

Big Push in Distorted Economies*

Francisco Buera

Washington University in St. Louis

Yongseok Shin

Washington University in St. Louis

Hugo Hopenhayn

UCLA

Nicholas Trachter

Federal Reserve Bank of Richmond

October 27, 2020

Abstract

Why don't poor countries adopt more productive technologies? Is there a role for policies that coordinate technology adoption? To answer these questions, we develop a model that features complementarity in firms' technology adoption decisions: The gains from adoption goes up when more firms adopt. When this complementarity prevails, multiple equilibria and hence coordination failures are possible. However, even without equilibrium multiplicity, the model elements that cause the complementarity can substantially amplify the negative impact of distortions. Our quantitative analysis calibrated to the establishment size distributions in the US and India successfully generates the gap in income per capita between the two countries, because of the larger adoption costs and distortions in India, not coordination failures even though India is in the region of multiplicity. Furthermore, the negative impact of idiosyncratic distortions on aggregate productivity can be highly non-linear. Over an intermediate range of distortions, a small reduction in distortions can disproportionately increase the number of adopters and hence aggregate productivity, with or without equilibrium multiplicity. This is what we call the big push in distorted economies.

*Buera: fjbuera@wustl.edu. Hopenhayn: hopen@econ.ucla.edu. Shin: yshin@wustl.edu. Trachter: trachter@gmail.com. We thank Andy Atkeson, Ariel Burstein, Ezra Oberfield, and participants at several seminars and conferences for their feedback. We thank 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? The standard view emphasizes the role of distortions or barriers to technology adoption (e.g., Parente and Prescott, 1999; Hsieh and Klenow, 2014; Cole et al., 2016; Bento and Restuccia, 2017). According to this view, eliminating the distortions is the obvious policy response. An alternative view emphasizes the role of multiple equilibria and coordination failures: Firms in poor countries use unproductive technologies because other firms use unproductive technologies, even though they will all benefit from a coordinated decision to adopt. According to this view, there is a role for a policy that coordinates 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. We emphasize the model elements that cause the complementarity in firms' technology adoption decisions, which is necessary for the equilibrium multiplicity central to the second view. Our quantitative analysis shows that these model elements substantially amplify the negative effect of distortions on aggregate productivity and GDP, sometimes through equilibrium multiplicity but also when the equilibrium is unique.

We develop a model of entry and technology adoption by ex-ante heterogeneous firms, which are subject to idiosyncratic distortions and 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. Firms produce differentiated intermediate goods.

We first theoretically analyze the conditions under which the complementarity in firms' technology adoption decisions supports multiple equilibria, which is a generalization of the existing theoretical results. We find that the complementarity and the existence of multiple equilibria are determined by the nature of adoption costs, the intermediate input intensity in production, and the degree of ex-ante heterogeneity in firm productivity and idiosyncratic distortions. We also show that the amplification of the impact of taxes/subsidies is possible even without equilibrium multiplicity.

Our theoretical analysis begins with a simple version without the entry margin, in which the firms are homogeneous, labor is the only input of the firms' production, and the adoption costs are in units of intermediate goods only. In this setup, the technology adoption of other

firms has two opposing effects on a firm's gains from adopting the modern technology. First, when more firms adopt, the price of the intermediate goods falls, reducing the marginal adopter's output price and profit, even though it also reduces the cost of adoption. On the other hand, when more firms operate the modern technology, more intermediate goods are produced, raising the demand for the marginal adopter's output and hence its profit. Unless the elasticity of substitution across the differentiated intermediate goods is too high, the second effect dominates and the marginal adopter's gains from adoption is increasing in the number of adopters, generating two stable equilibria: One in which all firms adopt the modern technology, because the marginal adopter's gains from adoption is positive when everyone else has adopted, and the other extreme in which no firm adopts, because the marginal adopter's gains from adoption is negative when nobody has adopted. If production requires intermediate goods as well as labor, the first negative effect on the marginal adopter's profit becomes smaller, because the lower price of intermediate goods reduces the cost of production as well. As a consequence, a higher reliance of firms' production on intermediate goods relaxes the condition needed for multiplicity.

When we introduce heterogeneity in firm productivity, we show that a higher degree of ex-ante heterogeneity diminishes the possibility of multiple equilibria. 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. Multiple equilibria can only arise when a sizeable mass of firms switch their decisions based on the number of adopters in the economy. This explains why all models of multiple equilibria in the literature have abstracted from firm heterogeneity. We do not shy away from it because firm heterogeneity is an important element of our quantitative strategy. In the same vein, idiosyncratic distortions that tax productive firms and subsidize unproductive ones increase the incidence of multiple equilibria, because they compress the dispersion in ex-ante *profitability* across firms. In addition, we analyze the conditions under which the impact of taxes and subsidies is amplified through equilibrium effects, which are shown to be weaker than those necessary for multiplicity. That is, the model can amplify the impact of distortions even in the absence of multiple equilibria, a result confirmed by our richer quantitative model.

Next, we use aggregate and micro-level data from the US and India and provide a quantitative analysis. The US is an undistorted benchmark, and India is a large developing country for which relevant micro-level data is available. The full version of our model has several layers, but it is tractable enough that most of its 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 interior equilibrium where adopters and non-

adopters coexist, the values of the key model parameters are obtained in closed form from the establishment size distribution, once we identify the empirical counterparts to the adoption/non-adoption threshold and the productivity gap between the modern and the traditional technology in the model.

The analysis gives us three main results. First, although our calibration targets the moments from the establishment size distribution but not the income level of either country, the model generates nearly as large an income gap between the US and India—a factor of seven—as in the data, because of the higher adoption costs and distortions in India. Second, the calibrated parameters place the US in the unique equilibrium region but India in the multiplicity region. However, India is found to be in the good equilibrium, and hence coordination failures do not explain why India is so much poorer than the US.

The final, and the most important, result is that the effect of idiosyncratic distortions on aggregate productivity and GDP is highly non-linear. When the complementarity in firms' adoption decisions prevails, multiple equilibria appear above a threshold degree of idiosyncratic distortions, but eventually only the bad equilibrium survives at even higher degrees of distortions. This implies that, for a heavily distorted economy in a unique bad equilibrium, a reform that reduces the distortions just enough to place it in the multiplicity region gives it a chance to coordinate to the good equilibrium with more adoptions and vastly higher GDP. Even when the model has a unique equilibrium for any degree of distortions, the model elements causing the complementarity still amplify the impact of distortions. A small reduction in distortions can disproportionately increase the number of adopters and hence aggregate productivity, several times more than the gains from distortion reductions of similar magnitudes in standard models (e.g., [Hsieh and Klenow, 2009](#); [Hopenhayn, 2014](#); [Restuccia and Rogerson, 2017](#)). These positive non-linear effects with or without multiplicity are the big push in distorted economies.

We draw two broader conclusions from our quantitative analysis. First, the powerful amplification of the impact of distortions by complementarity suggests that our model can account for the huge income differences across countries with reasonable degrees of distortions. Second, it offers an explanation of why some distortion-reducing reforms are more successful than others: In the region of non-linear effects, even a small reform can unleash massive improvements.

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 monopolistic competition and aggregate demand spillovers and by [Ciccone](#)

(2002) in a model with differences in intermediate input intensities across technologies.¹

Some empirical works have applied the idea of multiple equilibria and coordination failures to historical contexts. [Davis and Weinstein \(2002, 2008\)](#) examine the effect of the Allied bombing of Japanese cities and industries during World War II; [Redding et al. \(2011\)](#) use the division of Germany as an exogenous shock that relocated the air travel hub from Berlin to Frankfurt; [Kline and Moretti \(2014\)](#) study the long-run effects of the Tennessee Valley Authority; and [Lane \(2019\)](#) study the persistent impact of the Korean heavy and chemical industry drive in the 1970s. The empirical evidence so far is mixed, suggesting that the possibility of multiple equilibria depends on the details of the economic environment, a theme emphasized in our paper.

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 in [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 related 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.

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 data.² More important, we find that, even in the absence of multiplicity, these model elements amplify the impact of distortions on aggregate productivity.

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, which makes possible multiple equilibria and coordination failures.³ Our modeling choice is partly motivated by the evidence in [Holmes and Stevens](#)

¹[Krugman \(1992\)](#) and [Matsuyama \(1995\)](#) review the earlier generation of 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.

²[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 that literature abstracts from micro-level heterogeneity.

³[Yeaple \(2005\)](#) and [Bustos \(2011\)](#) also consider firms' technology choice decisions, but they either ignore possible multiplicity or make assumptions that happen to guarantee uniqueness. The small- vs. large-scale sector choice in the entrepreneurship model with financial frictions of [Buera et al. \(2011\)](#) can also be thought of as a technology choice, but that model also has a unique equilibrium.

(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 that operate the productive modern technology, with the technology choice driven by and reinforcing the underlying heterogeneity in firm-level productivity.

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

Finally, following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), we introduce idiosyncratic distortions, which stand in for various types of frictions, including barriers to technology adoption as modeled by Parente and Prescott (1999) and Cole et al. (2016). The interaction between distortions and technology adoption in our model is related to the impact of distortions on productivity-enhancing investment in Bento and Restuccia (2017) and Bhattacharya et al. (2013). Our emphasis is the amplification of the effect of distortions by the complementarity in firms' adoption decisions, which is responsible for 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 mass $L = 1$ of workers and measure one of potential firms, each of them producing a differentiated good j . 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,

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

where η is the elasticity of substitution. As we shall highlight in our analysis, this parameter governs how complementary are differentiated goods in the production of the intermediate aggregate. The intermediate aggregate can be transformed with a linear technology to produce the final consumption good and the intermediate goods to be used by firms.

Firms are heterogeneous in their productivity z , drawn from a cumulative distribution $F(z)$. Based on their productivity, firms choose to be active or inactive. An active firm with idiosyncratic productivity z must incur κ_e units of labor to enter the market and operate. An active firm produces using technology $i \in \{t, m\}$, labor l , and intermediate aggregate x ,

as described by the following expression:

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

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 $A_m / [\nu_m^{\nu_m} (1 - \nu_m)^{1 - \nu_m}] > A_t / [\nu_t^{\nu_t} (1 - \nu_t)^{1 - \nu_t}]$ and $\nu_m \geq \nu_t$. The advanced technology requires κ_a units of an adoption good. The adoption good is produced by a representative firm in the adoption good sector by combining labor and the intermediate aggregate using a Cobb-Douglas production function, where $1 - \gamma$ is the labor factor elasticity. We denote by P_a the price of the adoption good. Notice that we assume a more complex structure for the adoption cost relative to the entry cost. These assumptions allow us to focus on the equilibrium multiplicity that stems from the technology adoption choice rather than the entry choice.

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 a scale parameter. With $\xi > 0$, low productivity firms are subsidized and high productivity firms are taxed.

In our quantitative analysis, we assume only the following parameters vary across countries: the labor cost of entry κ_e , the cost of adopting the modern technology κ_a , the productivity of the traditional technology A_t , the degree of distortions ξ , and the budget-balancing scale parameter τ .

As is standard, the demand for differentiated good j is

$$y_j = \left(\frac{P}{p_j} \right)^\eta X,$$

where we define the price index of intermediate aggregate

$$P = \left[\int p_j^{1 - \eta} dj \right]^{\frac{1}{1 - \eta}}.$$

Since each firm produces one differentiated variety, the above expression is also the effective demand of the firm producing the differentiated good j .

The labor market is assumed to be perfectly competitive and we denote the equilibrium wage by w .

2.1 The Problem of a Firm

An active firm with productivity z producing with technology i collects operating profits $\pi_i^o(z)$. The firm must choose the price p for its differentiated good, and the amount of labor l and intermediate input m required for production. Its problem is given by

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

subject to

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

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

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

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

and

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

Using these expressions we obtain an expression for operating profits,

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

which is increasing in z provided $\eta(1 - \xi) - 1 > 0$. Finally, the profits of a firm with productivity z given the optimal entry and adoption decisions are

$$\pi(z) = \max_{\text{inactive, active}} \left\{ 0, \max_{t,m} \{ \pi_t^o(z), \pi_m^o(z) - P_a \kappa_a \} - w \kappa_e \right\}. \quad (2)$$

It is easy to verify that 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 $z \geq z_a$.

2.2 Adoption Good Sector

The representative firm producing the adoption good solves

$$\max_{L_a, M_a} \frac{1}{\gamma^\gamma (1 - \gamma)^{1 - \gamma}} P_a L_a^{1 - \gamma} X_a^\gamma - w L_a - P X_a. \quad (3)$$

From the first order conditions of this problem, and substituting the demand for the adoption goods from the differentiated goods producers, we obtain expressions for the price of the adoption good and input demands

$$P_a = P^\gamma w^{1-\gamma}$$

$$X_a = (1 - F(z_a)) \gamma \left(\frac{w}{P}\right)^{1-\gamma} \kappa_a$$

and

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

where z_a is the threshold for adoption.

2.3 Equilibrium

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

Definition 1. *A symmetric equilibrium is composed of entry and adoption decisions by producers of differentiated goods, 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 the intermediate aggregate clear:*

$$\begin{aligned} \int_{z_e}^{z_a} l_t(z) dF(z) + \int_{z_a}^{\infty} l_m(z) dF(z) + (1 - F(z_e)) \kappa_e \\ + (1 - F(z_a)) (1 - \gamma) \left(\frac{P}{w}\right)^\gamma \kappa_a = L \end{aligned}$$

and

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

Manipulating the equilibrium conditions, we can characterize the equilibrium by three equations in three variables, z_e , z_a and P/w . Of these, the following equation relating the ratio of the two thresholds to the ratio of the adoption and the entry costs will be used for identifying the parameters (Section 4).

$$\kappa_e \left(\frac{z_a}{z_e}\right)^{\eta(1-\xi)-1} = \frac{\left(\frac{P}{w}\right)^\gamma}{\frac{A_m^{\eta-1}}{A_t^{\eta-1}} \left(\frac{w}{P}\right)^{(\nu_m-\nu_t)(\eta-1)} - 1} \kappa_a \quad (4)$$

3 Understanding Amplification and Multiplicity

An important insight of this paper is that the impact of distortions is greatly amplified with or without multiplicity, when we combine several model elements that are typically explored in isolation in the literature. These elements include: complementarity as governed by the elasticity of substitution η , the difference in productivity A_i and intermediate input intensity ν_i between technologies, the cost of adoption κ_a , the degree of productivity heterogeneity across firms $F(z)$, and the degree of idiosyncratic distortions that firms face, ξ and τ . When complementarity is sufficiently strong, the interaction of these elements results in multiple equilibria, which is the ultimate form of amplification. In this section we analyze the role of these ingredients in generating amplification and multiplicity using simple examples.

3.1 Simple Model I: Complementarity and Coordination

Consider the case with no heterogeneity, $z_j = 1$ for all j , no intermediate input in the production of differentiated goods, $\nu_i = 0$, and an exogenous population of active firms, i.e., no entry margin. There is no idiosyncratic distortion. Let $a \in [0, 1]$ be the measure of adopters, since we cannot speak of a threshold without heterogeneous z . This simple example captures the traditional model of coordination failures in technology adoption (Murphy et al., 1989; Matsuyama, 1995).⁴

In this particular case, there is a simple closed form expression for the relative factor prices and the quantity of the intermediate aggregate:

$$\frac{w}{P} = \left(\frac{P_a}{P}\right)^{\frac{1}{1-\gamma}} = \left(\frac{\eta-1}{\eta}\right) \underbrace{[A_t^{\eta-1}(1-a) + A_m^{\eta-1}a]^{\frac{1}{\eta-1}}}_{\equiv \tilde{A}(a)} \quad (5)$$

and

$$X = \tilde{A}(a)L - (1-\gamma) \left(\frac{\eta}{\eta-1}\right)^\gamma \tilde{A}(a)^{1-\gamma} a \kappa_a,$$

which shows that the relative price of the adoption good, the real wage, and the intermediate aggregate are functions of a generalized mean of the productivity of firms $\tilde{A}(a)$ and, more important, are increasing in the fraction of adopters a .

We are interested in understanding how profits of any of the homogeneous firms depend on the fraction of adopters a . We denote by $\pi_i^o(a)$ the operating profits of a firm using

⁴The traditional model also assumes that there exists a fringe firm that can produce the goods with a productivity that is a fraction $\lambda_i < 1$ of the monopolistically competitive firm, with $\lambda_i \geq (\eta-1)/\eta$. In this case, the price charged by the leader equals $p_i = w/(\lambda_i A_i)$. If we further assume $A_t = \lambda_t = 1$, $\lambda_m = 1/A_m$ and $\eta = 1$ we obtain the specification in Matsuyama (1995).

technology i as a function of the fraction of adopters.

Using the expression for the relative price of the adoption good and the real wage, the net gain from adoption, in units of the adoption good, is given by a simple function of the fraction of adopters,

$$\begin{aligned}
& \frac{1}{P_a} [\pi_m^o(a) - \pi_t^o(a) - P_a \kappa_a] \\
&= \frac{1}{\eta} \left(\frac{\eta - 1}{\eta} \right)^{\eta-1} \left(\frac{P}{w} \right)^{\eta-\gamma} X (A_m^{\eta-1} - A_t^{\eta-1}) \\
&= \frac{\tilde{A}(a)^{1+\gamma-\eta}}{\eta^\gamma (\eta - 1)^{1-\gamma}} (A_m^{\eta-1} - A_t^{\eta-1}) - \frac{1 - \gamma}{\eta - 1} \frac{a (A_m^{\eta-1} - A_t^{\eta-1})}{\tilde{A}(a)^{\eta-1}} \kappa_a - \kappa_a. \tag{6}
\end{aligned}$$

An individual firm adopts the modern technology if the net gain from adoption is positive and does not otherwise. We can use this expression to study the presence of multiple equilibria. If the net gain is decreasing in the fraction adopters a , there is a unique equilibrium. In this case the net gains are either always negative, always positive, or cross zero once (from above), as a function of a . Then, the unique equilibrium exhibits no adoption, full adoption, or partial adoption, respectively. Therefore, a necessary condition for multiple equilibria is that the net gains from adoption increase with the fraction of adopters a . A necessary condition for this positive relationship between the net gains from adoption and the fraction of adopters is

$$0 < 1 + \gamma - \eta. \tag{7}$$

This condition guarantees that the first term in (6) is an increasing function of a . Notice that the second term in that expression is a decreasing function of a .

At an intuitive level, understanding the possibility of multiplicity requires first describing how adoption by other firms affect the net gains from adoption of an individual firm j . There are two competing forces.

The first force leads to a decline in firm j 's net gains when the fraction of adopting firms increases: As more firms adopt, more firms become productive and compete away firm j 's profit, lowering the price of the intermediate goods relative to the wage, P/w . As seen in the second line of equation (6), the strength of this effect depends on the value of $\eta - \gamma$. On the one hand, a higher η makes goods better substitutes, increasing the negative competition effects on firm j 's profit. On the other hand, a higher γ makes the production of the adoption goods more intensive in the intermediate aggregate, mitigating the negative effect of a lower price of the intermediate goods relative to the wage, as $P_a/w = (P/w)^\gamma$.

The second force, solely operating through the quantity of the intermediate aggregate, X , increases firm j 's net gains from adoption when the fraction of adopting firms increases:

As more firms adopt, the production of intermediate goods increases for a given amount of labor supplied, raising the demand for firm j 's output. As seen in (6), this positive effect of adoption is partially dampened as the labor available to produce the intermediate goods is reduced as more labor is devoted to the adoption goods production.

The strength of the net effects and hence the possibility of multiplicity depend on η and γ . In particular, the elasticity of the net gains from adoption with respect to the fraction of adopters is a function of the ratio $(\gamma - \eta + 1) / (\eta - 1)$, which explains condition (7).⁵

Let a_0 denote the fraction of adopters such that the marginal adopter's net gains from adoption are 0. A simple sufficient condition for multiplicity is that an interior a_0 exists such the net gains from adoption are negative when the fraction of adopter is less than a_0 and positive when the fraction is greater than a_0 . In this case, the interior a_0 represents an unstable equilibrium. If the adoption costs are in units of intermediate goods only, i.e. $\gamma = 1$, a sufficient condition for multiplicity is given by (7) and

$$\frac{A_t^{2-\eta} (A_m^{\eta-1} - A_t^{\eta-1})}{\eta} < \kappa_a < \frac{A_m^{2-\eta} (A_m^{\eta-1} - A_t^{\eta-1})}{\eta}, \quad (8)$$

where the term to the left of the first inequality is the gross gain from adoption when $a = 0$ and the term to the right of the second inequality is the gross gain from adoption when $a = 1$. In the η - κ_a space, the region of multiple equilibria is non-empty.

In this simple example, as long as multiple equilibria exist, small policies can have disproportionate effects on outcomes. One way of understanding this is by thinking about the best response of a firm to the fraction of adopters a in the economy. For $a < a_0$, the best response of all firms is to use the traditional technology. For $a > a_0$, the best response is to adopt the modern technology. This implies that in the neighborhood of the unstable equilibrium a_0 , a policy that increase the fraction of adopters from $a_0 - \epsilon$ to $a_0 + \epsilon$, where $\epsilon > 0$ is arbitrarily close to 0, has an extreme effect on adoption and output. We will expand on the connection between amplification and multiplicity as we enrich the model in the following sections.

Comparison with the Planner's Allocation To further clarify the source of multiplicity, it is instructive to compare the equilibrium outcome with the solution to the planner's problem. To simplify the analysis, we will study the planner's problem for the $\gamma = 1$ case

⁵The elasticity of the net gain from adoption with respect to the fraction of adopters equals

$$\frac{\partial \frac{\pi_m^o(a) - \pi_t^o(a)}{P_a}}{\partial a} \frac{a}{\frac{\pi_m^o(a) - \pi_t^o(a)}{P_a}} = \frac{\gamma - \eta + 1}{\eta - 1} \frac{a (A_m^{\eta-1} - A_t^{\eta-1})}{\tilde{A}(a)^{\eta-1}} \frac{L}{L - L_a} - \frac{A_t^{\eta-1}}{\tilde{A}(a)^{\eta-1}} \frac{L_a}{L - L_a}.$$

(goods only adoption costs). The planner chooses the allocation of labor to traditional and modern firms producing differentiated goods, l_t and l_m , and the fraction of firms adopting the modern technology, a , in order to maximize final consumption

$$\max_{l_t, l_m, a} \left[(A_t l_t)^{1-\frac{1}{\eta}} (1-a) + (A_m l_m)^{1-\frac{1}{\eta}} a \right]^{\frac{\eta}{\eta-1}} - a \kappa_a$$

subject to the resource constraint for labor

$$(1-a) l_t + a l_m = L = 1.$$

We first use the first order conditions with respect to l_t and l_m in order to obtain a simpler version of the planner's problem,

$$\max_a \tilde{A}(a) - a \kappa_a.$$

We make two points here. The first is that when $0 < 2 - \eta$, which is the necessary condition (7) when $\gamma = 1$, the planner's problem exhibits increasing returns in adoption. In this case, the planner will choose a unique extreme value for the fraction of adopters: either no firm adopts or all firms adopt, depending on $A_m - \kappa_a \lesseqgtr A_t$. The second is that the marginal incentives in the equilibrium and the planner's problem can coincide. The social marginal value of adoption equals

$$\frac{1}{\eta-1} \tilde{A}(a)^{2-\eta} [A_m^{\eta-1} - A_t^{\eta-1}] - \kappa_a.$$

In the case of $\gamma = 1$, it is easy to see in equation (6) that there is a strictly positive difference between the social marginal value of adoption and the private value,

$$\left(\frac{1}{\eta-1} - \frac{1}{\eta} \right) \tilde{A}(a)^{2-\eta} (A_m^{\eta-1} - A_t^{\eta-1}) > 0.$$

The planner's incentives and an individual firm's incentives to adopt in the equilibrium are different because the latter faces a higher price of the adoption goods. If we introduce a revenue subsidy $\tau = \eta/(\eta-1)$ to undo the markup distortions, the marginal gains from adoption in the planner's problem and in the equilibrium are aligned. Even without the revenue subsidy, if $(A_t/A_m)^{2-\eta} < (\eta-1)/\eta$, for the adoption cost satisfying (8), the social marginal value of adoption and the private value are both strictly negative when $a = 0$ and strictly positive when $a = 1$. This is what we refer to as the coincidence of the marginal incentives in the equilibrium and the planner's problem. The crucial difference between the

planner's solution and the equilibrium outcome is not that individual firms do not internalize all the effects of their choices, but rather that the planner is able to coordinate the action of many firms (that is, choosing the fraction a , rather than taking it as given) to obtain a global optimum, which is not an option for an individual firm in an economy stuck in the bad (no-adoption) equilibrium.

3.2 Simple Model II: Heterogeneity

In this section we introduce ex-ante heterogeneity in firm productivity into the model. We have two objectives. First, we will analyze how heterogeneity affects the incidence of multiple equilibria. Second, we will precisely define amplification and study whether the amplification of the effect of taxes/subsidies is possible regardless of equilibrium multiplicity.

We assume that the intermediate goods producers' idiosyncratic productivity z follows a Pareto distribution, $z \in [1, \infty) \sim 1 - z^{-\zeta}$. We retain all other simplifying assumptions of the previous example: no entry margin, no idiosyncratic distortions ($\xi = 0$ and $\tau = 1$), no intermediate input in the production of intermediate goods ($\nu_i = 0$), and adoption costs in units of the intermediate goods only ($\gamma = 1$).

3.2.1 Multiplicity of Equilibria

Utilizing the mapping between the marginal adopter's productivity and the fraction of adopters $a = z_a^{-\zeta}$, we can derive a simple expression for the relative factor price

$$\frac{w}{P} = \frac{\eta - 1}{\eta} \left(\frac{\zeta}{\zeta - (\eta - 1)} \right)^{\frac{1}{\eta - 1}} \underbrace{\left[A_t^{\eta - 1} \left(1 - a^{1 - \frac{\eta - 1}{\zeta}} \right) + A_m^{\eta - 1} a^{1 - \frac{\eta - 1}{\zeta}} \right]^{\frac{1}{\eta - 1}}}_{\equiv \tilde{A}(a)} \quad (9)$$

as a function of the modified generalized mean of the firms' productivity $\tilde{A}(a)$, which itself is an increasing function of the fraction of adopters a (and hence decreasing in z_a), in the relevant parameter space $\zeta > \eta - 1$.

We want to analyze how the net gains from technology adoption for the marginal adopter vary with the fraction of adopters. Again denoting by $\pi_i^o(a)$ the operating profit of operating technology i for the marginal adopter with $z_a = a^{-\frac{1}{\zeta}}$, the net gains from adoption in units of the intermediate aggregate and also the adoption goods are:

$$\frac{1}{P} [\pi_m^o(a) - \pi_t^o(a) - P\kappa_a] = \left(\frac{\zeta}{\zeta - (\eta - 1)} \right)^{\frac{2 - \eta}{\eta - 1}} \frac{1}{\eta} \tilde{A}(a)^{2 - \eta} a^{-\frac{\eta - 1}{\zeta}} (A_m^{\eta - 1} - A_t^{\eta - 1}) - \kappa_a. \quad (10)$$

As in the previous example, if the net gains are decreasing in a , the equilibrium is unique.

Taking the derivative of the net gains with respect to a , we obtain the following condition for the net gains to be increasing in a :

$$0 < 2 - \frac{1}{1 + \zeta} - \eta, \quad (11)$$

which is a necessary condition for multiple equilibria. As heterogeneity vanishes ($\zeta \rightarrow \infty$), this condition coincides with the necessary condition without heterogeneity, which is $0 < 2 - \eta$ when $\gamma = 1$ —see (7). For finite values of ζ , it is a stronger condition, and more so when there is more heterogeneity in firm productivity z (smaller ζ).

Comparing equations (6) and (10), one finds that heterogeneity in z introduces the $a^{-\frac{\eta-1}{\zeta}}$ terms into the net gains, which reflect the fact that a higher fraction of adopters inherently means the marginal adopter's productivity z_a is lower. When the fraction of adopters is sufficiently close to zero, the productivity of the marginal adopter goes to infinity, implying that the net gains from adoption are unboundedly large and a strictly decreasing function of the fraction of adopters. As a consequence, any equilibrium must feature a positive fraction of adopters $a > 0$. When the fraction of adopters is close to 1, condition (11) guarantees that the net gains from adoption are a strictly increasing function of the fraction of adopters. With more heterogeneity, i.e., a smaller ζ , an increase in a translates into a larger fall in z_a (the elasticity of z_a with respect to a is $-1/\zeta$), and hence a stronger degree of complementarity (smaller η) is necessary for the net gains to be increasing in a and hence for equilibrium multiplicity.

3.2.2 Complementarity and Amplification

To formally define amplification, we consider the effect of a subsidy that reduces the adoption costs by a fraction s . The condition determining the adoption threshold z_a for a given fraction of adopters a is

$$\Delta\pi^o(z_a, a) \equiv \pi_m^o(z_a, a) - \pi_t^o(z_a, a) = (1 - s)P(a)\kappa_a. \quad (12)$$

We define the direct effect of the subsidy s on the marginal adopter z_a (and through $a = 1 - F(z_a)$, on the fraction of adopters) by

$$\frac{\partial z_a}{\partial s} = -\frac{1}{f(a)} \frac{\partial a}{\partial s} = -\frac{P(a)\kappa_a}{\Delta\pi_z^o}.$$

Totally differentiating (12), we obtain an expression for the total effect on the marginal adopter z_a (and on the fraction of adopters) of a change in the subsidy s , which includes the

equilibrium effects:

$$\frac{dz_a}{ds} = -\frac{1}{f(z)} \frac{da}{ds} = -\frac{P(a)\kappa}{(\Delta\pi_z^o(z, a) + (P'(a)(1-s)\kappa - \Delta\pi_a^o(z, a))f(z))}.$$

We define the policy multiplier as the ratio of the total effect to the direct effect:

$$\begin{aligned} \frac{\frac{dz_a}{ds}}{\frac{\partial z_a}{\partial s}} &= \frac{\Delta\pi_z^o(z, a)}{\Delta\pi_z^o(z, a) + (P'(a)(1-s)\kappa - \Delta\pi_a^o(z, a))f(z)} \\ &= \frac{1}{1 - \frac{2-\eta}{\eta-1}\Delta\pi^o(z, a)zf(z)}. \end{aligned}$$

Thus, a necessary and sufficient condition for the multiplier to be greater than one (i.e., amplification through equilibrium effects) is $\eta < 2$, as $\Delta\pi^o(z, a) > 0$. This condition for amplification is weaker than the necessary condition (11) for the existence of multiple equilibria with Pareto heterogeneity. In other words, even when the model parameters do not support multiplicity, the model can amplify the effect of taxes/subsidies. Such amplification without multiplicity will be an important feature of our richer quantitative model.

As in Section 3.1, we can study how amplification relates to the best response of a firm, which we label by a' , to the aggregate adoption rate a . Here, we obtain that

$$\frac{da'}{da} = \frac{2-\eta}{\eta-1}\Delta\pi^o(z, a)zf(z),$$

so that the best response of a firm is increasing in the aggregate adoption rate as long as $\eta < 2$, which is the same condition we obtained above to obtain a multiplier greater than one. In other words, amplification requires coordination—in terms of adoption decisions—by firms in the economy.

3.3 Simple Model III: Distortions

We now introduce idiosyncratic output distortions to the simple example with heterogeneity above, and explore their role in generating multiple equilibria. We retain all other simplifying assumptions: no entry margin, $\nu_i = 0$, and $\gamma = 1$. Combining the price index and the labor market clearing condition, and again denoting by $\pi_i^o(a)$ the operating profit for the marginal adopter when fraction a of the firms has adopted, the net gains from adopting the modern

technology for the marginal adopter can be written as follows.

$$\begin{aligned}
& \frac{1}{P} [\pi_m^o(a) - \pi_t^o(a) - P\kappa_a] \\
&= \frac{1}{\eta} \left(\frac{\eta-1}{\eta} \right)^{\eta-1} \tau^\eta \left(\frac{P}{w} \right)^{\eta-1} X a^{-\frac{\eta(1-\xi)-1}{\zeta}} (A_m^{\eta-1} - A_t^{\eta-1}) - \kappa_a \\
&= \frac{\tau}{\eta} \left(\frac{\zeta}{\zeta - (1-\xi)(\eta-1)} \right)^{\frac{1}{\eta-1}} \frac{\zeta - (\eta(1-\xi) - 1)}{\zeta} \\
&\quad \times \frac{\left[A_t^{\eta-1} \left(1 - a^{1-\frac{(1-\xi)(\eta-1)}{\zeta}} \right) + A_m^{\eta-1} a^{1-\frac{(1-\xi)(\eta-1)}{\zeta}} \right]^{\frac{1}{\eta-1}} a^{-\frac{\eta(1-\xi)-1}{\zeta}}}{A_t^{\eta-1} \left(1 - a^{1-\frac{\eta(1-\xi)-1}{\zeta}} \right) + A_m^{\eta-1} a^{1-\frac{\eta(1-\xi)-1}{\zeta}}} (A_m^{\eta-1} - A_t^{\eta-1}) - \kappa_a \quad (13)
\end{aligned}$$

If there were no distortions ($\xi = 0$), the exponent of a in the square brackets $\zeta - (1-\xi)(\eta-1)$ is the same as the exponent of a in the denominator $\zeta - \eta(1-\xi) + 1$, and equation (13) becomes the same as equation (10). The output distortions introduce heterogeneous markups, resulting in different powers of z (and hence a) coming from the price index and in the labor market clearing condition. In addition, distortions have a direct effect on the profitability of the marginal adopter, negative for productive firms and potentially positive for unproductive ones. This direct effect mitigates the (negative) sensitivity of the marginal adopter's profit to the fraction of adopters. With small distortions, $\xi \approx 0$, an increase in distortions has no effect on the quantity of the intermediate aggregate X , and only has a positive general equilibrium effect on the net gains of adoption, since it raises the price of the intermediate aggregate relative to the wage. This positive effect is countervailed by the direct negative effect. When there are few adopters, the direct negative effect dominates, while the opposite is true when the fraction of adopters is close to one. All in all, distortions tilt the net gains profile with respect to the fraction of adopters a counterclockwise.

As in the previous simple examples, if the marginal adopter's net gains are decreasing in a , the equilibrium is unique. Taking the derivative of the right-hand side of equation (13), we derive a necessary condition for the net gains to be increasing in a for A_t close to 0:

$$0 < 2 - \frac{1}{1 + \frac{\zeta}{1-\xi}} - \eta. \quad (14)$$

This condition is identical to condition (11) when $\xi = 0$, but is weaker when $\xi > 0$. When distortions are higher, i.e., ξ is larger, the sensitivity of the aggregate price index to individual productivity is dampened by a factor $1-\xi$. This mitigates the negative effect of heterogeneity on the sensitivity of the net gains from adoption with respect to the fraction of adopters. Indeed, when distortions are larger, there is less effective heterogeneity, i.e., $(1-\xi)/\zeta$ is

lower.

Amplification It is also possible to analyze the condition for amplification as in Section 3.2.2. We find the condition for amplification to be different from condition (14), although we cannot show that it is weaker for all parameter values. The equations get complicated and do not permit many clean results, but we find that an increase in the degree of distortions (ξ) leads to bigger amplification when the fraction of adopters is close to 0, and to smaller amplification when the fraction of adopters is close to 1. That is, amplification is possible without multiplicity, and distortions can amplify the amplification of the impact of taxes/subsidies through equilibrium effects in poor countries with few adopters.

3.4 Discussion of the Role of Additional Elements

In this section we briefly describe other ingredients present in our model that help strengthen complementarity or match relevant aspects of the data. We first discuss the role of intermediate inputs in production and then the role of the entry margin.

Intermediate Inputs in Production In the model, the differentiated goods produced by the heterogeneous firms are aggregated and used both for the production of the adoption goods and for the production of the differentiated goods themselves as intermediate input (i.e., round-about production). In the case where the intermediate input elasticity does not vary with the technology i , the parameter $\nu_i = \nu$ for all i denotes the intermediate input elasticity. In this case, for the simple model developed in Section 3.1, where ν was zero, the necessary condition for multiplicity generalizes to

$$0 < \gamma - (\eta - 1)(1 - \nu) = 1 + \gamma - \eta + \nu(\eta - 1). \quad (15)$$

Comparing conditions (7) and (15), we see that the latter poses a weaker condition for multiplicity as long as $\nu > 0$ and that an η exceeding 2 is now compatible with multiplicity, unlike in the previous examples.

Intermediate input use in the production relaxes the necessary condition for multiplicity for two reasons. First, the decline in the price of the intermediate goods reduces a firm's production cost, dampening the negative competition effect on its profits. When $\nu = 1$, the necessary condition is always satisfied for any positive γ (intermediate input share in adoption goods production), independently of the value of η .

The second reason is as follows. As in the simple example, the strength of the net effect of the fraction of adopters on the gains from adoption depends on the degree of complementarity

in the production of the intermediate aggregate determined by η . When intermediate aggregate is itself an input to the production of differentiated goods, the strength of the net effect also depends on ν , through a standard intermediate input multiplier $1/(1-\nu)$.⁶ When $\nu > 0$, as firms adopt the modern technology, the demand for firm j 's output increases not only because the demand for the adoption goods increases as in Section 3.1, but also because the adopting firms become more productive and demand more of firm j 's output as intermediate input.

Finally, when the intermediate input elasticity varies with the technology choice, $\nu_m \geq \nu_t$, as is the case in our benchmark model, these positive effects of other firms' adoption on firm j 's incentive to adopt are further amplified. The 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. In this case, multiplicity can arise even when the production of the adoption goods requires only labor ($\gamma = 0$), which shares some similarity with the result in Ciccone (2002).

Entry Margin Our benchmark model incorporates the firm entry margin: A firm must pay a fixed labor cost $\kappa_e > 0$ to become active. If firms' productivity follows a Pareto distribution, the entry margin is summarized by the productivity of the marginal entrant z_e given by

$$z_e = \max \left\{ 1, \left[1 + \frac{(\eta-1)\zeta}{\zeta - (\eta-1)} \left[1 + \left(\left(\frac{A_m}{A_t} \right)^{\eta-1} - 1 \right) a^{1-\frac{\eta-1}{\zeta}} \right] \right]^{\frac{1}{\zeta}} \kappa_e^{\frac{1}{\zeta}} \right\}. \quad (16)$$

The productivity of the marginal entrant and, therefore, the average size of firms z_e^ζ are an increasing function of the fraction of adopters a , the cost of entry κ_e , and the relative productivity of the modern and the traditional technology A_m/A_t . The other elements of the equilibrium are characterized as before with the lower bound of the integrals replaced by z_e and the labor used for entry $z_e^{-\zeta}\kappa_e$ netted out from the labor that can be used for the production of the intermediate goods and the adoption goods.

⁶The elasticity of the net gains from adoption with respect to the fraction of adopters equals

$$\begin{aligned} \frac{\partial \frac{\pi_m^o(a) - \pi_t^o(a)}{P_a}}{\partial a} \frac{a}{\frac{\pi_m^o(a) - \pi_t^o(a)}{P_a}} &= \frac{\gamma - (\eta(1-\nu) + \nu) + 1}{(\eta-1)(1-\nu)} \frac{a \left(A_m^{\eta-1} - A_t^{\eta-1} \right)}{\tilde{A}(a)^{\eta-1}} \\ &+ \frac{\frac{\gamma-\eta+1}{(\eta-1)(1-\nu)} a \left(A_m^{\eta-1} - A_t^{\eta-1} \right) - A_t^{\eta-1}}{\tilde{A}(a)^{\eta-1}} \frac{L_a}{L - L_a}, \end{aligned}$$

where $L = 1$ is the mass of workers and L_a is the fraction of workers engaging in adoption good production.

4 Identification

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

To keep things as simple as possible, we provide our constructive argument for the case where the intermediate input elasticity is the same for the modern and the traditional technology, i.e., $\nu_t = \nu_m = \nu$, and the adoption goods production only uses intermediate input ($\gamma = 1$). In addition, following the common practice in the literature, we assume that firm-level productivity follows a Pareto distribution with a tail parameter ζ and its c.d.f. denoted by F . We set the elasticity of substitution across differentiated goods η from outside of the model, using estimates common in the literature on demand or profit share estimation. We also assume that there is no distortion $\xi = 0$. However, our identification strategy also holds without these parameter restrictions.

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

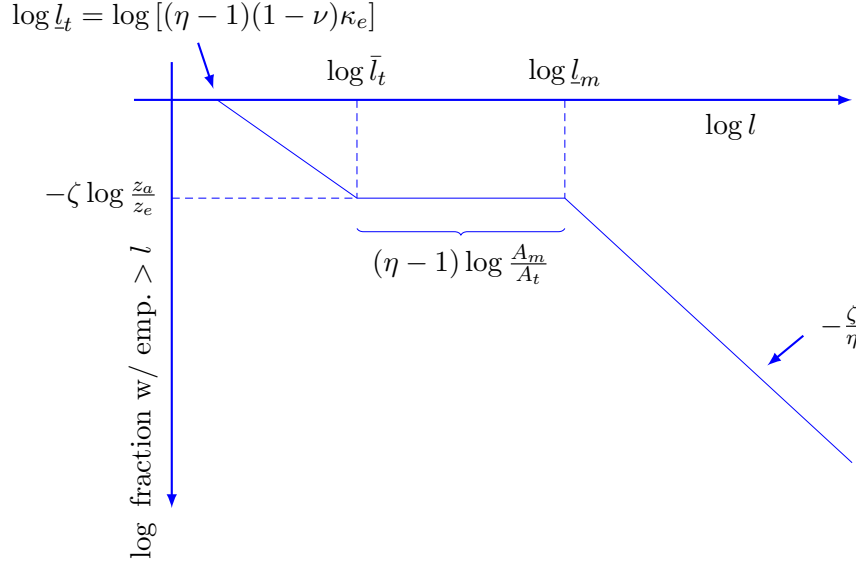
We begin by showing that, for a given η , we can identify the intermediate input elasticity ν directly from the intermediate input share in the data. Straightforward calculations show:

$$\nu = \frac{\eta}{\eta - 1} \times \text{intermediate input share.} \quad (17)$$

The other four parameters are identified from the establishment size distribution in the data. In particular, we rely on the implications of the theory for the relationship between the log of employment ($\log l$) and the log of the fraction of establishments with size larger than l as illustrated in Figure 1. We hereafter refer to this relationship simply as the log-log relationship. Our identification strategy relies on the three thresholds in the figure, which are guaranteed to 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 $\underline{l}_m \equiv l_m(z_a)$.

Figure 1: Identification from the establishment size distribution



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

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

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 :

$$\underline{l}_t = [(\eta - 1)(1 - \nu)\kappa_e] \Rightarrow \kappa_e = \frac{\underline{l}_t}{(\eta - 1)(1 - \nu)}. \quad (18)$$

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.⁷ 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 the productivity of the

⁷This is akin to the concept of missing middle in Tybout (2000).

traditional technology A_t :

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

Finally, to identify the adoption cost κ_a we use equation (4) with $\nu_m = \nu_t$ and $\gamma = 1$:

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

where we use the fact that the fraction of adopters equals $G(\bar{l}_t) = (z_a/z_e)^{-\zeta}$ and that the relative price $\frac{P}{w}$ is the following function of the other identified parameters:

$$\begin{aligned} \frac{P}{w} = & z_e^{\frac{\zeta+1-\eta}{(\eta-1)(1-\nu)}} \left[\frac{\eta}{\eta-1} \left(\frac{w}{1-\nu} \right)^{1-\nu} \left(\frac{1}{\nu} \right)^{\nu} \right]^{\frac{1}{1-\nu}} \\ & \times \left(\left(\frac{\zeta}{\zeta+1-\eta} \right)^{\frac{1}{\eta-1}} \left[A_t^{\eta-1} + (A_m^{\eta-1} - A_t^{\eta-1}) \left(\frac{z_a}{z_e} \right)^{-(\zeta+1-\eta)} \right]^{\frac{1}{\eta-1}} \right)^{-\frac{1}{1-\nu}}. \end{aligned}$$

5 Calibration

We use data from the US and India to calibrate the parameters of the model. We think of the US as an economy that is largely distortion-free ($\xi = 0$) and therefore informative about the tail parameter of the firm productivity distribution ζ , among others. India is a large developing economy with relatively good data availability and is informative about the extent of idiosyncratic distortions ξ in particular.

For our benchmark model without the simplifying assumptions of Section 3 or 4, the following 11 parameters need to be calibrated: the elasticity of substitution among differentiated goods η ; the share of intermediate input in the production of adoption goods 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 distortion elasticity ξ and the scale parameter τ .

Six of the eleven parameters are assumed to be the same for the US and India: η , γ , ν_m , ν_t , A_m and ζ . Four of these six are fixed outside of the model. We maintain the normalization of $A_m = 1$. We set $\eta = 3$, which is considered to be on the lower side, as discussed in [Hsieh and Klenow \(2009\)](#). As in Section 4, we set $\gamma = 1$, i.e., the only input of the adoption goods

production is the intermediate goods. 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 multiple equilibria more likely as explained in Section 3, but the conclusions from our quantitative analysis do not rest on these assumptions. In Section 6.3, we show the result with $\gamma = 0$ —i.e., labor is the only input of adoption goods production. In Appendix A, we present a sensitivity analysis with different values of η and ν_t .

One of the remaining common parameters, ν_m , is then calibrated to match the intermediate input share in the US data, yielding $\nu_m = 0.70$.⁸ The other, the tail parameter of the productivity distribution, is calibrated to match the tail of the establishment size distribution for the US in the Census Bureau's 2007 Business Dynamics Statistics (BDS), giving $\zeta = 2.42$.

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

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

$$\frac{\zeta}{\eta(1-\xi)-1} = -\frac{\partial \log(1-G(l))}{\partial \log l}. \quad (19)$$

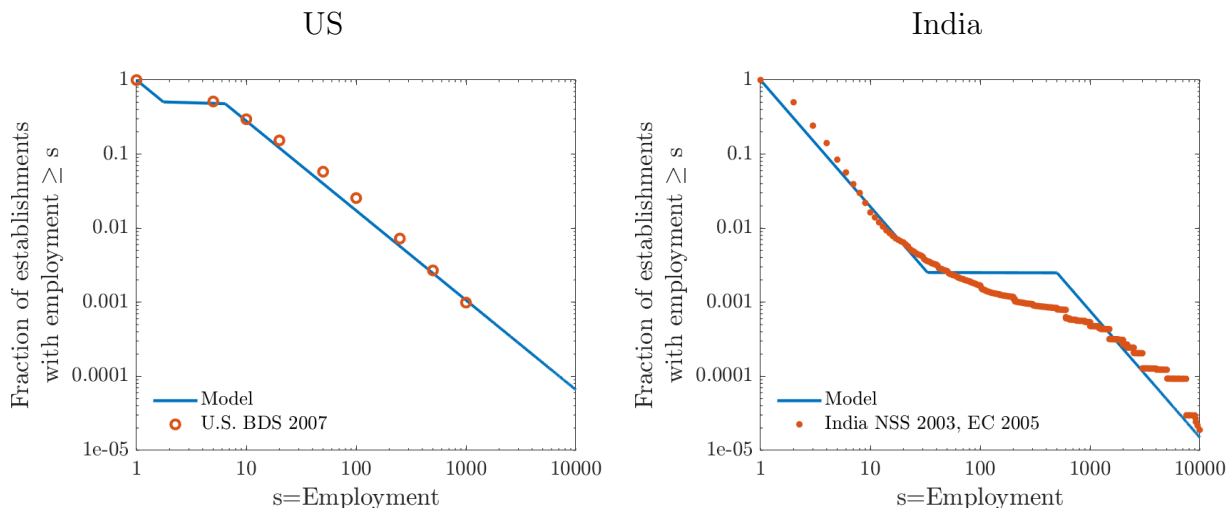
This is essentially what we discussed in Section 4 for the identification of the tail parameter, but allowing for distortions. We obtain $\xi = 0.19$, and $\tau = 2.14$ makes the distortions revenue-neutral for India.

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

⁸The 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 (17), we obtain $\nu_m = 0.70$.

relationship, and 26 points for India. That is, our estimation model is over-identified.

Figure 2: Establishment Size Distribution: Model and Data



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

In Figure 2 we present the log-log plots of the establishment size distribution from the calibrated model for the US and India, together with the empirical log-log relationship. The US data again comes from the 2007 BDS. To produce the figure for India, we use the size distribution of establishments for the entire Indian economy constructed by [Buera et al. \(2020\)](#), who combine data from the Fifth Economic Census in 2005 and the 2003 Survey of Land and Livestock Holdings carried out in 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 2004.⁹

As shown in the left panel, the model calibrated to the US (solid line) generates a flat region that is the size gap between firms using the traditional technology and those using the modern technology, in order to match the concavity of the log-log relationship in the data (circles) for small establishments. The vertical location of the flat region shows that roughly half of all establishments adopt the modern technology. The Indian calibration in the right panel also shows both the traditional and the modern technology in use in equilibrium,

⁹By comparison, the agricultural employment share was only 1.4 percent in 2007 in the US.

separated by a flat region (solid line). However, in India a much larger fraction of firms uses the traditional technology, as evidenced by the fact that less than 1 percent (0.25 percent, to be precise) of establishments are to the right of the flat region in the calibrated economy. The calibrated model captures the conspicuously flat region over intermediate establishment sizes in the Indian data (dots): a missing middle.¹⁰

Table I: Calibrated Parameters

Parameter	US	India
Elasticity of substitution, η		3
Intermediate aggregate share in adoption good production, γ		1
Productivity distribution Pareto tail parameter, ζ		2.42
Modern technology productivity, A_m		1
Modern technology intermediate input elasticity ν_m		0.70
Traditional technology intermediate input elasticity, ν_t		0
Entry cost, κ_e	0.50	0.50
Traditional technology, A_t	0.43	0.07
Adoption cost, $P\kappa_a$	9.36	809
Degree of distortions, ξ	0	0.19
Distortion scale parameter, τ	0	2.14

In Table I we report the calibrated parameters, for both the US and India. Some remarks are in order. First, the US and India have the same entry cost κ_e , which is identified from the size of the smallest establishment using (18). Because the smallest establishment is of the same size in both countries (one employee) and we assume that η and ν_t are the same for both countries, so is κ_e . However, this does not mean that the entry rate of firms is the same in the two countries. In fact, as shown in Table II, the fraction of firms that enter in India is three times that in the US. Second, the traditional technology parameter A_t is six times as high in the US as in India. Given that both countries have the same productivity level of the modern technology A_m by assumption, the technology gap between the modern and the traditional technology is six times as high in India. Third, the cost of adoption is more than 80 times higher in India. The cost of adoption in India must be higher in order to rationalize the minuscule fraction of firms adopting the modern technology in spite of the large productivity gains from doing so.¹¹ Finally, the Indian calibration exhibits significant

¹⁰This seemingly contradicts the finding in Hsieh and Olken (2014), who do not find evidence of a missing middle in Indian manufacturing. Buera et al. (2020) incorporate information from all sectors of the economy, including agriculture. They find a missing middle within the agricultural sector.

¹¹The 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

idiosyncratic distortions, as given by $\xi = 0.19$. The tail of the establishment size distribution in India is thinner, implying a steeper log-log relationship in the right tail. Our identification scheme matches this empirical pattern with idiosyncratic distortions as shown in (19).¹²

Table II: Statistics from the Calibrated Economy

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

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

This last result is a success for the model. Even though the calibration is based on the difference in the establishment size distribution between the US and India and does not use any information on the income or productivity gap between the two, the model generates a huge income gap. In the Penn World Tables the GDP per worker of India is 6 percent of the US level in 2005, while in the model this figure is 15 percent. That is, the model accounts for 90 percent = $(1 - 0.15)/(1 - 0.06)$ of the US-India income gap.¹³

technology, financial constraints, and bureaucratic or anti-competitive barriers to adoption.

¹²One way to make sense of the magnitude of $\xi = 0.19$ is to set $\xi = 0$ and $\tau = 1$ for India, holding all other parameters fixed. The resulting GDP is 2.2 times the calibrated Indian GDP. This impact on GDP is in line with that in Hsieh and Klenow (2009), where reducing the degree of distortions in India to the US level roughly doubles the GDP if capital accumulated in response to the aggregate TFP gains of 40-60 percent.

¹³We use the output-side real GDP at chained PPPs and the number of persons engaged from the Penn World Tables 9.0.

6 Quantitative Explorations

In this section, we first ask whether our calibrated model features multiple equilibria and, if so, coordination failures explain the income difference between the US and India. We then explore more broadly the role of the various model elements in generating complementarity and amplifying the impact of distortions with or without multiplicity. We also discuss what a big push in distorted economies is.

6.1 Multiple Equilibria and the US-India Income Difference

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

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

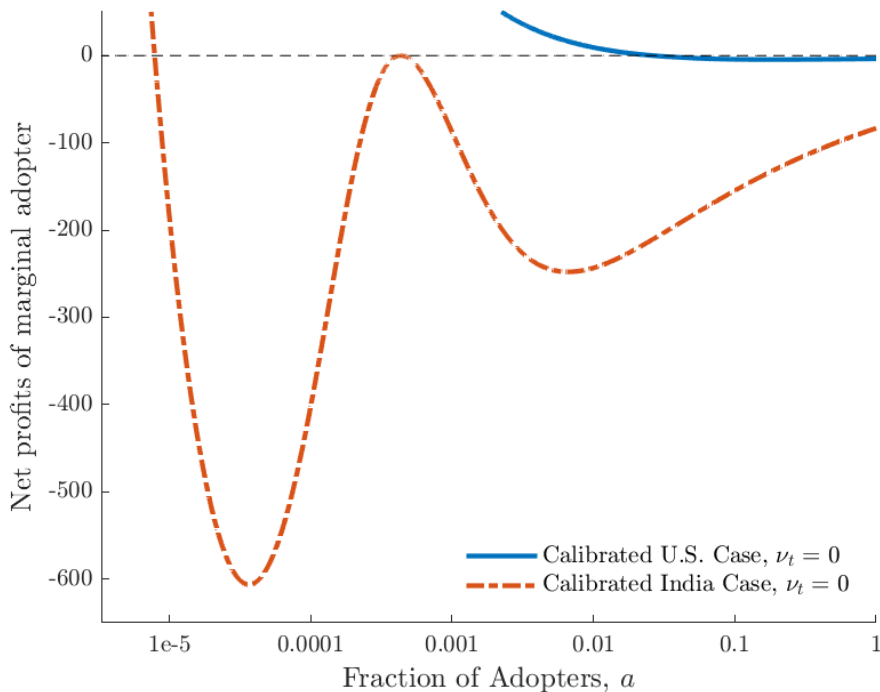
which must be 0 in an equilibrium. Figure 3 shows this object for the US and India. As we vary z_a , the price of the intermediate aggregate and the wage adjust to clear markets.

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

If coordination failures do not account for the income difference between the US and India, then what does? The two countries in our model have different productivity of the traditional technology A_t , adoption costs κ_a , and distortions ξ , all identified only from their establishment size distribution.¹⁴ In Table III, we calculate the contribution of each of these elements to the US-India gap in per-capita consumption. To do so, we compute the hypothetical aggregate consumption of the US by replacing one of the parameters with its

¹⁴The entry cost is also country-specific, but the calibrated κ_e ’s for the US and India coincide. As ξ changes, τ adjusts accordingly to maintain revenue neutrality of the distortions.

Figure 3: Net Gain of the Marginal Adopter, US and India



Note: The figure shows, for the US and India, the gains for the marginal adopter from operating the modern technology, net of adoption costs.

value in the Indian calibration, holding all others constant. This result is in the first column, where we replace κ_a , ξ , and A_t one by one. In the second column, we do the reverse: Starting from the Indian calibration, we replace one of the parameters with its value in the US calibration. All per-capita consumption is normalized by the US level in the benchmark calibration.

Table III: Explaining Consumption Difference

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

The first row shows that the model generates a factor of 7 difference between the US and Indian consumption ($=1/0.14$).¹⁵ Starting from the US calibration, we see that the adoption

¹⁵The gap in consumption is slightly larger than the gap in output, because entry costs and adoption costs

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: Introducing idiosyncratic distortions of Indian proportions ($\xi = 0.19$) to the US economy reduces the consumption by a factor of 2.5 ($=1/0.41$). The last row of the first column shows that, if we replace the traditional technology productivity A_t of the US with the lower value from India, the US consumption actually rises modestly. This is because the very low A_t leads to more adoption of the modern technology.¹⁶

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

To summarize, the model nearly replicates the large income gap between the US and India in the data, without directly targeting the income or productivity level of either country. Coordination failures turn out not to be part of the story despite the existence of multiple equilibria, and adoption costs and distortions explain most of the income gap. Finally, the adoption cost difference has a larger effect when the productivity gap between the modern and the traditional technology is larger.

6.2 Multiplicity, Amplification and Distortions

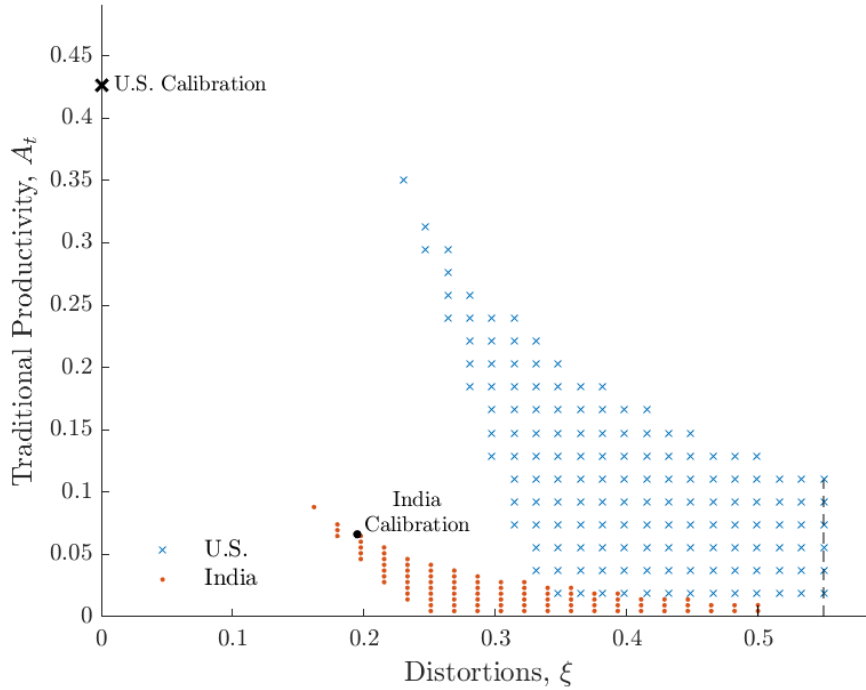
In this section we further explore the role of adoption costs κ_a , relative technology productivity A_m/A_t , and idiosyncratic distortions ξ . We first identify the set of these three parameter values that generates multiple equilibria, holding fixed the other parameters as calibrated. Second, we show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

Region of Multiplicity Figure 4 shows, for a low adoption cost economy (the US, cross) and a high adoption cost economy (India, dot), the combination of the distortion parameter

are a larger fraction of the output in India than in the US.

¹⁶In this case, the number of entrants is nearly halved, but all the entrants adopt the modern technology.

Figure 4: Region of Multiple Equilibria, (ξ, A_t) Space



Note: The figure shows, for the US (cross) and India (dot), the combination of the distortion parameter ξ and the traditional technology productivity A_t for which multiple equilibria exist. The vertical dashed lines in the far right are the upper bound of ξ we considered. The larger cross and dot correspond to the calibrated US and India respectively.

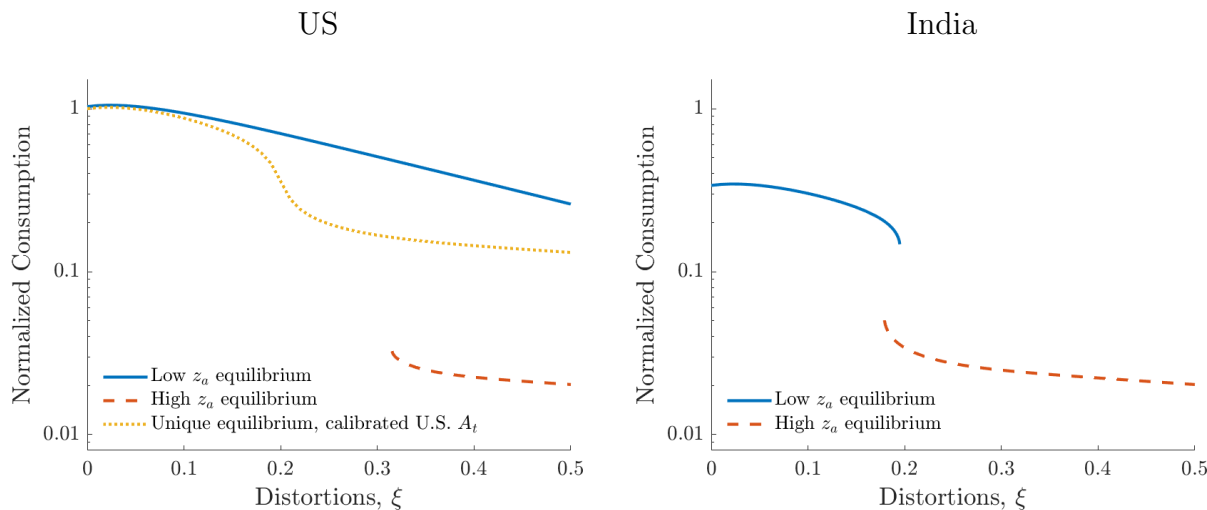
ξ and the traditional productivity A_t that generates multiple equilibria. (A_m for both is normalized to 1.) We hold all other parameters fixed at their respective calibrated values, except that we adjust τ so that distortions are revenue neutral. The larger cross and dot represent the calibrated US and Indian economies, respectively. We see that India is in the region of multiplicity but the US is far from it.

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

(India, dot). The unique equilibrium with A_t close to 0 has a small number of entrants, nearly all of whom adopt the modern technology: With a useless traditional technology, entry also implies adopting the modern technology, which effectively raises the cost of entry and results in few entrants. Finally, the region of multiple equilibria is smaller for the high adoption cost economy. In this case, the model features a unique equilibrium with few adopters in most of the (ξ, A_t) space. However, multiplicity can occur for smaller degrees of idiosyncratic distortions (as low as $\xi = 0.17$) than in the low adoption cost economy.

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

Figure 5: Distortions and Consumption per Capita



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

In Figure 5 we show the equilibrium consumption per capita as we vary the degree of distortions ξ for the US (left panel) and India (right panel). The consumption per-capita in

the vertical axis (log scale) is normalized by the per-capita consumption in the US calibration.

For the US, the equilibrium is unique for any value of ξ (dotted line). There are two notable features. First, the impact of distortions is large, reducing consumption by nearly 90 percent for large values of ξ . Second, for intermediate values of ξ , small changes in the degree of distortions have a highly nonlinear effect on consumption. That is, even without multiplicity, distortions can have an amplified impact. We discuss these features more rigorously in Section 6.3.

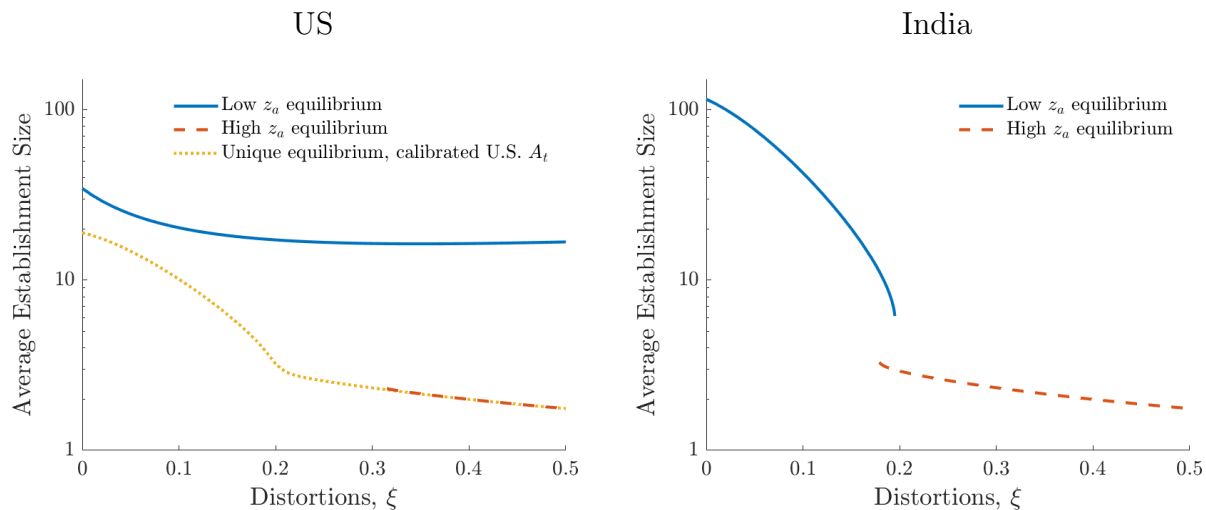
We now turn to the modified US case (i.e., with Indian A_t) in the left panel and the India case in the right panel, both of which exhibit multiplicity. When the distortions are small, the equilibrium is unique in both countries. The solid line is the per-capita consumption in the equilibrium with a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

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

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

Figure 6 illustrates the corresponding effect of distortions on the average size of firms, measured by the number of employees. The vertical axis is in log scale. For the modified US and the India cases, the solid lines trace the average size of firms in the good equilibrium. In this equilibrium, there are fewer entrants but many of them adopt the productive modern

Figure 6: Distortions and Average Firm Size



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

technology, resulting in large firms. As distortions get bigger, there is more entry but less adoption, bringing down the average firm size, which is more pronounced in the high adoption cost economy (right panel). The dashed lines, which appear when distortions are high enough, are the average firm size in the bad equilibrium. In this equilibrium, the number of entrants are large but very few adopt the modern technology, which implies that firms are small on average, in line with equation (16).¹⁷

6.3 Unpacking the Mechanisms

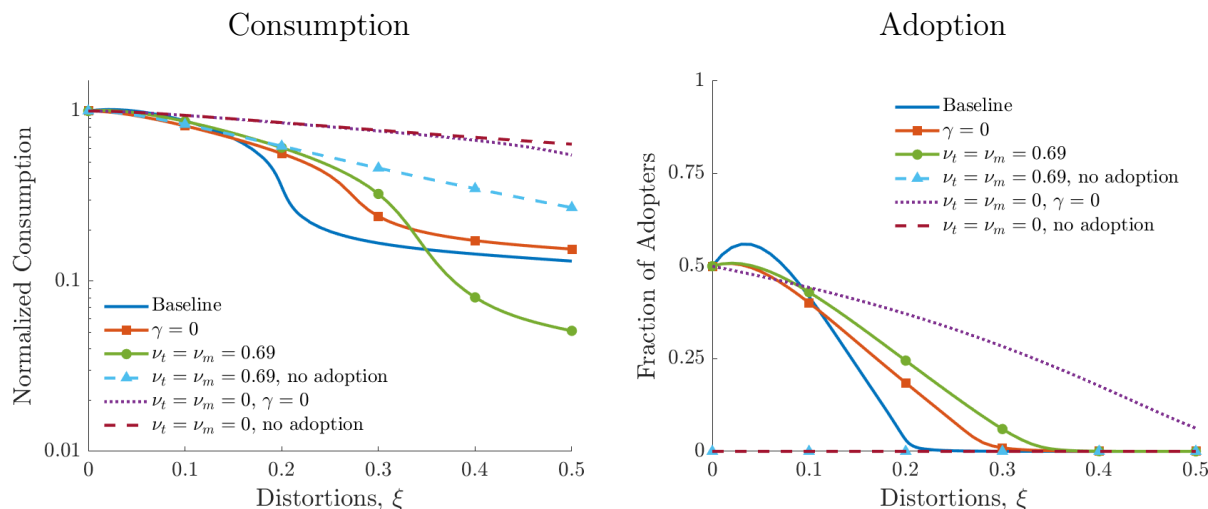
Our model combines several elements whose importance in explaining cross-country income differences has been studied in isolation in the literature. In this section we illustrate the role of each element, in comparison with the findings in the literature. We start with the basic model in the distortion literature that abstracts from technology choices or input-output linkages, as in Restuccia and Rogerson (2008) for example. We then introduce input-output linkages in the form of round-about production but still no technology choice. This is our

¹⁷The modified US case uses India's A_t , which is smaller than its value in the US calibration. As a result, the modified US with no distortion ($\xi = 0$) has more adopters and a larger average firm size than the benchmark calibration (dotted line). The direct negative effect of the traditional technology productivity on the average firm size can be seen in equation (16).

adaptation of Jones (2011). Next, we consider a case with a technology choice, but this time without intermediate inputs or round-about production, which is similar to Bento and Restuccia (2017). The fourth one has a technology choice and round-about production as in our benchmark case, but with the modification that the adoption costs are in units of labor only instead of the intermediate goods. Finally, we consider a case with technology choices, round-about production, adoption costs in goods, except that both the modern and the traditional technologies have the same intermediate input intensity ($\nu_t = \nu_m$), unlike in the benchmark case with $\nu_t \ll \nu_m$.

For each exercise, we re-calibrate the model parameters to the same set of target moments as in the benchmark case and then calculate the effect of idiosyncratic distortions. Below we present the results for the US economy, which does not feature multiple equilibria. The left panel of Figure 7 shows the effect of distortions (ξ) on consumption per capita, and the right panel the effect on the fraction of adopters.

Figure 7: Unpacking the Mechanisms



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

The solid line reproduces the effect of distortions in our benchmark economy (the dotted line in the left panel of Figure 5). It confirms that idiosyncratic distortions have a large negative effect on aggregate consumption, which is particularly pronounced around $\xi = 0.2$, close to its value in the Indian calibration ($\xi = 0.19$). Consumption is down by 60 percent and the fraction of adopters collapses to 0 at $\xi = 0.2$.

The first alternative we consider is the basic model without intermediate inputs or

technology adoption ($\nu_t = \nu_m = 0$ and $A_t = A_m = 0.69$, with re-calibration), shown by the dashed line. This specification is the polar opposite of our benchmark model. Consumption falls almost linearly in the semi-log scale and by much less than in the baseline case. At $\xi = 0.2$, consumption goes down by less than 20 percent from its no-distortion level. Even with $\xi = 0.5$, the loss in consumption from distortions is only 30 percent.

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

Next, the dotted line instead introduces technology adoption to the basic model but takes out the round-about production. Consistent with the literature, for example, [Bento and Restuccia \(2017\)](#), introducing the technology adoption by itself makes the effect of distortions on consumption only marginally bigger and only at extreme degrees of distortions (ξ near 0.5): The dotted line and the dashed line are nearly indistinguishable in the left panel, and the reduction in consumption at $\xi = 0.2$ is again less than 20 percent.

The solid line with squares is the modified benchmark case where the adoption costs are in units of labor only, i.e., $\gamma = 0$ instead of $\gamma = 1$. We see that the effect of distortions on consumption is smaller than, yet still comparable to, that in the benchmark, except for intermediate values of ξ between 0.2 and 0.3. The same is true for the impact of distortions on the fraction of adopters in the right panel. At $\xi = 0.2$, consumption is about 50 percent lower than in the no-distortion case (compared to 60 percent in the benchmark), and about 20 percent of active firms adopt the modern technology (compared to nearly 0 in the benchmark). The difference between the $\gamma = 0$ and the $\gamma = 1$ cases shows the quantitative relevance of the feedback effect of adoption on the price of the adoption goods, as discussed in [Section 3.1](#).

Finally, the solid line with circles is the modified benchmark case where the two technologies have the same intermediate input intensities, $\nu_t = \nu_m = 0.69$, instead of $\nu_t \ll \nu_m$. (We reinstate $\gamma = 1$.) The effect of distortions on consumption and technology adoption is more measured until ξ becomes large enough ($\xi > 0.35$). This highlights another feedback mechanism operating in the benchmark model: As more firms adopt, the lower is the price of the intermediate goods relative to labor, and therefore the higher the profitability of the modern technology that uses the intermediate input more intensively. Because this feedback mechanism is absent in this modified model, the negative effect of distortions on adoption and consumption is smaller than in the benchmark. At $\xi = 0.2$, consumption is 40 percent lower than in the no-distortion case, and nearly a quarter of active firms adopt the modern

technology. On the other hand, when distortions are large enough that the fraction of adopters approaches 0 ($\xi > 0.35$), the negative effect on consumption is considerably larger than in the benchmark that has $\nu_t = 0$. This is because the dearth of adopters makes the intermediate input expensive, but the traditional technology in the modified case is still dependent on the intermediate input ($\nu_t = 0.69$), reducing its effective productivity.

Overall, the analysis in this section emphasizes the interaction among our model elements that is more than simply additive. Round-about production, technology adoption, and the nature of the adoption cost (i.e., labor or goods) can jointly explain the vast cross-country income differences at plausible degrees of idiosyncratic distortions, even without resorting to multiplicity or coordination failures.

6.4 Big Push

“Big Push” is the name [Rosenstein-Rodan \(1961\)](#) gave to the idea that a minimum scale of investment is necessary for economic development. The rationales are indivisibilities in the production function, especially social overhead capital, and complementarities across sectors. Under these conditions, individual firms may not find it profitable to industrialize alone, even though all firms are better off industrializing together. It 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.” [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, possibly through encouragement and persuasion. Alternatively, the government can promise compensations for losses from unilateral technology adoption, which will go unclaimed because this policy will lead to the good equilibrium where all firms profit from the adoption.

If all we need is the coordination of firms’ decisions so that they all become better off, why do so many countries still remain unindustrialized and poor? Our framework helps address this question in two ways. First, the heterogeneity across firms implies that not all firms are better off in the good equilibrium. As discussed in our explanation of [Figure 6](#), 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, opens doors for 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 could be a necessary ingredient of successful development policies. As the Indian calibration in Figure 5 shows, if we are in a situation with a degree of idiosyncratic distortions ξ just above 0.2 (unique bad equilibrium), a reform that reduces ξ slightly to the left into the multiplicity region provides a role for policies coordinating the economy into the good equilibrium, discontinuously raising consumption by an order of magnitude. This highlights the complementarity between distortion-reducing reforms and coordination policies. In addition, further distortion reductions lead the economy to the region of unique good equilibrium, obviating the need for coordination altogether. In fact, even in the absence of multiplicity, the effects of distortions get amplified in our framework, as shown by the nonlinear effect in Figure 7 around $\xi = 0.2$.

We view this novel result that small reforms can have a magnified effect with or without multiplicity as the big push in distorted economies. In this view, the role of the government is to reduce distortions, identifying and exploiting the “big push region,” where the returns to economic reforms are discontinuously high.

6.5 Big Push and Industrial Policy

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

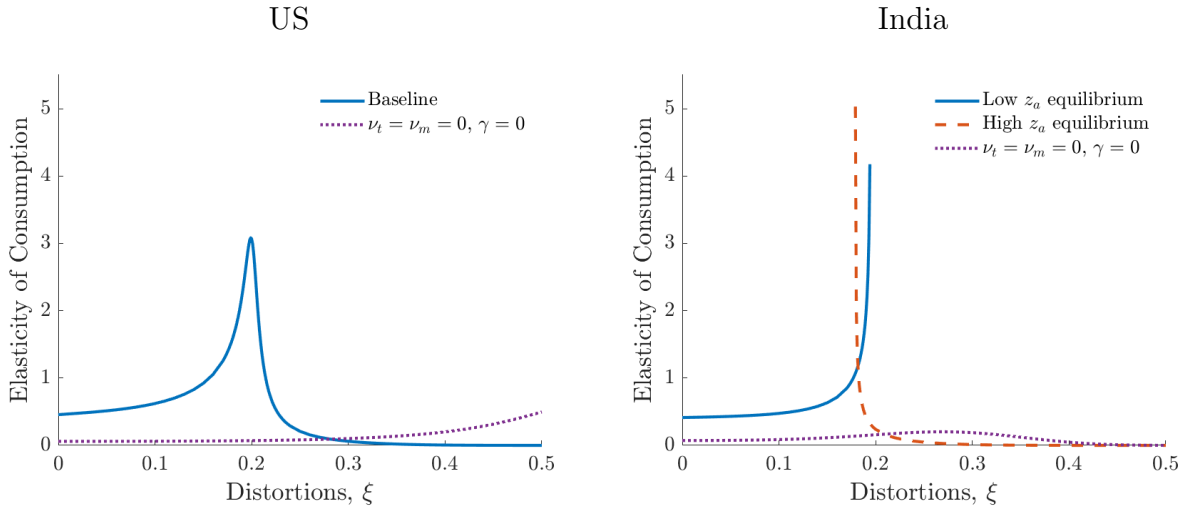
Here we study such a policy in our model: subsidies for technology adoption. We show that the success of this policy hinges on the extent of distortions in the economy: In the big push region, the policy propels development, but less so in other reasons. Furthermore, we consider the problem of a constrained planner that chooses the optimal technology adoption taking as given the distortions and the set of active firms. It shows that the effect of such an optimal adoption subsidy policy can be comparable in size to, but still smaller than, the

effect of eliminating distortions, even in the big push region.

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. The profit of a potential entrant with productivity z presented in 2 is now:

$$\pi(z) = \max_{\text{inactive, active}} \left\{ 0, \max_{t,m} \{ \pi_t^o(z), \pi_m^o(z) - (1-s)P_a\kappa_a \} - w\kappa_e \right\}.$$

Figure 8: Elasticity of Aggregate Consumption to Adoption Subsidy

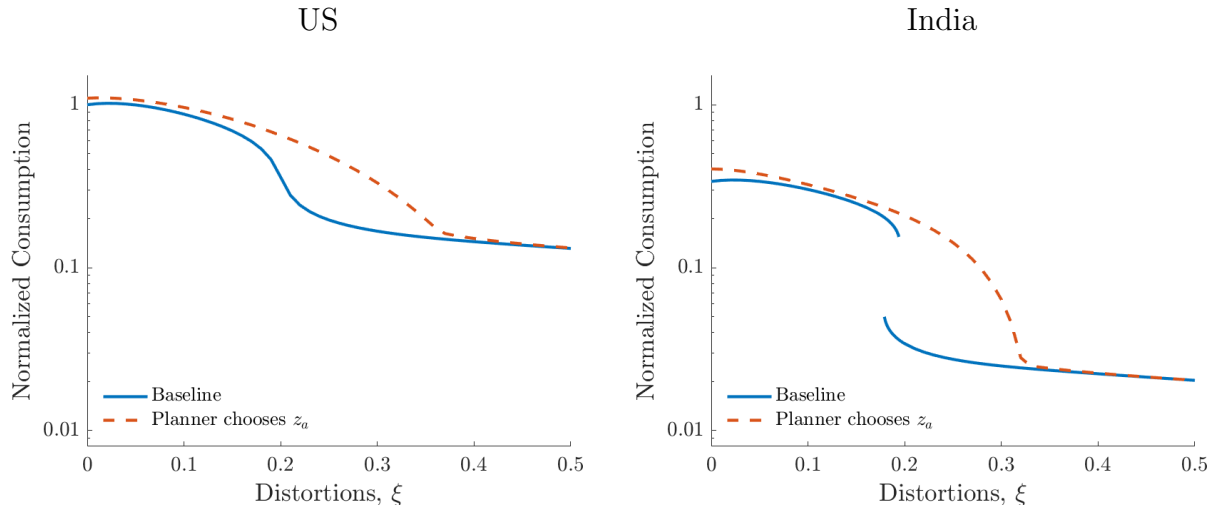


Notes: The elasticity of aggregate consumption to adoption subsidy for the US and India calibration, as the distortion parameter ξ goes from 0 to 0.5. For the Indian case, the solid line is the high-adoption (low adoption threshold z_a) equilibrium and the dashed line is the low-adoption (high adoption threshold z_a) equilibrium. The dotted lines are for the simple model with no intermediate input in production or adoption costs.

We begin by calculating the elasticity of aggregate consumption to the subsidy for both the US and India calibration, as we vary the degree of distortions ξ . For comparison, we also compute the elasticity for a simpler model without the round-about production ($\nu_m = \nu_t = 0$) and with the adoption costs in labor only ($\gamma = 0$). As explained in Section 3, this shuts down the feedback channels that are responsible for amplification and multiplication. Figure 8 shows two noteworthy results. First, in our benchmark model with the feedback effects, the elasticity of aggregate consumption to the subsidy is high in the big push region, especially in the multiplicity region for India, but relatively low outside of it. This can potentially rationalize why some industrial policies succeed but others fail: It is a matter of whether the distortions in the economy places it in or outside the big push region. Second, the elasticity is uniformly low in the version with no feedback effect and there is no big push

region, implying that the existence of the feedback channels is a necessary condition for any successful industrial policy.¹⁸

Figure 9: Constrained Planner Allocation



Notes: The figure presents the aggregate consumption in the constrained planner allocation for the US and India calibration (dashed line). The planner chooses the adoption threshold z_a , taking as given the degree of distortions and the set of active firms. The equilibrium consumption levels are shown with solids lines. Consumption is normalized by the equilibrium consumption in the undistorted US economy and is in log scale.

We now solve for the optimal adoption subsidy s , by solving the problem of a constrained planner, who chooses the adoption threshold z_a taking as given the distortions and the set of active firms.¹⁹ The aggregate consumption from this constrained planner allocation is shown with dashed lines in Figure 9, while the solid lines reproduce the equilibrium outcomes from Figure 5. Consistent with the elasticity results above, the largest gains from influencing firms' adoption decisions are in the big push region, where the dashed and the solid lines are farthest apart.

Furthermore, the India case attests to the power of coordination achieved by the policy: Although the adoption subsidy has small effects on consumption for ξ greater than 0.2 (Figure 8), the planner can generate massive consumption gains by coordinating the economy away from the low adoption equilibrium to the high adoption equilibrium, as long as the degree of distortions is not too high.

¹⁸The still significant elasticity with no distortions ($\xi = 0$) is the amplification characterized in Section 3.2.2.

¹⁹We also worked out the case in which the planner chooses both the entry and the adoption thresholds. The entry margin is found to play a minor role, except at high degrees of distortions.

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

7 Concluding Remarks

This paper provides answers to the following three questions: (i) Can economic development be explained solely by coordination failures? (ii) Can economic development be a story of coordination failures and distortions? and (iii) Are there large non-linear effects of distortions and policies even without multiplicity, which can explain the large income differences across countries? We find that the US calibration gives a unique equilibrium and that the calibrated Indian economy is in the multiplicity region but is in the good equilibrium, which leads to a negative answer to the first question. We find that small changes in idiosyncratic distortions can move the economy in and out of the region of multiplicity, resulting in discontinuously large changes in the aggregate output. More important, even without multiplicity, the feedback channels in our model creates big push regions, where small changes in distortions and policies have disproportionate effects. This is the affirmative answer to the second and the third question.

This paper presents a framework for both theoretical and quantitative analysis of the role of coordination failures and distortions in economic development. The framework is multi-layered but still allows for a sharp characterization of the role of the various model elements that generate amplification and multiplicity and for a transparent identification.

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 moderately 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. Technical report, Washington University in St. Louis.
- 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? *The 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. *The Quarterly Journal of Economics*, 103(3):441–463.
- 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. *The Quarterly Journal of Economics*, 124(4):1403–1448.
- Hsieh, C.-T. and Klenow, P. J. (2014). The life cycle of plants in India and Mexico. *The Quarterly Journal of Economics*, 129(3):1035–1084.
- Hsieh, C.-T. and Olken, B. A. (2014). The missing “missing middle”. *Journal of Economic Perspectives*, 28(3):89–108.
- 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.
- Kline, P. and Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *The Quarterly Journal of Economics*, 129(1):275–331.
- Krueger, A. O. (1997). Trade policy and economic development: How we learn. *The American Economic Review*, 87(1):1–22.
- Krugman, P. (1992). Toward a counter-counterrevolution in development theory. *The 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. *The 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. *The 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. *The Economic Journal*, 53(210/211):202–211.
- Rosenstein-Rodan, P. N. (1961). Notes on the theory of the ‘big push’. In Ellis, H. S., editor, *Economic Development for Latin America*, International Economic Association Series, pages 57–81. Palgrave Macmillan, London.
- Schaal, E. and Taschereau-Dumouchel, M. (2019). Coordinating business cycles. Technical report, Centre de Recerca en Economia Internacional.
- Tybout, J. R. (2000). Manufacturing firms in developing countries: How well do they do, and why? *Journal of Economic Literature*, 38(1):11–44.
- Valentinyi, A., Herrendorf, B., and Waldmann, R. (2000). Ruling out multiplicity and indeterminacy: The role of heterogeneity. *Review of Economic Studies*, 67:295–307.
- World Bank (2005). *Economic Growth in the 1990s: Learning from a Decade of Reform*. The World Bank.
- Yeaple, S. R. (2005). A simple model of firm heterogeneity, international trade, and wages. *Journal of International Economics*, 65(1):1–20.

Appendix

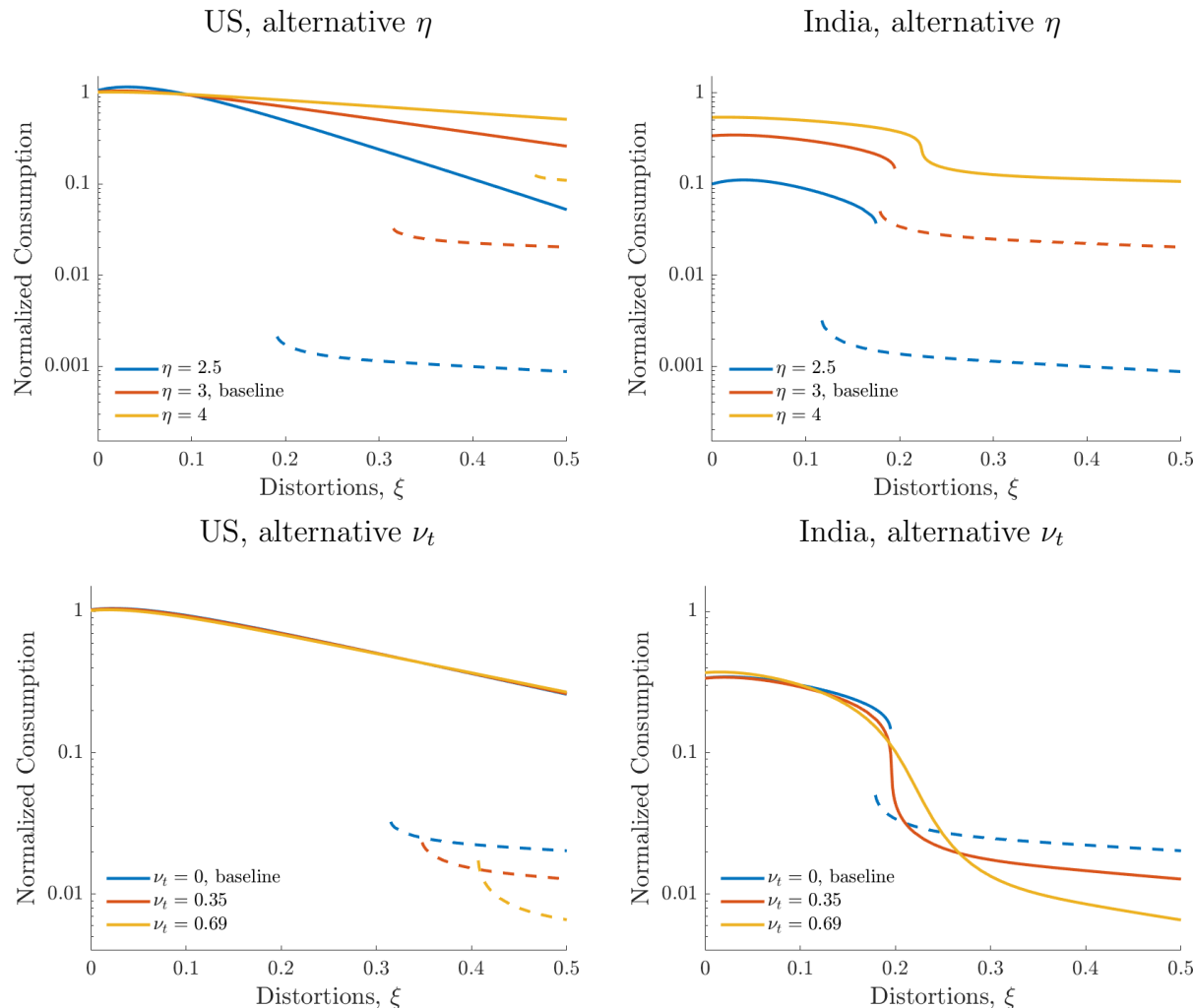
A Sensitivity Analysis

Because the literature lacks precise estimates of the elasticity of substitution across intermediate goods η and the intermediate input share for the traditional technology ν_t , we fixed $\eta = 3$ and $\nu_t = 0$ outside of the model. While $\eta = 3$ falls within the standard range in the literature, there is no available estimate of ν_t , beyond the fact that it is smaller than the intermediate input share for the modern technology ν_m (Chenery et al., 1986; Blaum et al., 2018).

In this section we explore the role of η and ν_t . As we vary η or ν_t , we re-calibrate the model and then re-do the exercises that produced Figure 5. We consider two values of η , 2.5 and 4, one on either side of the baseline value $\eta = 3$. We also consider two cases for ν_t : one where $\nu_t = 0.35$ is larger than the baseline value of zero but smaller than ν_m , and the

other where ν_t and ν_m are both 0.69, which is the highest value for ν_t given the restriction $\nu_t \leq \nu_m$ and the overall intermediate input share. Figure 10 shows consumption as we vary the idiosyncratic distortion parameter ξ on the horizontal axis. The two top panels are the cases with different η 's for the US and India. The two bottom panels are for alternative ν_t 's. In all cases, consumption is normalized by the US consumption in the equilibrium with no distortions.

Figure 10: Sensitivity of GDP and Equilibrium Multiplicity to η and ν_t



Notes: Equilibrium consumption per capita of the US and India as the distortion parameter ξ goes from 0 to 0.5. GDP per capita is normalized by the US consumption (no distortion) and in log scale.

The top panels of Figure 10 show that a smaller elasticity of substitution η increases the consumption difference between good and bad equilibria and widens the set of the distortion parameter ξ that generates multiple equilibria. This is consistent with the analysis

in Section 3.1: A lower η makes goods less substitutable and firms' adoption decisions more complementary. In addition, the income gap between the US and India is larger with a smaller η . We draw the conclusion that a small elasticity of substitution across goods are conducive to explaining huge income gaps across countries, with or without multiplicity.

The bottom panels show that a lower intermediate input elasticity of the traditional technology ν_t has two effects. On the one hand, a lower ν_t enlarges the set of ξ 's generating multiple equilibria. This is consistent with the discussion in Section 3.4. On the other hand, a lower ν_t compresses the consumption gap between good and bad equilibria. Intuitively, when ν_t is small, it is less costly to use the traditional technology in a world where few firms adopt the modern technology and the intermediate aggregate is expensive. The two effects run in opposite directions when it comes to explaining cross-country income differences.

Another robust result in the plots for India is that the model features a very non-linear effect of distortions, either through or independently of equilibrium multiplicity. The non-linearities are more pronounced when η is low or when ν_t is high.