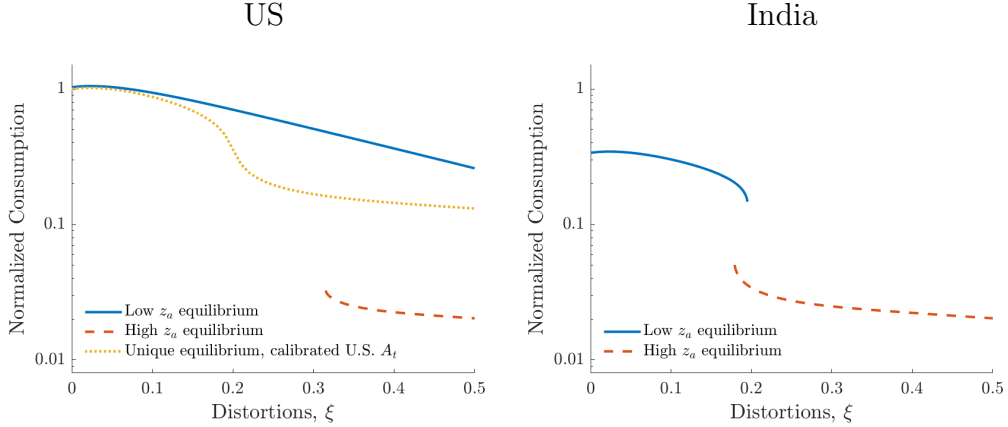
Note: Equilibrium consumption per capita as ξ goes from 0 to 0.5. Consumption is normalized by its level in the US calibration and in log scale. The dotted line in the left panel is the no-multiplicity result of the US. For the modified US and the India cases, the solid lines are the high adoption equilibrium (low adoption threshold z_a) and the dashed lines are the low adoption equilibrium (high z_a).

Figure 6: Distortions and Consumption per Capita

modern technology: With a useless traditional technology, entry also implies adopting the modern technology, which effectively raises the cost of entry and results in few entrants. Finally, the region of multiple equilibria is smaller for the high adoption cost economy. In this case, the model features a unique equilibrium with few adopters in most of the (ξ, A_t) space. However, multiplicity can occur for smaller degrees of idiosyncratic distortions (as low as $\xi = 0.17$) than in the low adoption cost economy.

Nonlinear Impact of Distortions   We now explore the role of distortions in generating large differences in income levels, with or without multiple equilibria. We start with the US and Indian calibration, and vary the degree of distortions ξ, holding fixed the other parameters at their respective calibrated values, except that τ adjusts to balance the budget. In addition, since there is no multiplicity for any ξ in the US calibration, we also consider a modified US case that has the lower traditional technology productivity A_t of India.

In Figure 6 we show the equilibrium consumption per capita as we vary the degree of distortions ξ for the US (left panel) and India (right panel). The consumption per-capita in the vertical axis (log scale) is normalized by the per-capita consumption in the US calibration.

For the US, the equilibrium is unique for any value of ξ (dotted line). There are

31

two notable features. First, the impact of distortions is large, reducing consumption by nearly 90 percent for large values of ξ. Second, for intermediate values of ξ, small changes in the degree of distortions have a highly nonlinear effect on consumption. That is, even without multiplicity, distortions can have an amplified impact. We discuss these features more rigorously in Section 6.3.

We now turn to the modified US case (i.e., with India's A_t) in the left panel and the Indian case in the right panel, both of which exhibit multiplicity. When the distortions are small, the equilibrium is unique in both countries. The solid line is the per-capita consumption in the equilibrium with a higher fraction of adopters, or the good equilibrium. As we increase distortions, a second equilibrium, one with a small fraction of adopters emerges (dashed line). This is the bad equilibrium. For India, both equilibria exist over a short interval of intermediate values of ξ, above which only the bad equilibrium survives.

Tracing either the good (solid line) or the bad (dashed line) equilibrium, idiosyncratic distortions have moderate to large effects on consumption. The effect is even larger, however, since distortions can make the economy jump between the two lines. Near the boundaries of the region of multiplicity, the effect of distortions are extremely disproportionate. Once the economy enters the multiplicity region from left, coordination failures can send the economy to the bad equilibrium. On the other hand, even without better coordination, a small reduction in distortions can push the economy from the bad equilibrium region to the unique good equilibrium region, which discontinuously increases consumption. This happens around $\xi = 0.3$ in the modified US case and around $\xi = 0.2$ in the Indian case. Our calibrated Indian economy is in the good equilibrium with $\xi = 0.19$, near the end of the solid line. If ξ were to rise past the narrow interval of multiplicity, its per-capita consumption will shrink by a factor of 5, without coordination failures playing any role.

These results highlight the potentially disproportionate gains from reducing idiosyncratic distortions. Multiplicity is an extreme form of amplification. However, as the no-multiplicity US case shows (dotted line), even without multiplicity, our model elements that are responsible for complementarity in adoption decisions amplifies the impact of distortions to a magnitude unattainable in conventional models, a statement we make precise in Sections 6.3 and 6.5.

Figure 7 illustrates the corresponding effect of distortions on the average size of firms, measured by the number of employees. The vertical axis is in log scale. For

US
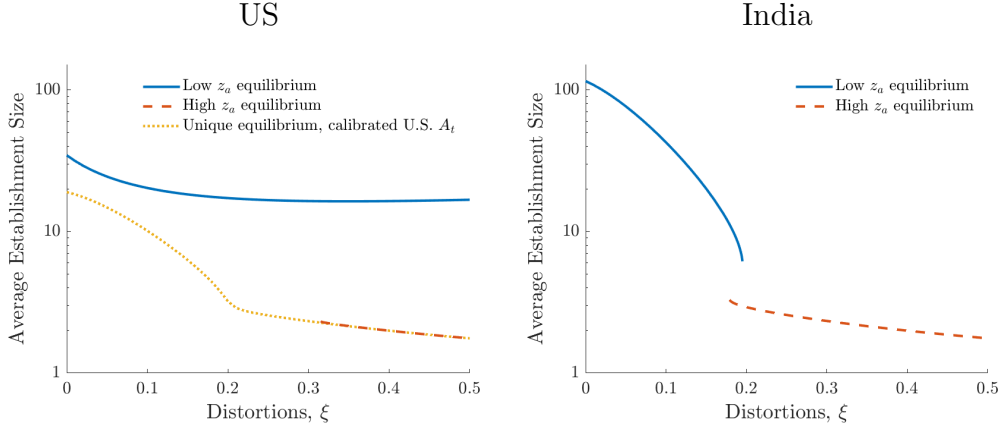
India

Note: Average firm size (the number of employees, in log scale) as ξ goes from 0 to 0.5. The dotted line in the left panel is the no-multiplicity result of the US. For the modified US and the India cases, the solid lines are the high adoption equilibrium (low adoption threshold z_a) and the dashed lines are the low adoption equilibrium (high z_a).

Figure 7: Distortions and Average Firm Size

the modified US and the India cases, the solid lines trace the average size of firms in the good equilibrium. In this equilibrium, there are fewer entrants but many of them adopt the modern technology, resulting in large firms. As distortions get bigger, there is more entry but less adoption, bringing down the average firm size, which is more pronounced in the high adoption cost economy (right panel). The dashed lines, which appear when distortions are high enough, are the average firm size in the bad equilibrium. In this equilibrium, the number of entrants is large but very few adopt the modern technology, which implies that firms are small on average, in line with equation (22).[21]

6.3   Unpacking the Mechanisms

Our model combines several elements whose importance in explaining cross-country income differences has been studied in isolation in the literature. In this section we illustrate the role of each element, in comparison with the findings in the literature. In Section 6.3.1, we examine how the model elements amplify the impact of idiosyncratic distortions. In Section 6.3.2, we analyze which elements are responsible for the large

---

[21]The modified US case uses India's A_t, which is smaller than its value in the US calibration. As a result, the modified US with no distortion ($\xi = 0$) has more adopters and a larger average firm size than the benchmark calibration (dotted line). The direct negative effect of the traditional technology productivity on the average firm size can be seen in equation (22).

income gap between the US and India that our calibrated model generates.

6.3.1 Nonlinear Impact of Idiosyncratic Distortions

We start with the basic model in the distortion literature that abstracts from technology choices and input-output linkages, as in Restuccia and Rogerson (2008) for example. Next, we consider a case with a technology choice, but without intermediate input or round-about production, which is our rendition of Bento and Restuccia (2017). We then introduce input-output linkages in the form of round-about production but remove the technology choice. This is our adaptation of Jones (2011). The fourth one has a technology choice and round-about production as in our benchmark model, but with the modification that the adoption costs are in units of labor only instead of goods. Finally, we consider a case with technology choices, round-about production, adoption costs in goods, except that both the modern and the traditional technologies have the same intermediate input intensity ($\nu_t = \nu_m$), unlike in our benchmark with $\nu_t \ll \nu_m$.

For each model, we re-calibrate the parameters to the same set of target moments as in our US benchmark (Section 5) and calculate the effect of idiosyncratic distortions. Below we present the results for the US calibration of each model, which does not feature multiple equilibria for any degree of distortions ξ. The left panel of Figure 8 shows the effect of distortions (ξ) on consumption per capita, and the right panel the effect on the fraction of adopters.

The solid line reproduces the effect of distortions in our benchmark economy (the dotted line in the left panel of Figure 6). It confirms that idiosyncratic distortions have a large negative effect on aggregate consumption, which is particularly pronounced around $\xi = 0.2$, close to its value in the Indian calibration ($\xi = 0.19$). Consumption is down by 60 percent and the fraction of adopters collapses to nearly 0 at $\xi = 0.2$.

The first comparison model we consider is the basic model in the distortion literature that abstracts from intermediate inputs and technology adoption ($\nu_t = \nu_m = 0$ and $A_t = A_m = 0.69$, with re-calibration), shown by the dashed line. This specification should be considered the polar opposite of our benchmark model. For this model, consumption falls almost linearly with the distortion parameter ξ in the semi-log scale and by much less than in the baseline model. At $\xi = 0.2$, consumption goes down by less than 20 percent from its no-distortion level. Even with $\xi = 0.5$, the loss in consumption is only 30 percent.

Consumption   Adoption
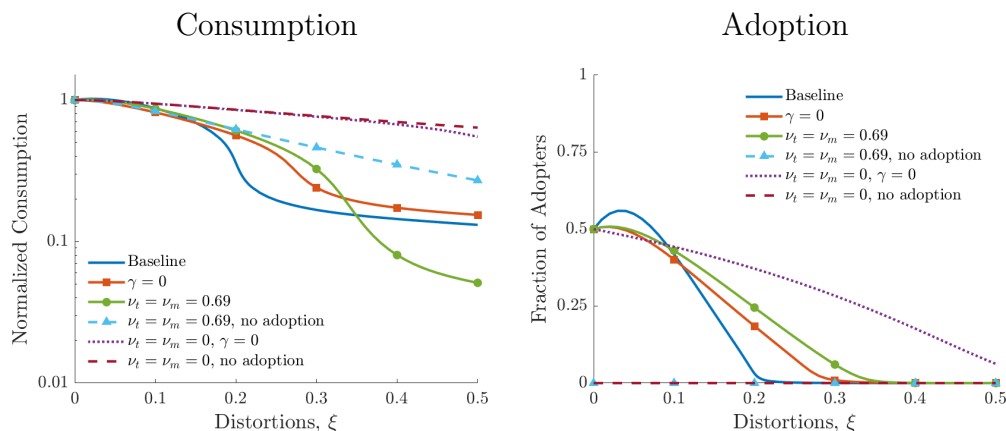
Notes: The figure presents the equilibrium consumption per capita and fraction of adopters in the US as ξ goes from 0 to 0.5. Consumption is normalized by the consumption in the re-calibrated, no-distortion US economy and is in log scale.

Figure 8: Unpacking the Mechanisms

Next, the dotted line introduces technology adoption to the basic model but without intermediate input. Consistent with the literature, for example, Bento and Restuccia (2017), introducing the technology adoption by itself makes the effect of distortions on consumption only marginally bigger and only at extreme degrees of distortions (ξ near 0.5): The dotted line and the dashed line are nearly indistinguishable in the left panel, and the reduction in consumption at $\xi = 0.2$ is again less than 20 percent.

The dashed line with triangles instead adds round-about production (intermediate input) to the basic model, but with no technology choice. Round-about production more than doubles the effect of distortions on consumption, which decreases by almost 40 percent as ξ goes to 0.2. However, the effect here is nearly linear in the semi-log scale with respect to ξ and are still significantly smaller than those in our benchmark.

The solid line with squares is a modified benchmark. It has the technology choice and round-about production, but the adoption costs are in units of labor only, i.e., $\gamma = 0$ instead of $\gamma = 1$. We see that the effect of distortions on consumption is smaller than, yet comparable to, that in the benchmark, except for intermediate values of ξ between 0.2 and 0.3. The same is true for the impact of distortions on the fraction of adopters in the right panel. At $\xi = 0.2$, consumption is about 50 percent lower than in the no-distortion case (compared to 60 percent in the benchmark), and about 20 percent of active firms adopt the modern technology (compared to nearly 0 in

the benchmark). The difference between the $\gamma = 0$ and the $\gamma = 1$ cases shows the quantitative relevance of the feedback effect of adoption on the price of the adoption goods, as discussed in Section 3.1.

Finally, the solid line with circles is the modified benchmark case where the two technologies have the same intermediate input intensities, $\nu_t = \nu_m = 0.69$, instead of $\nu_t \ll \nu_m$. (We reinstate $\gamma = 1$.) The effect of distortions on consumption and technology adoption is more measured than in the benchmark until ξ becomes large enough ($\xi > 0.35$). This highlights another feedback mechanism operating in the benchmark model: As more firms adopt, the lower is the price of the intermediate goods relative to labor, and therefore the higher the profitability of the modern technology that uses the intermediate input more intensively. Because this feedback mechanism is absent in this modified model, the negative effect of distortions on adoption and consumption is smaller than in the benchmark. At $\xi = 0.2$, consumption is 40 percent lower than in the no-distortion case, and nearly a quarter of active firms adopt the modern technology. On the other hand, when distortions are large enough that the fraction of adopters approaches 0 ($\xi > 0.35$), the negative effect on consumption is considerably larger than in the benchmark that has $\nu_t = 0$. This is because the dearth of adopters makes the intermediate input expensive, but the traditional technology in this modified model is still dependent on the intermediate input ($\nu_t = 0.69$), reducing its effective productivity.

Overall, the analysis in this section emphasizes the interaction among our model elements that is more than simply additive. Technology adoption, round-about production and the nature of the adoption cost (i.e., labor or goods) jointly explain the large, nonlinear effect of idiosyncratic distortions, even in the absence of multiplicity or coordination failures.

6.3.2 Income Gap between the US and India

Our benchmark model calibration generated a large income gap between the US and India, a factor of nearly 7, accounted for by the high adoption cost κ_a and the high degree of idiosyncratic distortions ξ in India (Table III). Since other models in the literature rarely generate such a large income gap, it is natural to ask which model elements of ours are responsible for this result.[22]

---

[22]Francesco Caselli suggested the exercise in this section, for which we are grateful.

		A_t	κ_a	$P_a\kappa_a$	GDP p.c.	Ratio
Case 1: $\nu = 0$, no adoption	US	1	0	0	1.89	$\frac{1}{0.86}$
	India	1	0	0	1.62	
Case 2: $\nu_t = \nu_m = 0$, $\gamma = 0$	US	0.54	2.17	2.17	1.64	$\frac{1}{0.51}$
	India	0.28	321	321	0.83	
Case 3: $\nu = 0.69$, no adoption	US	1	0	0	5.33	$\frac{1}{0.63}$
	India	1	0	0	3.37	
Case 4: Benchmark	US	0.43	15.9	9.35	4.34	$\frac{1}{0.15}$
	India	0.07	271	810	0.66	
Case 5: $\gamma = 0$	US	0.43	8.74	8.74	3.33	$\frac{1}{0.11}$
	India	0.06	1123	1123	0.34	
Case 6: $\nu_t = \nu_m = 0.69$	US	0.54	12.1	7.00	4.15	$\frac{1}{0.12}$
	India	0.29	272	1034	0.49	

Notes: The price of the adoption good is $P_a = P^\gamma w^{1-\gamma}$. With $\gamma = 0$, the price of the adoption good is $P_a = w = 1$, while with $\gamma = 1$ it is $P_a = P$.

Table IV: Alternative Model Results

In Section 6.3.1, we calibrated various comparison models to the US data with $\xi = 0$ and calculated the impact of higher degrees of distortions ξ. In this section, we calibrate each of the comparison models separately to the US and the Indian data, following the procedure in Section 5. We report the resulting GDP per capita of the two economies in Table IV, which also shows the calibrated productivity of the traditional technology A_t, the adoption cost parameter κ_a, and the adoption cost in units of labor $P_a\kappa_a$. One thing to note is that the degree of distortions ξ is assumed to be 0 for the US and is identified from the right tail of the establishment size distribution for India. Since the model elements that vary across the comparison models do not affect the right tail of the size distribution, across all these calibrations the ξ for India remains at 0.19.

We start with the basic model in the distortion literature that has no intermediate input nor technology adoption ($\nu = 0$, $A_t = A_m$ and $\kappa_a = 0$; case 1 in Table IV). For this model, the only parameters we can use to match the size distribution in either economy is the entry cost κ_e, the Pareto tail parameter ζ, and the distortion parameter ξ. The US right tail pins down ζ and the Indian right tail ξ, given the assumptions of a common ζ and $\xi = 0$ for the US. Unsurprisingly, this simple model has a hard time matching the size distribution of either economy. The result is that the GDP per capita of India is only 14 percent lower than that of the US, a magnitude

comparable to the corresponding model result in Figure 8 (dashed line, $\xi = 0.19$).

The second comparison model has a technology choice with labor-only adoption costs, but still has no intermediate input (case 2 in Table IV). To match the establishment size distribution, the adoption costs become vastly different between the two economies. However, the gap in the traditional technology productivity A_t becomes smaller than in the benchmark model (case 4 in the table). The resulting GDP per capita gap is now larger: India's is half that of the US. This gap is larger than what we see in Figure 8 (dotted line), but much smaller than what our benchmark model generates.

The third comparison model has intermediate input, but no technology adoption (case 3 in the table). Like the first comparison model, without the technology choice and the related parameters, this model cannot closely match the establishment size distributions either. However, the linkage in the form of the round-about production generates a meaningful GDP per capita gap between the two economies: India's is nearly 40 percent smaller than that of the US, which is nearly identical to the effect of ξ shown in Figure 8 (dashed line with triangles).

When the technology choice and the linkages through intermediate input uses are both incorporated, as in our benchmark, the model generates a much larger GDP per capita gap between the two countries. The exact size of the gap can vary depending on the nature of the adoption costs (goods vs. labor, case 4 vs. 5 in the table) and the intermediate input intensity of the two technologies (case 6), but what does not change is the insight that these model elements interact and generate a cross-country GDP per capita gap larger than the sum of their respective individual effects.

6.4   Big Push

"Big Push" is the name Rosenstein-Rodan (1961) gave to the idea that a minimum scale of investment is necessary for economic development. The rationales are indivisibilities in the production function, especially social overhead capital, and complementarities across sectors. Under these conditions, individual firms may not find it profitable to industrialize alone, even though all firms are better off industrializing together.

Murphy et al. (1989) note that government investments in infrastructure do not automatically solve the coordination problem. In fact, if unaccompanied by

firms' coordinated decision to industrialize and utilize the infrastructure, the modern infrastructure becomes a classic "white elephant." In this regard, the role of the government is to promote a coordinated, collective decision.

If all we need is the coordination of firms' decisions so that they all become better off, why do so many countries still remain unindustrialized and poor? Our framework helps address this question in two ways. First, the heterogeneity across firms implies that not all firms are better off in the good equilibrium. As discussed in our explanation of Figure 7, many firms that would be active (and make profits) in the bad equilibrium are inactive (and make no profit) in the good equilibrium. Although we have not specified preferences or welfare criteria, it is easy to see that the presence of losers, as well as winners, can validate the explanations that vested interests block the adoption of better technologies (Olson, 1982; Parente and Prescott, 1999).

Second, our framework introduces another dimension to the notion of Big Push, beyond the coordination over multiple equilibria. Reforms that reduce idiosyncratic distortions are necessary elements of successful development policies. As Section 6.2 showed, the effects of reducing distortions get amplified in our framework with or without multiplicity, which we call Big Push in distorted economies. In this view, the role of the government is to reduce distortions, identifying and exploiting the "Big Push region," where the returns to economic reforms are discontinuously high.

6.5   Big Push and Industrial Policy

The view of economists and policymakers on industrial policy has evolved over time. In the early years, government planning, public investment, and protectionist trade policies were the dominant development strategy, but the results were more often than not disappointing (Krueger, 1997). The mounting evidence of "government failures" ushered in the wave of market-fundamentalist liberalization of the 1990s, with equally, if not more, disappointing results (World Bank, 2005). Renewed thinking on industrial policy emphasizes governments' coordination of innovation and technology adoption (e.g., Rodrik, 2004), the very elements central to our framework.

Here we calculate the effect of such a policy: subsidies for technology adoption. In addition, we solve the problem of a constrained planner that chooses the optimal technology adoption taking as given the distortions and the set of active firms.

The industrial policy we implement subsidizes a fraction s of the cost of adopting

US   India
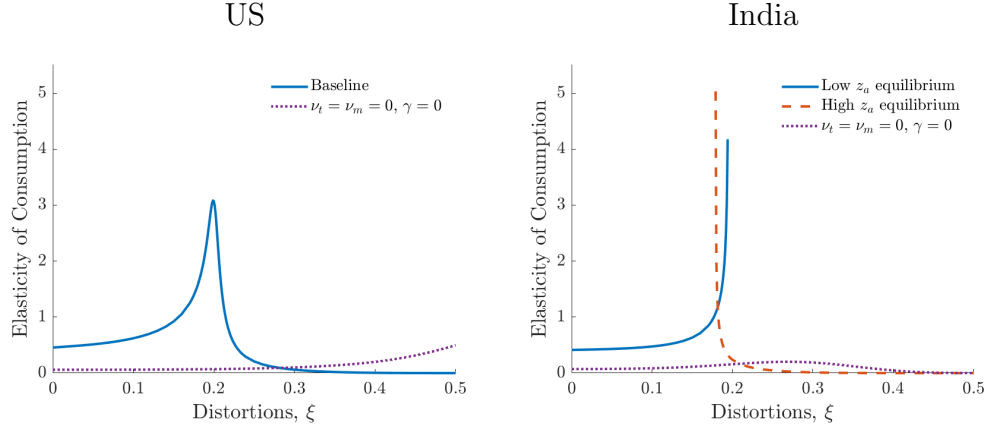
Notes: The elasticity of aggregate consumption to adoption subsidy for the US and India calibration, as the distortion parameter ξ goes from 0 to 0.5. For the Indian case, the solid line is the high adoption (low adoption threshold z_a) equilibrium and the dashed line is the low adoption (high z_a) equilibrium. The dotted lines are for the simple model without intermediate input in production and with labor only adoption costs.

Figure 9: Elasticity of Aggregate Consumption to Adoption Subsidy

the modern technology, financed by a lump-sum tax on consumers.

We show the elasticity of aggregate consumption to the subsidy for both the US and the India calibrations, as we vary the degree of distortions ξ. For comparison, we also compute the elasticity from a simpler model without the round-about production ($\nu_m = \nu_t = 0$) and with labor only adoption costs ($\gamma = 0$), which corresponds to the dotted lines in Figure 8 or case 2 in Table IV. Figure 9 presents two noteworthy results. First, in our benchmark model with the feedback effects, the elasticity of aggregate consumption to the subsidy is high in the Big Push region, especially in the multiplicity region for India, but relatively low outside of it. This can potentially explain why some industrial policies are more successful than others: It is a matter of whether the economy is in or outside the Big Push region. Second, the elasticity is uniformly low in the version with no feedback effect and there is no Big Push region, implying that complementarity and amplification are necessary conditions for successful industrial policy.

We now calculate the optimal adoption subsidy s, by solving the problem of a constrained planner, who chooses the adoption threshold z_a taking as given the distortions and the set of active firms.[23] The aggregate consumption from this

---

[23]We also worked out the case in which the planner chooses both the entry and the adoption thresholds. The entry margin is found to play a minor role, except at high degrees of distortions.

US   India
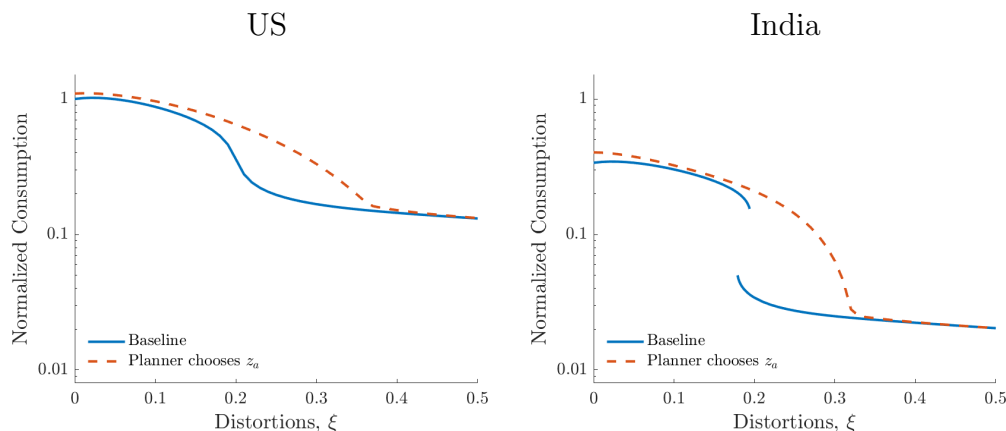
Notes: The aggregate consumption in the constrained planner allocation for the US and India (dashed line). Consumption is normalized by the equilibrium consumption in the undistorted US economy and is in log scale.

Figure 10: Constrained Planner Allocation

constrained planner allocation is shown with dashed lines in Figure 10, while the solid lines reproduce the equilibrium outcomes from Figure 6. Consistent with the elasticity results above, the largest gains from influencing firms' adoption decisions are in the Big Push region, where the dashed and the solid lines are farthest apart.

Furthermore, the India case attests to the power of coordination achieved by the policy: Although the adoption subsidy has small effects on consumption for ξ greater than 0.2 in Figure 9, the planner can generate massive consumption gains by coordinating the economy away from the low adoption equilibrium to the high adoption equilibrium, as long as the degree of distortions is not too high.

Nevertheless, the figure shows that the gains from reducing distortions tend to be larger than the gains from optimal adoption subsidies, and that adoption subsidies are not very effective at the low and the high end of the degree of distortions. In other words, not only do the idiosyncratic distortions determine the effectiveness of the industrial policy subsidizing technology adoption, but are also an important source of underdevelopment themselves.

7 Concluding Remarks

This paper provides answers to the following three questions: (i) Can economic development be explained solely by coordination failures? (ii) Can economic

development be a story of coordination failures and distortions? and (iii) Are there large non-linear effects of distortions and policies even without multiplicity, which can explain the huge income differences across countries? We find that the US calibration gives a unique equilibrium and that the calibrated Indian economy is in the multiplicity region but is in the good equilibrium, which says no to the first question. We find that small changes in idiosyncratic distortions can move the economy in and out of the region of multiplicity, resulting in discontinuously large changes in the aggregate output. More important, even without multiplicity, the feedback channels in our model creates Big Push regions, where small changes in distortions and policies have disproportionate effects. The answer to the second and the third questions is in the affirmative.

A promising avenue for future research is the exploration of an asymmetric input-output structure of production—for example, a multi-sector extension, in which sectors differ in adoption costs and forward/backward linkages. We conjecture that this extension will feature clusters of amplification and multiplicity. Another is a dynamic extension of the model, where only a subset of firms make entry and adoption decisions each period. In this extension, coordination failures may show up as multiple steady states and history dependence. Whether policies that subsidize adoption or reforms that reduce idiosyncratic distortions can move the economy from bad to good steady states is an open question.

References

Bento, P. and Restuccia, D. (2017). Misallocation, Establishment Size, and Productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303.

Bhattacharya, D., Guner, N., and Ventura, G. (2013). Distortions, Endogenous Managerial Skills and Productivity Differences. *Review of Economic Dynamics*, 16(1):11–25.

Blaum, J., Lelarge, C., and Peters, M. (2018). The Gains from Input Trade with Heterogeneous Importers. *American Economic Journal: Macroeconomics*, 10(4):77–127.

Buera, F., Fattal Jaef, R. N., Laski, R., and Trachter, N. (2020). The Missing Middle in a Broad Cross-Section of Indian Establishments. Unpublished manuscript.

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.

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review*, 101(1):304–340.

Caucutt, E. M. and Kumar, K. B. (2008). Africa: Is aid an answer? *B.E. Journal of Macroeconomics*, 8(1).

Chenery, H. B., Robinson, S., and Syrquin, M. (1986). *Industrialization and Growth: A Comparative Study*. Oxford University Press.

Ciccone, A. (2002). Input Chains and Industrialization. *Review of Economic Studies*, 69(3):565–587.

Cole, H. L., Greenwood, J., and Sanchez, J. M. (2016). Why Doesn't Technology Flow From Rich to Poor Countries? *Econometrica*, 84(4):1477–1521.

Cooper, R. and John, A. (1988). Coordinating Coordination Failures in Keynesian Models. *Quarterly Journal of Economics*, 103(3):441–463.

Crouzet, N., Gupta, A., and Mezzanotti, F. (2020). Shocks and Technology Adoption: Evidence from Electronic Payment Systems. Unpublished manuscript.

Davis, D. R. and Weinstein, D. E. (2002). Bones, Bombs, and Break Points: The Geography of Economic Activity. *American Economic Review*, 92(5):1269–1289.

Davis, D. R. and Weinstein, D. E. (2008). A Search for Multiple Equilibria in Urban Industrial Structure. *Journal of Regional Science*, 48(1):29–65.

Graham, B. S. and Temple, J. R. W. (2006). Rich Nations, Poor Nations: How Much Can Multiple Equilibria Explain? *Journal of Economic Growth*, 11(1):5–41.

Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale University Press.

Holmes, T. J. and Stevens, J. J. (2014). An Alternative Theory of the Plant Size Distribution, with Geography and Intra- and International Trade. *Journal of Political Economy*, 122(2):369–421.

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

Hopenhayn, H. A. (2014). Firms, Misallocation, and Aggregate Productivity: A Review. *Annual Review of Economics*, 6(1):735–770.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *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.

Jones, C. I. (2011). Intermediate Goods and Weak Links in the Theory of Economic Development. *American Economic Journal: Macroeconomics*, 3(2):1–28.

Jovanovic, B. (1989). Observable Implications of Models with Multiple Equilibria. *Econometrica*, 57(6):1431–1437.

Kaplan, G. and Menzio, G. (2016). Shopping Externalities and Self-Fulfilling Unemployment Fluctuations. *Journal of Political Economy*, 124(3):771–825.

Kim, M., Lee, M., and Shin, Y. (2020). The Plant-Level View of an Industrial Policy: The Korean Heavy Industry Drive of 1973. Unpublished manuscript.

Kline, P. and Moretti, E. (2014). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*, 129(1):275–331.

Krueger, A. O. (1997). Trade Policy and Economic Development: How We Learn. *American Economic Review*, 87(1):1–22.

Krugman, P. (1992). Toward a Counter-Counterrevolution in Development Theory. *World Bank Economic Review*, 6(S1):15–38.

Lane, N. (2019). Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea. Technical report, Monash University.

Matsuyama, K. (1991). Increasing Returns, Industrialization, and Indeterminacy of Equilibrium. *Quarterly Journal of Economics*, 106(2):617–650.

Matsuyama, K. (1995). Complementarities and Cumulative Processes in Models of Monopolistic Competition. *Journal of Economic Literature*, 33(2):701–729.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Murphy, K. M., Shleifer, A., and Vishny, R. W. (1989). Industrialization and the Big Push. *Journal of Political Economy*, 97(5):1003–1026.

Okuno-Fujiwara, M. (1988). Interdependence of Industries, Coordination Failure and Strategic Promotion of an Industry. *Journal of International Economics*, 25(1-2):25–43.

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

Owens, R. E., Rossi-Hansberg, E., and Sarte, P.-D. G. (2018). Rethinking Detroit. Opportunity and Inclusive Growth Institute Working Papers 11, Federal Reserve Bank of Minneapolis.

Parente, S. L. and Prescott, E. C. (1999). Monopoly Rights: A Barrier to Riches. *American Economic Review*, 89(5):1216–1233.

Redding, S. J., Sturm, D. M., and Wolf, N. (2011). History and Industry Location: Evidence from German Airports. *Review of Economics and Statistics*, 93(3):814–831.

Restuccia, D. and Rogerson, R. (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Plants. *Review of Economic Dynamics*, 11(4):707–720.

Restuccia, D. and Rogerson, R. (2017). The Causes and Costs of Misallocation. *Journal of Economic Perspectives*, 31(3):151–74.

Rodríguez-Clare, A. (1996). The Division of Labor and Economic Development. *Journal of Development Economics*, 49(1):3–32.

Rodrik, D. (1996). Coordination Failures and Government Policy: A Model with Applications to East Asia and Eastern Europe. *Journal of International Economics*, 40(1-2):1–22.

Rodrik, D. (2004). Industrial Policy for the Twenty-First Century. Discussion Papers 4767, CEPR.

Rosenstein-Rodan, P. N. (1943). Problems of Industrialisation of Eastern and South-Eastern Europe. *Economic Journal*, 53(210/211):202–211.

Rosenstein-Rodan, P. N. (1961). Notes on the Theory of the 'Big Push'. In Ellis, H. S., editor, *Economic Development for Latin America*, International Economic Association Series, pages 57–81. Palgrave Macmillan, London.

Schaal, E. and Taschereau-Dumouchel, M. (2019). Coordinating Business Cycles. Technical report, Centre de Recerca en Economia Internacional.

Tybout, J. R. (2000). Manufacturing Firms in Developing Countries: How Well Do They Do, and Why? *Journal of Economic Literature*, 38(1):11–44.

Valentinyi, A., Herrendorf, B., and Waldmann, R. (2000). Ruling Out Multiplicity and Indeterminacy: The Role of Heterogeneity. *Review of Economic Studies*, 67:295–307.

World Bank (2005). *Economic Growth in the 1990s: Learning from a Decade of Reform*. The World Bank.

Yeaple, S. R. (2005). A Simple Model of Firm Heterogeneity, International Trade, and Wages. *Journal of International Economics*, 65(1):1–20.

Online Appendix (Not for Publication)

A   Equilibrium Conditions

In Section 2, we note that the equilibrium conditions can be represented by three equations in three unknowns. They are:

$$\kappa_e = \frac{\left(\frac{P}{w}\right)^\gamma}{\frac{A_m^{\eta-1}}{A_t^{\eta-1}}\left(\frac{w}{P}\right)^{(\nu_m-\nu_t)(\eta-1)}-1}\kappa_a\left(\frac{z_e}{z_a}\right)^{\eta(1-\xi)-1},$$

$$\left(\frac{w}{P}\right)^{\eta-1} = \left(\frac{(\eta-1)\tau}{\eta}\right)^{\eta-1}\left[\left(\frac{w}{P}\right)^{\nu_t(\eta-1)}A_t^{\eta-1}\int_{z_e}^{z_a}z^{(\eta-1)(1-\xi)}dF(z)\right.$$

$$\left.+\left(\frac{w}{P}\right)^{\nu_m(\eta-1)}A_m^{\eta-1}\int_{z_a}^{\infty}z^{(\eta-1)(1-\xi)}dF(z)\right],$$

$$\kappa_e = \frac{1}{\eta-1}\frac{L-(1-F(z_e))\kappa_e-(1-F(z_a))\left(\frac{P}{w}\frac{1-\gamma}{\gamma}\right)^\gamma\kappa_a}{Z^{\eta(1-\xi)-1}}\left(\frac{P}{w}\right)^{\nu_t(1-\eta)}A_t^{\eta-1}z_e^{\eta(1-\xi)-1},$$

where

$$Z^{\eta(1-\xi)-1} \equiv (1-\nu_t)A_t^{\eta-1}\left(\frac{P}{w}\right)^{\nu_t(1-\eta)}\int_{z_e}^{z_a}z^{\eta(1-\xi)-1}dF(z)$$

$$+(1-\nu_m)A_m^{\eta-1}\left(\frac{P}{w}\right)^{\nu_m(1-\eta)}\int_{z_a}^{\infty}z^{\eta(1-\xi)-1}dF(z).$$

B   Sensitivity Analysis

Because the literature lacks precise estimates of the elasticity of substitution across intermediate goods η and the intermediate input elasticity for the traditional technology ν_t, we fixed $\eta = 3$ and $\nu_t = 0$ outside of the model. While $\eta = 3$ falls within the standard range in the literature, there is no available estimate of ν_t, beyond the fact that it is smaller than the modern technology elasticity ν_m (Chenery et al., 1986; Blaum et al., 2018; Kim et al., 2020). In this section we explore the role of η and ν_t. As we vary η or ν_t, we re-calibrate the model and re-do the exercises that produced Figure 6. We consider two values of η, 2.5 and 4, one on either side of the benchmark value $\eta = 3$. We also consider two values for ν_t: one where $\nu_t = 0.35$ is larger than the benchmark value of zero but smaller than ν_m, and the other where ν_t and ν_m are both 0.69, which is the highest value possible for ν_t given the restriction

$\nu_t \leq \nu_m$ and the overall intermediate input share. Figure 11 shows consumption per capita as we vary the idiosyncratic distortion parameter ξ. The two top panels are the cases with different η's for the US and India. The two bottom panels are for alternative ν_t's. In all cases, consumption is normalized by the US consumption in the equilibrium with no distortions.
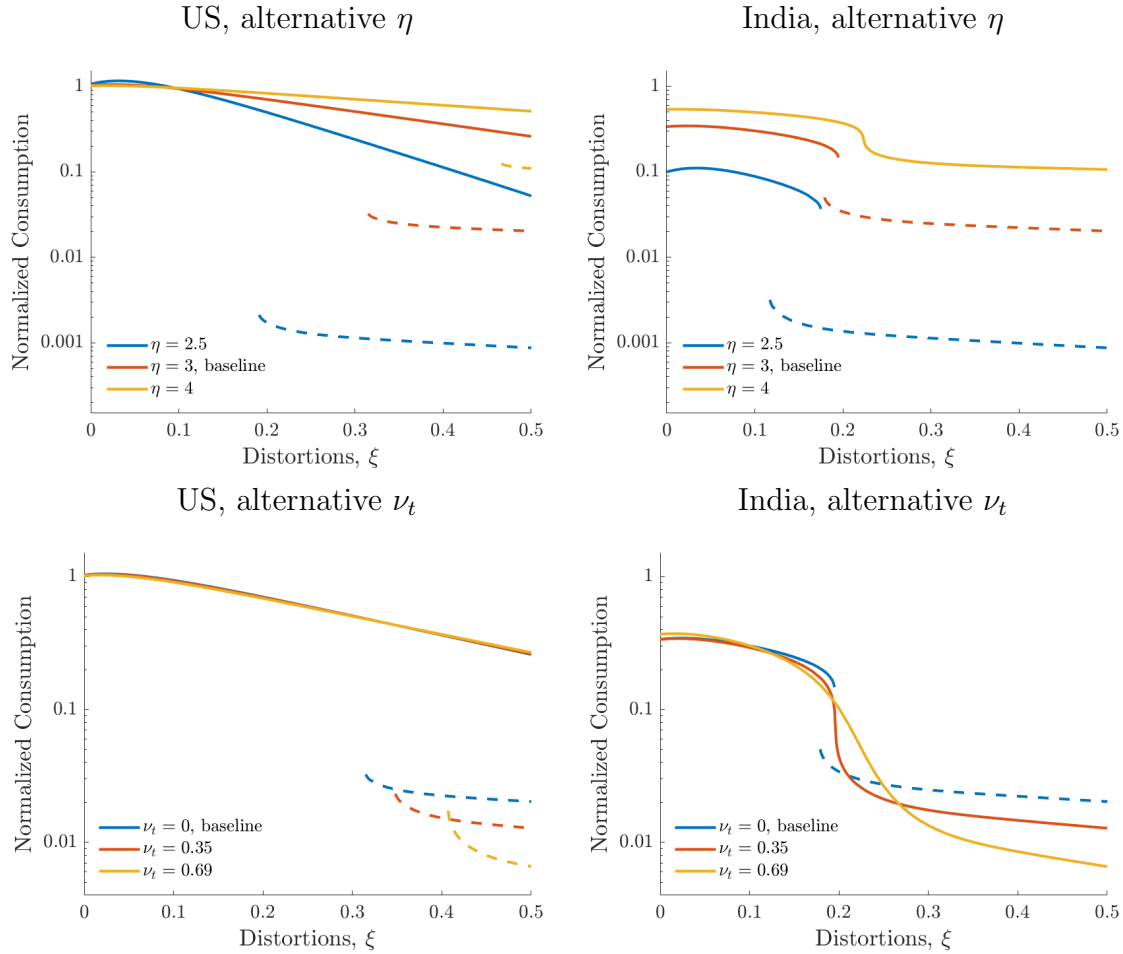
Notes: Equilibrium consumption per capita of the US and India as the distortion parameter ξ goes from 0 to 0.5. Consumption per capita is normalized by the US consumption (no distortion) and in log scale.

Figure 11: Sensitivity of Consumption and Equilibrium Multiplicity to η and ν_t

The top panels of Figure 11 show that a smaller elasticity of substitution η increases the consumption difference between good and bad equilibria and widens the set of the distortion parameter ξ that generates multiple equilibria. This is consistent with the analysis in Section 3.1: A lower η makes goods less substitutable and firms'

adoption decisions more complementary. In addition, the income gap between the US and India is larger with a smaller η. We draw the conclusion that a small elasticity of substitution across goods are conducive to explaining the income differences across countries, with or without multiplicity.

The bottom panels show that a lower intermediate input elasticity of the traditional technology ν_t has two effects. On the one hand, a lower ν_t enlarges the set of ξ's generating multiple equilibria. This is consistent with the discussion in Section 3.3. On the other hand, a low ν_t compresses the consumption gap between good and bad equilibria. Intuitively, when ν_t is small, it is less costly to use the traditional technology in a world where few firms adopt the modern technology and the intermediate aggregate is expensive. The two effects run in opposite directions when it comes to explaining cross-country income differences.

Another robust result in the plots for India is that the model features a very non-linear effect of distortions, either through or independently of equilibrium multiplicity.

C   Idiosyncratic Distortion

We show the magnitude of idiosyncratic distortions in the Indian calibration with $\xi = 0.19$. We model distortions as effective taxes and subsidies on firms' revenue, and the tax rate is $1 - \tau z^{-\xi}$. The solid line is the cumulative distribution of tax rates across firms in the Indian calibration, while the dashed line is the same object, but for the re-computed US economy using India's ξ. For a given tax rate on the vertical axis, we show the fraction of firms with higher tax rates on the horizontal axis in log scale. In both cases, a small fraction of firms pays taxes and subsidizes the rest of the economy. In India, the most productive firms (those at the top 10^{-5} percentile of the active firm productivity distribution) face tax rates of about 45 percent, and nearly half the firms receive subsidies of at least 60 percent of their revenue. For the US case with India's ξ, the whole tax schedule shifts up for two reasons. First, the active firms in the US are more productive, because the entry threshold z_e is higher than in India. As a result, holding constant the $\bar{\tau}$ of India, the US firms face a higher tax schedule. Second, in the US equilibrium with India's distortions, $\bar{\tau}$ is higher. The top US firms face higher taxes than top Indian firms—a tax rate of almost 60 percent—and half the firms receive subsidies of at least 30 percent of their revenue. Overall, we conclude that the magnitude of taxes to large firms and subsidies to small
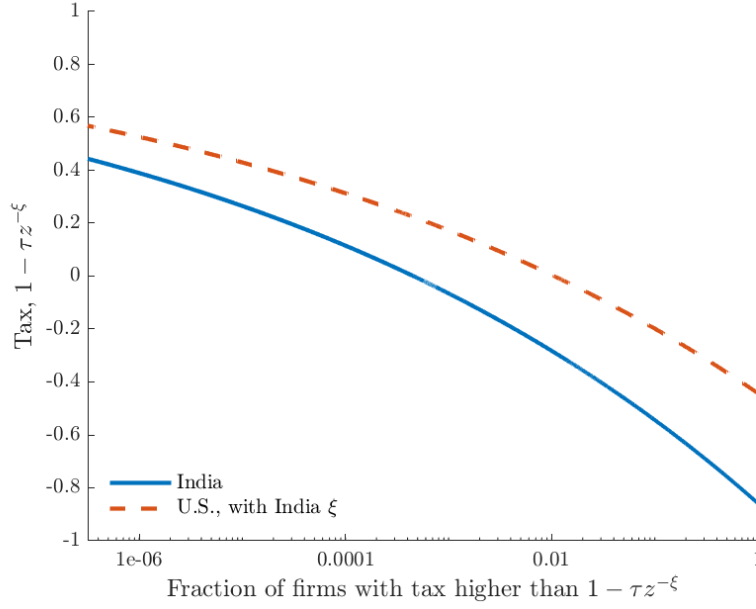
Note: The figure shows the cumulative distribution of taxes across firms. The solid line is for the Indian calibration. The dashed line corresponds to the re-computed US economy using India's ξ.

Figure 12: Idiosyncratic Distortions across Firms

firms is not implausible, relative to the numbers in the misallocation literature.