The Micro and Macro Productivity of Nations*

Stephen Ayerst
IMF

Duc Nguyen
University of Toronto

Diego Restuccia
University of Toronto
and NBER

February 2024

Abstract

We examine the disparity in aggregate productivity across nations using cross-country firm-level panel data and a quantitative model featuring production heterogeneity with distortions and entry and operation decisions of firms. Empirically, we find that less developed countries feature higher distortions and larger dispersion in firm-level productivity, mostly resulting from the prevalence of unproductive firms compared to developed countries. Quantitatively, variation in the distribution of firm-level productivity across countries accounts for about two-thirds of international productivity differences. Variation in static misallocation also plays a quantitatively important role, albeit smaller. Measured differences in correlated distortions across countries, the elasticity of distortions with respect to firm productivity, generates the bulk of the empirical patterns.

Keywords: Firms, productivity, size, distortions, misallocation.
JEL classification: O11, O14, O4.

*Preliminary and incomplete. All remaining errors are our own. Restuccia gratefully acknowledges the support from the Canada Research Chairs program and the Bank of Canada Fellowship program. The views expressed herein are those of the authors and should not be attributed to the Bank of Canada or its Governing Council, nor the IMF, its Executive Board, or its management. Contact: stephen.b.ayerst@gmail.com, ducm.nguyen@mail.utoronto.ca, and diego.restuccia@utoronto.ca.

1 Introduction

There are large disparities in aggregate productivity across countries which are at the core of international differences in GDP per capita (Klenow and Rodriguez-Clare, 1997; Prescott, 1998; Hall and Jones, 1999). Cross-country differences in aggregate productivity are linked to distortions in the allocation of resources across firms within sectors (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Bartelsman et al., 2013). Whereas the misallocation literature emphasizes the aggregate productivity gains from factor reallocation across a given set of producers, producer-level data also reveal substantial differences in the productivity distribution across countries (Hsieh and Klenow, 2009; Gal, 2013; Andrews et al., 2015). In this paper, we follow Restuccia and Rogerson (2017) in linking observed firm-level TFP distributions to policies and institutions that misallocate resources across firms. In particular, we examine the role of distortions on aggregate productivity across nations using cross-country firm-level panel data and a quantitative model featuring production heterogeneity with distortions and entry and operation decisions of firms.

We construct a panel firm-level financial dataset across countries using ORBIS, collected and standardized by Bureau Van Dijk. We focus on the period 2000-2019. Our final dataset contains 28 countries with an average of 1.3 million firm-year observations and covers the a wide range of the world income distribution. We construct a measure of firm-level total factor productivity (TFP) using data on revenues and inputs. We also construct a firm-level measure of distortions, a model-based measure of idiosyncratic distortions faced by the firm. Using this data we document the following facts: (1) firm-level TFP is more dispersed in less developed countries, (2) larger TFP dispersion in less developed countries arises mostly due to the prevalence of low productivity firms in poor countries, (3) dispersion of idiosyncratic distortions is higher in less developed countries, with relatively flat dispersion in relative input wedges (capital to labor), and (4) distortions are more highly correlated with firm productivity in less developed countries. We also note that average firm size is lower in less developed countries (larger number of firms per capita) (Bento and Restuccia, 2017, 2021).

To address the quantitative role of distortions on factor misallocation and differences in the distribution of firm-level TFP across countries, we develop a model of production heterogeneity with distortions and entry and operation decisions by firms building on Hopenhayn (1992) and Restuccia and Rogerson (2008). The quantitative framework allows for productivity enhancing investment broadly capturing costly activities that firms undertake to improve productivity or in the adoption of more advanced technologies. Production of a homogeneous good takes place in establishments with access to a decreasing returns technology on labor input, but establishments are subject to idiosyncratic distortions and fixed operation costs. Establishments are also subject to a transitory shock to productivity after production decisions are made. New establishments can be created at a cost but the productivity of entering establishments is determined by a costly investment process. In particular, entering establishments draw an idiosyncratic investment ability χ which determines the cost of productivity investment and hence idiosyncratic distortions affect this investment process. We parameterize distortions to feature a systematic component that relates to firm-level productivity (ρ) and a random component drawn from a log normal distribution, which provides an excellent fit of measured distortions in the data within and across countries. We calibrate a distorted benchmark economy to micro and aggregate observations for France. Critical are the parameters of idiosyncratic distortions, investment ability, and operating costs that are jointly restricted to match moments for the French data on measured distortions, dispersion in firm-level TFP and employment, and average firm size.

We first examine the quantitative effect of changes in the TFP elasticity of distortions (ρ) relative to the benchmark economy. Increasing ρ from the calibrated value of 0.525 to 0.90 reduces aggregate output by 77 percent. While this drop is all due to the increase ρ in the model, it is useful to decompose this change in terms of the channels of resource misallocation and the change in the productivity distribution. To this end, we measure the change in the productivity distribution by quantifying the change in aggregate output in the efficient allocation. This change only depends on the change in firm-level productivity since aggregate employment remains constant. We find that efficient output drops by 44 percent, hence this component alone accounts for 40 percent of the drop in aggregate output ($\log 0.56 / \log 0.23$) . The remaining 60 percent is the change in allocative efficiency (factor misallocation). Interestingly, half of this change in allocative efficiency arises because of the interaction between the shift in the productivity distribution and the change in the allocation of resources across firms, an often unappreciated component of the cost of misallocation in survey data. Overall, abstracting from the shift in the productivity distribution misses two-thirds of the aggregate output loss across distorted economies, that is, static misallocation (i.e., the change in resource misallocation associated with a constant productivity distribution) accounts for only one-third of the aggregate output loss. Our results indicate that firm-level TFP dispersion increases from 1.06 in the benchmark economy to 1.57 in the most distorted economy, consistent with our empirical findings of larger dispersion in TFP in poorer countries. We find that our experiments are also consistent with larger dispersion in TFP due to lack of selection at the bottom of the TFP distribution. Similarly, the increase in correlated distortions (ρ) generates a reduction in average firm size consistent with cross-country data (Bento and Restuccia, 2017). Our results also indicate that whereas ρ increases from 0.525 to 0.90, the TFP elasticity of distortions among operating firms increases more narrowly, indicating a sample selection bias in common statistics about cross-country changes in misallocation.

We also perform other experiments allowing for changes in the dispersion of productivity shocks and the random component of distortion in addition to the change in correlated distortions. These additional experiments are motivated by the fact that we observe in the cross-country data more dispersion in log TFP and in employment than in poor countries than what is implied by the change in ρ alone. We report the results of these experiments along with the cross-country data for a set of statistics of interest. While the bulk of the empirical patterns is accounted for by the model with changes in ρ, variations in σ_v and σ_ϵ allow the model to better fit the cross country data in terms of the dispersion of firm-level distortions, TFP, and employment.

Our paper relates closely to the broad literature on production heterogeneity and misallocation (Restuccia and Rogerson, 2008; Guner et al., 2008; Hsieh and Klenow, 2009) and the related literature on producer dynamics, technology adoption, and aggregate productivity (Bhattacharya et al., 2013; Hsieh and Klenow, 2014; Bento and Restuccia, 2017; Ayerst, 2022; Parente and Prescott, 1994; Comin and Mestieri, 2018). We build on this literature by providing a systematic assessment of the role of distortions on aggregate productivity using cross-country producer-level data. Our work follows Restuccia and Rogerson (2017) in connecting factor misallocation with effects on the productivity distribution. In particular, Restuccia and Rogerson (2017) highlight the potential role of selection (the type of producers in operation in a country) and technology (the level of producer productivity that may be affected by the diffusion of technology or productivity enhancing investments) in accounting for aggregate productivity differences across countries. Empirical evidence of this connection include: technology upgrading (Bustos, 2011), export quotas (Khandelwal et al., 2013), trade reform (Pavcnik, 2002), among many others. We also join recent work exploiting Orbis data, such as Andrews et al. (2015) and Alviarez et al. (2023). Andrews et al. (2015) emphasize the growing disparity over time between frontier and laggard firms, whereas we focus on the disparity in the entire productivity distribution across nations. Alviarez et al. (2023) analyze the role of firm-embedded productivity for international income differences, using evidence from Orbis on multinational enterprises. Our work is closely related in that we assess the role of firm-level productivity differences across nations, but we quantify the role of distortions on these differences and aggregate productivity in a general equilibrium model.

The paper proceeds as follows. In the next section, we describe the data and present the main empirical observations from the cross-country data. Section 3 describes the model and characterizes the qualitative role of distortions on firm-level TFP. In section 4, we calibrate a distorted benchmark economy to micro and aggregate data for France and quantify the effect of distortions on the distribution of firm-level TFP and other outcomes. We conclude in section 5.

2 Stylized Facts

2.1 Data

We use firm-level financial data from Orbis collected and standardized by Bureau Van Dijk as the main dataset for our analysis. Given that our goal is to characterize cross-country facts on productivity and misallocation, we focus on assembling available data for as many countries as possible in our final dataset. We restrict to countries where at least 5,000 observations are available with sufficient data to construct measures of firm-level productivity. We focus on the period from 2000 to 2019 since earlier periods tend to have fewer observations in many countries and later periods coincide with the COVID-19 pandemic that may affect cross-firm and cross-country statistics.

Within countries, we restrict to firms in the manufacturing sector and drop firm-year observations that are missing sufficient information to construct productivity, are inactive, or are duplicate observations. We trim the remaining variables for extreme values based on output at the top and bottom 0.1%, employment greater than 100,000 workers, we drop the bottom 1% of firms based on labor share or firms where the wage bill is greater than revenue or value added. We correct employment for firm-year observations that are likely incorrectly reported by replacing employment at the top and bottom 1% of firms based on the wage bill-per-employee with the wage bill-implied employment. We also trim observations by the top and bottom 2% of the productivity and wedge (described below) distribution in each year to limit the influence of outliers. Appendix A provides a detailed description of the data and cleaning procedure.

Our final dataset contains data on 28 countries with an average of 1.3 million firm-year observations. This ranges widely across countries, with around 9,000 observations in Netherlands and around 4.5 million observations in France and Spain. The dataset covers a wide range of the world income distribution with India, Vietnam, and Morocco among the low income countries and France, Germany, and Austria at the high end of the income distribution.

2.2 Variable Construction

Our goal is to use the data to describe the distribution of firm-level productivity and misallocation. We construct two variables that measure firm-level labor productivity and distortion. We refer to the latter as the firm's wedge since it is a model-based measure of the difference between the firm's realized market allocation and the hypothetical first-best allocation, in which wedges are equalized across firms. In this regard, the measure is the same as the marginal product of factor inputs in Hsieh and Klenow (2009). We derive model-based measures of labor productivity and the output wedge as:

$$\mathrm{TFP}_{i,t} = \frac{y_{i,t}}{n_{i,t}^{\gamma}}, \qquad \qquad wedge_{i,t} = \frac{y_{i,t}}{n_{i,t}}. \tag{1}$$

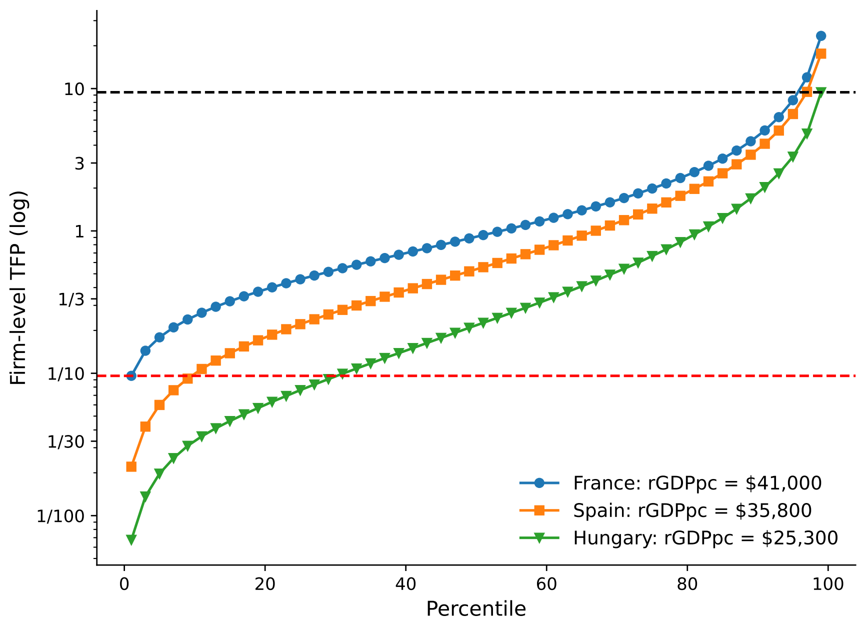
We construct measures of output y and employment n. We measure output as the firm operating revenue, and sales when operating revenue is unreported. We do not use value added because we find that material costs are not widely reported outside of Europe and this limits the final distribution of countries. Employment is measured as the number of employees hired by the firm. In cases where the number of employees is unavailable, we back out a measure using the wage bill of the firm and a constructed average wage rate for that firm's country sector (four-digit SIC) year.

Appendix B reports the robustness of the main results to alternative model-implied measures of productivity and wedges in equation (1). In particular, we show that the main cross-country observations hold if we construct total factor productivity that adjusts for capital inputs, value added as the measure of firm output, if we use a constant elasticity of substitution model as in Hsieh and Klenow (2009), or if we weight observations on the relative share of firms using national statistics data.

2.3 Cross-Country Productivity Distribution

We start by looking at how the firm productivity distribution varies across countries at different stages of development. For illustration, we compare the productivity distributions of three European countries in 2005 that differ in terms of the level of development. Figure 1 reports the average productivity of firms within a percentile of the productivity distribution and does not detrend productivity such that the levels of productivity are comparable across countries, where we note that all values are reported in US dollars.

Figure 1: Productivity Distribution of Operating Firms in 2005

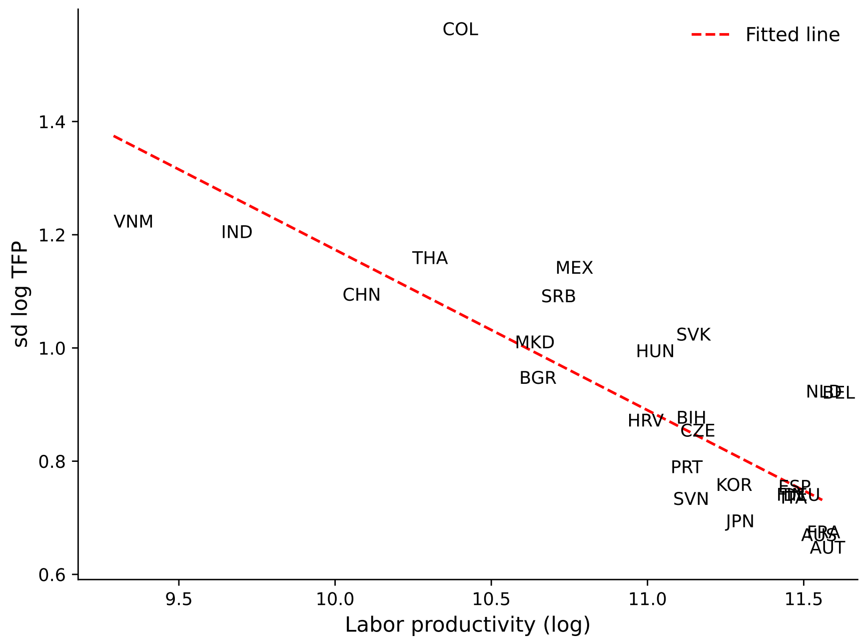


Notes: For ease of illustration the figure reports values for 50 percentile points of the firm-level TFP distribution, from percentile one (p1) to percentile 99 (p99).

Figure 1 highlights several differences between the countries. First, the higher income country such as France appears to have less dispersed productivity. Second, the productivity distributions in other countries (Spain and Hungary) appear to "fan out" from the top end of the distribution. The most productive firms in Spain and France have relatively similar productivity but the productivity gap at the bottom percentile of firms is much larger. For instance, more than 30 percent of firms in Hungary have lower TFP than the p1 firm in France, whereas slightly more than 5 percent of firms in France have higher TFP than the p99 firm in Hungary. Third, there is more of a level shift in the entire distribution in Hungary relative to France as compared to that in Spain.

Next, we use wider ranging cross-country data to draw broader conclusions on these observations. In the comparisons that follow, we detrend productivity and wedges by regressing each variable on country by year by sector fixed effects. In this regard, we are comparing the distribution of relative firm productivities excluding the level, unlike Figure 1.

Figure 2: Cross-Country Productivity Dispersion

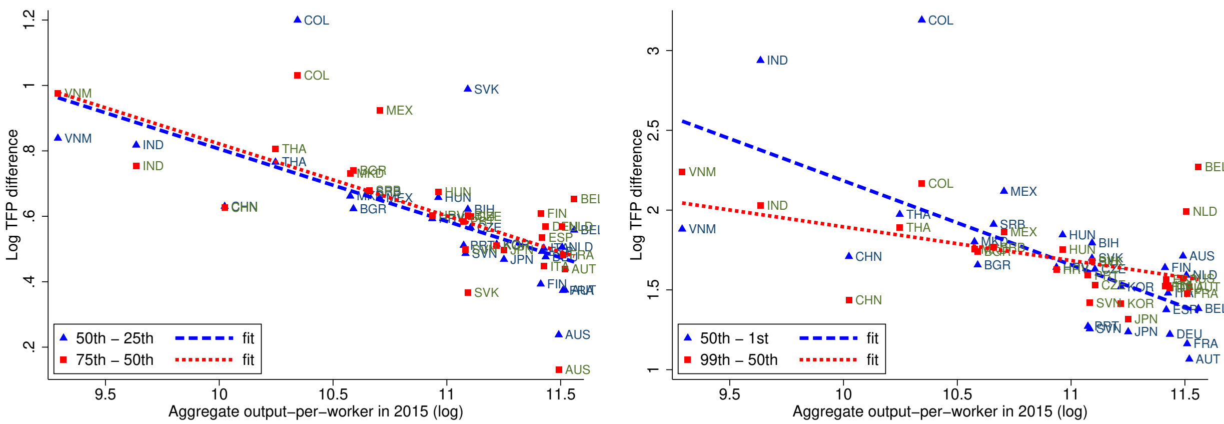


Notes: Productivity dispersion is measured by the standard deviation of log TFP across firms in each country. Each observation is the estimated value for the indicated country. Aggregate labor productivity in logs.

Figure 2 reports the standard deviation of firm-level productivity in countries against development, measured by aggregate labor productivity. Data is pooled at the country level, noting that measures of productivity and wedges are demeaned in each sector-year implying that the standard deviation does not capture between year differences. Figure 2 shows that productivity tends to be more dispersed in less productive economies. This is also confirmed by similar patterns in other measures of dispersion, such as the inter-quartile or inter-decline range, that place less weight on outliers.

Figure 3 reports the comparison of firms at different percentiles of the productivity distribution. Panel (a) reports the difference between the 75th percentile and the 50th percentile and between the 50th percentile and the 25th percentile. Panel (b) documents the difference between the 99th percentile and the 50th percentile and between the 50th percentile and the 1st percentile. Unlike with the standard deviation, these measures provide information on the shape of the overall distribution of productivity and whether, in general, the productivity distribution fans out or shifts down in countries at lower levels of development.

Figure 3: Comparison of Productivity Distribution



(a) 75th to 25th percentile

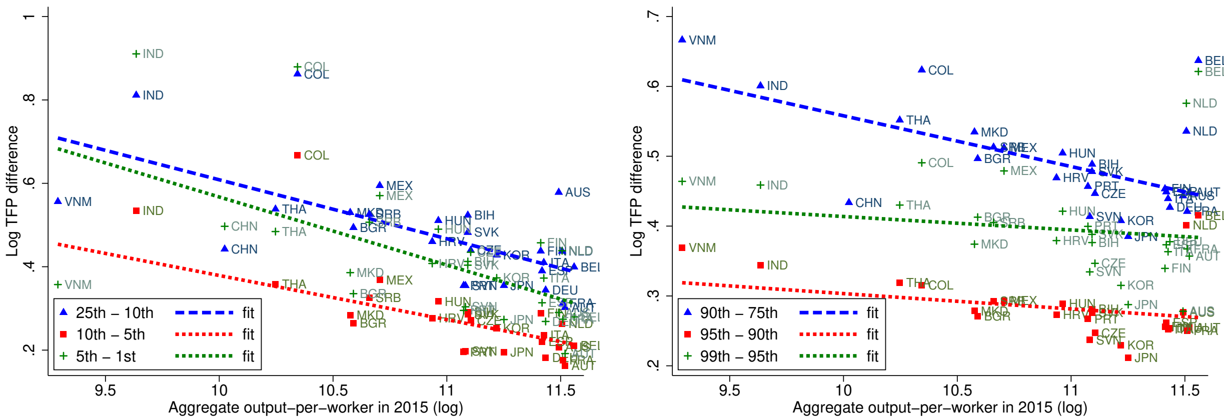
(b) 99th to 1st percentile

Notes: Each observation is the estimated value for the indicated country.

Figure 3 corroborates the result documented earlier that the productivity distribution tends to narrow in more developed countries. This is much stronger in the bottom of the distribution than in the top of the distribution. That is, the narrowing productivity distribution is driven much more by an improvement in the relative position of less productive firms compared to the median firm rather than the median firm relative to the most productive firms. This is consistent with the fanning out pattern as the level shift would imply a flat profile throughout the distribution. Panel (b) shows that this is stronger at more extremes. The gap between the 99th and 50th percentile firms are almost flat across the distribution whereas the gap between the 50th and 1st percentile firms falls by half over the range of countries in the sample.

Figure 3 highlights that the gap between the top end of the distribution and the median firm experiences less differences across countries. Figure 4 extends this fact by decomposing the changes in the 50th to 1st percentile (Panel a) and the 99th and 50th percentile (Panel b).

Figure 4: Comparison of the Top and Bottom Productivity Distribution



(a) Bottom of distribution

(b) Top of distribution

Notes: Each observation is the estimated value for the indicated country.

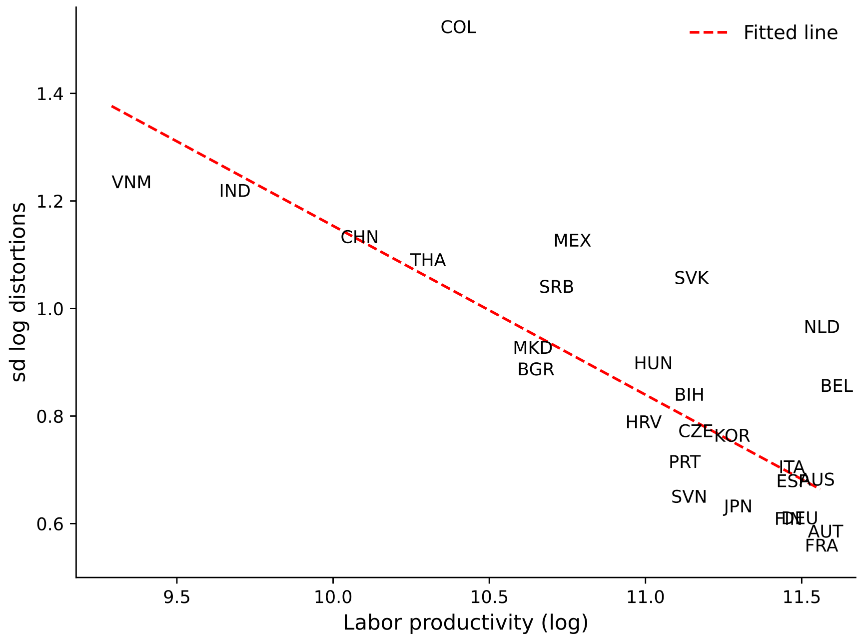
The figure highlights two main results. First, the decline in the the gap between the 50th and 1st percentile is relatively stable throughout the distribution. Second, at higher percentiles of the distribution the narrowing of the distribution reverses. In particular, the gap between the 99th and 95th percentile firms increases in more developed countries.

2.4 Cross-Country Wedge Distribution

We next use the data to examine how the wedge distribution varies across countries at different levels of development. We focus on two commonly reported moments: the standard deviation of wedges and the elasticity of wedges with respect to measures of productivity.

Figure 5 reports the standard deviation in the output, labor, and capital wedges in the data. The figure shows that lower income countries tend to have more dispersed wedges. Figure 6 reports the relationship between firm-level wedges and productivity. Panel (a) reports the elasticity of wedges with respect to firm-level labor productivity. Panel (b) reports the elasticity of employment with respect to firm-level labor productivity. In both cases,

Figure 5: Cross-Country Dispersion in Distortions



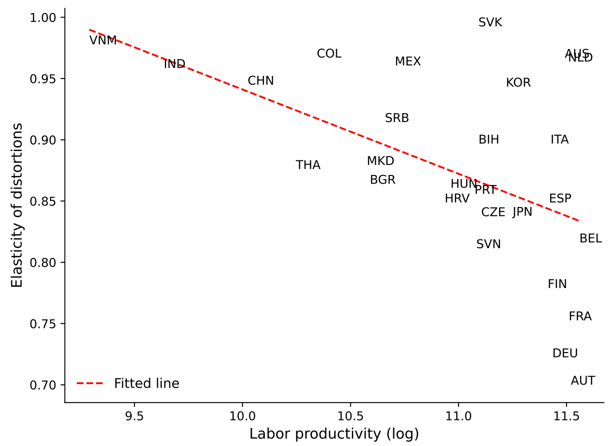
Notes: Dispersion in distortions is measured by the standard deviation of log wedge across firms in each country. Each observation is the estimated value for the indicated country. Aggregate labor productivity in logs.

lower income countries tend to be more distorted. In Appendix B we show that the negative relationship between the elasticity of distortions and firm-level productivity is robust to alternative models, variable constructions, and weighting observation by firm shares. However, the magnitude of these elasticities is sensitive to these choices. For example, we find that using the Hsieh and Klenow (2009) setup results in elasticities that are between 0.3 and 0.6, around half of the values in Figure 6.

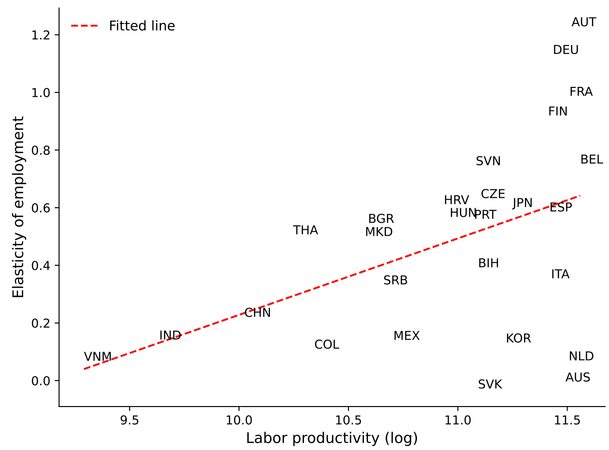
3 Model

We develop an otherwise standard model of production heterogeneity with distortions and entry and operation decisions by firms building on Hopenhayn (1992) and Restuccia and Rogerson (2008). We extend the framework to allow for productivity enhancing investment and focus on a stationary competitive equilibrium of the model.

Figure 6: Cross-Country Elasticity of Distortions



(a) Elasticity of distortions



(b) Elasticity of employment

Notes: Elasticity of distortions measured by the slope coefficient of a regression between log wedge and log TFP in each country. Aggregate labor productivity in logs.

3.1 Economic Environment

Technologies. At each date, a homogeneous good is produced by firms indexed by i. Firms have access to a decreasing-return-to-scale technology,

$$y_i = v_i {z_i}^{1-\gamma} {n_i}^{\gamma}, \quad \gamma \in (0, 1),$$

where y_i is output, n_i is the labor input, and $v_i {z_i}^{1-\gamma}$ is the firm total factor productivity. The term $z_i^{1-\gamma}$ is a permanent component of total factor productivity which is the result of a firm's investment decision while v_i is a transitory component of total factor productivity with $\mathbb{E}v_i = 1$ that is drawn each period from iid cdf distribution $H(v)$ after production decisions are made.

To attain productivity z_i, a firm incurs a convex investment cost of $\psi \frac{z_i^{\phi}}{\chi_i}$, where $\phi > 1$ and χ_i is a firm-specific innovation ability, drawn from an iid cdf distribution $G(\chi)$. Firms are subject to an operating fixed cost c_f per period in units of labor.

Entry and exit. Firms exit at an exogenous rate λ every period. Entering firms are subject to an entry cost c_e units of labor. Firms draw the firm-specific innovation ability χ_i after paying the entry cost.

Households. There is a representation household of measure one with standard preferences on consumption $u(C) = \log(C)$. The household is endowed with one unit of productive time each period that is supplied inelastically to the market.

3.2 Market Structure

Firms face idiosyncratic distortions which for simplicity we model as proportional revenue taxes τ_i. Following Bento and Restuccia (2017) and Restuccia (2019), we assume that idiosyncratic distortions feature a component related with firm's productivity and a random component, as a result, we generally denote distortions by the function $\tau_i(z_i, \epsilon_i)$. Specifically, we assume:

$$(1 - \tau_i) = \left(z_i^{-\rho} \epsilon_i\right)^{1-\gamma},$$

where ρ is the elasticity of distortions with respect to firm's permanent TFP, determining the systematic component of distortions, and ϵ_i is the random component of distortions drawn from an iid cdf distribution $F(\epsilon)$. Intuitively, ρ distorts the productivity gradient of firm size, whereas ϵ captures an effect of distortions on firm size that is independent of firm's productivity.

3.3 Equilibrium

We consider a stationary competitive economy in which households and firms take prices as given, prices are constant, and the distribution of resource allocations and firm types are stationary. The price of the output good is normalized to one and the price of labor is denoted by w.

Incumbent firms. An incumbent firm is characterized by productivity z and distortion τ. The firm chooses the optimal labor n to maximize expected per-period profit $\pi(z,\tau)$:

$$\begin{aligned}\pi(z_i,\tau_i) &= \max_{n\geq 0} \quad \mathbb{E}_v\left[v_i(1-\tau_i)z_i^{1-\gamma}n^{1-\gamma} - wn - c_f w\right], \\ &= \max_{n\geq 0} \quad (1-\tau_i)z_i^{1-\gamma}n^{1-\gamma} - wn - c_f w.\end{aligned}$$

In the above expression, the period transitory TFP shock v drops out of the firm problem since $\mathbb{E}v = 1$. The solution to the firm's problem implies that the labor demand and optimal output are given by

$$\begin{aligned}n(z_i,\tau_i) &= (1-\tau_i)^{\frac{1}{1-\gamma}} z_i \left(\frac{\gamma}{w}\right)^{\frac{1}{1-\gamma}}, \\ y(z_i,v_i,\tau_i) &= (1-\tau_i)^{\frac{\gamma}{1-\gamma}} v_i z_i \left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\gamma}}.\end{aligned}$$

Expected operating profits are equal to

$$\pi(z_i,\tau_i) = \Omega(1-\tau_i)^{\frac{1}{1-\gamma}} z_i - c_f w, \quad \text{where} \quad \Omega \equiv \left(\frac{\gamma}{w}\right)^{\frac{\gamma}{1-\gamma}}(1-\gamma).$$

The expected value of a firm can be written as the expected discounted per-period profit stream. The expected value of an incumbent firm $W(z_i,\tau_i)$ is:

$$\begin{aligned}W(z_i,\tau_i) &= \max\left\{\pi(z_i,\tau_i) + (1-\lambda)\frac{1}{1+r}W(z_i,\tau_i), 0\right\}, \\ &= \max\left\{\frac{\Omega(1-\tau_i)^{\frac{1}{1-\gamma}} z_i - c_f w}{1-R}, 0\right\},\end{aligned}$$

where $R = (1-\lambda)/(1+r)$, noting that firms with negative profit would not operate and hence have zero value.

Entering firms. A firm that enters the market draws an idiosyncratic innovation ability χ_i from distribution $G(\chi)$ and the random component of the distortion ϵ_i from distribution $F(\epsilon)$.

The firm then decides the level of productivity z at a cost. The firm chooses productivity to maximize the value of an incumbent firm net of productivity investment cost:

$$V(\chi_i, \epsilon_i) = \max_{z \geq 0} \left[W(z, \tau(z, \epsilon_i)) - \psi \frac{z^\phi}{\chi_i} \right],$$

where $W(z, \tau)$ is the value of an incumbent firm with productivity z and $\tau(z, \epsilon_i)$ is the distortion faced by the firm given the choice of z and the random component ϵ_i, as described above. We denote by the function $z(\chi, \epsilon)$ the optimal productivity level from this problem. Note that even though there is an optimal productivity level associated with every χ, only a fraction of firms with such χ operate in the market, a decision that depends on the random component of distortions.

Optimal productivity z for an entrant drawing (χ_i, ϵ_i) is given by:

$$z(\chi_i, \epsilon_i) = \left(\frac{(1-\rho)\tilde{\Omega}\chi_i\epsilon_i}{\psi\phi} \right)^{\frac{1}{\phi+\rho-1}}, \quad \text{where} \quad \tilde{\Omega} \equiv \frac{\Omega}{1-R}. \tag{2}$$

Note that χ and ϵ affect productivity in the same proportion, which is affected by the systematic component of distortions ρ.

Using this optimal productivity and substituting for the value of an incumbent firm, the value of an entrant firm drawing (χ_i, ϵ_i) is given by:

$$\begin{aligned}
V(\chi_i, \epsilon_i) &= \max \left\{ \tilde{\Omega} z_i^{1-\rho} \epsilon_i - \psi \frac{z_i^\phi}{\chi_i} - \frac{c_f w}{1-R}, 0 \right\}, \\
&= \max \left\{ \frac{\phi+\rho-1}{\phi} \tilde{\Omega} z^{1-\rho} \epsilon_i - \frac{c_f w}{1-R}, 0 \right\}, \\
&= \max \left\{ \frac{\phi+\rho-1}{\phi} \tilde{\Omega} \left(\frac{(1-\rho)\chi_i \tilde{\Omega} \epsilon_i}{\psi\phi} \right)^{\frac{1-\rho}{\phi+\rho-1}} \epsilon_i - \frac{c_f w}{1-R}, 0 \right\}, \\
&= \max \left\{ \Gamma \chi_i^{\frac{1-\rho}{\phi+\rho-1}} \epsilon_i^{\frac{\phi}{\phi+\rho-1}} - \frac{c_f w}{1-R}, 0 \right\},
\end{aligned}$$

where

$$\Gamma \equiv \frac{\phi+\rho-1}{\phi}\tilde{\Omega}\left(\frac{(1-\rho)\tilde{\Omega}}{\psi\phi}\right)^{\frac{1-\rho}{\phi+\rho-1}}.$$

As firms only operate when their value is non-negative, the decision to operate for a firm drawing (χ_i, ϵ_i) can be characterized as:

$$o(\chi_i, \epsilon_i) = \begin{cases} 1 & \text{if} \quad \Gamma \chi_i^{\frac{1-\rho}{\phi+\rho-1}} \epsilon_i^{\frac{\phi}{\phi+\rho-1}} \geq \frac{c_f w}{1-R}, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

At the beginning of each period, there is a mass of potential entrants and the value of entry V_e is given by,

$$V_e = \mathbb{E}V(\chi, \epsilon) - c_e w \leq 0,$$

where the expected value is taken with respect to values of χ and ϵ.

Firm distribution. The firm distribution is straightforward to characterize since we abstract from firm dynamics other than entry, exit, and the random productivity shock v. In particular, the firm distribution over productivity levels can be determined from the distribution of firms over innovation ability χ and the random distortion ϵ. The law of motion for the distribution of firms $\mu(\chi, \epsilon)$ is given by:

$$\mu'(\chi, \epsilon) = (1-\lambda)\mu(\chi, \epsilon) + Eo(\chi, \epsilon)dF(\epsilon)dG(\chi).$$

In a stationary equilibrium, the distribution is then:

$$\mu(\chi, \epsilon) = \frac{E}{\lambda} o(\chi, \epsilon) dF(\epsilon) dG(\chi), \tag{4}$$

and the mass (number) of firms is

$$N = \int_\chi \int_\epsilon d\mu(\chi, \epsilon) = \frac{E}{\lambda} \int_\chi \int_\epsilon o(\chi, \epsilon) dF(\epsilon) dG(\chi). \tag{5}$$

Definition of equilibrium. A stationary competitive equilibrium comprises a wage w; decision functions for firms: labor demand $n(z, \tau)$, profits $\pi(z, \tau)$, value of incumbent firm $W(z, \tau)$, productivity $z(\chi, \epsilon)$, operating decision $o(\chi, \epsilon)$, net value of firm $V(\chi, \epsilon)$, value of entry V_e, a distribution of firms $\mu(\chi, \epsilon)$, mass of firms N and entrants E; and allocation C for households such that:

(i) Given w, allocation C solves the household's problem.

(ii) Given w, decision function $n(z, \tau)$ solves the incumbent's firm problem, determining per-period profits $\pi(z, \tau)$ and the value of incumbent firms $W(z, \tau)$.

(iii) Given w, entrants choose productivity $z(\chi, \epsilon)$ and operating decision $o(\chi, \epsilon)$ to maximize the net value of the firm $V(\chi, \epsilon)$.

(iv) Zero profit entry condition $V_e = 0$.

(v) Invariant distribution of firms μ given by equation (4), which implies the mass of firms is constant and given by equation (5).

(vi) The goods and labor markets clear:

$$E \int_v \int_\chi \int_\epsilon v z(\chi, \epsilon)^{1-\gamma} n(\chi, \epsilon)^\gamma dF(\epsilon) dG(\chi) dH(v) = C + E \int_\chi \int_\epsilon \psi \frac{z(\chi, \epsilon)^\phi}{\chi} dF(\epsilon) dG(\chi)$$

and

$$1 = \int_\chi \int_\epsilon n(\chi, \epsilon) o(\chi, \epsilon) \mu(\chi, \epsilon) d\epsilon d\chi + E c_e + N c_f.$$

Equilibrium solution. The stationary competitive equilibrium is straightforward to compute. Given a wage rate w, all firm's decision functions can be solved and since V_e is a strictly declining function of w, the zero profit entry condition solves for w (see Proposition 1). The labor market clearing condition solves for the mass of entry E which in turn determines all other variables such as the invariant distribution and number of firms.

Proposition 1. *The equilibrium wage rate is determined by the zero profit entry condition:*

$$\int_{\chi}\int_{\epsilon}\max\left\{\Gamma(\rho,w)\chi_i^{\frac{1-\rho}{\phi+\rho-1}}\epsilon_i^{\frac{\phi}{\phi+\rho-1}}-\frac{c_f w}{1-R},0\right\}\mu(\chi,\epsilon)d\chi d\epsilon-c_e w=0. \tag{6}$$

The equilibrium wage rate w is decreasing in the elasticity of distortion ρ.

The left hand side (LHS) of equation (6) represents the expected value of potential entrants which must be zero in an equilibrium with positive entry as in our framework. Since $\Gamma(w,\rho)$ is decreasing in w and ρ, the LHS is a decreasing function of w and ρ. Given a w, when ρ increases, the LHS decreases. As a result, when ρ increases, w has to decrease for the zero profit entry condition to hold.

3.4 Model Implications

Our empirical facts in Section 2 emphasized the distributional outcomes for firm-level productivity and wedges across countries. We now discuss how the model relates with these observations. Measured firm-level productivity and wedge of firm i in the model are given by:

$$\text{TFP}_i=\frac{y_i}{n_i^{\gamma}}=z(\chi_i,\epsilon_i)^{1-\gamma}v_i, \tag{7}$$

$$wedge_i=\frac{y_i}{n_i}=\left(\frac{w}{\gamma}\right)\frac{v_i}{1-\tau_i}. \tag{8}$$

The model has two important implications for the measurement of firm productivity and wedges. First, the model implies that measured productivity in equation (7) varies endogenously with firm technology choices z. Firm technology choice from equation (2) can be rewritten as:

$$\log\left(z(\chi_i,\epsilon_i)\right) = \frac{1}{\phi+\rho-1}\log\left(\frac{(1-\rho)\tilde{\Omega}}{w\psi\phi}\right) + \frac{1}{\phi+\rho-1}\left[\log(\chi_i)+\log(\epsilon_i)\right]. \tag{9}$$

We are interested in characterizing how technology choice, which is a function of innovation ability χ, is affected by distortions. First, as noted earlier, the random component of distortions ϵ affects technology choice in the same proportion as χ. Hence, this factor alone generates some dispersion in z for a given χ. Second, technology choice is affected by the systematic component of distortions in two ways. The first term in equation (9) represents a constant in the relationship between technology choice z and χ, whereas the second term represents the gradient of χ differences on technology choice. We focus on this second term abstracting from random distortions. A higher elasticity of distortions (ρ), reduces the χ-gradient of technology choice, reducing differences in technology choice across firms. Hence, on this factor, the model would imply lower dispersion of firm-level TFP in higher ρ economics, in contrast to the empirical fact in Section 2.

The cutoff condition for the operation decision of firms in equation (3) can be written as:

$$\frac{1-\rho}{\phi+\rho-1}\log(\chi) + \frac{\phi}{\phi+\rho-1}\log(\epsilon) \geq \log\left[\frac{c_f w}{(1+R)\Gamma(w,\rho)}\right]. \tag{10}$$

The LHS of this equation is decreasing in ρ, whereas the effect on the RHS is ambiguous. Hence, the effect on selection is ambiguous, depending on the quantitative impacts of ρ and w. It follows from the above expression that a higher RHS (more selection) implies that, for a given χ, operating firms have a higher random component of distortions ϵ. In turn, this implies that higher selection is associated with more negative measured covariance between χ and ϵ, which we denote by $cov(\chi,\epsilon|o)$, as well as lower variances of χ and ϵ conditional on operating, denoted as $\sigma^2_{\chi|o}$ and $\sigma^2_{\epsilon|o}$. Propositions 2 and 3 highlight the impact of these factors on productivity dispersion and the elasticity of productivity and wedges in economies that vary in ρ.

Proposition 2. *Dispersion in productivity across firms is given by:*

$$\begin{aligned} var(TFP) &= (1-\gamma)^2\sigma^2_{z|o} + \sigma^2_v, \\ &= \left(\frac{1-\gamma}{\phi+\rho-1}\right)^2 (\sigma^2_{\chi|o} + \sigma^2_{\epsilon|o} + 2\ cov(\chi, \epsilon|o)) + \sigma^2_v. \end{aligned}$$

Productivity dispersion is decreasing in the elasticity of distortions ρ and in the extent of selection.

Proposition 2 shows that the measured variance of TFP across firms is decreasing in the elasticity of distortions ρ. In the absence of selection (i.e., when $cov(\chi, \epsilon|o) = 0$, $\sigma^2_{\chi|o} = \sigma^2_\chi$, and $\sigma^2_{\epsilon|o} = \sigma^2_\epsilon$), more distorted economies have lower productivity dispersion, in contrast with the data. Our empirical facts point to stronger selection in higher income countries, offsetting the impact of technology choice on the variance of measured productivity.

Proposition 3. *The elasticity of measured wedges with respect to productivity is given by*

$$elas(TFP, wedge) = \frac{\rho(1-\gamma)^2\sigma^2_{z|o} + \sigma^2_v - (1-\gamma)^2 cov(z, \epsilon|o)}{(1-\gamma)^2\sigma^2_{z|o} + \sigma^2_v}. \tag{11}$$

Proposition 3 shows three biases in the estimated elasticity of measured productivity and wedges. First, the ex-post random component of productivity v creates a mechanical relationship between the measured wedge and productivity that increases the elasticity. Since firms cannot adjust inputs in response to this shock, the shock has the same impact on both the measured wedge and productivity creating a positive bias. Second, selection decreases the covariance of firm-level productivity and wedges creating a positive bias in the estimated elasticity. Intuitively, this is because unproductive (low χ), high distortion (high τ) firms select out of the economy, making it more likely that observed unproductive firms have relatively low distortions (low τ), biasing upwards the measured elasticity. Third, firms choose technology z based on their draw of random distortions ϵ creating a mechanical relationship between measured wedges and productivity, biasing downwards the estimated elasticity.

It follows that the measured elasticity of productivity and wedges accurately reflect the underlying elasticity of distortions ρ only when these biases are zero, such that firms have full ex-ante information ($\sigma_v^2 = 0$), there is no selection ($cov(\chi, \epsilon|o) = 0$), and technology is exogenous. The actual magnitude of these biases and their net impact across countries is a quantitative question that we explore in the next section.

4 Quantitative Analysis

We proceed in three steps. First, we calibrate a distorted benchmark economy to micro and aggregate data for France. Second, we examine the micro and macro effects of changes in distortions, in particular we focus on changes in the TFP elasticity of distortions ρ. Third, we examine the implications of the model when we vary ρ, as well as σ_ϵ and σ_v across a variety of observations in the cross-country data.

4.1 Calibration

We calibrate a distorted benchmark economy to micro and aggregate data for France. There are 11 parameters to calibrate in the model: decreasing returns to scale γ, firm's exit rate λ, real interest rate r, dispersion in innovation ability σ_χ, dispersion transitory ex-post productivity shock σ_v, level and curvature parameters of innovation cost function ϕ and ψ, fixed costs of entry c_e and operation c_f, and the parameters of idiosyncratic distortions ρ and σ_ϵ.

A set of 6 parameters are either normalized or assigned values from outside evidence. These are: $\gamma = 0.8$ commonly used in the misallocation literature (Guner et al., 2008; Restuccia and Rogerson, 2008), exit rate $\lambda = 0.1$ (Davis et al., 1998), real interest rate $r = 0.04$, curvature of investment cost function $\phi = 2$ (Acemoglu et al., 2018). We normalize productivity investment cost $\psi = 1$ and the cost of entry $c_e = 1$.

The remaining 5 parameters are jointly calibrated so that the model matches moments from the French data. The parameters we calibrate are ρ, σ_χ, σ_ϵ, σ_v, and c_f to jointly to match the following moments: (1) the distortion-productivity elasticity, (2) the standard deviation of log employment, (3) the standard deviation of log distortions, (4) the standard deviation of log TFP, and (5) average firm size.

Table 1: Calibration of Distorted Benchmark Economy

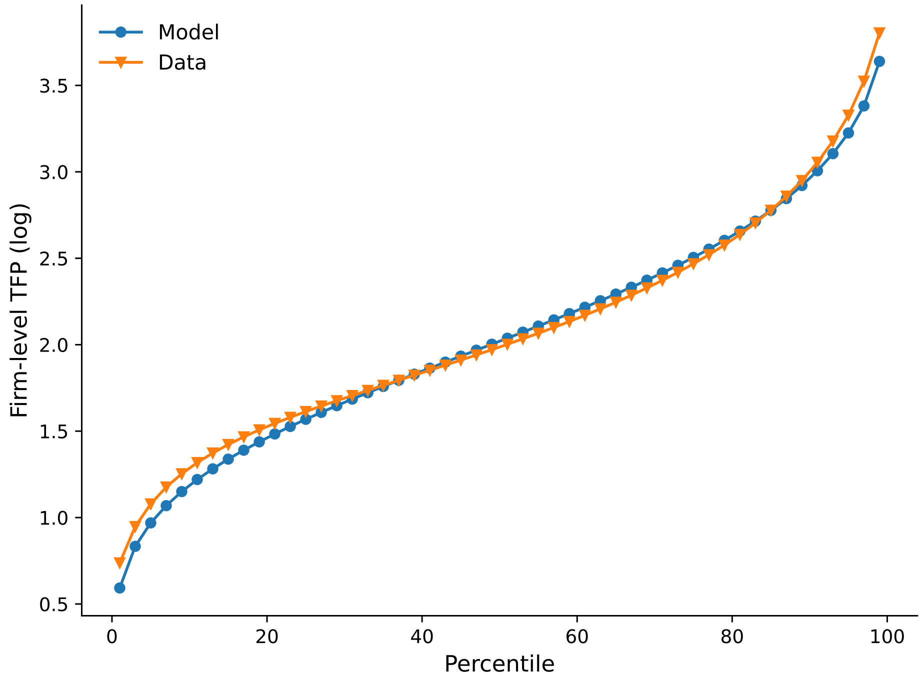
Parameter	Value	Targeted moments	Model	Data
ρ	0.525	Elasticity of distortions	0.75	0.75
σ_χ	11.0	sd log employment	1.31	1.31
σ_ϵ	1.4	sd log distortions	0.55	0.55
σ_v	0.2	sd log TFP	0.68	0.66
c_f	0.14	Average firm size	14.7	14.9

Table 1 reports the resulting calibrated parameter values as well as the values of sample data and model implied moments. We note that the calibrated parameter values for distortions ($\rho = 0.525$ and $\sigma_\epsilon = 1.4$) imply values of distortion moments that are consistent with estimates found in other studies (Hsieh and Klenow, 2009). We also note that whereas the calibrated value of ρ is 0.525, the implied TFP elasticity of distortions across firms in the model is 0.75 as measured in the French data. The reason for the difference between the population parameter ρ and the sample moment is a strong selection of firms in operation in the model which raises the correlation between TFP and distortions across firms in the sample of operating firms.

We also report in Figure 7 the implied percentile distribution of firm-level TFP in the model compared to the French data. Despite the model assuming a log normal distribution of innovation abilities and the calibration only targeting the standard deviation of log TFP in the data, the resulting distribution of firm-level TFP matches quite closely that of the data for France.

The calibrated benchmark economy features strong selection of firms in operation, that is, the distribution of innovation ability features large dispersion in potential firm-level TFP but many of these potential firms do not operate in equilibrium. For instance, the percentile one (p1) TFP firm in the calibrated benchmark economy would be a percentile 80 (p80)

Figure 7: Firm-level TFP Distribution



Notes: Data refers to the distribution of firm-level TFP in France from Orbis, whereas Model refers to the calibrated distribution of productivity in the distorted benchmark economy. For ease of illustration, the figure only plots 50 percentile points of the distributions, from percentile one (p1) to percentile 99 (p99).

TFP firm in the economy without an operation cost.

4.2 Experiments

We examine the effects of changes in the TFP elasticity of distortions ρ on micro and aggregate outcomes. We consider changes in ρ from the benchmark value of 0.525 to 0.65, 0.80 and 0.90 consistent with the range of values for the sample elasticity found in our empirical section for countries such as Vietnam and China. We also report an economy with $\rho = 0$ for reference, although we not that this economy is not an undistorted economy since the random component of distortions σ_ϵ is always present at the calibrated value in the experiments.

The results are reported in Table 3. The main result is that increasing ρ from 0.525 to 0.00 reduces aggregate output and TFP substantially by 77% percent. Even though in the model the entire drop in aggregate output is due to the change in correlated distortions

Table 2: Experiments of Benchmark Economy with Alternative ρ Values

	Value of ρ				
	0.00	0.525	0.65	0.80	0.90
Aggregate output	1.49	1.00	0.75	0.41	0.23
Efficient output	1.34	1.00	0.88	0.70	0.56
Misallocation:					
Allocative efficiency	0.85	0.76	0.65	0.45	0.32
Static misallocation	1.15	1.00	0.91	0.76	0.64
Firm-level TFP distribution:					
SD (log TFP)	0.43	0.68	0.79	0.96	1.07
p95 (log TFP)	4.14	3.23	2.88	2.39	2.00
p5 (log TFP)	2.71	0.97	0.27	−0.77	−1.52
SD (log employment)	1.95	1.31	1.13	0.90	0.80
Average firm size	266.9	14.70	7.22	3.53	2.56
Elasticity of distortions	0.35	0.75	0.84	0.92	0.96
SD (log distortions)	0.31	0.55	0.68	0.89	1.03
Entry and selection:					
Mass of operating firms	0.003	0.06	0.12	0.25	0.35
Operation rate	0.004	0.05	0.10	0.22	0.31

Notes: Allocative efficiency is aggregate output relative to efficient output, which is aggregate output among the set of operating producers under the efficient allocation of aggregate resources. Static misallocation refers to the misallocation of resources in the distorted economy under the same productivity distribution of the benchmark economy ($\rho = 0.525$). The operation rate is the fraction of entering firms that operate.

(ρ), we can use the model to separate the change in aggregate output due to different features such as the change in allocative efficiency (factor misallocation) as emphasized in the literature and the change in the firm-level TFP distribution. Allocative efficiency is the ratio of aggregate output to efficient output, that is aggregate output among the set of operating producers under the efficient allocation of aggregate resources. Allocative efficiency is 0.76 in the benchmark economy and drops to 0.32 in the $\rho = 0.90$ economy. As a result, aggregate output due to factor misallocation (the change in allocative efficiency) falls to 0.42 (0.32/0.76) which represents about 60 percent of the overall reduction in aggregate output ($\log 0.42/\log 0.23$). Interstingly, the change in allocative efficiency can be further decomposed into the effect of static misallocation (misallocation of resources among the same set producers) and the effect of misallocation given the change in the TFP distribution. Static misallocation accounts for around 50 percent ($\log 0.64/\log 0.42$) of the reduction in allocative efficiency, leaving the remaining 50 percent attributed to the impacts on changes in firm-level TFP distribution. In terms of the overall decline in aggregate productivity, static misallocation is responsible for only about 30 percent ($\log(0.76)/\log(0.41)$), leaving the remaining 70 percent to be accounted for by changes in the TFP distribution.

A useful summary metric of the contribution of changes in the firm-level TFP distribution is the efficient aggregate output which only depends on firm-level TFP's. Relative to the benchmark economy aggregate efficient output falls to 0.70 in the $\rho = 0.80$ economy. Note also the asymmetry in the change of efficient output for an increase in ρ compared with a decrease in ρ from the benchmark economy. The reason for this asymmetric response of aggregate efficient output to changes in correlated distortions is that productivity investments disproportionately fall as ρ increases with a maximized effect as ρ approaches one.

We also note from the results in Table 3 that the extent of misallocation, measured by the standard deviation of log distortions in the model, which changes from 0.51 in the benchmark economy to 1.03 in the $\rho = 0.90$ economy, is consistent with empirical findings in the literature (Hsieh and Klenow, 2009) as well as with our own empirical results from ORBIS. But when the productivity cost of misallocation is calculated taking the distribution of firm-level productivity as given, the size of this effect is relatively small (Restuccia and Rogerson, 2017). For instance, the static reallocation gain, that is efficient to actual output given the distribution of firm-level TFP, is only 32 percent ($1/0.76$) in the $\rho = 0.80$ economy relative to the benchmark economy. What this misallocation effect misses is the fact that distortions also affect the distribution of TFP in the economy both because of selection of operating firms as well as other investments such as productivity improvements and technology adoption (Bento and Restuccia, 2017; Ayerst, 2022; Ayerst et al., 2020).

Figure 8: Firm-level TFP Distribution across Different ρ Economies

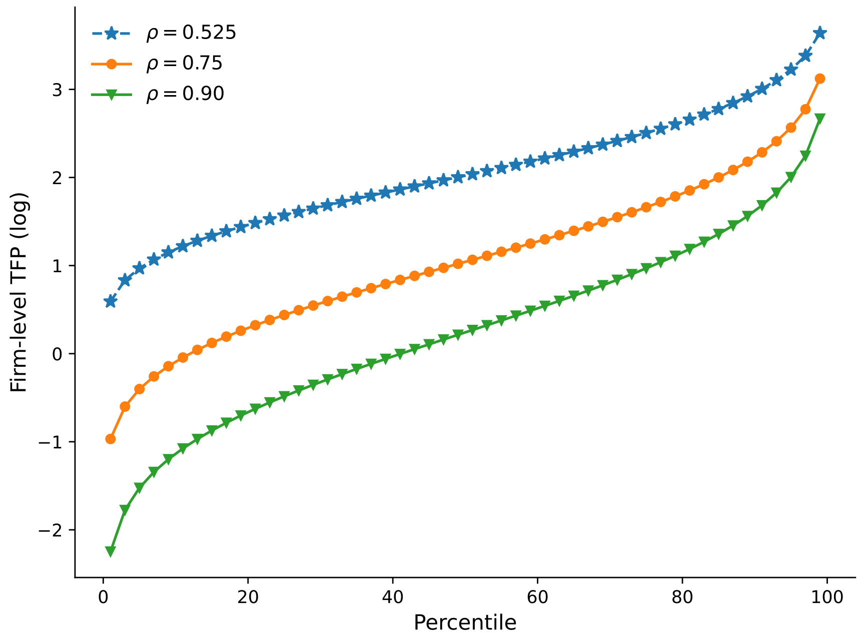


Table 3 also reports statistics about the TFP distribution in the benchmark and experiment economies. Increases in ρ increases the dispersion in firm-level TFP, measured by the standard deviation of log TFP, increases from 0.68 to 1.07, whereas in our empirical finding this moment increases up to 1.5 in less developed countries. Hence, just increases in correlated distortions can go a long way in accounting for the increased dispersion in firm-level TFP observed in the less developed countries. The experiments are also consistent with our empirical finding that the larger dispersion in TFP is due to lack of selection at the bottom of the TFP distribution. While the p95 of the log TFP distribution decreases by 1.23 log points from the benchmark economy to the $\rho = 0.90$ economy, TFP at the bottom (p5) decreases by more, 2.49 log points. We also illustrate the change in the TFP distribution by plotting the distributions for the benchmark economy and other ρ economies in Figure 8. Consistent with our empirical findings, the distributions are much closer to each other at the top than at the bottom of the distribution.

We also note two other interesting quantitative properties of the experiments. First, the increase in correlated distortions ρ generates a reduction in average firm size from about 19 workers per firm in the benchmark economy to 2.56 workers in the $\rho = 0.90$ economy. This range is consistent with cross-country data in average firm size in the manufacturing sector (Bento and Restuccia, 2017). The decrease in average firm size is the result of an increase in the number of firms operating in the economy. Second, whereas ρ increases from 0.525 to 0.90 in our experiments, the TFP elasticity of distortions in the model (the elasticity of distortions among operating firms) increases more narrowly from 0.75 to 0.96. This illustrates that differences in sample selection can bias common statistics about cross-country changes in misallocation, which tends to narrow the difference in measured misallocation by this elasticity.

To summarize, plausible variations in correlated distortions ρ from the benchmark economy generate implications on misallocation and the firm-level TFP distribution consistent with the cross-country data including a variation in aggregate output that is three times the change in aggregate output from static factor misalloction alone. We noted that variation in correlated distortions alone does not account for the entire range of dispersion in the TFP distribution and in terms of employment, correlated distortions alone naturally reduce dispersion in distortions across firms. Next, we evaluate variation in correlated distortions ρ along with other parameters in the model against the backdrop of cross-country data.

4.3 Cross-Country Experiments

We now perform 3 sets of experiments and compare the implications of each on outcomes of interest in the cross-country data. First, we report economies where only ρ varies as discussed in the previous section. Second, we report economies where both ρ and the random component of distortions σ_ϵ are changing. Third, we report economies with changes in ρ, σ_ϵ, and σ_v. We have also experimented with changing the fixed cost of operation c_f with no relevant variation across economies. The specific values we use for changes in the parameters across the experiments are reported in Table 3.

We now report the results of the various experiments for statistics of interest against

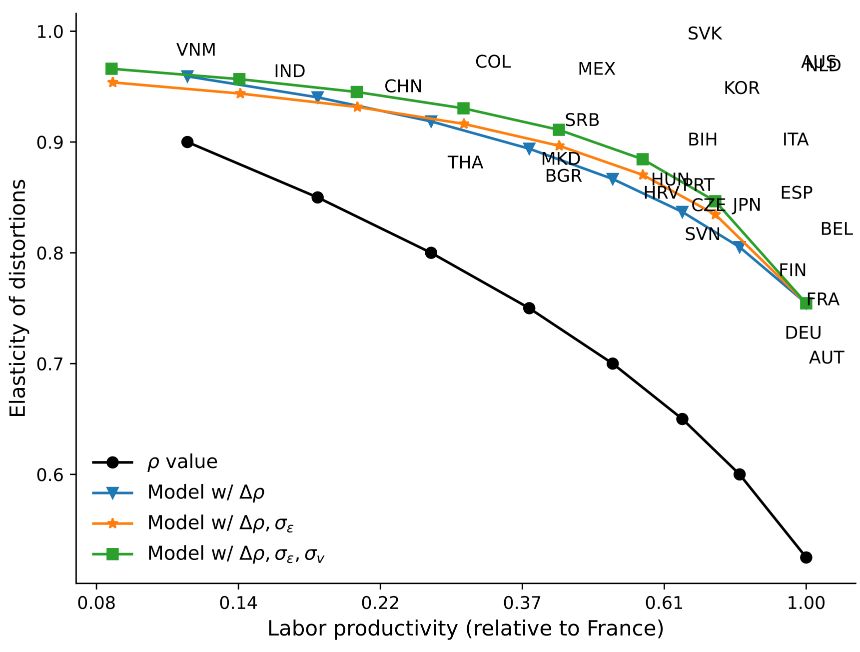
Table 3: Experiments of with Alternative ρ, σ_ϵ, and σ_v Values

ρ	0.525	0.60	0.65	0.75	0.80	0.85	0.90
σ_ϵ	1.40	1.84	2.13	2.43	2.72	3.01	3.60
σ_v	0.20	0.30	0.37	0.43	0.50	0.57	0.70

Notes: Parameter values of alternative experiments from the benchmark economy values ($\rho = 0.525$).

the corresponding cross-country data. For each statistic, we report the implications against aggregate labor productivity in the data and the model. This is because measures of aggregate TFP in the data can be subject to important measurement issues, especially related to measures of capital. In the model, the implied aggregate labor productivity based on a simple mapping between aggregate TFP and labor productivity implied by the standard neoclassical growth model, in particular, assuming a capital income share of one third, differences in labor productivity are 1.5 times the differences in aggregate TFP.

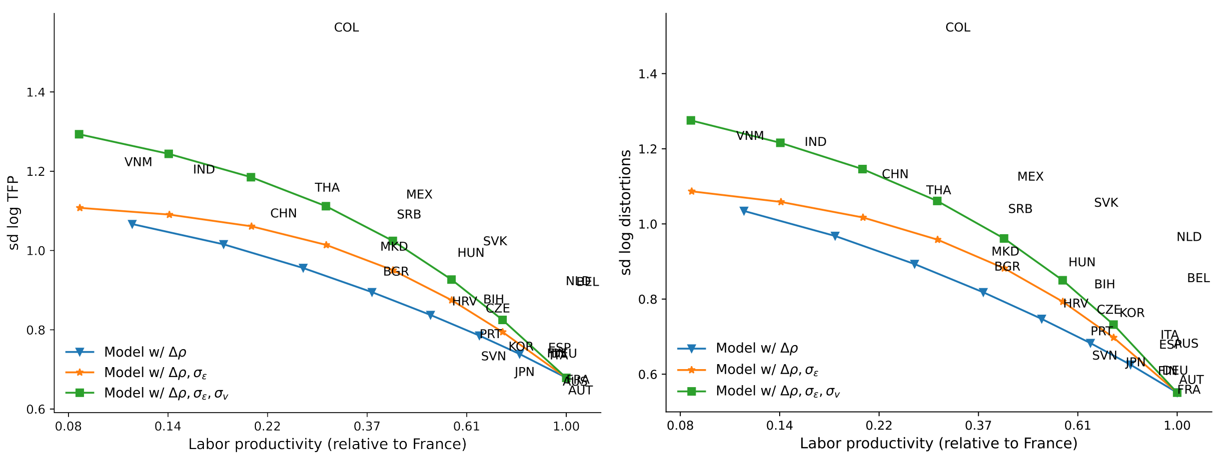
Figure 9: Elasticity of distortions



Notes: The circles-black line represents the values of ρ across experiments, whereas the triangle-blue line represent the sample elasticity in the model with only variation in ρ. The star-orange and square-green lines report the same but for economics that also differ on σ_ϵ, and σ_ϵ and σ_v, respectively, according to the values reported in Table 3.

Figure 9 documents the relationship between the elasticity of distortions and aggregate labor productivity across countries and the various experiment economies. We emphasize two important patterns that arise. First, as indicated earlier, the values of ρ in the model are smaller than the elasticity of distortions among operating firms, indicating a substantial upward bias in the estimated elasticity. Importantly, the bias is stronger for richer economies as there is more selection of operating firms and smaller in poorer economies where the selection of operating firms among the less productive firms is weaker. Second, while all the experiments fit the pattern in the data well, the results across experiments indicate that the bulk of the empirical relationship between the elasticity and aggregate labor productivity is generated by differences in correlated distortions ρ across economies.

Figure 10: Standard deviation of log TFP and log distortions

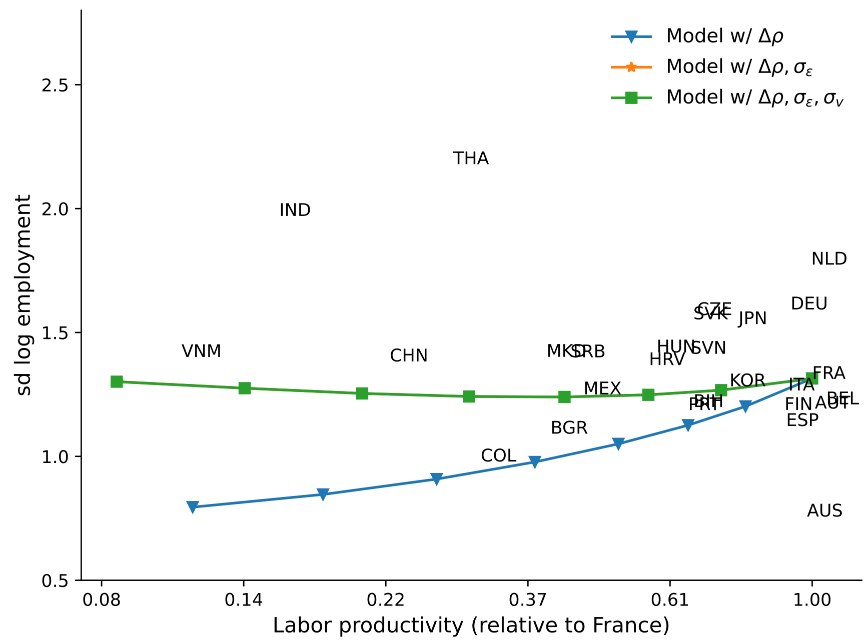


(a) Standard deviation of log TFP (b) Standard deviation of log distortions

Notes: The blue-triangular, orange-starred, and green-squared lines correspond to experiments: (1) varying ρ, (2) varying ρ and σ_ϵ, and (3) varying ρ, σ_ϵ, and σ_v, respectively, according to values reported in Table 3.

Figure 10 reports the standard deviation of log TFP and log distortions across countries and experiments by aggregate labor productivity. As discussed earlier in the empirical section 2, lower-income countries display greater variability in both productivity and measured distortions. Although the experiment with only variation in correlated distortions ρ can account for a substantial portion of the larger variability in TFP and distortions in the cross-country data, there is excess variability in the data to be explained. Adding variation in ϵ and v in our experiments lead to larger dispersion in productivity and distortions, aligning the model better with the cross-country data, especially among lower-income countries. Our finding is that larger dispersion in ϵ and especially v, contribute to accounting for about one-third of the dispersion in productivity and distortions, with the remaining two-thirds variation accounted for by variation in correlated distortions ρ.

Figure 11: Employment Dispersion across Firms



Notes: Employment dispersion is measured by the standard deviation of log employment across firms. The blue-triangular, orange-starred, and green-squared lines correspond to experiments: (1) varying ρ, (2) varying ρ and σ_ϵ, and (3) varying ρ, σ_ϵ, and σ_v, respectively, according to values reported in Table 3. Experiments (2) and (3) are perfectly aligned because v does not affect employment decisions.

As noted earlier, correlated distortions tend to compress the employment distribution across firms when the productivity distribution is constant. Quantitatively, we found that even when productivity dispersion increases in our experiments, the resulting dispersion in employment is smaller whereas in the data we find a roughly constant employment dispersion across developed and less developed countries. Figure 11 reports the dispersion in employment, measured by the standard deviation of log employment, across countries and experiments by aggregate labor productivity. The data does not indicate a systematic relationship between the dispersion in firm size and aggregate productivity across countries. Whereas variation in correlated distortions ρ alone results in lower dispersion in firm size across firms, more dispersed distortions, via increases in σ_ϵ, generates patterns that align more closely with the data. This suggests that while the dispersion in firm size is primarily driven by productivity dispersion in economies with low correlated distortions ρ where the allocation of employment across firms is closer to the efficient levels, in less developed countries, a larger component of the dispersion in firm size is driven by larger dispersion in size distortions.

5 Conclusions

We examine the disparity in aggregate productivity across nations using cross-country firm-level panel data and a quantitative model featuring production heterogeneity with distortions and entry and operation decisions of firms. Empirically, we find that less developed countries feature higher distortions and larger dispersion in firm-level productivity, mostly resulting from the prevalence of unproductive firms compared to developed countries. Quantitatively, we find that measured distortions in the form of higher productivity elasticity of distortions generates large aggregate output losses, two-thirds of which are accounted for by changes in the productivity distribution. About half the change in allocative efficiency is attributed to the change in misallocation due to changes in the productivity distribution and the other half is static misallocation.

Our quantitative analysis provides a connection of policies and institutions that generate static misallocation to explain lower allocative efficiency and large aggregate output losses in more distorted and less developed countries. Our analysis provides a parsimonious modeling of changes in firm-level productivity. Future work is needed to investigate the channels that may be important in productivity differences such as differential management practices, lags in technology diffusion, or barriers to foreign multinationals. Similarly, more work is needed to identify the specific policies that are relevant in accounting for the distortion patterns in less developed countries, an analysis that may require more specific country contexts.

References

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, reallocation and growth. *American Economic Review*, 126(4):1374–1443.

Alviarez, V., Ramondo, N., and Cravino, J. (2023). Firm-embedded productivity and cross-country income differences. *Journal of Political Economy*, 131(9).

Andrews, D., Criscuolo, C., and Gal, P. N. (2015). Frontier firms, technology diffusion and public policy: Micro evidence from oecd countries.

Ayerst, S. (2022). Distorted technology adoption. Technical report.

Ayerst, S., Brandt, L., and Restuccia, D. (2020). Market constraints, misallocation, and productivity in vietnam agriculture. *Food Policy*, 94:1–16.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *American economic review*, 103(1):305–334.

Bento, P. and Restuccia, D. (2017). Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303.

Bento, P. and Restuccia, D. (2021). On average establishment size across sectors and countries. *Journal of Monetary Economics*, 117:220–242.

Bhattacharya, D., Guner, N., and Ventura, G. (2013). Distortions, endogenous managerial skills and productivity differences. *Review of Econoic Dynamics*.

Bustos, P. (2011). Trade liberalization, exports and technology upgrading: Evidence on the impact of mercosur on argentinean firms. *American Economic Review*.

Comin, D. and Mestieri, M. (2018). If technology has arrived everywhere, why has income diverged? *American Economic Journal: Macroeconomics*, 10(3).

Davis, S. J., Haltiwanger, J. C., Schuh, S., et al. (1998). Job creation and destruction. *MIT Press Books*, 1.

Gal, P. N. (2013). Measuring total factor productivity at the firm level using oecd-orbis.

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

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *The quarterly journal of economics*, 114(1):83–116.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150.

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.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., and Yesiltas, S. (2023). How to construct nationally representative firm level data from the orbis global database: New facts and aggregate implications. Technical report, National Bureau of Economic Research.

Khandelwal, A. K., Schott, P. K., and Wei, S.-J. (2013). Trade liberalization and embedded institutional reform: Evidence from chinese exporters. *American Economic Review*, 103(6).

Klenow, P. J. and Rodriguez-Clare, A. (1997). The neoclassical revival in growth economics: Has it gone too far? *NBER macroeconomics annual*, 12:73–103.

Parente, S. L. and Prescott, E. C. (1994). Barriers to technology adoption. *Journal of Political Economy*, 102(2):298–321.

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *The Review of economic studies*, 69(1):245–276.

Prescott, E. C. (1998). Lawrence r. klein lecture 1997: Needed: A theory of total factor productivity. *International economic review*, pages 525–551.

Restuccia, D. (2019). Misallocation and aggregate prodcutivity across time and space. *Canadian Journal of Economics*, 52(1):5–32.

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *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.

On-line Appendix

A Data Details

We describe the details of the construction for the final dataset.

A.1 Variables, Sample Selection, and Data Cleaning

Our final dataset is constructed for the period 2000 to 2019 where an observation is at the firm-year (i, t) level. In most countries, the number of observations increases substantially around 2000 and starts to decline in more recent periods. We additionally drop 2020 and later periods to avoid including the covid-19 pandemic. We use two-digit SIC codes (denoted by s) as our definition for the firm's sector, which we base on the firm's primary sector of operations. Kalemli-Ozcan et al. (2023) discuss the advantages and disadvantages of the Orbis dataset and provide a comparison of aggregate outcomes with national statistics.

Variables. Our baseline model requires us to construct measures of firm-level output $y_{i,t}$ and labor $n_{i,t}$. We measure output as the firms reported sales or operating revenues, when sales is unavailable. We use this measure instead value added since material costs are not systematically available in most non-European countries in our dataset. We measure employment by the reported count of employees at the firm. We also use capital $k_{i,t}$ to construct our measure of firm-level TFP in the baseline analysis. We measures capital as the total book value of firm fixed assets. For some robustness checks we also construct measures of firm-level value added, as an alternative measure of output, in which we subtract material costs from sales (or operating revenue) data.

For firm-year observations that are missing employment data, we construct employment using the wagebill. Firm we construct the average wage bill of firms within a sector year as $\bar{w}_{s,t} = \sum_{i \in \mathcal{I}_{s,t}} W_{i,t}/n_{i,t}$ for firms that report both the wagebill $W_{i,t}$ and employment. The set $\mathcal{I}_{s,t}$ is the set of firms in sector s and year t. We then use this value to construct the firm's employment as the firm's reported wagebill divided by the constructed wage rate, i.e., $\hat{n}_{i,t} = \frac{W_{i,t}}{\bar{w}_{i,t}}$.

Dropped observations. Firm-year observations reported in the Orbis dataset cover a broader range than what we require for our analysis. Additionally, we exclude some observations to improve the comparability across countries and to reduce the influence of outliers. In particular, We drop observations based on the following criteria:

- **Missing data.** Our baseline measures of productivity and wedges require information on sales and employment. We also consider alternative models that require capital and material costs as well as data on the previous and next period variables. We exclude firms for which we cannot construct labor productivity - i.e., firms missing either output or employment data.

- **Inactive firms.** We include only firms that are listed as active in the current year, excluding firms with unknown status, in bankruptcy, dissolved, in liquidation, or inactive.

- **Sectors.** We focus on the manufacturing sector, four-digit SIC between 1000 and 6000. We exclude all firms that are not in the manufacturing sector.

- **Missing data.** We drop observations that are missing data needed to construct measures of TFP, i.e., negative or missing sales, employment, capital, or material costs.

- **Data trimming.** We trim the top and bottom 2% based on the measure of productivity and wedges. This is done after removing sector and year fixed effects such that the data trimming is not impacted by year trends or sectoral differences.

Multiple observations. Many firms report multiple filings within a year for various reasons. We remove multiple observations based on:

- **Consolidated financial records**. Firms may report financial records for either unconsolidated, consolidated, or both. In the case where both are reported, we default to using the unconsolidated records.

- **Filing type**. Firms may report financial records as "annual reports" or "local registry filings". In the case where both are reported, we use the annual reports.

- **Other duplicates**. Other instances of firms reporting multiple filings are relatively rare and for the most part represent duplicated data. For these duplicates, we choose between the maximum and minimum observed values based on which values minimize the absolute error with output and employment in the previous period.

Time and sector trends. We regress each nominal variable on year-by-sector fixed effects and report summary statistics for the residualized variable. That is, for variable $\tilde{X}_{f,t}$ we estimate $\tilde{X}_{f,t} = \Gamma_{s,t} + X_{f,t}$ and then construct the detrended variable as $X_{i,t} = \tilde{X}_{i,t} - \Gamma_{s,t}$.

A.2 Firm Weights

An issue with the Orbis data is that it tends to oversample large firms and undersample small firms in some countries. Kalemli-Ozcan et al. (2023) report that European countries tend to reflect the size distribution of firms reported in Eurostat (based on national statistics). However, less is known about the coverage outside of European countries. We construct firm weights using national statistics to allow us to reweight the data to match the true distribution. We show that our results are robust to this reweighting in Appendix B

We denote our final firm weights as $\omega_{n,t}$ and denote $h^{D}_{[\underline{n},\overline{n}],c,t}$ as the share of firms with between $\underline{n}$ and $\overline{n}$ employees from dataset D. For example, the share of firms in France in 2013 with between 10 and 19 employees in the Orbis dataset is denoted by $h^{Orbis}_{[10,19],FR,2013}$. We construct the final firm weights as

$$\omega_{[\underline{n},\overline{n}],c,t} = \frac{h^{D}_{[\underline{n},\overline{n}],c,t}}{h^{Orbis}_{[\underline{n},\overline{n}],c,t}}.$$

We discuss the construction of $h^{D}_{[\underline{n},\overline{n}],c,t}$ using nationally representative data below.

Eurostat data. We use the distribution of firms by employment size as reported by Eurostat in the business demography (BD) and structural business statistics (SBS) datasets to construct observation weights. The BD dataset reports more granular data for smaller business sizes and separates non-employer businesses. The BD dataset reports firms in employment bins $\{0, 1-4, 5-9, 10+\}$. The SBS dataset is more granular at higher employment levels but lumps non-employers into the smallest size bin. The SBS dataset reports firms in employment bins $\{0-9, 10-19, 20-49, 50-249, 250+\}$.

We exclude non-employer businesses from the final dataset. We construct the final Eurostat bins as

$$\begin{aligned}
h^{ES}_{[1,4],c,t} &= \frac{h^{SBS}_{[0,9],c,t} - h^{BD}_{0,c,t}}{1 - h^{BD}_{0,c,t}} \times \frac{h^{BD}_{[1,4],c,t}}{h^{BD}_{[1,4],c,t} + h^{BD}_{[5,9],c,t}} \\
h^{ES}_{[5,9],c,t} &= \frac{h^{SBS}_{[0,9],c,t} - h^{BD}_{0,c,t}}{1 - h^{BD}_{0,c,t}} \times \frac{h^{BD}_{[5,9],c,t}}{h^{BD}_{[1,4],c,t} + h^{BD}_{[5,9],c,t}} \\
h^{ES}_{[10,19],c,t} &= \frac{h^{SBS}_{[10,19],c,t}}{1 - h^{BD}_{0,c,t}} \\
h^{ES}_{[20,49],c,t} &= \frac{h^{SBS}_{[20,49],c,t}}{1 - h^{BD}_{0,c,t}} \\
h^{ES}_{[50,249],c,t} &= \frac{h^{SBS}_{[20,49],c,t}}{1 - h^{BD}_{0,c,t}} \\
h^{ES}_{[250,\infty],c,t} &= \frac{h^{SBS}_{[250,\infty)}}{1 - h^{BD}_{0,c,t}}
\end{aligned}$$

We extend the firm shares $h^{ES}_{[\underline{n},\overline{n}],c,t}$ to earlier and late periods by assuming that firm shares are the same as in the closest period. For example, if the earliest period with sufficient data for Austria is 2005 then we assume the weights $h^{ES}_{[\underline{n},\overline{n}],AT,2005}$ also apply to the period 2000 to 2004. Firm shares tend to be relatively stable over time and this allows us to maximize the usable data. We also interpolate data missing in intermediate periods as a linear combination of the two surrounding periods.

OECD data. The OECD database reports the firm size distribution divided into either three or five size bins. The three size bin categories are $\{1-19, 20-249, 250+\}$ and the five size bin categories are $\{1-9, 10-19, 20-49, 50-249, 250+\}$. We follow same procedure as with the Eurostat data to fill in missing data.

We also use the OECD data to construct weights for countries without alternative sources. For these countries, we first construct the expected share of firms in the $20-249$ size bin by regressing $h_{[20,249],c,t} = \alpha \ln GDP/Capita_{c,t} + F_t + \epsilon_{c,t}$, where F_t is a year fixed effect. The coefficient α captures the relationship between a countries output per worker and the size distribution, where $\alpha > 0$ ($\alpha < 0$) implies that wealthier countries have more (fewer) firms in this size bin.

Individual country data. We supplement the above information with data on Vietnam, Mexico, and Korea. The Vietnam data, from the Vietnamese Statistical Yearbook, groups firms into the size categories $\{1-4, 5-9, 10-49, 50-199, 200-299, 300+\}$ and is available from 2004 to 2019. The Mexico data groups firms into the size categories $\{1-10, 11-50, 51-250, 250+\}$ and is available every five years between 2004 to 2019. The Korea data groups firms into the the size categories $\{1-4, 5-9, 10-49, 50-99, 100-199, 200-299, 300+\}$ and is available from 2011 to 2019.

Table A.1: Final Dataset

Country	Observations	Firm Distribution (Source)
Austria (AUT)	27,657	Eurostat
Australia (AUS)	12,567	OECD
Bosnia and Herzegovina (BIH)	46,971	Eurostat
Belgium (BEL)	113,690	Eurostat
Bulgaria (BGR)	156,530	Eurostat
China (CHN)	1,987,483	Estimate
Colombia (COL)	6,169	Estimate
Czech Republic (CZE)	168,581	Eurostat
Germany (DEU)	195,840	Eurostat
Spain (ESP)	1,262,738	Eurostat
Finland (FIN)	143,123	Eurostat
France (FRA)	1,046,480	Eurostat
Croatia (HRV)	127,810	Eurostat
Hungary (HUN)	309,065	Eurostat
India (IND)	138,793	Estimate
Italy (ITL)	1,675,006	Eurostat
Japan (JPN)	537,463	OECD
Korea (KOR)	1,195,930	Statistics Korea
Mexico (MEX)	6,834	Mexico Economic Census
Montenegro (MNE)	6,527	Estimate
North Macedonia (MKD)	34,438	Eurostat
Netherlands (NLD)	15,987	Eurostat
Portugal (PRT)	414,366	Eurostat
Serbia (SRB)	171,442	OECD
Slovenia (SVN)	90,740	Eurostat
Slovakia (SVK)	11,956	Eurostat
Thailand (THA)	11,956	Estimate
Vietnam (VNM)	128,837	Vietnam Statistical Yearbook

A.3 Overview of Final Dataset

B Empirical Analysis

We examine the robustness of the main empirical results to alternative constructions of productivity and wedges. In addition, for each version of the model, we compare outcomes when factor inputs are only labor (tfp^{lp}) or a Cobb-Douglas aggregate of capital and labor (tfp^{cd}).

- **Value Added**: The baseline results use gross output to construct statistics since this improves the representation across countries. In the alternative model, we construct output as sales $s_{i,t}$ subtract material costs $m_{i,t}$. The measures of productivity are then:

$$tfp^{lp}_{i,t} = \frac{s_{i,t} - m_{i,t}}{n^{\gamma}_{i,t}}, \qquad tfp^{cd}_{i,t} = \frac{s_{i,t} - m_{i,t}}{(k^{\alpha}_{i,t} n^{1-\alpha}_{i,t})^{\gamma}}.$$

- **Constant Elasticity of Substitution**: Hsieh and Klenow (2009) construct a model in which firms have constant returns to scale and face constant elasticity of substitution against products produced by other firms. In this version of the model, productivity can be constructed as

$$tfp^{lp}_{i,t} = \frac{(p_{i,t} y_{i,t})^{\frac{\sigma}{\sigma-1}}}{n_{i,t}}, \qquad tfp^{cd}_{i,t} = \frac{(p_{i,t} y_{i,t})^{\frac{\sigma}{\sigma-1}}}{k^{\alpha}_{i,t} n^{1-\alpha}_{i,t}}.$$

- **Population Weighting**: The final version of the model that we report is identical to the baseline model but we weight the results by the population weights constructed in Appendix A.

Wedges are the same in each version of the model since wedges do not rely on the structure of the production function. We construct wedges based on the Cobb-Douglas inputs, the labor input, and the capital input. In the case where the distortion is on firm revenues, as opposed to factor inputs, these three wedges are theoretically equivalent. This is not necessarily the case when distortions tend to impact one factor more than the other, such as, if credit constraints limit capital inputs more than employment.

$$wedge^{y}_{i,t} = \frac{y_{i,t}}{k^{\alpha}_{i,t}\ell^{1-\alpha}_{i,t}}, \qquad wedge^{\ell}_{i,t} = \frac{y_{i,t}}{\ell_{i,t}}, \qquad wedge^{k}_{i,t} = \frac{y_{i,t}}{k_{i,t}}.$$

where in the value added approach the numerator is sales minus material costs.

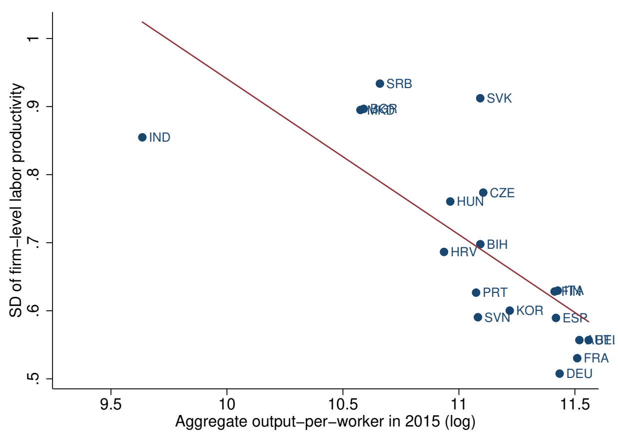
B.1 Firm Productivity Distribution

Figure B.1 reports the standard deviation of firm productivity across the six different models described above.

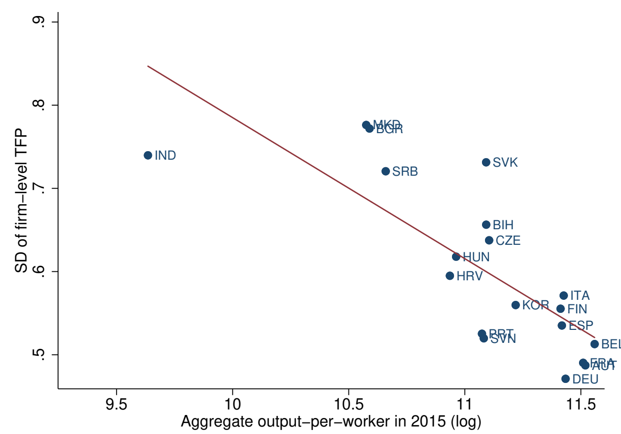
B.2 Elasticity of Wedges

Figure B.1: Standard Deviation of Firm Productivity

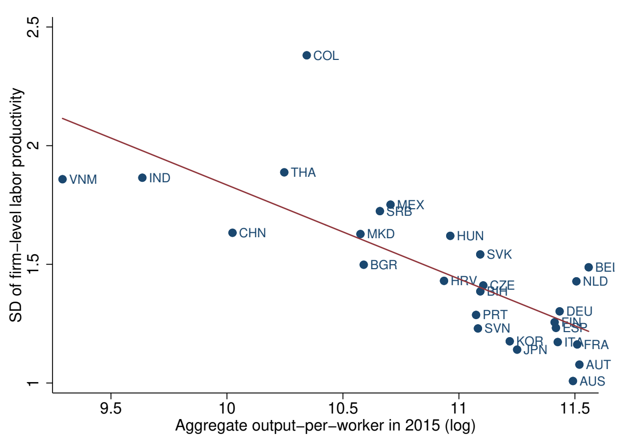
(a) Value Added Labor Productivity



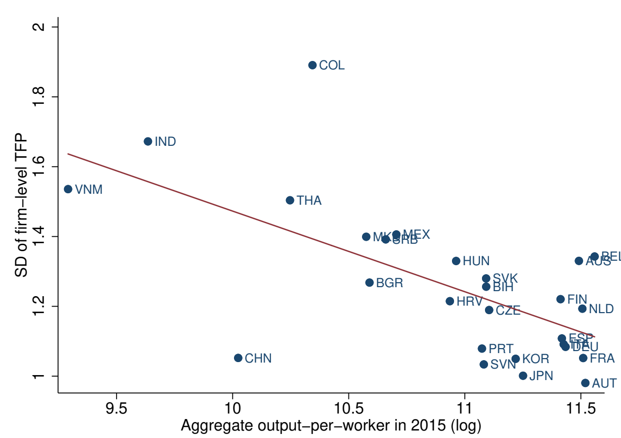
(b) Value Added TFP



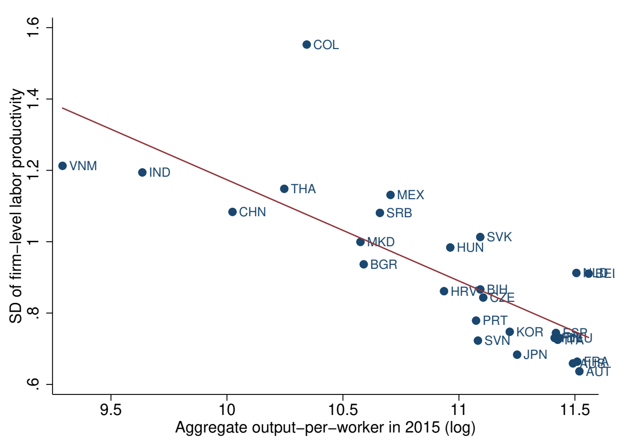
(c) CES Labor Productivity



(d) CES TFP



(e) Weighted Labor Productivity



(f) Weighted TFP

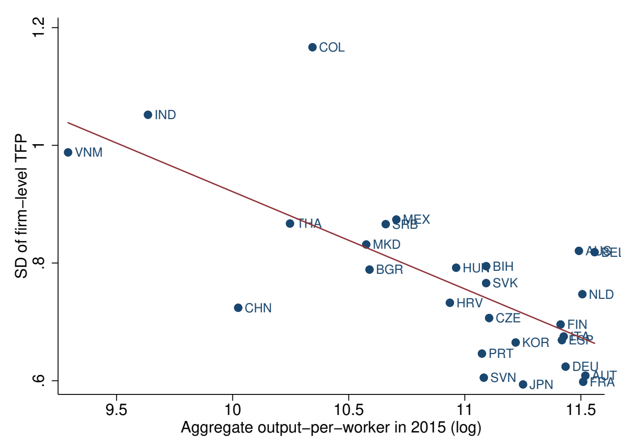


Figure B.2: Firm Productivity Distribution

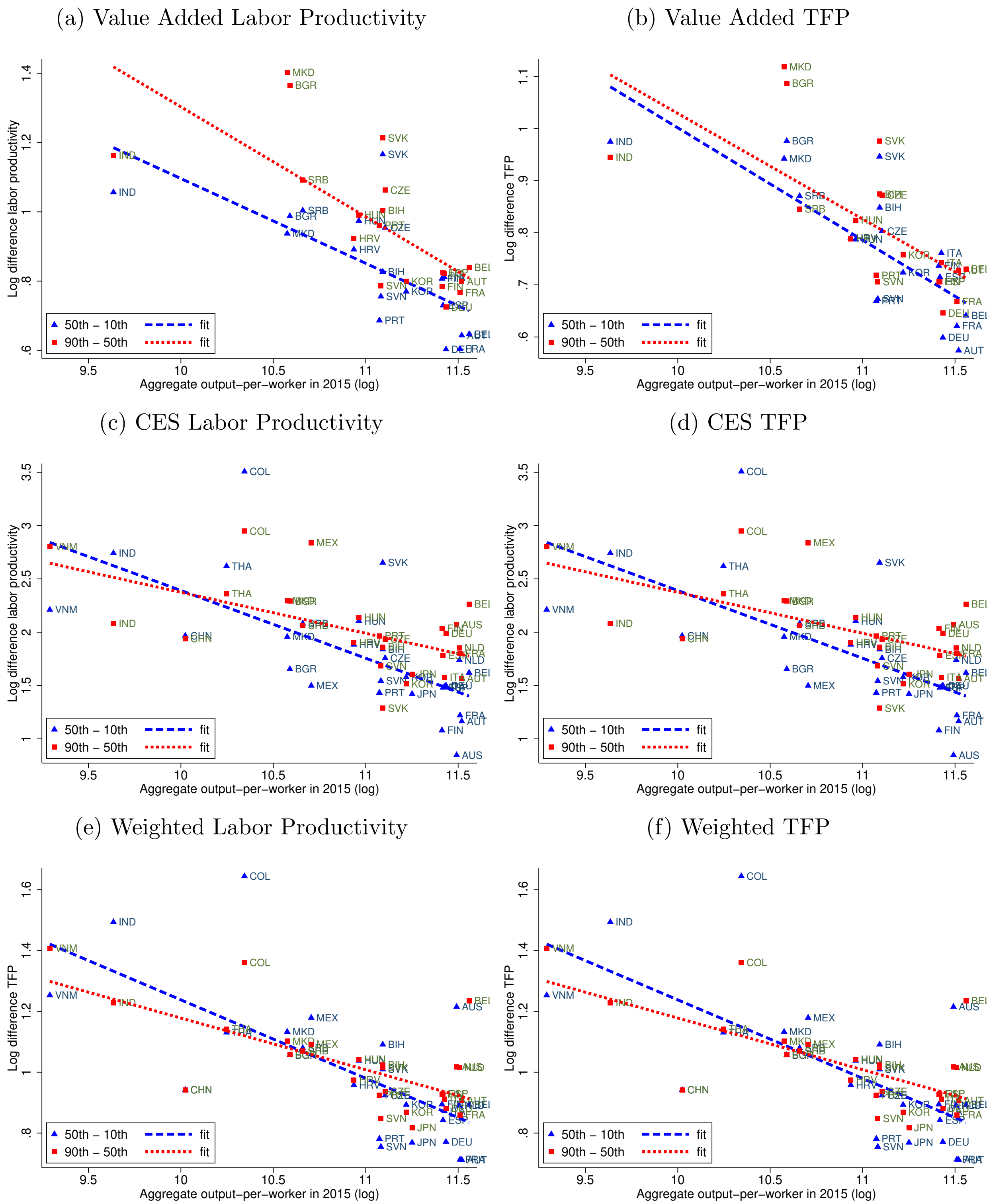
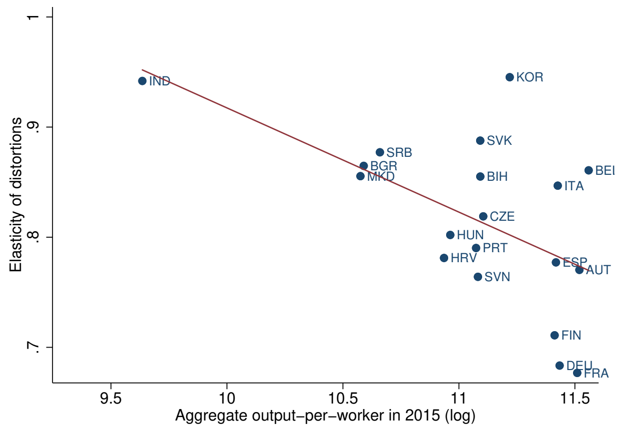
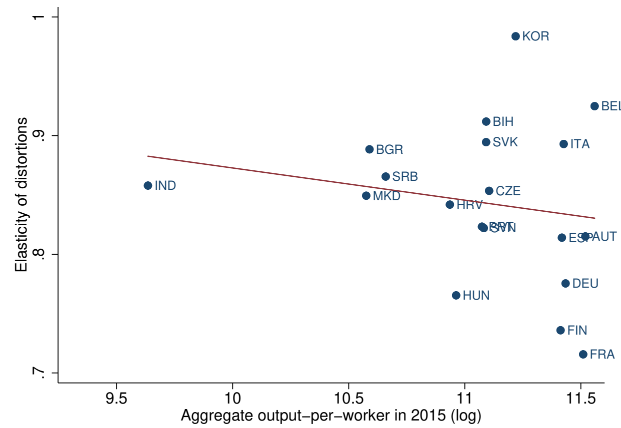


Figure B.3: Elasticity of Wedges

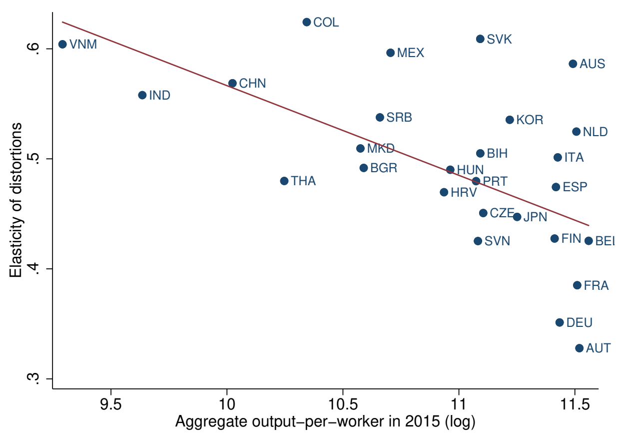
(a) Value Added Labor Productivity



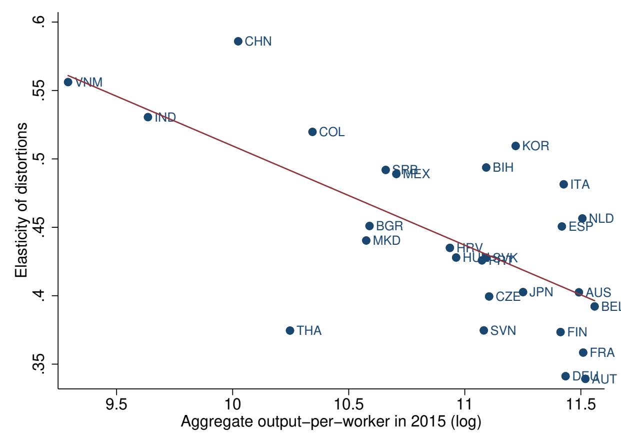
(b) Value Added TFP



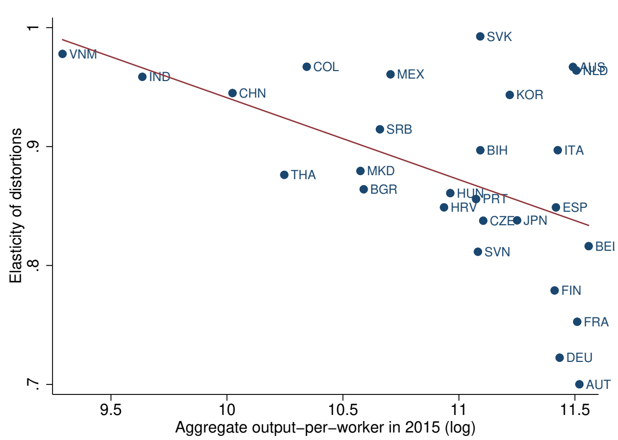
(c) CES Labor Productivity



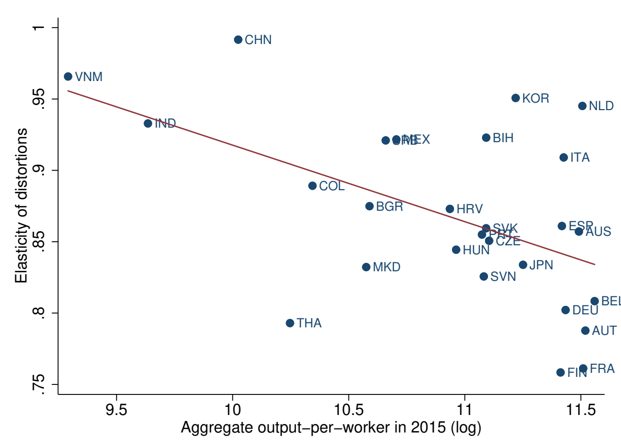
(d) CES TFP



(e) Weighted Labor Productivity



(f) Weighted TFP



C Model Details

We provide the proofs of propositions in the paper.

Proof of Proposition 1. Simply follows from the zero-profit entry condition.

Proof of Proposition 2. Productivity can be written as

$$\begin{aligned}\ln TFP_i &= (1-\gamma)\ln z_i + \ln v_i \\ &= \frac{1-\gamma}{\phi+\rho-1}\left[\ln \chi_i + \ln \epsilon_i\right] + \ln v_i.\end{aligned}$$

The standard deviation of the above is given by

$$\sigma^2_{TFP} = (1-\gamma)^2\sigma^2_{z|o} + \sigma^2_v,$$

where this follows from the first expression above and the fact that $cov(z, v) = 0$. Define a variable $\ln \check{x}$ as $\ln x - \ln \bar{x}$ where $\ln \bar{x} = \mathbb{E}\ln x$. Then, going further

$$\begin{aligned}\sigma^2_{z|o} &= \mathbb{E}\left[\frac{1}{\phi+\rho-1}(\ln \chi_i + \ln \epsilon_i) - \ln \bar{z} \mid o\right]^2 \\ &= \left(\frac{1}{\phi+\rho-1}\right)^2 \mathbb{E}\left[(\ln \check{\chi}_i + \ln \check{\epsilon}_i) \mid o\right]^2 \\ &= \left(\frac{1}{\phi+\rho-1}\right)^2 \mathbb{E}\left[(\ln \check{\chi}_i)^2 + (\ln \check{\epsilon}_i)^2 + \ln \check{\chi}_i \ln \check{\epsilon}_i \mid o\right] \\ &= \left(\frac{1}{\phi+\rho-1}\right)^2 \mathbb{E}\left[\sigma^2_{\chi|o} + \sigma^2_{\epsilon|o} + cov(\chi, \epsilon|o)\right]\end{aligned}$$

Substituting into the expression for σ^2_{TFP} confirms the result in the main text.

Proof of Proposition 3. The elasticity of the wedge with respect to firm-level labor productivity is given by

$$elas(wedge_i, TFP_i) = \frac{\mathbb{E}[\ln w\check{edge}_i \ln T\check{F}P_i]}{\sigma^2_{TFP}}$$

The numerator is equal to

$$\begin{aligned}
\mathbb{E}[\ln w\check{edge}_i \ln T\check{F}P_i] &= \mathbb{E}\left[((1-\gamma)\ln \check{z}_i + \ln \check{v}_i)(\ln \check{v}_i + \rho(1-\gamma)\ln \check{z}_i - (1-\gamma)\ln \check{\epsilon}_i)\right] \\
&= \mathbb{E}\left[\begin{array}{l} (1-\gamma)\ln \check{z}_i \ln \check{v}_i + \rho(1-\gamma)^2(\ln \check{z}_i)^2 - (1-\gamma)^2 \ln \check{z}_i \ln \check{\epsilon}_i \\ +(\ln v_i)^2 + \rho(1-\gamma)\ln \check{z}_i \ln \check{v}_i - (1-\gamma)\ln \check{v}_i \ln \check{\epsilon}_i \end{array}\right] \\
&= \rho(1-\gamma)^2\sigma^2_{z|o} + \sigma^2_v - (1-\gamma)^2 cov(z, \epsilon|o).
\end{aligned}$$

The last line follows from $\mathbb{E} \ln \check{z}_i \ln v_i = 0$ and $\mathbb{E} \ln v \ln \check{\epsilon}_i = 0$. Along with $\sigma^2_{TFP} = (1-\gamma)^2\sigma^2_{z|o} + \sigma^2_v$, the above expression confirms the result in the main text.