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

April 5, 2023

Abstract

Why don't poor countries adopt more productive technologies? Is this all because of distortions? Is there a role for policies that coordinate technology adoption? To answer these questions, we develop a quantitative model with heterogeneous firms and idiosyncratic distortions, in which the gains from technology adoption are larger when more firms adopt. When this complementarity is strong enough, multiple equilibria and hence coordination failures are possible. More important, even without equilibrium multiplicity, the model elements responsible for the complementarity amplify the effect of distortions substantially. In what we call the Big Push region, the impact of distortions is four times as large as in models without such complementarity. The Big Push region exists for a broad, empirically-relevant range of parameter values, whereas multiplicity is confined to a smaller segment of the parameter space.

*Buera: fjbuera@wustl.edu. Hopenhayn: hopen@econ.ucla.edu. Shin: yshin@wustl.edu. Trachter: trachter@gmail.com. We thank Andy Atkeson, Ariel Burstein, Chad Jones, Ezra Oberfield, Michael Peters, and participants at many seminars and conferences for comments and suggestions. We thank Samira Gholami, Eric LaRose, Reiko Laski and James Lee for outstanding research assistance. The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

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 and hence economic development? 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 to deficient adoption and low productivity. However, most quantitative studies show that distortions can generate only modest development gaps, a small fraction of the actual gap between poor and rich countries. An alternative view emphasizes the role of complementarity and coordination failures: No individual firm in poor countries wants to adopt unilaterally, even though the gains from adoption increase with the number of adopters so that it is profitable for all firms to adopt together. According to this view, the role of policies is to coordinate 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 the equilibrium is unique, 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. Indeed, in what we call the Big Push region, the impact of idiosyncratic distortions on welfare is four times as large as in models without such complementarity, even though coordination failures do not play any 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. A rise in the mass of adopters has opposing effects on the gains from adoption for the marginal firm. First, the aggregate price falls, and the competition effect implies lower prices, profits, and gains from adoption for the marginal firm. On the other hand, the lower aggregate price comes with a larger aggregate output and hence a larger demand for the marginal firm's output. The lower aggregate price also reduces the cost of the intermediate input for production and the goods portion of the adoption costs. These latter forces increase the net gains from adoption for the marginal firm. If the positive effects dominate the negative competition effect, the gains from adoption is increasing in the mass of adopters—that is, we have complementarity and amplification.

The key determinants of the relative strengths of these countervailing effects are equilibrium elasticities that capture the feedback effect of how the adoption by marginal firms affects the aggregate output and price. These elasticities in turn depend crucially on the joint distribution of productivity and idiosyncratic distortions across firms.

More heterogeneity in firm productivity implies that the density of firms near the adoption margin is smaller and that the productivity gap between the marginal and the average adopters is larger. As a result, the feedback effect from the marginal firms' adoption on the aggregate economy is small. This partly explains why multiplicity results in the literature typically come from models with homogeneous firms.[1] However, firm heterogeneity is essential for our analysis. It allows us to decouple the notion of amplification from multiplicity, so that we can analyze the role of complementarity in the absence of multiplicity. In addition, it is at the core of our quantitative strategy.

Next, idiosyncratic distortions that are correlated with firm productivity (i.e., more productive firms subjected to higher "taxes") make taxed high-productivity firms and subsidized low-productivity firms behave more similarly. In other words, such distortions compress the effective heterogeneity across firms, strengthening the

[1] Intuitively, firms that are much more productive than others will adopt the modern technology and those that are much less productive than others will not adopt, *regardless of* what other firms do. This makes multiplicity less likely.

feedback effect. At the same time, distortions reduce the mass of adopters and hence their importance in the aggregate economy, which weakens the feedback effect. As a result, complementarity and amplification are non-monotonic in the degree of distortions.

For a simpler version of our model, we derive the conditions for amplification and equilibrium multiplicity. Consistent with the above discussion, what contribute to amplification are: a small elasticity of substitution across differentiated products, a high intermediate input intensity of the modern technology, adoption costs in units of goods rather than labor, small heterogeneity in firm productivity, and an intermediate degree of idiosyncratic distortions correlated with firm productivity. Multiplicity is an extreme form of amplification, and indeed the condition for the existence of multiple equilibria is stronger than that for amplification.

For our quantitative analysis, we use aggregate and micro-level data from the US and India. 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 can be obtained in closed form from the establishment size distribution.

Our main quantitative result is on the amplification of the impact of distortions. Starting from the undistorted US economy, the effect of idiosyncratic distortions on aggregate productivity is large and highly non-linear, even though equilibrium multiplicity is not present. In particular, in an empirically plausible range of distortions, a small change in distortions can have a disproportionate impact on technology adoption and aggregate productivity. We call this range the Big Push region. When the degree of distortions reaches the level estimated from the Indian data, which happens to be in the middle of the Big Push region, the aggregate consumption is 60 percent lower than that in the undistorted US economy. This effect of idiosyncratic distortions is several times larger than what one finds in standard models without complementarity (e.g., Hsieh and Klenow, 2009; Hopenhayn, 2014; Restuccia and Rogerson, 2017). We provide a thorough decomposition of the relative contribution of various model elements to this large effect. We also show that the Big Push region exists for a broad, empirically-relevant range of technology parameter values.

The next quantitative result is on the income gap between the US and India. Whereas we only vary the degree of idiosyncratic distortions for the main result above, here we allow technology parameters to be different between the two countries as well. Although our calibration targets the empirical moments from the establishment size distribution but not the income level of either country, the model generates a huge income gap between the US and India, nearly a factor of seven, due to the higher adoption costs and degrees of idiosyncratic distortions in India, whose impact is amplified by the complementarity. Again, multiple equilibria and coordination failures do not play any role.

Our final quantitative result is on the empirical relevance of multiple equilibria and coordination failures. Our calibrated US and Indian economies are outside the region of multiplicity for any degree of idiosyncratic distortions, but India is close enough to it that we do not dismiss multiplicity as empirically irrelevant. However, when we inspect the economy in the multiplicity region that is closest to the Indian economy, we find that the overall effect of idiosyncratic distortions is comparable in magnitude to the one in our Indian economy, where equilibrium is unique. Multiplicity is an extreme form of amplification, but not a necessary condition for the impact of distortions to be large. We conclude that the focus should be on how big amplification is rather than on whether or not multiple equilibria exist.

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 disproportionate 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

[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.

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, 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.

[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.

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 Bhattacharya et al. (2013) and Bento and Restuccia (2017). 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 $L \equiv 1$ 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 substitutable differentiated goods are 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 distribution $F(z)$ for $z \geq 1$, with density $f(z) = \partial F(z)/\partial z$. We assume that F is Pareto, with tail parameter ζ. 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

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

$i \in \{t, m\}$, 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, where $\gamma \in [0, 1]$ denotes the intermediate aggregate share. 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 4.

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

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

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 x required for production. The problem is:

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

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

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

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

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

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

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

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

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

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

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 Appendix A.

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 determines the amplification of the impact of distortions and policies and also the possibility of multiple equilibria and coordination failures.[6]

For our exposition, we consider a policy that subsidizes technology adoption. For simplicity, we assume that the adoption good is produced only with the intermediate aggregate, i.e., $\gamma = 1$. This implies that the price of the intermediate aggregate and the price of the adoption good coincide, i.e. $P_a = P$.

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

$$D(z; a) \equiv \pi_m^o(z; a) - \pi_t^o(z; a) - P(a)\kappa_a \ , \tag{10}$$

where $\pi_i^o(z; a)$ is the operating profit of a firm with productivity z using technology $i = t, m$. We made explicit the dependence of π_i^o and P on the mass of adopters a.

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. 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

$$D(z_a; a) = 0 \ . \tag{11}$$

We now consider the effect of subsidizing the cost of adopting the modern technology. We define the *direct* effect as the impact of the adoption cost on z_a holding fixed the mass of adopters a at the initial equilibrium level,

$$\left[\frac{dz_a}{d\kappa_a}\right]^{direct} = -\frac{D_{\kappa_a}(z_a; a)}{D_z(z_a; a)}. \tag{12}$$

The *total* effect includes the equilibrium response of the mass of adopters a:

$$\left[\frac{dz_a}{d\kappa_a}\right]^{total} = -\frac{D_{\kappa_a}(z_a; a)}{D_z(z_a; a) - D_a(z_a; a) f(z_a)} \ .$$

The ratio of the total to the direct effect is

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

which can be interpreted as a multiplier measuring the feedback or equilibrium effect.

[6] The discussion in this Section is in part related to the analysis of Cooper and John (1988).

When this feedback ratio is greater than one, the feedback effect is amplifying the direct effect, an indication of complementarity in adoption decisions. We call the term in the denominator of equation (13) the amplification rate:

$$r\left(z_a, a\right) \equiv D_a\left(z_a; a\right) \frac{f\left(z_a\right)}{D_z\left(z_a; a\right)} \ . \tag{14}$$

Amplification or complementarity (i.e., a feedback ratio greater than one) requires a positive amplification rate.

The feedback ratio in equation (13) plays an important role in the analysis of equilibrium determination and the possibility of multiplicity. An equilibrium for a given adoption cost κ_a consists of a threshold z_a and $a = 1 - F\left(z_a\right)$ satisfying equation (11). It is useful to define the following mapping:

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

which is a best response mapping that gives the fraction of adopters $T(a)$ when firms behave optimally in response to the equilibrium prices that prevail with a given mass of adopters a. An equilibrium is given by a fixed point $a = T\left(a\right)$. The slope of this mapping is:

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

with z_a satisfying $a = 1 - F(z_a)$, which is precisely the amplification rate $r(z_a, a)$. The equilibrium is unique if $T'\left(a\right) < 1$ for all a. Multiple equilibria are possible if $T'\left(a\right) > 1$ is followed by $T'\left(a\right) < 1$ over some interval of a, as shown in the left panel of Figure 1. In the right panel, we zoom in on the right-most equilibrium of the left panel. We show the effect of a change in the subsidy that reduces the adoption cost, breaking it into the direct and the total effects. Note that the condition for multiplicity, $r(z_a, a) > 1$, is stronger than the condition for amplification, $r(z_a, a) > 0$.

3.1 The Amplification Rate

We can rewrite the amplification rate in equation (14) as:

$$r\left(z_a, a\right) \equiv \frac{D_a\left(z_a; a\right) a}{\Delta\pi\left(z_a; a\right)} \left[\frac{D_z\left(z_a; a\right) z_a}{\Delta\pi\left(z_a; a\right)}\right]^{-1} \frac{f\left(z_a\right) z_a}{a} \ . \tag{15}$$

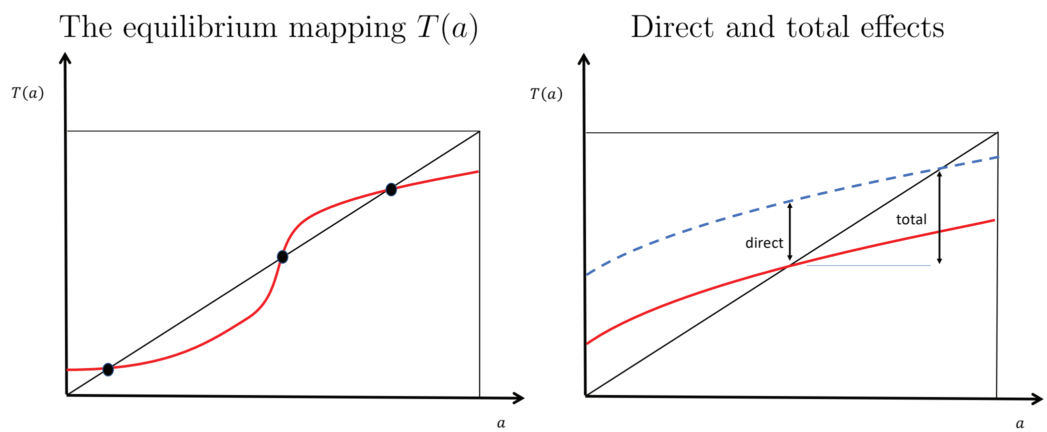


Figure 1: Equilibrium Mapping and Amplification: An Example

We will study $r(z_a, a)$ analytically for the case where $\nu_t = \nu_m \equiv \nu$ (the same intermediate input intensity for the modern and the traditional technology) and there is a fixed set of firms. However, our quantitative analysis confirms all predictions for the general case with $\nu_t \neq \nu_m$ and endogenous entry.

The expression for the amplification rate has three distinct terms. The first term, $\frac{D_a(z_a;a)a}{\Delta\pi(z_a;a)}$, which we refer to as the incentive elasticity, controls the rate at which changes in the aggregate adoption rate a increase the net gains from adoption for the marginal adopter. The second term is the reciprocal of the elasticity of the net gains from adoption to changes in the productivity of the marginal adopter, $\frac{D_z(z_a;a)z_a}{\Delta\pi(z_a;a)}$. With the assumptions in Section 2, $\frac{D_z(z_a;a)z_a}{\Delta\pi(z_a;a)} = \eta(1-\xi) - 1$. The demand elasticity η controls the curvature of the profit function. A more elastic demand implies that profits increase at a faster rate with the productivity of the firm. The distortion parameter ξ by construction dampens the response of profits to the productivity of the firm. The third term, $\frac{f(z_a)z_a}{a} = \zeta$ for the Pareto distribution, accounts for the elasticity of the mass of adopters a with respect to changes in the productivity of the marginal adopter z_a. Note that this third term, and hence the amplification rate, is larger when firms are less heterogeneous (a higher ζ).

We can thus simplify the amplification rate to:

$$r(z_a, a) \equiv \overbrace{\frac{D_a(z_a;a)\,a}{\Delta\pi(z_a;a)}}^{\text{Incentive elasticity}} \frac{\zeta}{\eta(1-\xi)-1}, \tag{16}$$

where the incentive elasticity can be written as

$$\frac{D_a\left(z_a;a\right)a}{\Delta\pi\left(z_a;a\right)}=\varepsilon_{X_a}-\left[\eta-\nu\left(\eta-1\right)-1\right]\varepsilon_{P_a}\ , \tag{17}$$

with $\varepsilon_{X_a}\equiv\frac{\partial X}{\partial a}\frac{a}{x}$ denoting the elasticity of the intermediate aggregate X to the mass of adopters a, and $\varepsilon_{P_a}\equiv-\frac{\partial P}{\partial a}\frac{a}{P}$ the elasticity of the aggregate price level P to a. ε_{X_a} has a one-to-one effect on the incentive elasticity, because profits are proportional to the intermediate aggregate X. The effect of ε_{P_a} on the incentive elasticity is mediated by the parameters η and ν through three channels. First, more adoption lowers the price of the intermediate aggregate and hence the price of adopters' output through competition, depressing the gains from adoption. A higher demand elasticity makes this channel, $-\eta\varepsilon_{P_a}$, stronger. Second, a lower price of the intermediate aggregate raises firms' profits because it means a lower input cost. The strength of this channel, $\nu\left(\eta-1\right)\varepsilon_{P_a}$, is pinned down by the intermediate input share in production ν and the elasticity of profits to marginal cost $\eta-1$. Third, a lower intermediate aggregate price means a lower adoption cost. With the assumption that the adoption good is produced only with the intermediate aggregate, this channel has a one-to-one effect, ε_{P_a}, on the incentive elasticity.

Equation (17) shows that ε_{P_a} weakens the incentive elasticity and hence the amplification rate. In a partial equilibrium model or one with limit pricing as in Murphy et al. (1989), one has $\varepsilon_{P_a}=0$ and the amplification rate will be larger, holding other things equal.

We can express the elasticities ε_{P_a} and ε_{X_a}, which are equilibrium objects, as:

$$\varepsilon_{P_a}=\frac{1}{1-\nu}\frac{1}{\eta-1}\left(1-\frac{\eta-1}{\zeta}+\left(\eta-1\right)\frac{\xi}{\zeta}\right)\mathcal{M}^v(a)\left[1-\left(\frac{A_t}{A_m}\right)^{\eta-1}\right]\ , \tag{18}$$

$$\varepsilon_{X_a}=\varepsilon_{P_a}\left[1+\left(\eta-1\right)\left(1-\nu\right)\right]-\left(1-\frac{\eta-1}{\zeta}+\eta\frac{\xi}{\zeta}\right)\mathcal{M}^l(a)\left[1-\left(\frac{A_t}{A_m}\right)^{\eta-1}\right]\ , \tag{19}$$

where $\mathcal{M}^v(a)$ and $\mathcal{M}^l(a)$ denote the share of value-added and employment, respectively, belonging to modern firms for a given mass of adopters a. These shares satisfy:

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

That is, in a distorted economy ($\xi > 0$), the modern firms' share of value-added is larger than their share of employment. This follows from the fact that firms' prices, output, and labor demands satisfy $p \propto z^{-(1-\xi)}$, $y \propto z^{\eta(1-\xi)}$, and $l \propto z^{\eta(1-\xi)-1} = z^{(\eta-1)(1-\xi)-\xi}$. The key difference is that, while firms' effective (or after-tax) profitability is distorted, their true productivity is not.

Next we explain how the parameters of the model pin down the equilibrium elasticities, ε_{P_a} and ε_{X_a}. We first explore the case with no distortions, and then discus how distortions affect the results.

No distortions, $\xi = 0$. In this case we have $\mathcal{M}^l(a) = \mathcal{M}^v(a) \equiv \mathcal{M}(a)$, and the two equilibrium elasticities coincide:

$$\varepsilon_{X_a} = \varepsilon_{P_a} = \frac{1}{1-\nu}\left(\frac{1}{\eta-1} - \frac{1}{\zeta}\right)\mathcal{M}(a)\left[1 - \left(\frac{A_t}{A_m}\right)^{\eta-1}\right].$$

The intermediate input share ν affects the elasticities through the standard input-output multiplier $(1-\nu)^{-1}$: A higher ν implies stronger linkages across firms. The Pareto tail parameter ζ also has a positive effect on these elasticities. A higher ζ implies a more compressed productivity distribution, so it dampens the negative selection effect that occurs when more firms adopt. The elasticity of substitution η has two opposing effects on the elasticities. First, a higher η implies that more of the output is already produced by highly productive firms, so the contribution from marginal adopters would be smaller. Second, a higher η amplifies the differences between the two technologies. Finally, these equilibrium elasticities are increasing in the relative size and the productivity advantage of the modern sector, $\mathcal{M}(a)$ and A_m/A_t, respectively.

Distortions, $\xi > 0$. Distortions affect the elasticities ε_{P_a} and ε_{X_a}, and hence the incentive elasticity, through both a direct and an indirect channel. The derivative of ε_{P_a} with respect to the degree of distortions ξ is as follows.

$$\frac{d\varepsilon_{P_a}}{d\xi} = \frac{1-\left(\frac{A_t}{A_m}\right)^{\eta-1}}{(1-\eta)(1-\nu)}\left[\overbrace{\frac{\eta-1}{\zeta}\mathcal{M}^v(a)}^{\text{direct channel}} + \overbrace{\frac{\zeta-(\eta-1)(1-\xi)}{\zeta}\frac{d\mathcal{M}^v(a)}{d\xi}}^{\text{indirect channel}}\right]$$

The direct channel captures how distortions partly offset the negative selection that occurs when more firms adopt. As explained above, a higher degree of distortions ξ implies that the prices and output of the marginal adopters are less sensitive to their productivity, which is lower than the productivity of other adopters. As a result, the aggregate price is more sensitive to the mass of adopters.

The indirect channel captures how distortions affect ε_{P_a} through its effect on $\mathcal{M}^v(a)$. As long as more distortions imply less adoption, $\mathcal{M}^v(a)$ is decreasing in ξ,[7] and the indirect channel has a negative sign.

In summary, the direct and the indirect channels affect the elasticity ε_{P_a} in opposite directions. We obtain a similar result for the intermediate aggregate elasticity with respect to the mass of adopters ε_{X_a}.[8]

Since the amplification rate $r(z_a, a)$ is a linear combination of these two equilibrium elasticities—see equation (17), it may not change monotonically with respect to the degree of distortions ξ. If one of the channels were to dominate the other, the largest amplification should occur at either the lowest or the highest value of ξ. If the relative strengths of the channels vary with ξ, the amplification rate may attain its maximum at an intermediate degree of distortions, which turns out to be the case in our quantitative exploration.

3.2 Amplification and Multiplicity

We further analyze the connection between amplification and multiplicity. To keep things simple, we maintain the assumption that $\nu_t = \nu_m = \nu$, and we only consider the case with no distortions. With $\xi = 0$, $\mathcal{M}^l(a) = \mathcal{M}^v(a) = \mathcal{M}(a)$ and hence $\varepsilon_{X_a} = \varepsilon_{P_a}$. Combining equations (16) and (17), the amplification rate is

$$r(z_a, a) = \frac{1}{1-\nu}\left(\frac{2-\eta}{\eta-1}+\nu\right)\left(\frac{\zeta-(\eta-1)}{\eta-1}\right)\mathcal{M}(a)\left[1-\left(\frac{A_t}{A_m}\right)^{\eta-1}\right] . \tag{20}$$

For the effect of policies to be amplified, the amplification rate must be positive. In this case, the amplification rate is positive if and only if

$$\eta < \frac{2-\nu}{1-\nu} , \tag{21}$$

[7]A sufficient, but not necessary, condition is $da/d\xi < 0$.

[8]The added complexity is that ε_{X_a} also depends on the wedge between $\mathcal{M}^v(a)$ and $\mathcal{M}^l(a)$. This wedge is large for intermediate degrees of distortions. A larger wedge between $\mathcal{M}^v(a)$ and $\mathcal{M}^l(a)$ implies a larger ε_{X_a} and hence a larger incentive elasticity.

since the first term and the last three terms in the right-hand side of equation (20) are strictly positive.[9] When η is small, the negative effect of competition on the gains from adoption is small, and the positive effect of a larger mass of adopters on the gains from adoption dominates, generating a positive amplification and hence a multiplier effect.

As discussed above, multiplicity requires that the amplification rate be greater than one. Because the last two terms in the right-hand side of equation (20) are strictly less than one, a necessary condition for multiplicity is

$$\eta < \frac{2 + \frac{1-\nu}{1+\zeta} - \nu}{1 + \frac{1-\nu}{1+\zeta} - \nu} \; . \tag{22}$$

The necessary condition for multiplicity is stricter than the condition for amplification, implying that amplification is possible even in the absence of multiple equilibria. When the heterogeneity across firms vanishes ($\zeta \to \infty$), the necessary condition for multiplicity becomes weaker to coincide with the one for amplification. This partly explains why most multiplicity results in the literature come from models with homogeneous producers and, more important, why the literature cannot separate the notion of amplification from equilibrium multiplicity.

3.3 The Role of Intermediate Input Intensity

To simplify our theoretical analysis of amplification and multiplicity in this section, we assumed that the intermediate input intensity of the modern and the traditional technologies is the same, $\nu_m = \nu_t$. We also assumed that the adoption good is produced only with the intermediate aggregate, $\gamma = 1$. We relax these assumptions in our quantitative exercises.

When the modern technology uses intermediate input more intensively, $\nu_m > \nu_t$, as in our quantitative model, there is an additional positive effect of the mass of adopters on firms' gains from adopting the modern technology. More adoption implies a lower price of the intermediate input relative to labor, raising the relative profitability of the technology that uses the intermediate input more intensively.

The same logic applies to the intensity of the intermediate aggregate in the production of the adoption good, γ. More adoption reduces the adoption costs

[9] For aggregates to be finite, $\zeta - (\eta - 1)$ must be positive.

by more when a larger fraction of the costs is in units of goods rather than labor. However, amplification and multiplicity can arise even when the production of the adoption good requires only labor ($\gamma = 0$), provided $\nu_m > \nu_t$, a result related to the one in Ciccone (2002).

In summary, both the intermediate input intensity difference between the two technologies and the high intermediate input intensity in the production of the adoption good contribute to amplification and multiplicity, but neither is necessary for such outcomes.

4 Calibration

Section 3 shows that several features of the model may 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. We propose a calibration strategy that provides point identification of model parameters by comparing the size-distribution of establishments implied by the model with that in the data. Our key assumption for this procedure to work is that traditional and modern establishments coexist in the data. However, we do 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. We provide a constructive proof of our approach in Appendix B.

For our benchmark model, 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 (for the undistorted US economy).

Six of these 11 parameters are fixed outside of the model: A_m, η, γ, ν_t, ξ, and τ. We maintain the normalization of $A_m = 1$. We set $\eta = 3$, which is considered to be on the lower side if considering the elasticity of substitution within a narrowly defined industry, as discussed in Hsieh and Klenow (2009). However, in our model, η is supposed to be the elasticity of substitution across all differentiated goods, rather than within an industry, so it is not an unreasonable value. As in Section 3, 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 5.1, 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 D reports a sensitivity analysis with different values of η and ν_t.

In addition, we think of the US as an economy that is distortion-free ($\xi = 0$), which means the budget-balancing scale parameter is $\tau = 1$.

The remaining five parameters are chosen to match relevant empirical moments. The intermediate input elasticity of the modern technology, ν_m, is calibrated to match the intermediate input share in the US data, yielding $\nu_m = 0.70$.[10]

The other four are pinned down by the size distribution of establishments, with the number of employees as the size measure, for which our model has a distinct implication, as explained in detail in Appendix B.

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

The other three parameters, A_t, κ_e, and κ_a, are chosen to match the features of the empirical establishment size distribution (the log-log relationship), given the distinct non-linearities of the log-log relationship that our model implies, especially the size gap between the traditional technology establishments and the modern technology establishments.

Figure 2 is the log-log plot of the establishment size distribution from the calibrated model (solid line) and its empirical counterpart from the 2007 BDS (circles). The calibrated model (solid line) places the flat region that is the size gap between the largest traditional-technology establishment and the smallest

[10] The intermediate input share in the US in 2007 was 0.46, calculated from the BEA input-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 (24), we obtain $\nu_m = 0.70$.

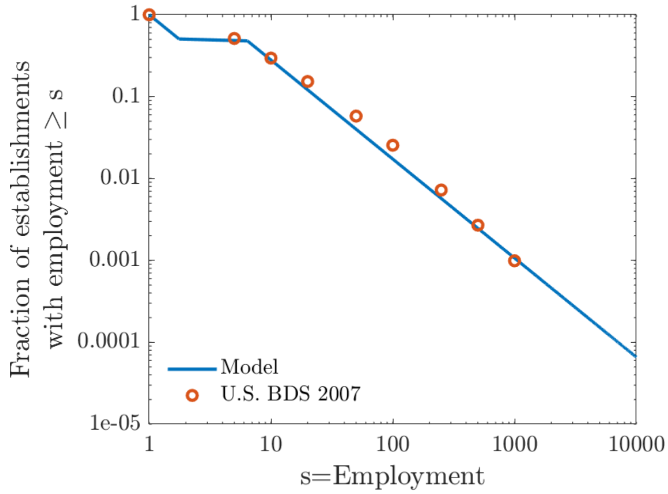


Figure 2: Establishment Size Distribution for the US: Model and Data

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

Table I: Calibrated Parameters

modern-technology establishment in a way that matches 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.

Table I summarizes the calibrated parameter values and how they were chosen.

5 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, accounting for the US-India income gap, and generating multiple equilibria and coordination failures.

5.1 Amplified Impact of Idiosyncratic Distortions

The goal of this section is to quantify the amplified effect of idiosyncratic distortions in our model and to explore the contribution of each of our model elements to this effect.

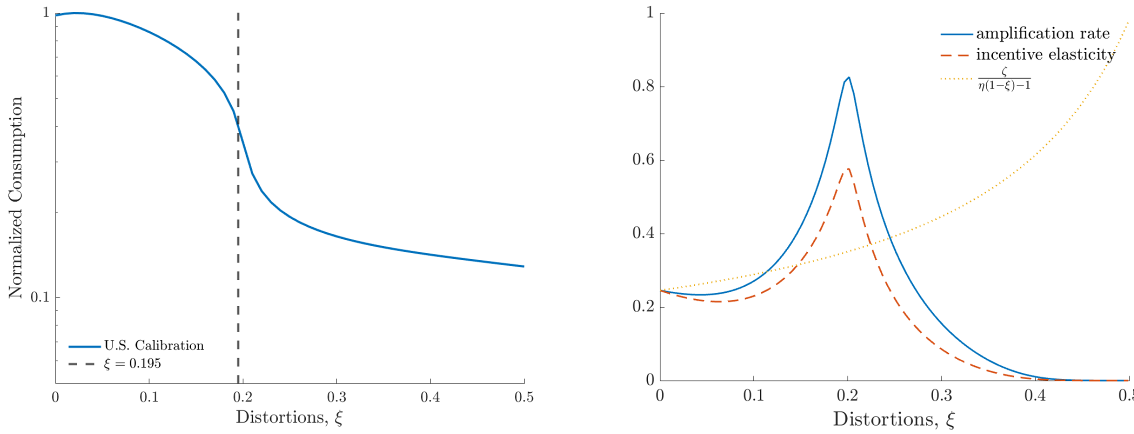
Starting with the distortion-free US economy ($\xi = 0$), we increase the degree of idiosyncratic distortions that are correlated with establishment-level productivity z. When we increase ξ, we hold fixed the other parameters at their respective calibrated values, except that τ adjusts to balance the budget. The solid line in the left panel of Figure 3 plots per-capita consumption for each ξ, normalized by its level in the US economy. (The vertical axis is in log scale.) There is only one solid line because, with our calibration, multiple equilibria do not exist for the range of ξ we consider.

We see that idiosyncratic distortions have a large, nonlinear effect on consumption, with an especially pronounced effect for intermediate values of ξ. At $\xi = 0.2$, per-capita consumption is only 40 percent of the US consumption. Larger values of ξ can reduce consumption by nearly 90 percent from the undistorted US benchmark. As we show below, these magnitudes are unattainable in conventional models.

If we assume that differences between the US and India are driven only by distortions, we can back-out the degree of distortions in India by comparing the tail of the size distribution of establishments in India (distorted) with that in the US (undistorted). Following this procedure yields $\xi = 0.195$ for India. Thus, we focus on the degree of distortions around 0.2. More details on this, and on a more general calibration exercise for India, are in Section 5.2. Finally, we note that the establishment-level revenue taxes and subsidies implied by $\xi = 0.2$ are not extreme. As we show in Appendix C, the establishment at the top 10^{-4} percentile of the active establishment productivity distribution is taxed at about 40 percent, while the maximum subsidy is 45 percent for the least productive establishment (the marginal entrant with $z = z_e$).

The solid line in the right panel of Figure 3 plots the amplification rate defined in equation (16).[11] The positive values show that there is amplification, consistent with the large effect of idiosyncratic distortions in the left panel. Furthermore, the

[11] As explained in Section 3, the amplification rate determines how large the total effect of reducing the adoption cost—perhaps through a subsidy—is relative to its direct effect.



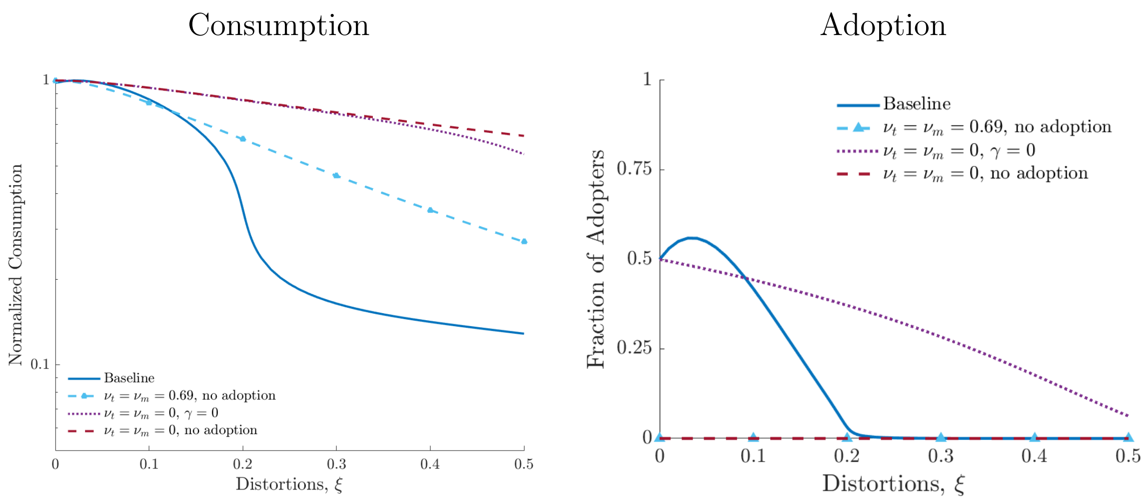
Note: The left panel shows the effect of idiosyncratic distortions indexed by ξ on consumption. Consumption is normalized by the US (no distortion) level and in log scale. The right panel shows the amplification rate (solid line) and two of its components for given degrees of distortions ξ.

Figure 3: Effect of Distortions

amplification rate spikes up around $\xi = 0.2$, precisely where the distortions have a highly local nonlinear effect. This is mostly driven by the incentive elasticity component, which is the effect of the mass of adopters on the net gains from adoption for the marginal adopter (the dashed line resembling the solid line), rather than the reciprocal of the elasticity of the net gains from adoption to changes in the productivity of the marginal adopter, which was the second term in equation (15) (the monotonically increasing dotted line). As we see in the right panel of Figure 4, $\xi = 0.2$ is also where the fraction of adopters approaches zero. This helps us make sense of the non-monotonicity of the incentive elasticity with respect to ξ. As ξ rises and the fraction of adopters get smaller, the equilibrium prices and quantities, and hence the gains from adoption for the marginal adopter, are more sensitive to small changes in the fraction of adopters, explaining why the incentive elasticity rises with ξ to the left of $\xi = 0.2$. However, once the distortions are severe enough and the fraction of adopters is close enough to 0, small changes in the fraction of adopters can have less of an effect on the marginal adopter facing high taxes.

We call the interval of ξ around 0.2 the Big Push region, where the amplification rate is particularly high and hence a small change in distortions has a disproportionate effect on consumption.[12]

[12] In the standard notion of Big Push, aggregate outcomes can change significantly even without



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 4: Unpacking the Mechanisms

Unpacking the Mechanisms The 60-percent reduction in consumption as the degree of distortions goes from the US level to the Indian level (ξ from 0 to 0.195) is more than four times the effect in a basic model in the literature. To place this result in a sharper context and to explain how our model generates such an amplified effect, we contrast our model with alternative models that lack some of our model elements.

For each model, we re-calibrate the parameters to the same set of target moments as in our US benchmark (Section 4) and calculate the effect of idiosyncratic distortions. None of these alternative models features multiple equilibria for any degree of distortions ξ.

The left panel of Figure 4 shows the per-capita consumption against the degree of idiosyncratic distortions ξ. The right panel shows how the fraction of establishments that adopt the modern technology varies with ξ. In both panels, the solid lines are for our benchmark model.

The first comparison model we consider is the basic model in the distortion literature—for example Restuccia and Rogerson (2008)—that abstracts from intermediate inputs and technology adoption ($\nu_t = \nu_m = 0$ and $A_t = A_m$, with

any change in fundamentals, as there are multiple equilibria. We extend this notion to situations where a small change in fundamentals, e.g., the degree of distortions, has a disproportionately large effect on aggregate outcomes.

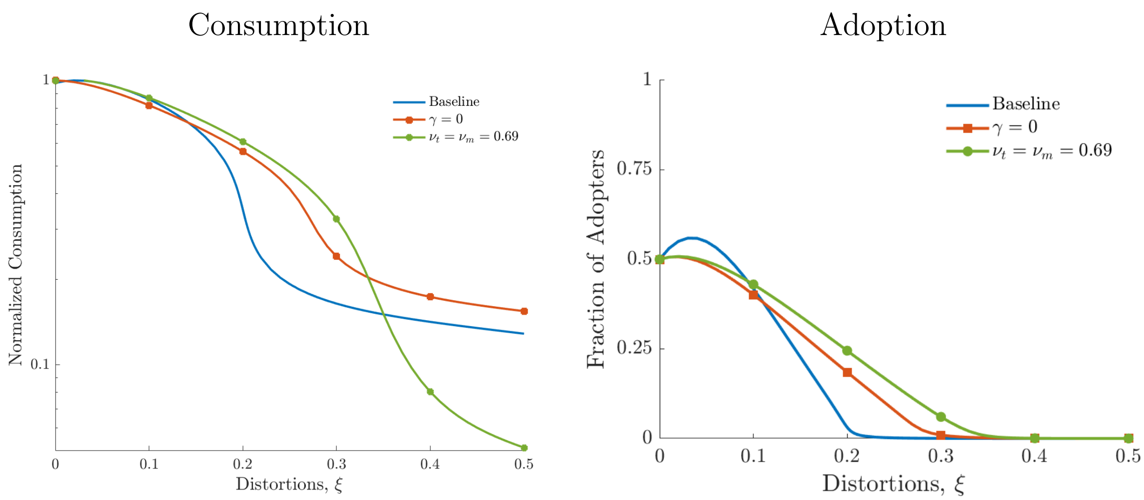
re-calibration), shown by the dashed line in the left panel of Figure 4. This specification should be considered as the polar opposite of our benchmark model. For this model, consumption falls almost linearly with the distortion parameter ξ in the semi-log scale and by much less than in the baseline model. At $\xi = 0.2$, consumption goes down by only 14 percent from its no-distortion level. Even with $\xi = 0.5$, the loss in consumption is only 30 percent.[13]

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 only 15 percent. The dotted line in the right panel shows that the adoption rate declines much more gently with the degree of distortions than in our benchmark model: It does not collapse to zero even with $\xi = 0.5$, whereas the adoption rate essentially hits zero when the degree of distortions reaches the empirically relevant value of 0.2 in our benchmark model with intermediate input.

The dashed line with triangles in the left panel instead adds round-about production (intermediate input) to the basic model, but with no technology choice. This is our adaptation of Jones (2011). Round-about production more than doubles the effect of distortions on consumption, which decreases by 37 percent as ξ goes to 0.2. However, the effect here is nearly linear in the semi-log scale with respect to ξ and is still significantly smaller than those in our benchmark.

We further examine the role of the model elements that are more specific to our model in Figure 5. The solid line is our benchmark result and the other two are for modified benchmark models. The solid line with squares is for a model with technology choice and round-about production, but one in which the adoption costs are in units of labor only, i.e., $\gamma = 0$. In the benchmark, the adoption costs are in units of the intermediate aggregates, i.e., $\gamma = 1$. 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

[13]Since this model has no technology adoption, the dashed line is flat at 0 in the right panel.



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 5: Unpacking the Mechanisms

the benchmark), and about 20 percent of active establishments adopt the modern technology (compared with nearly zero in the benchmark). The adoption rate goes near zero with $\xi = 0.3$, where the gap between the solid line and the solid line with squares narrows again. 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 3.3.

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 $\nu_t = 0 \ll \nu_m = 0.7$ in the benchmark. (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 38 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 zero ($\xi > 0.35$),

Model	Modification	Consumption (US = 1)
Benchmark		0.41
Labor only adoption cost	$\gamma = 0$	0.49
Same intermediate good share	$\nu_t = \nu_m$	0.62
Intermediate input w/o adoption	$\nu_t = \nu_m$, ~~adopt~~	0.63
Adoption w/o intermediate input	$\nu = 0, \gamma = 0$	0.85
No adoption or intermediate input	$\nu = 0$, ~~adopt~~	0.86

Table II: Impact of Distortions ($\xi = 0.2$) on Consumption for Alternative Models

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.

Table II summarizes the above results. Starting with the benchmark model, it shows the per-capita consumption level of each alternative model when the degree of distortions is $\xi = 0.2$. Consumption is normalized by its value in the undistorted US economy. It shows how much various model elements contribute to the negative effect of distortions on consumption in the benchmark model. In particular, the consumption reduction in our benchmark model is more than the simple product of the effect in the technology adoption model (second to last row) and that in the model with intermediate input but no technology choice (fourth row): at $\xi = 0.2$, the consumption level in our benchmark is 0.41, which is smaller than $0.85 \times 0.63 = 0.54$. This suggests a meaningful interaction between the two model elements rather than the two being simply additive.

The analysis in this section shows that technology adoption, round-about production, and the nature of the adoption cost (i.e., labor or goods) jointly generate the large, nonlinear effect of idiosyncratic distortions, even in the absence of multiple equilibria. This result highlights the potentially disproportionate gains from reducing idiosyncratic distortions.

Robustness The highly-nonlinear nature of the effect of idiosyncratic distortions in our model raises a question. Does this local nonlinearity exist for a wide range of parameter values or is this an artifact of the way we calibrate the model that will disappear with a small perturbation of the parameter values? In Appendix D we report how consumption per capita changes in response to the degree of idiosyncratic distortions for a wide range of technological parameters. The result shows that the highly-nonlinear and large effects of idiosyncratic distortions are a robust feature of our model.

5.2 Income Gap between the US and India

According to the results in the previous section, even if countries have the same technology and cost parameters and the only difference across them is the degree of distortions, our model can generate a huge gap in per-capita consumption, a factor of nearly 10 between $\xi = 0$ (no distortion) and $\xi = 0.5$, the upper bound of our quantitative exercise.

As mentioned in the previous section, to compute the ξ for India we assume that the underlying productivity distribution is the same as the (distortion-free) US distribution and calibrate the degree of distortions to match the tail of the establishment size distribution, according to

$$\frac{\zeta}{\eta\,(1-\xi)-1} = -\frac{\partial \log(1-G(l))}{\partial \log l}\ , \tag{23}$$

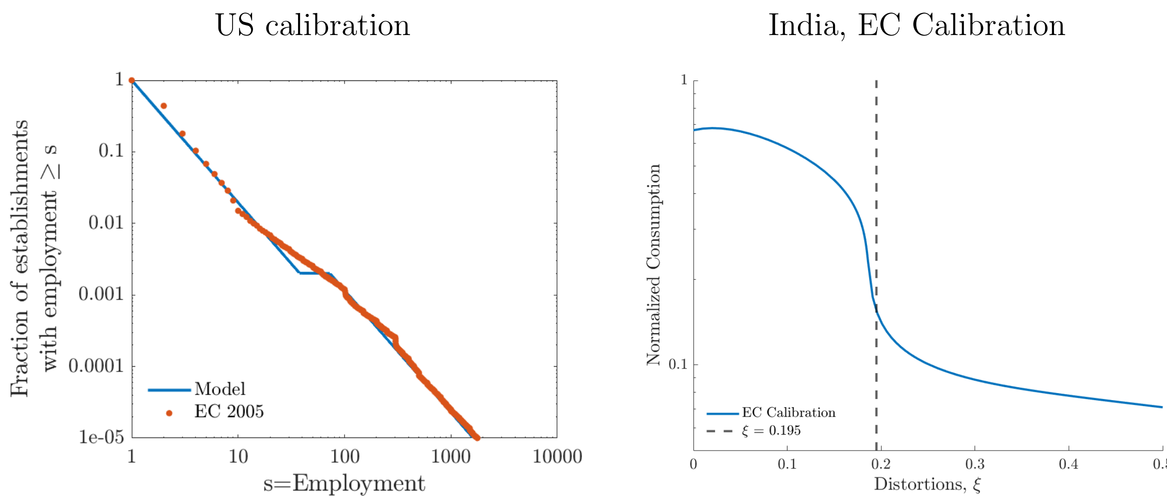
which is the distortion-augmented version of equation (25) in the appendix. The tail of the establishment size distribution in India is thinner and hence the log-log relationship in the right tail is steeper, resulting in $\xi = 0.195$.[14]

As we saw in the previous section, if the only difference between the US and India is the degree of idiosyncratic distortions, the per-capita consumption in India is 60 percent lower than that in the US. This is a large effect relative to that in more conventional models in the literature, but obviously falls short of the actual income gap between the two countries, which is a factor of 16.7.[15]

A natural question is then whether our model can come closer to accounting for the actual income gap if we allow countries to differ in other dimensions as well. For this purpose, we allow India's entry cost (κ_e), adoption cost (κ_a), and traditional sector productivity (A_t) to differ from the US calibration, in addition to the degree of distortions ξ. As with the US calibration, the target moments are from the Indian

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

[15] We use the output-side real GDP at chained PPPs and the number of persons engaged from the Penn World Tables 9.0.



Note: The left panel is the log-log plot from the calibrated model (solid line) and the data (dots) for India. The right panel shows the effect of idiosyncratic distortions for the Indian calibration. Per-capita consumption in log scale is normalized by its level in the US calibration.

Figure 6: Indian Calibration and the Effect of Distortions

establishment size distribution—26 points chosen from the establishment size log-log relationship, to be precise.

The left panel of Figure 6 is the log-log plot from the model (solid line) and the data (dots) for India.[16] The Indian calibration of the model shows both the traditional and the modern technology in use in equilibrium, separated by a flat region. The location of the flat region shows that only a small fraction of firms, 0.25 percent, uses the modern technology.

Table III shows the parameters that differ between the US and India. A few remarks are in order. First, the US and India have the same entry cost in units of labor κ_e, which is identified from the size of the smallest establishment using equation (26) in the appendix. This does not mean that the entry rate of firms is the same in the two countries. The fraction of potential entrants that do enter in India is six times that in the US. Second, the traditional technology parameter A_t for the US is

[16]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 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.

Parameter	US	India
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	0.20
Distortion scale parameter, τ	1	2.14

Table III: Calibrated Parameters for India

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 large 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 small fraction of firms adopting the modern technology in spite of the enormous productivity gains from doing so.[17] When measured in units of labor ($P\kappa_a$, since adoption costs are all in units of goods, $\gamma = 1$), 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.

The right panel of Figure 6 shows the impact of idiosyncratic distortions for the Indian calibration. Similar to the US benchmark (Figure 3), it shows the same large nonlinear effect of idiosyncratic distortions, although the consumption levels (normalized by the US level) are lower everywhere. There is only one line because the Indian calibration does not generate multiple equilibria for any ξ either. At $\xi = 0.195$, where the vertical line is, the per-capita consumption in India is only 15 percent of the US level. In other words, the US-India income gap is a factor of 6.7 in our model if some technology parameters, in addition to idiosyncratic distortions, are allowed to differ between the two countries. This last result is remarkable. 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, accounting for 73 percent of the US-India income gap in the data.[18]

[17] 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.

[18] The income gap is a factor of 16.7 in the Penn World Tables and 6.7 in our model. 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.

	US w/ Indian Parameters	India w/ US Parameters
Benchmark	1.0	0.15
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 IV: Explaining Consumption Difference

In Table IV, we calculate the contribution of each of the country-specific parameters to the US-India gap in the model. To do so, we compute the hypothetical US consumption by replacing one of the parameters with its value in the Indian calibration, holding all others constant. This result is in the first column. 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 6.7 difference between the US and Indian consumption (=1/0.15). Starting from the US calibration, the adoption cost difference has the largest impact: Giving 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. 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.[19]

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 adoption cost of the US while holding all other parameters constant results in a nearly five-fold increase in consumption, which is much larger than the factor of 2.7 in the first column: The rise in adoption gives a larger increase in productivity because the traditional technology is less productive in India. Eliminating idiosyncratic distortions in India raises consumption by a factor of 2.3 (=0.34/0.15), which is similar 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.

To summarize, the model generates a large income gap between the US and India, without directly targeting the income or productivity level of either country. Their

[19]In this case, the number of entrants is nearly halved, but all the entrants adopt the modern technology.

difference in adoption costs and distortions explains most of the income gap.

We can go one step further and ask how large an income gap can be generated by simpler versions of our model lacking certain features, when the same calibration strategy is followed. This exercise is analogous to the one in Section 5.1 that studied the impact of distortions in simpler versions of our model. The result is reported in Appendix E.[20]

5.3 Multiplicity and Amplification

In our quantitative analysis so far, multiple equilibria did not emerge. Does this mean that multiplicity does not exist in the empirically relevant regions of the parameter space? In this section, we first identify the combinations of technology and distortion parameters that generate multiple equilibria. We then discuss the effect of distortions with and without multiplicity.

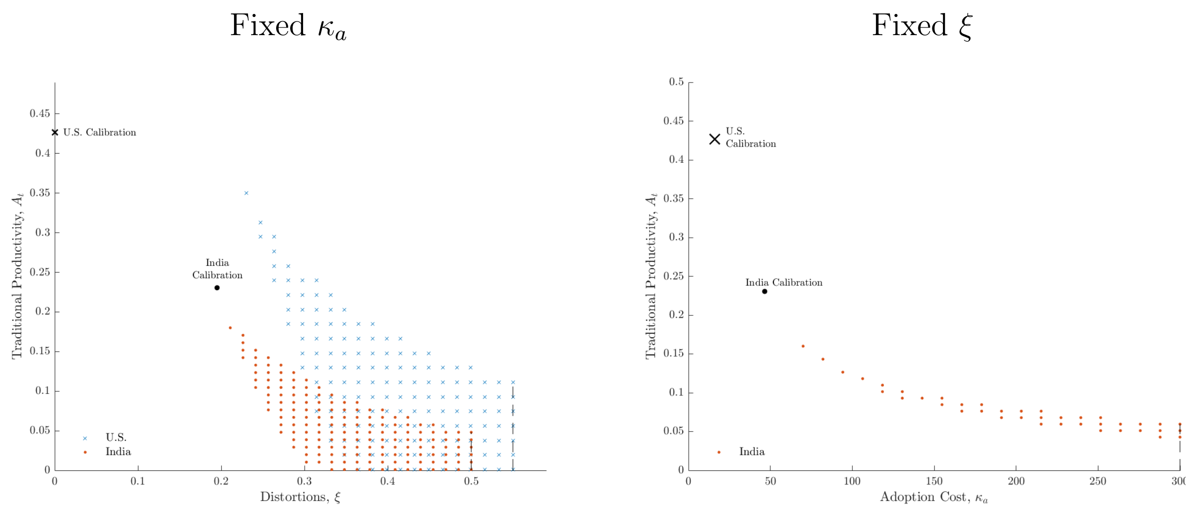
Region of Multiplicity Figure 7 shows the combination of parameters that generates multiple equilibria. In the left panel, starting with the US calibration (bold cross), we vary only the traditional technology productivity A_t and the degree of distortions ξ and mark with crosses the region of the $\xi - A_t$ space where multiple equilibria exist.[21] Similarly, starting with the India calibration (bold dot), we vary A_t and ξ to find the region of multiple equilibria, which is marked with dots.

We first note that, for both countries, increasing ξ alone will not result in multiplicity, consistent with our findings in the previous sections. Second, while the US is far removed from the region of multiplicity, India is close to it.

Multiple equilibria arise in economies with both high degrees of distortions and unproductive traditional technology. One interesting result is that, holding fixed the productivity of the traditional technology A_t, as we increase 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 ξ, 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. The unique equilibrium with A_t close to 0 has a small number of entrants, nearly all of whom adopt the modern technology: With a useless traditional technology, entry also

[20]Francesco Caselli suggested this comparison exercise, for which we are grateful.

[21]When we change ξ, we also change the scale parameter τ to balance the government budget.



Note: The figure shows, for the US (cross) and India (dot), the combination of κ_a, A_t and ξ for which multiple equilibria exist. The larger cross and dot correspond to the calibrated US and India respectively.

Figure 7: Region of Multiple Equilibria

implies adopting the modern technology, which effectively raises the cost of entry and results in few entrants.

The right panel of Figure 7 shows the corresponding exercise except that we vary the adoption cost κ_a and A_t, but not ξ. Again, India is close to the (narrow) region of multiplicity, but in this case the US calibration cannot generate multiple equilibria for any values of κ_a and A_t with ξ fixed at 0.

Multiplicity and Amplification To illustrate the role of multiplicity and amplification in the context of the effect of idiosyncratic distortions, we consider the point in the multiplicity region that is closest to the Indian calibration. For this purpose we assume $A_t = 0.18$, while holding fixed all other parameters to the Indian calibration. As can be seen in the left panel of Figure 7, for a very short interval of ξ, the model will now feature multiple equilibria.

Figure 8 shows in more detail how idiosyncratic distortions affect consumption per capita when multiple equilibria are present. Unlike in the Indian calibration, which is the unbroken solid line reproduced from the right panel of Figure 6, we see two separate lines, the solid line at the top and the dashed line at the bottom. When the degree of distortions ξ is low, the equilibrium is unique and has a high

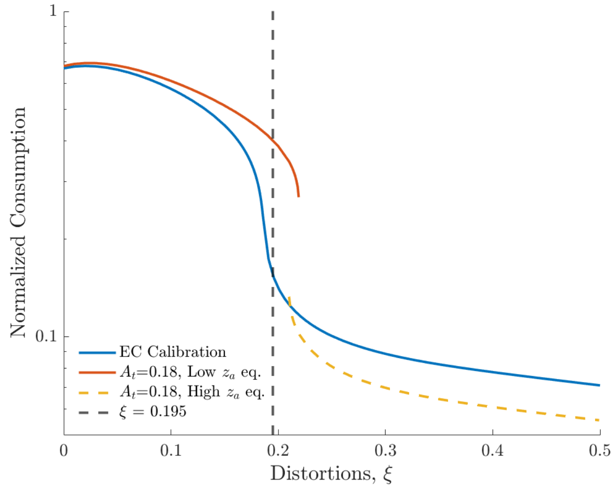


Figure 8: Effect of Distortions with and without Multiplicity

fraction of adopters, which is the good equilibrium. As we increase the degree of distortions, a second equilibrium, one with a small fraction of adopters, emerges (dashed line). This is the bad equilibrium with a low fraction of adopters. Both equilibria exist over a short interval of ξ around 0.22, beyond 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, discontinuously reducing consumption. However, coordination failures become irrelevant once ξ increases slightly more and push the economy to the region with the unique bad equilibrium.

It is notable that the overall effect of idiosyncratic distortions is comparable in magnitude whether or not multiple equilibria exist. As we saw in Section 5.1, even without multiplicity, idiosyncratic distortions have a highly nonlinear effect in some region, where complementarity in adoption decisions amplifies the impact of distortions. Multiplicity is an extreme form of amplification, but not a necessary condition for the impact of distortions to be large. The focus hence should be on how big the amplification is rather than whether or not we have multiple equilibria.

5.4 Big Push and Industrial Policy

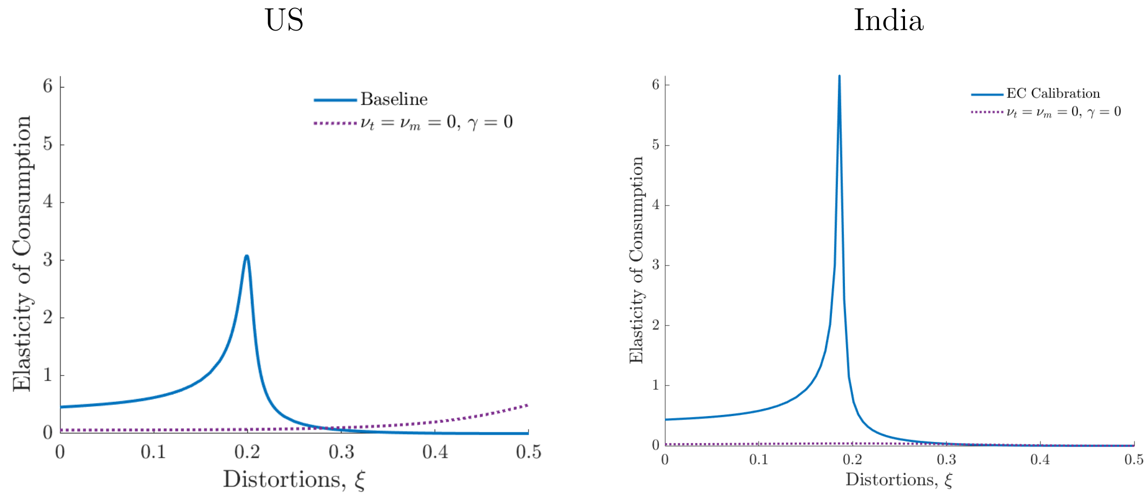
"Big Push" is the name Rosenstein-Rodan (1961) gave to the idea that a minimum scale of investment is necessary for economic development. The rationales are indivisibilities in the production function, especially social overhead capital, and complementarities across sectors. Under these conditions, individual firms may not find it profitable to industrialize alone, even though all firms are better off industrializing together.[22]

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. 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 in Section 5.3, the effects of reducing distortions get amplified in our framework with or without multiplicity. 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 that reduce distortions are discontinuously high, whether multiplicity is involved or not.

[22] This idea inherently presupposes multiple equilibria, and the proposed solution is an integrating, synchronizing force that coordinates toward the good equilibrium. For example, regarding the non-development of the British India in the nineteenth century, Rosenstein-Rodan (1961) noted that "an investment trust like the East India Company might have [made the investment], but the single firms approach of the City of London made this impossible."



Notes: The elasticity of aggregate consumption with respect to adoption subsidy for the US and India calibration, as the distortion parameter ξ goes from 0 to 0.5. 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

Industrial Policy 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 subsidies for technology adoption.

The industrial policy we implement subsidizes a fraction s of the cost of adopting 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 4. Figure 9 presents two noteworthy results. First, in our benchmark model with amplification, the elasticity of aggregate consumption to the subsidy is high in the Big Push region, especially for India, but relatively low outside of it. Second, the elasticity is much lower in the version lacking complementarity. One implication is that industrial policy can be highly successful if the economy is situated in a region of high degrees of complementarity and amplification.

In Appendix F, we consider optimal subsidy policy. To be specific, we solve the problem of a constrained planner, who chooses the adoption threshold z_a and the entry threshold z_e taking as given the idiosyncratic distortions. This exercise shows that the gains from reducing distortions tend to be larger than the gains from optimal 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 are the idiosyncratic distortions an important source of underdevelopment themselves, but also determine the effectiveness of the industrial policy subsidizing technology adoption.

6 Concluding Remarks

Can economic development be explained by multiple equilibria and coordination failures? The US and India calibration of our model gives a unique equilibrium, which says no to this question, although multiplicity cannot be dismissed as empirically irrelevant. However, our analysis emphasizes the role of complementarity and amplification, rather than the existence of multiple equilibria. Even without multiplicity, the complementarity in our model creates Big Push regions, where small changes in distortions and policies have disproportionately large effects. In fact, the overall effect of idiosyncratic distortions is comparable in magnitude whether or not multiple equilibria exist.

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.

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.

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.

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

Appendix

A Equilibrium Conditions

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

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

$$\left(\frac{w}{P}\right)^{\eta-1} = \left(\frac{(\eta-1)\tau}{\eta}\right)^{\eta-1}\left[\left(\frac{w}{P}\right)^{\nu_t(\eta-1)}A_t^{\eta-1}\int_{z_e}^{z_a} z^{(\eta-1)(1-\xi)}dF(z) + \left(\frac{w}{P}\right)^{\nu_m(\eta-1)}A_m^{\eta-1}\int_{z_a}^{\infty} z^{(\eta-1)(1-\xi)}dF(z)\right],$$

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

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 Identification of Model Parameters

To keep things simple, 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 the intermediate aggregate ($\gamma = 1$). In addition, we maintain the assumption that firm-level productivity distribution F is Pareto with tail parameter ζ. We set the elasticity of substitution across differentiated goods η from outside of the model. We also assume that there is no distortion $\xi = 0$. However, our identification strategy also holds without these 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, $A_m = 1$. For

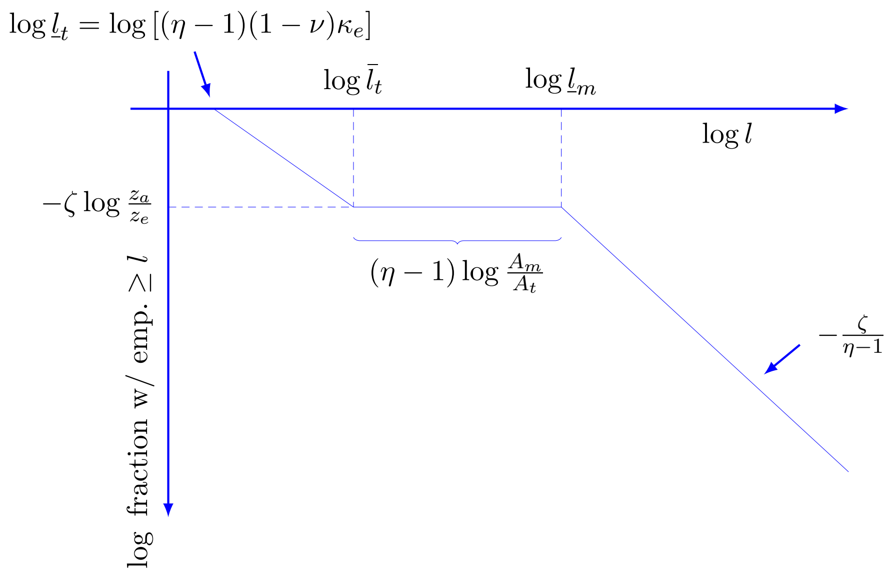


Figure 10: Identification from the Establishment Size Distribution

a given η, we can identify the intermediate input elasticity ν directly from the intermediate input share in the data:

$$\nu = \frac{\eta}{\eta-1} \times \text{intermediate input share}\ . \tag{24}$$

The other four parameters are identified from the establishment size distribution in the data, whose c.d.f. is $G(l)$. 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 greater than or equal to l as illustrated in Figure 10. 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:

$$\frac{\zeta}{\eta} = -\frac{d\log(1-G(l))}{d\log l} \;\Rightarrow\; \zeta = -\eta\frac{d\log(1-G(l))}{d\log l}\ . \tag{25}$$

Given the value of the intermediate input elasticity ν, the size of the smallest establishment is a simple function of the entry cost, pinning down κ_e:

$$\underline{l}_t = [(\eta - 1)(1-\nu)\kappa_e] \;\Rightarrow\; \kappa_e = \frac{\underline{l}_t}{(\eta-1)(1-\nu)}. \tag{26}$$

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.[23] 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:

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

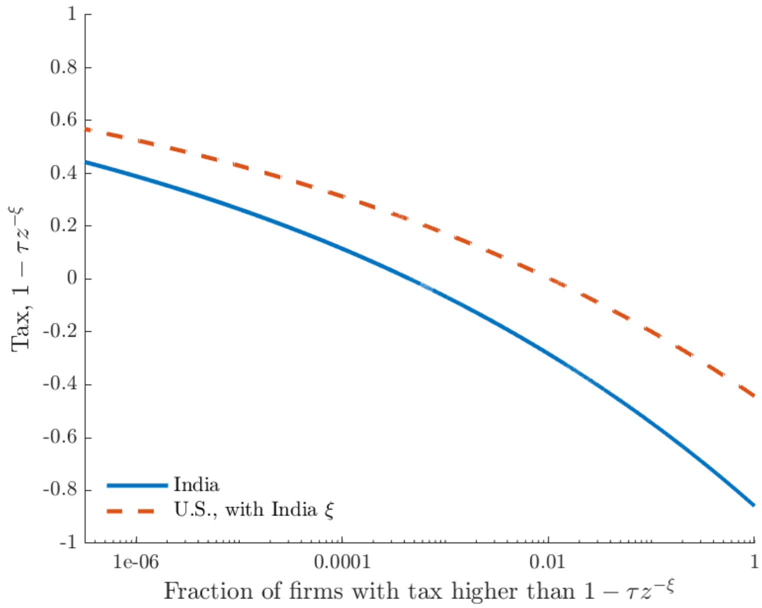
Finally, to identify the adoption cost κ_a we use the equilibrium condition relating the ratio of the two thresholds to the ratio of the adoption and the entry costs:

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

C Idiosyncratic Distortions

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 (Section 5.2), while the dashed line is the same object, but for the US economy using India's $\xi = 0.195$. 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

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



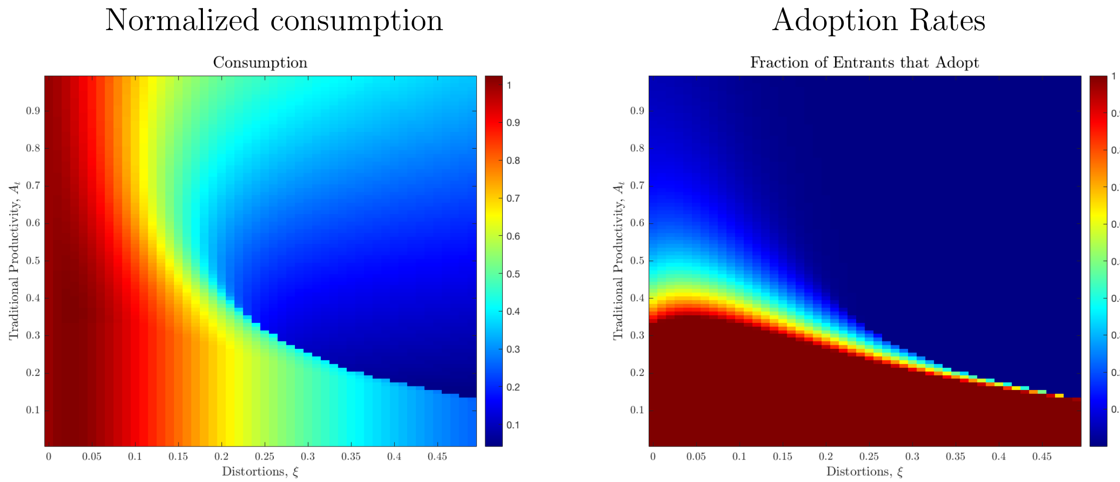
Note: The figure shows the cumulative distribution of taxes across firms. The solid line is for the Indian calibration with $\xi = 0.195$. The dashed line corresponds to the re-computed US economy using India's ξ.

Figure 11: Firm-specific Taxes and Subsidies implied by Idiosyncratic Distortions

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

D Sensitivity Analysis

We explore whether the highly-nonlinear effect of idiosyncratic distortions is a model feature that is present in a large swath of the parameter space. Figure 12 shows the effect of the degree of distortions ξ on consumption and adoption rates for different productivity levels of the traditional technology A_t. All other parameter values are fixed at the US calibration. The nonlinear effect is present in the middle interval of A_t. When A_t is too low, essentially all entrants adopt and hence the adoption margin and the complementarity of adoption become unimportant. When A_t is too close



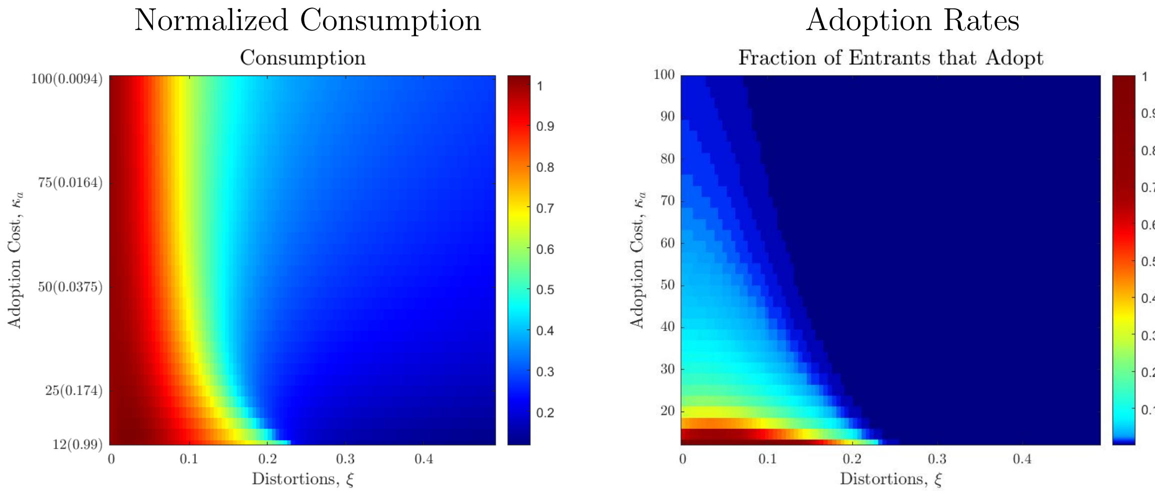
Notes: Heat maps of consumption per capita (left panel) and adoption rates (right panel) for a given degree of distortions (horizontal axis) and traditional technology productivity (vertical axis). Consumption is normalized by the equilibrium consumption in the undistorted US economy.

Figure 12: Consumption and Adoption

to $A_m = 1$, the benefit of adoption becomes small enough that the adoption margin becomes unimportant.

Figure 13 shows the effect of the degree of distortions ξ on consumption and adoption rates for different levels of adoption costs κ_a. All other parameter values are fixed at the US calibration. In the parentheses along the vertical axis is the adoption rate among active firms in the no-distortion ($\xi = 0$) equilibrium. The heat map shows that the nonlinear effect of distortions is present for nearly all values of the adoption costs we consider.

We now consider the role of the elasticity of substitution across differentiated products η and the intermediate input elasticity of the traditional technology ν_t, which we fixed at $\eta = 3$ 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). As we vary η or ν_t, we re-calibrate the model to the US data and re-do the exercises that show the effect of distortions (indexed by ξ) on consumption. 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



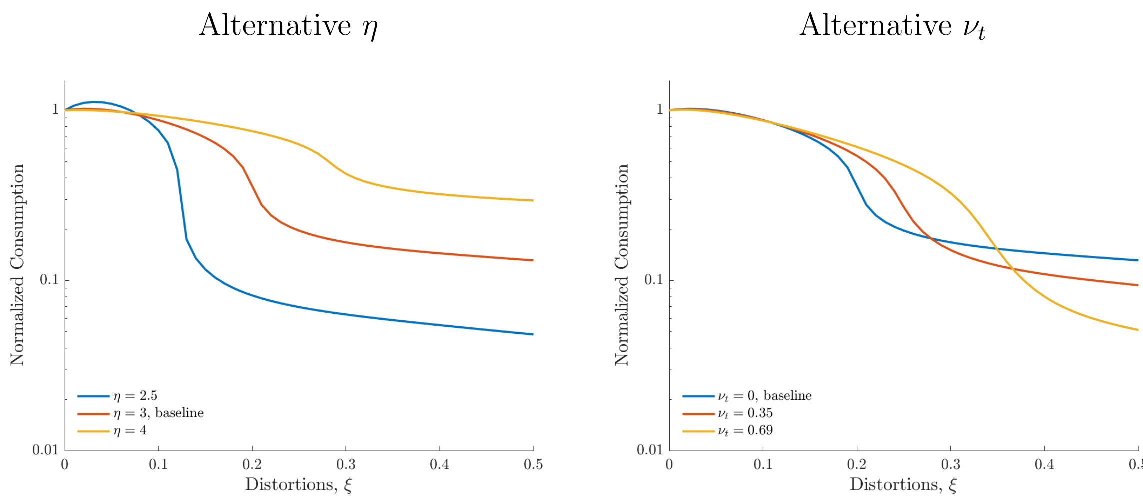
Notes: Heat maps of consumption per capita (left panel) and adoption rates (right panel) for a given degree of distortions (horizontal axis) and adoption costs (vertical axis). Consumption is normalized by the equilibrium consumption in the undistorted US economy.

Figure 13: Consumption and Adoption

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. This last one was shown in Figure 5. Figure 14 shows consumption per capita as we vary the idiosyncratic distortion parameter ξ. In all cases, consumption is normalized by the US consumption in the equilibrium with no distortions.

The left panel of Figure 14 show that a smaller elasticity of substitution η increases the impact of idiosyncratic distortions on consumption. A lower η makes goods less substitutable and firms' adoption decisions more complementary.

The right panel shows that the intermediate input elasticity of the traditional technology ν_t has two effects. Up to intermediate degrees of distortions, a higher ν_t means a smaller difference between the modern and the traditional technology and hence a smaller impact of distortions on consumption. However, with severe distortions, a higher ν_t means that even the traditional technology is dependent on the endogenously expensive intermediate input, leading to a larger impact of distortions on consumption.



Notes: Equilibrium consumption per capita of the US economy as the distortion parameter ξ goes from 0 to 0.5. Consumption per capita is normalized by its level with $\xi = 0$ and in log scale.

Figure 14: Effect of Distortions on Consumption with Alternative η and ν_t

E Accounting for the US-India Income Gap

To better understand how our model generates such a large income gap between the US and India, we calculate and compare the income gap in various alternative models that lack some of our model features.

In Section 5.1, we calibrated various comparison models to the US data with $\xi = 0$ and calculated the impact of higher degrees of distortions ξ. Here, we calibrate each of the comparison models separately to the US and the Indian data. We report the resulting GDP per capita of the two economies in Table V, 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 exercises the ξ for India remains at 0.195.

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 V). 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

		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 V: Alternative Model Results

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. 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.

The second comparison model has a technology choice with labor-only adoption costs, but still has no intermediate input (case 2 in the table). 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.

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.

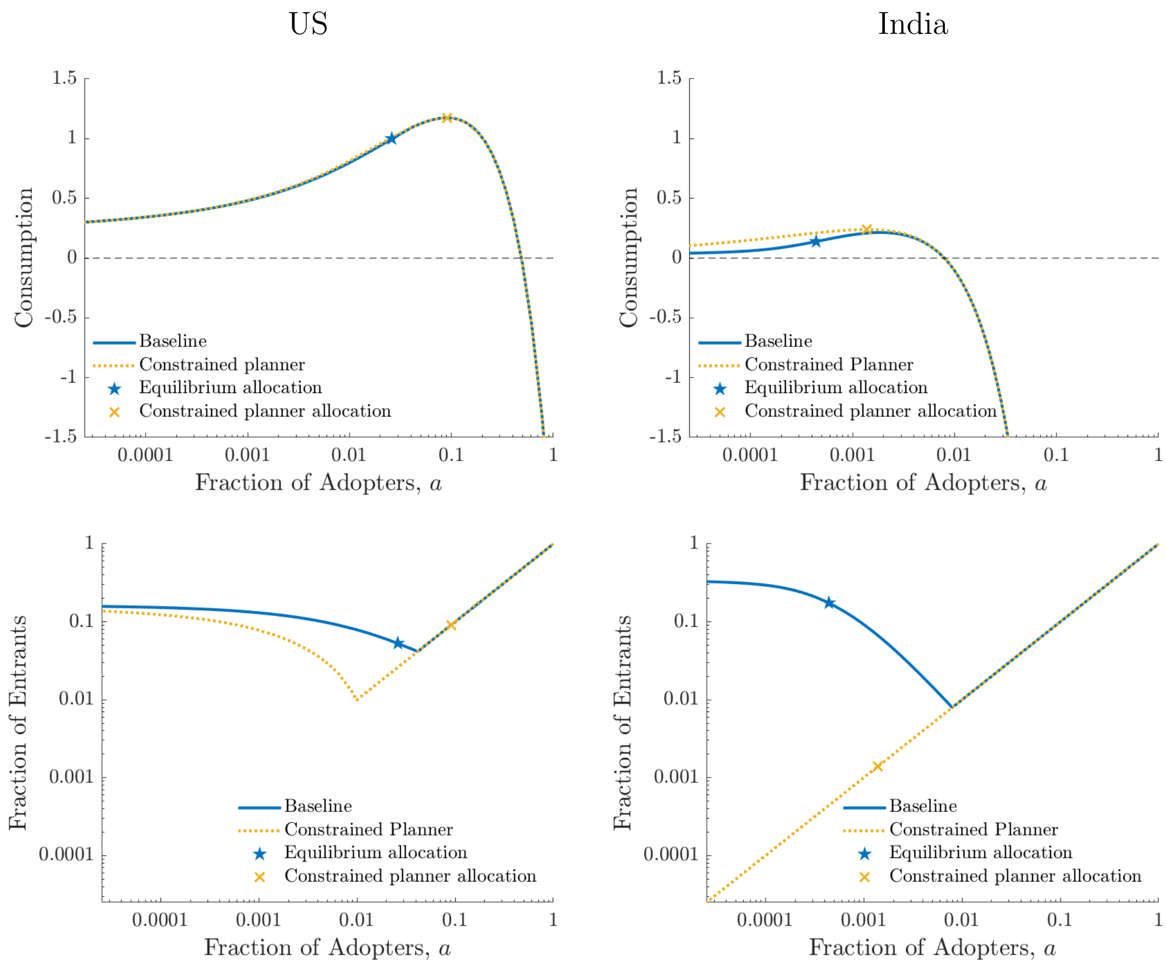
When the technology choice and the linkages through intermediate input use are both incorporated, as in our benchmark, the model generates a much larger GDP per capita gap between the two countries. The exact size of the gap can vary depending on the nature of the adoption costs (goods vs. labor, case 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 income gap larger than the sum of their respective individual effects.

F Constrained Planner Problem

To assess the potential gains from subsidizing entry and technology adoption, we solve the problem of a constrained planner, who chooses the adoption threshold z_a and the entry threshold z_e taking as given the idiosyncratic distortions.

The results for the US and the Indian calibration is shown in Figure 15. In the US, the planner allocation (star) has a slightly higher entry rate than the equilibrium allocation (cross), with all the entrants adopting the modern technology, as shown in the bottom left panel. (In the benchmark equilibrium, about half the entrants adopt the modern technology.) The planner allocation has a consumption level about 20 percent higher than that in the equilibrium (top left panel).

In India, the planner has much fewer firms entering but has all entering firms adopt the modern technology (bottom right panel). Since all entrants adopt, the high adoption costs imply high entry costs effectively, reducing the number of entrants. This planner allocation leads to a consumption level that is about 60 percent higher than that in the equilibrium (top right panel).



Notes: The top panels show the relationship between the fraction of adopters among all potential firms and the normalized consumption in the benchmark economy (solid line) and the constrained planner's problem (dotted line). The equilibrium is denoted with a star and the constrained planner allocation with a cross. The bottom panels show the relationship between the fraction of adopters among all potential firms and the fraction of entrants. Because the fraction of adopters among potential entrants cannot be larger than the fraction of entrants, the relationship is bounded from below by the 45-degree line.

Figure 15: Constrained Planner Allocation