Misallocation and the Selection Channel*

Jeremy Majerovitz†

Paper Under Active Revision: Click Here for Latest Version

February 2023

Abstract

An important determinant of aggregate productivity is the selection channel: the process by which less efficient firms are driven out of the market by more efficient firms. Conventional wisdom suggests that markets in developing countries are more sclerotic, allowing inefficient firms to survive that would have exited in a developed country. I provide a tractable model to examine the importance of the selection channel, and show how to calibrate it to panel data on firms. I use this model to show that the effect of the selection channel on aggregate productivity is approximately equal to the average difference in log productivity between stayers and exiters, which can be measured easily in firm panel data. Results for Indonesia, Spain, Chile, and Colombia suggest that Indonesia could raise its aggregate productivity by roughly 30% if its firm exit process became as selective as Spain's. However, cross-country estimates suggest that the selection channel is not an important explanation for cross-country differences in output per capita.

Keywords: Misallocation, Development, Aggregate Productivity, Firm Survival
JEL: D22, D24, E23, O40

*I am grateful to my advisors, David Atkin, Ben Olken, and Robert Townsend, for their thoughtful guidance and advice throughout this project. I thank Mert Demirer for helping me obtain the Chilean and Colombian data, and for sharing codes to clean the data. I also thank Laura Castillo-Martínez, Ezra Oberfield, Felipe Saffie, Karthik Sastry, Pari Sastry, and seminar participants at MIT for helpful comments. The views expressed in this paper are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Federal Reserve Bank of St. Louis; Email: jeremy.majerovitz@gmail.com

1 Introduction

A popular explanation for low aggregate productivity is that unproductive firms are able to survive in poor countries. In rich countries, the argument goes, only the most productive firms survive, and unproductive firms are competed out of existence by their more efficient counterparts. This model of development as a process of survival of the fittest has potential policy implications as well. If poor countries get rich by selecting only the most productive firms to survive, then policies promoting more competition could be important to raising living standards, and policies that prop up small, inefficient firms may actually hold back growth.

Is this "selection channel" important for aggregate productivity? And if so, can it help explain the large differences in output per capita that we observe across countries?

The answers, respectively, are yes and no.

Yes, the selection channel is important for aggregate productivity. In the Indonesian manufacturing sector, I find that the selection channel raises aggregate productivity by 12%, relative to a benchmark where firm exit is independent of productivity. If exit of Indonesian firms were as selective as in Spain, aggregate productivity would rise by roughly 30%.

But no, the selection channel cannot explain large differences in productivity across countries. Although the selection channel is an important component of aggregate productivity, I find that the strength of selection does not meaningfully covary with log GDP per capita. For example, although Indonesia has weaker selection than Spain, the selection in Chile and Colombia is similar to Spanish selection, despite both being substantially poorer than Spain.

These contrasting answers reflect the relative magnitudes of the selection channel and of cross-country differences in development. Indonesia could be 30% more productive if its selection channel was as strong as Spain. Yet Spain's GDP per capita is ten times that of Indonesia. The selection channel is important, but it is not up to the task of explaining cross-country differences.

To determine the importance of the selection channel, I first build a model of heterogeneous firms with misallocation and selective exit. The model features cohort exit dynamics: the probability of firm exit is a function of the firm's productivity and age, leading to a distribution of productivity within cohorts, and a steady-state distribution of firm ages. I capture misallocation of inputs across firms with a parsimonious model of correlated distortions: firms face larger distortions the more they try to produce, creating decreasing returns to scale and limiting the economy's ability to take advantage of heterogeneous productivities. The model is parsimonious, allowing me to calibrate it to data from many countries and easily perform cross-country comparisons.

In this model, I highlight that, for a range of plausible parameter values, the selection channel will be roughly equal to the difference in log productivity between surviving firms and exiting firms. This result is especially appealing because it means that we can take a complicated macroeconomic question about aggregate productivity in general equilibrium, and reduce it to a simple, single moment in the microdata. With representative panel data on firms, we can easily implement this by simply regressing log productivity on a dummy variable for firm exit.

I then turn to analyze firm microdata from a range of rich and poor countries. I focus first on Indonesia, where the annual manufacturing census gives me an especially long panel of all formal manufacturing plants with twenty or more employees. I document a number of basic facts, including the fact that exiting plants have 12% lower labor productivity than surviving plants, controlling for sector by year fixed effects. I repeat this analysis using data from Chile and Colombia, where I also have annual manufacturing censuses, and from Spain, where I use data from Orbis.

Although selection is stronger in Spain than in Indonesia, it is similar in Spain to Chile and Colombia, hinting that selection may not be systematically correlated with development. I thus turn to data from Bartelsman et al. (2009), who compute a similar statistic to my own in harmonized manufacturing censuses across a range of countries. Combining their estimates with mine shows that selection is not systematically correlated with development. The fact that selection cannot explain cross-country differences in output per capita makes sense: the selection channel is important, but modest relative to cross-country differences in GDP per capita.

Having analyzed the stylized facts in the data, I then explicitly calibrate the model to match the data. My baseline calibration is based on Indonesia, since the long panel allows me to estimate profiles of exit rates and selection by age. However, I re-estimate key parameters in Spain, Chile, and Colombia, and recalibrate the model to each of them. The parsimony of the model shines here: the parameters of the model map straightforwardly to the data, and so I can estimate and re-estimate for many countries.

I find that the back-of-the-envelope approximation (the selection channel equals the difference in log productivity between surviving firms and exiting firms) is very accurate, regardless of the calibration. This shows that my micro-level findings about selection patterns in different countries are also findings about the selection channel and its effect on aggregate productivity. Aggregate productivity depends on the strength of the selection, but cross-country differences in selection are not large enough to explain any sizable component of cross-country differences in output.

I also find an interesting interaction between the selection of which firms survive and the allocation of inputs across firms. I find that selection and allocation are substitutes: if inputs are sufficiently well allocated across firms, then the selection channel will be weakened. Intuitively, wants to not allocate inputs to unproductive firms; it can accomplish this by killing off these firms (selection) or by starving them of inputs (allocation). However, for the parameter values that I estimate in different countries, this interaction is not strong enough to meaningfully affect the validity of the back-of-the-envelope approximation.

Related Literature This study is related to a body of research on the importance of misallocation in explaining cross-country differences in aggregate productivity (see Restuccia and Rogerson 2008; Hsieh and Klenow 2009, 2014). Relative to this literature, I focus on the contribution of selective exit to aggregate productivity, and study how this interacts with static misallocation induced by distortions. This study follows a long line of research into models of firm dynamics and selection (Jovanovic, 1982; Hopenhayn, 1992; Cooley and Quadrini, 2001; Luttmer, 2007), as well as more recent research that integrates misallocation into models of firm dynamics (Bento and Restuccia, 2017; Acemoglu et al., 2018; Castillo-Martinez, 2020; Peters, 2020; Kochen, 2022; Asturias et al., 2023).

This study is most closely related to three recent papers that use a range of models and strategies to study the interaction of selection and misallocation. Yang (2021) incorporates an extensive margin into a model of misallocation, modeling the decision to operate as a one-shot game. Calibrating to Indonesian data, he finds that the cost of misallocation is over 40% larger than it would have been without an extensive margin.

Fattal Jaef (2018) studies a model of creative destruction under correlated distortions, and calibrates it to US data. He conducts quantitative experiments in which the economy begins at a distorted steady-state (with the degree of initial misallocation corresponding to estimates from Hsieh and Klenow, 2009) and then distortions are removed. He shows that a full accounting of the welfare benefits of removing distortions requires examining the transition path: the net present value of welfare gains can be twice as high as the rise in steady-state aggregate productivity.

Peters and Zilibotti (2021) examine a model of creative destruction, where firms can be run either by "subsistence entrepreneurs" who never innovate, or "transformative entrepreneurs" who have the skills to innovate and expand their business. They calibrate their model to three moments of the data in India and the United States: the entry rate of new firms, the share of firms that are small, and the relative size of older versus younger firms. Based on the fact that India has more small firms and slower growth in firm size (the entry rate is roughly 8% in both countries), they conclude that a smaller share of entrants in India are transformative, and that the transformative entrepreneurs in India are less efficient at innovation than their counterparts in the United States.

My study is complementary to these papers, but differs from them in a few respects. I model exit and firm dynamics differently than these papers. Yang (2021) models entry/exit as a one-shot game; I model exit as a dynamic process over the firm's life cycle. Fattal Jaef (2018) and Peters and Zilibotti (2021) follow a tradition of creative destruction models in which only the most productive firm will produce a given product line, and firms exit when they run out of product lines. In my model, firms will continue to produce, even if they are less productive than their competitors: this may be a more realistic description of markets, especially in developing countries where many unproductive firms sell similar products.

Beyond making different modeling choices, two key aspects of my paper distinguish it from the previous literature. First, I focus my calibration on the difference in productivity between exiting and surviving firms. I show that this is the critical moment for understanding the effect of the selection channel on aggregate productivity. This establishes a tight connection between the moments used to calibrate the model and the quantitative conclusions drawn. In the language of Nakamura and Steinsson (2018), I calibrate the model parameters using "identified moments," rather than unconditional moments like the size distribution of firms.

Second, I show that this moment is an approximately-sufficient for the strength of the selection channel, within the empirically relevant range of parameters. Crucially, this allows me to make comparisons across many countries, rather than just one or two. This is an important piece of the contribution: being able to measure the selection channel in many countries reveals that the selection channel is not meaningfully correlated with development, and allows me to conclude that it is not a quantitatively important contributor to cross-country differences in development.

Outline Section 2 of the paper presents the model, and highlights the back-of-the-envelope approximation. Section 3 introduces the data and the key stylized facts. Section 4 calibrates the model to the Indonesian data, as well as recalibrating it to data from Chile, Colombia, and Spain. Section 5 provides the results from the calibrated model. Section 6 concludes.

2 Model

How does the selective exit of firms affect aggregate productivity? To answer this question, I compare aggregate productivity under prevailing exit patterns to what aggregate productivity would have been if exit rates by age were held constant, but exit were random (uncorrelated with productivity). To provide intution, I begin by providing a back-of-the-envelope approximation; in Section 5 I find that this back-of-the-envelope is a very accurate approximation of the selection channel in the full model.

Back-of-the-Envelope Approximation

Much of the basic insight can be captured by a back-of-the-envelope approximation to the selection channel. Rather than focusing on aggregate productivity, we instead focus on how the average log productivity is affected by selection. The annual effect of selection on average log productivity is equal to the exit rate times the difference in average log productivity between survivors and exiters, or $\rho \times (\mu_{\text{Survivors}} - \mu_{\text{Exiters}})$. The total effect on average log productivity is the annual effect multiplied by the average age. We can simplify even further by noting that, if the exit rate is constant across ages, then the average age is $(1-\rho)/\rho$. If we assume ρ is not that large, then we have:

$$\begin{aligned} \text{Selection Channel} &\approx \underbrace{\overbrace{\rho}^{\text{Exit Rate}} \times (\mu_{\text{Survivors}} - \mu_{\text{Exiters}})}_{\text{Annual Selection Effect}} \times \text{Average Age} \\ &\approx (1-\rho) \times (\mu_{\text{Survivors}} - \mu_{\text{Exiters}}) \\ &\approx (\mu_{\text{Survivors}} - \mu_{\text{Exiters}}) \end{aligned}$$

Thus, the back-of-the-envelope approximation tells us that the selection channel is approximately equal to the difference in mean log productivity between survivors and exiters, a quantity that can be easily measured in the data by regressing log productivity on exit.

There are three main reasons why effect on aggregate productivity will differ from this back-of-the-envelope approximation. First, the approximation assumes that the exit rates and selection are the same at each age. If selection is stronger earlier then that would increase productivity, because cohorts would reach high productivity earlier and enjoy the benefits for longer. If mortality rates are higher earlier and then decline, then that will raise the average age, again increasing the strength of the selection channel. Second, the back-of-the-envelope focuses on the effect on the mean of log productivity, which is not the same thing as the log of mean productivity. Finally, and perhaps most importantly, mean productivity is not the same thing as aggregate productivity: markets will allocate more inputs to more productive firms, thus the average productivity will be higher than the mean. Later in the paper, I explore the implications this has for the selection channel, and whether allocation is a complement or substitute for selection.

Model Overview

To examine the importance of these issues, and to test the accuracy of the back-of-the-envelope approximation, I now turn to lay out a complete model of aggregate productivity. I begin by laying out a model with heterogeneous firms, that differ in their productivity and in their age. There is a continuum of firms corresponding to each cohort, but time (and thus age) is discrete. Firms enter and exit, and the probability of exit depends on age and productivity, inducing a selection channel that affects aggregate productivity. I study the stationary equilibrium of this economy: prices are constant, the joint distribution of age and productivity is constant, and entry and exit are equated. My primary interest is aggregate productivity, and how it is affected by firm's survival dynamics, i.e. the selection channel.

2.1 Firm's Static Problem

Each firm has a productivity, z, and an age, a. Firms produce homogeneous goods, using a linear production function with labor as their input. Specifically, production is given by:

$$q = zl$$

where q is output and l is labor.

The firm also faces a revenue wedge, which can equivalently be viewed as a revenue tax. The firm's profits are reduced by τ times revenue, and this wedge is increasing in the quantity produced. From the tax perspective, this is a progressive tax on revenue: firms with higher production face a higher marginal tax rate. The tax rate is given by:

$$1 - \tau(q) = \frac{c_\tau}{1-\gamma} q^{-\gamma}$$

where c_τ and γ are parameters, and $0 < \gamma < 1$. The parameter γ governs the slope of the wedge with respect to quantity, or, equivalently, the progressivity of the revenue tax. This increasing tax is similar to that in Bento and Restuccia (2017) and Hsieh and Klenow (2014). This generates "correlated distortions," where the most productive firms are the ones who face the highest wedges on the margin; Restuccia and Rogerson (2008) note that correlated distortions are much more important for misallocation than uncorrelated distortions.

I normalize the price of output to 1. The firm solves the static profit maximization problem:

$$\begin{aligned}\pi(z) &= \max_l \left(1 - \tau(q)\right) q - wl \\ &= \max_q \left(1 - \tau(q)\right) q - \frac{w}{z} q\end{aligned}$$

where w is the wage. The firm's solution is:

$$\begin{aligned}q(z) &= \left(\frac{c_\tau z}{w}\right)^{\frac{1}{\gamma}} \\ l(z) &= \frac{q(z)}{z} \\ \pi(z) &= \frac{\gamma}{1-\gamma} c_\tau^{\frac{1}{\gamma}} \left(\frac{z}{w}\right)^{\frac{1-\gamma}{\gamma}}\end{aligned}$$

Note the important role played by the wedge slope parameter γ. Taking logs, one can see that $\log q = \frac{1}{\gamma}\left(\log z + \log \frac{c_t}{w}\right)$, and thus γ is the inverse elasticity of output with respect to productivity. The parameter γ creates decreasing returns to scale, ensuring that each firm has a finite labor and output as the solution to their static problem. Moreover, it creates a distortion that induces misallocation.

2.2 Aggregate Productivity within Cohort

Within each cohort, I assume a continuum of firms, and a parametric distribution of productivity. I begin by assuming that productivity is log-normally distributed. I will later switch to a truncated log-normal distribution, where the same results will go through but with a correction term. Since the essence of the results is the same, I present the untruncated distribution to aid in the clarity of the exposition.

Let productivity for a given cohort, a, be distributed as follows:

$$\log z \mid a \sim N\left(\mu_a, \sigma^2\right)$$

where μ_a is the (age-specific) mean of log productivity, and σ^2 is the variance. From the solution of the firm's static problem, I have that output and labor are log-linear in the firm's productivity. Solving for aggregate productivity, I have:

$$\log\left(Z_a\right) := \log\left(\frac{\mathbb{E}\left[q \mid a\right]}{\mathbb{E}\left[l \mid a\right]}\right) = \underbrace{\mu_a + \frac{1}{2}\sigma^2}_{\log \mathbb{E}[z]} + \underbrace{\left(\frac{1}{\gamma} - 1\right)\sigma^2}_{\text{Benefits of Allocation}}$$

where Z_a is the aggregate productivity of a given cohort, defined as the quantity produced by that cohort divided by the labor employed by that cohort. The first two terms, $\mu_a + \frac{1}{2}\sigma^2$ reflect the average productivity of firms in the cohort. If all firms received equal inputs, then aggregate productivity would equal average productivity. However, the last term, $\left(\frac{1}{\gamma} - 1\right)\sigma^2$ reflects the benefits of allocation. Because more labor is allocated to more productive firms, aggregate productivity is higher. When γ is lower, the benefits of allocation are higher, and output is higher. As $\gamma \to 1$, the allocation of inputs converges to be uniform across firms, and the benefits of allocation disappear. As $\gamma \to 0$, all inputs are concentrated on the most productive firm. With the log-normal distribution of productivity, there is no upper bound on productivity, and so aggregate productivity explodes; with the truncated log-normal distribution, productivity converges to the upper bound.

Note also the important role played by σ^2. Holding μ_a constant, a higher σ^2 raises aggregate productivity. Part of this is mechanical: a higher σ^2 raises the average productivity of firms. But part of this also operates through the allocation channel: a higher σ^2 means more variance of (log) productivity, and thus more scope for benefits from allocating inputs to the most productive firms.

Finally, note that aggregate productivity does not depend on the average wedge or on the wage, and thus c_τ and w do not feature in the expression for aggregate productivity. This convenient result is a consequence of assuming constant returns to scale in the production function. If the firms did not have constant returns to scale, then aggregate productivity would also have a scale term: e.g. if firms had decreasing returns to scale then aggregate productivity would be higher if the average firm were smaller, while if firms had increasing returns to scale then the scale term would be increasing in the average firm size.

For the main model that I take to the data, I will instead assume a truncated log-normal distribution: the distribution of productivity is log-normal, but is truncated so that $\bar{z}$ is the upper bound on productivity. The results for the truncated log-normal are very similar, but with a truncation correction term: Appendix B shows how truncation modifies the untruncated formulas.

2.3 Exit Dynamics

At the end of each year, some firms exit the economy. I assume that a firm's probability of exit is a function of their age and their productivity, and is given by

$$1 - \Pr(\text{exit} \mid z, a) = \exp(c_{\text{exit},a}) \cdot z^{\delta_a}$$

where $\Pr(\text{exit} \mid z, a)$ is the probability of exit, and $c_{\text{exit},a}$ and δ_a are parameters that vary with age. Note that if $\delta_a > 0$, then more productive firms are more likely to survive than less productive firms.

This functional form is convenient, because it ensures that (truncated) log-normality is preserved. If $\log z \sim N(\mu_a, \sigma^2)$, with truncation at an upper bound of $\bar{z}$, then the distribution of productivity for surviving firms is given by $\log z \mid \text{survival} \sim N(\mu_a + \delta_a\sigma^2, \sigma^2)$, with truncation at $\bar{z}$ once again. Note that the exit function is the reason I must use a truncated distribution of productivity: if there were no upper bound to productivity, then sufficiently productive firms would have a negative probability of exit. An appropriately selected upper bound, $\bar{z}$, ensures this pathological behavior does not occur.

There are a variety of reasons why exit can depend on productivity. Foster et al. (2008) provide one microfoundation, in which exit is a function of expected profitability, and firms exit if they have negative net present value. If productivity is the only state variable, then firms will play a cutoff strategy, but if other things vary across firms (e.g prices and demand for their goods, operating costs, etc.) then this relationship will be continuous. Other microfoundations for exit are also possible, e.g. firms could exit when they are overly indebted or when they run out of cash. I do not take a stand on the microfoundation for why survival probabilities depend on productivity and age, but instead I merely note it as a robust feature of the data, and I treat it as a reduced form stand-in for the underlying structural process. My results are robust to various microfoundations of exit, as long as those microfoundations do not alter the firm's static problem.

2.4 Productivity Dynamics

For entrants, I assume that firms' initial productivity follows a truncated log-normal distribution, with mean and variance parameters μ_0 and σ^2 and with upper bound $\bar{z}$. Moreover, I assume that an individual firm's productivity does not change over time: for a given firm, $z_a = z_0 \forall a$.

Given the functional form of exit, it then follows that each successive cohort inherits the truncated log-normality of the previous cohort. Thus, we have:

$$\log z_a \sim N\left(\mu_a, \sigma^2\right)$$
$$\mu_a = \mu_0 + \sum_{s=0}^{a-1} \delta_s \sigma^2$$

with an upper bound at $\bar{z}$ under the truncated log-normal distribution.

Although I do not consider more complex processes for productivity, I expect that mean reversion will tend to attenuate the effect of selection on aggregate productivity. If survival is based on today's productivity, but today's productivity is only moderately predictive of tomorrow's productivity, then the selection channel will be weaker than in a world with constant productivity.

2.5 Closing the Model

To solve this model in general equilibrium, I would need to specify the labor supply function of the household and model the entry decision faced by new firms. For example, I could specify that labor supply is inelastic and normalized to 1, and that firms face some entry cost and enter until the expected net present value of the firm is equal to the entry cost. This would allow me to solve for the equilibrium wage and firm size. Identifying these entry and exit models, however, can be very difficult or even impossible: dynamic discrete choice models, in which there are entry costs, stochastic operating costs, and scrap values are typically not identified, and many counterfactuals with these models are also not identified (Kalouptsidi et al., 2021).

Fortunately, I am interested in aggregate productivity, and solving for aggregate productivity does not require solving the full model. Thanks to constant returns to scale, aggregate productivity does not depend on firm size. Thus, I do not need to solve for the wage. I next turn to compute aggregate productivity.

2.6 Aggregate Productivity

To solve for aggregate productivity, I need to take a labor-weighted average of cohort level productivities. I have already shown the expression for aggregate productivity within cohort; I now derive the expression for each cohort's share of labor. I will show the results under the assumption of a log-normal productivity distribution; the results for the truncated log-normal are similar but with a correction term, and are provided in Appendix B. Note that the average firm size in a cohort is:

$$\log \mathbb{E}\left[l \mid a\right] = \frac{\mu_a}{\gamma}\left(1-\gamma\right) + \frac{\left(1-\gamma\right)^2}{2\gamma^2}\sigma^2 + \frac{1}{\gamma}\log\left(\frac{c_\tau}{w}\right)$$

Multiplying by the number of firms in the cohort, N_a, and dividing by aggregate labor, we have the share of labor employed by the cohort:

$$\frac{N_a \cdot \mathbb{E}\left[l \mid a\right]}{\sum_s N_s \cdot \mathbb{E}\left[l \mid s\right]} = \frac{N_a \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_a\right)}{\sum_s N_s \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_s\right)}$$

Let ρ_a denote the exit rate of firms in cohort a. In a stationary equilibrium, $N_a = \left(1-\rho_{a-1}\right) N_{a-1} = N_0 \cdot \prod_{t=0}^{a-1}\left(1-\rho_t\right)$, where, for notational convenience, I define $\prod_{t=0}^{a-1}\left(1-\rho_t\right)$ to be 1 when $a=0$.

Thus, aggregate productivity is characterized as follows:

$$
\begin{aligned}
Z &= \frac{\sum_a Z_a \cdot N_a \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_a\right)}{\sum_s N_s \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_s\right)} \\
&= \frac{\sum_a \prod_{t=0}^{a-1}(1-\rho_t)\cdot \exp\left(\frac{1}{\gamma}\mu_a + \frac{1}{2}\sigma^2 + \left(\frac{1}{\gamma}-1\right)\sigma^2\right)}{\sum_s \prod_{t=0}^{s-1}(1-\rho_t)\cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_s\right)}
\end{aligned}
$$

We can thus compute aggregate productivity in the model without solving for the wage or for the equilibrium firm size or number of firms. To compute aggregate productivity in the stationary equilibrium, it is sufficient to know the within-cohort variance of log productivity, σ^2, the inverse elasticity of output with respect to productivity, γ, the exit rates by age, ρ_a, and the life-cycle of mean log productivity, μ_a. The mean log productivity by age can be computed using the previously described dynamics for productivity in subsection 2.4.

2.7 The Selection Channel

I define the selection channel as the difference between aggregate productivity under prevailing exit patterns and what aggregate productivity would be if exit rates by age were held constant, but exit were uncorrelated with productivity conditional on cohort. More formally:

Definition. *I define the selection channel as*

$$
\textit{Selection Channel} := \log Z - \log Z^*
$$

where Z^* is aggregate productivity under an economy where all parameters are the same as in the main economy, except that the new probability of exit is given by

$$
Pr^*\left(exit \mid z, a\right) = \mathbb{E}\left[Pr\left(exit \mid z, a\right) \mid a\right]
$$

that is, the exit rates by age remain the same, but the probability of exit and productivity are now independent of productivity, conditional on cohort.

The selection channel can be easily computed in a fully calibrated model. However, to provide further intuition, I first examine a simplified parametrization of the full model, in which exit patterns do not vary with age.

Simplified Model Approximation

To simplify, I examine a special case of the full model. Assume that ρ and δ are the same for each age. Ignoring the truncation issue, with log-normally distributed productivity:

$$
\begin{aligned}
Z &= \frac{\sum_a \prod_{t=0}^{a-1} (1-\rho_t) \cdot \exp\left(\frac{1}{\gamma}\mu_a + \frac{1}{2}\sigma^2 + \left(\frac{1}{\gamma}-1\right)\sigma^2\right)}{\sum_s \prod_{t=0}^{s-1} (1-\rho_t) \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_s\right)} \\
&\approx \frac{\int_0^\infty \exp(-\rho a) \cdot \exp\left(\frac{1}{\gamma}(\mu_0 + a \cdot \delta\sigma^2) + \frac{1}{2}\sigma^2 + \left(\frac{1}{\gamma}-1\right)\sigma^2\right) da}{\int_0^\infty \exp(-\rho a) \cdot \exp\left(\frac{1-\gamma}{\gamma}(\mu_0 + a \cdot \delta\sigma^2)\right) da} \\
&= \exp\left(\mu_0 + \frac{1}{2}\sigma^2 + \left(\frac{1}{\gamma}-1\right)\sigma^2\right) \cdot \left(1 + \frac{\delta\sigma^2}{\rho - \frac{1}{\gamma}\delta\sigma^2}\right) \\
\implies \log Z &\approx \underbrace{\mu_0 + \frac{1}{2}\sigma^2}_{\log \mathbb{E}[z|a=0]} + \underbrace{\left(\frac{1}{\gamma}-1\right)\sigma^2}_{\text{Benefits of Allocation}} + \underbrace{\log\left(1 + \frac{\delta\sigma^2}{\rho - \frac{1}{\gamma}\delta\sigma^2}\right)}_{\text{Selection Channel}}
\end{aligned}
$$

This is the same as the previous formula for aggregate productivity at the cohort level, but with the addition of a new term, representing the selection channel. Thus, noting that $\delta\sigma^2 = \rho \times (\mu_{\text{Survivors}} - \mu_{\text{Exiters}})$, the selection channel is:

$$
\begin{aligned}
\text{Selection Channel} &\approx \log\left(1 + \frac{\delta\sigma^2}{\rho - \frac{1}{\gamma}\delta\sigma^2}\right) \\
&= \log\left(1 + \frac{\mu_{\text{Survivors}} - \mu_{\text{Exiters}}}{1 - \frac{1}{\gamma}(\mu_{\text{Survivors}} - \mu_{\text{Exiters}})}\right)
\end{aligned}
$$

which, for γ close to 1 and for $(\mu_{\text{Survivors}} - \mu_{\text{Exiters}})$ small, is quite close to the back-of-the-envelope approximation.[1]

This approximation, like the back-of-the-envelope approximation, makes a myriad of assumptions, and the full model is necessary for a proper estimate of the selection channel. However, the two approximations establish a strong benchmark: in simple models, the selection channel is roughly equal to the difference in mean log productivity between survivors and exiters.

[1] I have assumed log-normal productivity as a useful simplification, but in order to ensure that the probability of exit remains non-negative, we must instead use a truncated log-normal. This will yield the same formulas as the log-normal, plus a correction term. As the term $\frac{1}{\gamma}(\mu_{\text{Survivors}} - \mu_{\text{Exiters}})$ grows large, the correction term becomes important. The fact that $-\frac{1}{\gamma}$ appears in the denominator is thus misleading: although the simplified model approximation makes it appear that the selection channel will be larger if γ is closer to zero, we will see later that in fact the selection channel is weaker when γ is closer to zero.

2.8 Extensions: Allowing for Growth

A natural extension would be to allow for growth in the model. This can be population growth, or uniform growth in the productivity of all firms. If productivity growth is uniform (all firms' productivity is raised proportionally by the same amount), then it has no effect on the model: everything in the model is log-linear, and so productivity can grow as long as $c_{\text{exit},a}$ adjusts accordingly to keep exit rates, ρ_a, constant.

Population growth requires a minor but straightforward tweaking of the model. Suppose that the population grows, annually, by a constant ratio $1 + g$. It follows that, in order to maintain a stationary equilibrium, where firm size is constant, the number of entrants must also grow by the same constant ratio. Thus, everything else in the model remains the same, but instead of $N_a = N_0 \cdot \prod_{t=0}^{a-1} (1 - \rho_t)$, we have $N_a = N_0 \cdot \prod_{t=0}^{a-1} \frac{1-\rho_t}{1+g}$. Population growth shifts the firm age distribution to be younger.

3 Empirics

Having laid out the model, I now begin to analyze the data. In this section, I introduce the relevant data sets and provide descriptive statistics. This lays the groundwork for Section 4, where I use the data to calibrate the model.

3.1 Indonesian Data

For the main analysis, I use data from 1975-2012 from Indonesia's Annual Manufacturing Survey. This survey provides annual panel data on all formal manufacturing establishments in Indonesia with twenty or more workers. In some early years of the survey, some establishments with fewer than twenty workers were included: I drop these observations to maintain consistency across years, and because the reason for including some smaller establishments in early years is not well documented. Although this is technically an establishment suvey, I will use the term "firm" throughout to refer to an observation.

The analysis requires me to measure a few key variables. My main measure of productivity is labor productivity, or value added per worker. I construct value added as total output minus total inputs and expenses. I use the definitions favored by the manufacturing survey in order to define these quantities. Total output is defined as the sum of value of goods produced, revenue from electricity sold, revenue from industrial services, change in inventory of semi-finished products, and profit from the sale of unprocessed goods, non-manufacturing services, and scrap waste. Total inputs and expenses is defined as the sum of electricity purchased, fuel and lubricant purchased, materials, and other expenses, excluding land rent, taxes, interest on loans, and donations. This definition of inputs does not include labor costs. I also construct a measure of total factor productivity using data on both labor and capital; I measure capital as fixed capital. However, this is not my main measure of productivity because data on capital is not available for all years or is reported as zero (which I interpret as missing) for many firms. The manufacturing survey also provides five digit industry codes, which I use in some parts of the analysis to generate industry-year fixed effects.

I also construct a measure of age and exit using the manufacturing survey. I consider a firm to have exited if the firm leaves the sample the following year and does not ever return to the sample. As a result, I do not observe exit for 2012, the last year of the data. I measure age as the number of years since the first time the firm is observed in the data, as a result, I cannot reliably measure age for the firms that are already present in 1975, when the data begins.[2] However, since the data span 38 years, I am able to use a wide range of ages. For example, I am able to measure exit rates for firms from age zero to thirty five (one year is dropped because I can only reliably measure age for firms that enter after 1975, and another year is dropped because I cannot measure exit in 2012).

3.2 Descriptive Statistics in Indonesia

In this subsection I summarize the key descriptive statistics for the Indonesian data. I will focus on the sorts of statistics that will eventually be key for calibrating the model: the distribution of productivity, the rate of exit, and the relationships between exit, productivity, age, and firm size.

Summary Statistics

I begin by providing summary statistics on some of the key variables of interest. Table 1 provides the mean, standard deviation, median, and 10th and 90th percentiles for five key variables: exit, age, log value added per worker (my main measure of productivity), log TFP per worker (I construct TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$), and the log number of workers. For both measures of log productivity, I demean the values by year before computing statistics: the mean is thus mechanically zero. Some variables are missing for some firms or some years: age is missing for firms that were already present in 1975, because I construct age based on how long the firm has been in the sample. Exit is missing for 2012, because it is the last year of the data. Other variables are missing for particular firms: in particular, data on capital is not available in all years, and even in years where the information is available, capital is recorded as zero for many firms. Regardless of whether these are true zeroes for capital or simply missing data, the result is that TFP will be missing since capital is in the denominator.

A few statistics are worth noting. The exit rate is 7.6%, implying frequent firm turnover: if exit rates did not vary with age, then firm life expectancy, and also the average firm age in steady-state,

[2]Unfortunately, the Indonesian manufacturing survey does not have reliable information on the birth year of establishments, which is why I rely on the first year in which they are observed in the sample.

	Mean	Standard Deviation	10th Percentile	Median	90th Percentile	Observations
Exit	0.076	0.265	0	0	0	610596
Age	7.068	6.691	0	5	17	537877
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	0	1.296	-1.424	-0.115	1.669	627829
log (Value Added TFP)	0	1.107	-1.199	-0.107	1.401	341305
log (Workers)	4.164	1.153	3.091	3.761	5.914	630777

Table 1: Summary Statistics for Indonesia

Notes: This table shows summary statistics for the Indonesian data. I define Value-added TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$. I demean log productivity by year, thus the mean is mechanically equal to zero. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

would be $(1-\rho)/\rho \approx 12$ years.[3] Yet this is above the average age in the data: this is mainly because I cannot compute age for the incumbent firms in 1975, and so I am missing much of the full age distribution, although it is also partly driven by growth in the annual number of entrants. The median number of workers (exponentiating the log) is 43, the 10th percentile is 22, and the 90th percentile is 370.

The percentiles of the log productivity distributions imply that the distribution of productivity is not quite lognormal: it is slightly skewed and more heavy tailed than the log-normal. In Appendix Figure A.1, I plot the log productivity distribution using kernel density estimates. Productivity is not quite lognormal, but it appears to be fairly close (the long tails may also be the result of measurement error, rather than a true deviation from lognormality).

Exit, Productivity, Age, and Size

Since my model treats age and productivity as the key state variables, I next show how outcomes depend on age and productivity. I control for industry-year fixed effects throughout, unless otherwise noted, in order to focus on variation within a given sector, at a given time. Formally, I run the regression:

$$Y_{it} = \alpha_{st} + X_{it}'\beta + \varepsilon_{it}$$

where Y_{it} is the outcome of interest for establishment i in year t, α_{st} denotes industry-year fixed effects (I use five digit industries), and X_{it} denotes the regressors of interest (usually there is only one regressor of interest, but in some regressions I include both age and productivity). Binned scatter plots control for industry-year fixed effects by partialling these fixed effects out of both the outcome and regressor before forming bins and plotting. All standard errors are clustered at the firm level. In the process, I replicate some well-known stylized facts: exit is decreasing in age and

[3]The fact that life expectancy and the steady state average age are the same may be puzzling to some readers, who are used to intuitions from humans, where the average age is typically lower than life expectancy. In addition to population growth, which lowers the average age, humans have rising mortality rates with age, which leads life expectancy to be longer than average age. With constant mortality rates, the two coincide, while if mortality rates fall with time (the relevant case for firms), the average age will actually be above life expectancy.

in productivity, and more productive and older firms employ more labor.

I begin with age. In Appendix Table A.1, I show how exit, productivity, and the number of workers depend on firm age. Exit declines with age: a firm that is ten years older is 2.6 percentage points less likely to exit (compare to the average exit rate of 7.6%). This relationship is depicted graphically in Appendix Figure A.2. Productivity is increasing in age, as is the log number of workers. Because of non-random exit, the estimated effect of age on productivity and on the number of workers mixes a direct effect of age with a selection effect: more productive, larger firms are more likely to survive.

I shows how exit and the number of workers depend on productivity in Appendix Table A.2. The results are quite similar whether productivity is measured as labor productivity (value-added per worker) or as total factor productivity. More productive firms employ more labor: the elasticity of labor to productivity is roughly 0.2 for both measures of productivity. I show this relationship graphically in Figure 1; in order to allow for easy comparison on the same graph, I flip the regression to make productivity the outcome variable.

More productive firms are also less likely to exit. I show this relationship graphically in Figure 2. Although the negative relationship between exit and productivity is robust and highly statistically significant, it is far from the sharp cutoff relationship predicted by simple theories of exit. A 100 log point increase in labor productivity leads to a 0.78 percentage point decline in exit (0.98 for TFP). Doubling a firm's productivity (a 69 log point increase) will only reduce its exit probability from 7.6% (the average exit rate) to 7.1%. Clearly, exit is only modestly correlated with productivity.

Finally, in Appendix Table A.3 I combine age and productivity in the same regression. The same patterns hold from before, and the point estimates are not changed much from the single variable regressions: although age and productivity both effect the outcomes of interest, the cross-sectional correlation of age and productivity is not that strong, and so the omitted variable bias is not too severe.

Estimates of β_{exit}

I finish the empirical analysis of Indonesia in this section by focusing on β_{exit}, the coefficient from a regression of log productivityon exit. As highlighted in Section 2, this parameter is crucial to the calibration of the model, and the selection channel will be approximately equal to β_{exit} under the back-of-the-envelope approximation. Here, I provide estimates of the parameter under different measures of productivity and under different specifications. My main specification is

$$\log\left(\text{Productivity}_{it}\right) = \alpha_{st} + \beta_{\text{exit}} \cdot \text{Exit}_{it} + \varepsilon_{it}$$

where $\log\left(\text{Productivity}_{it}\right)$ is a measure of productivity (my main measure is value added per worker) at establishment i in year t, α_{st} is an industry-year fixed effect (I use five digit industries) and Exit_{it} is an indicator for exit that is equal to one in the last year that the establishment

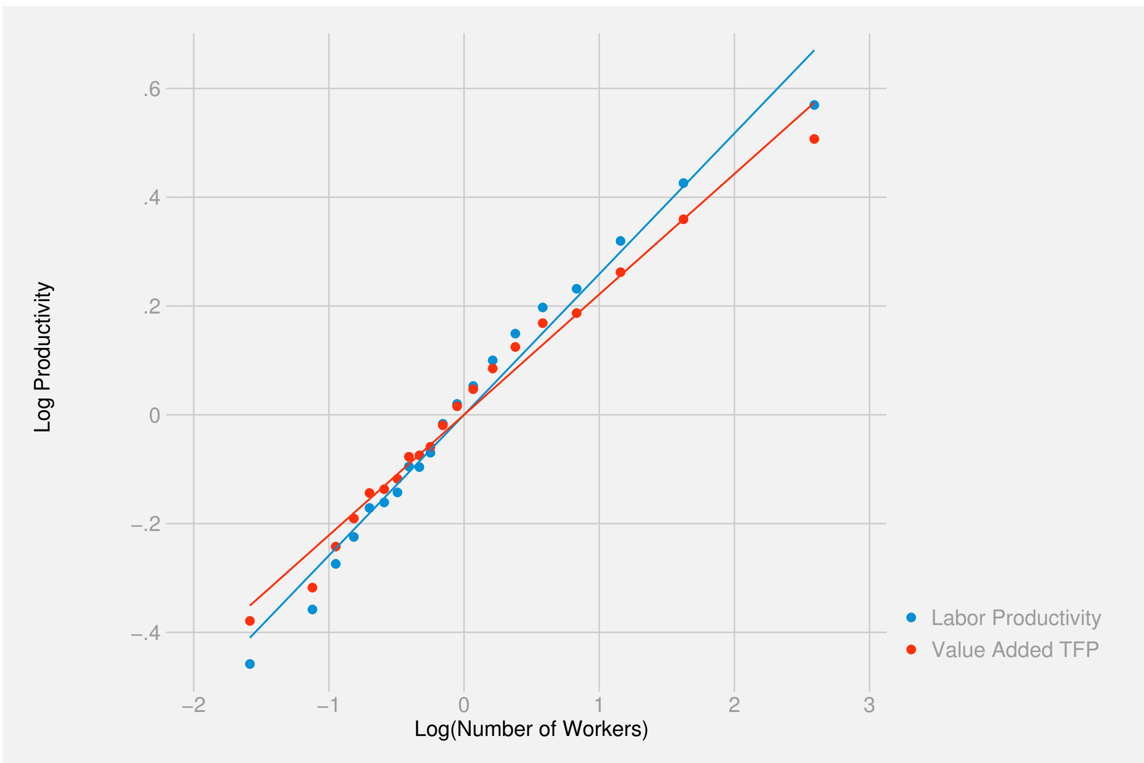
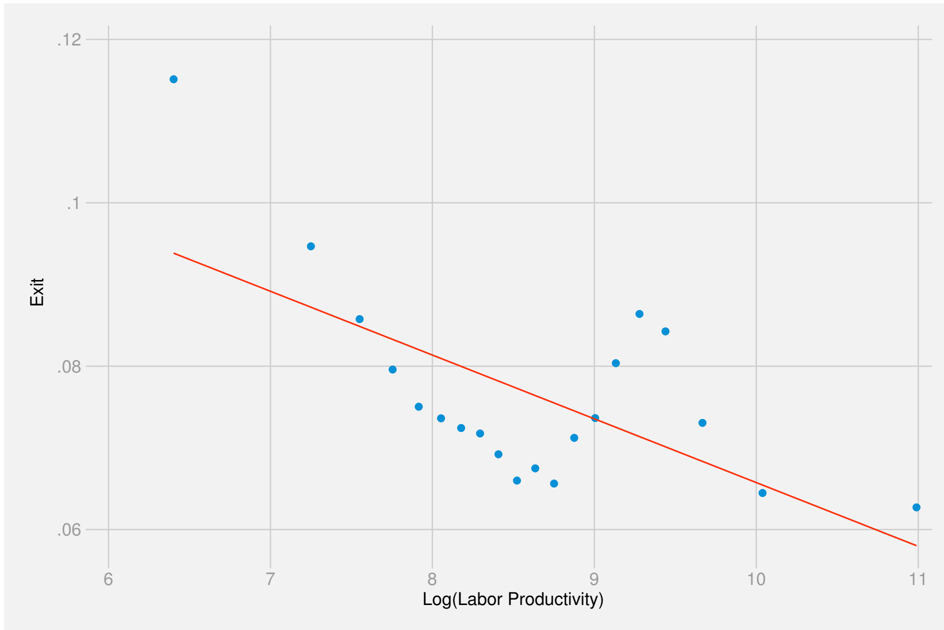
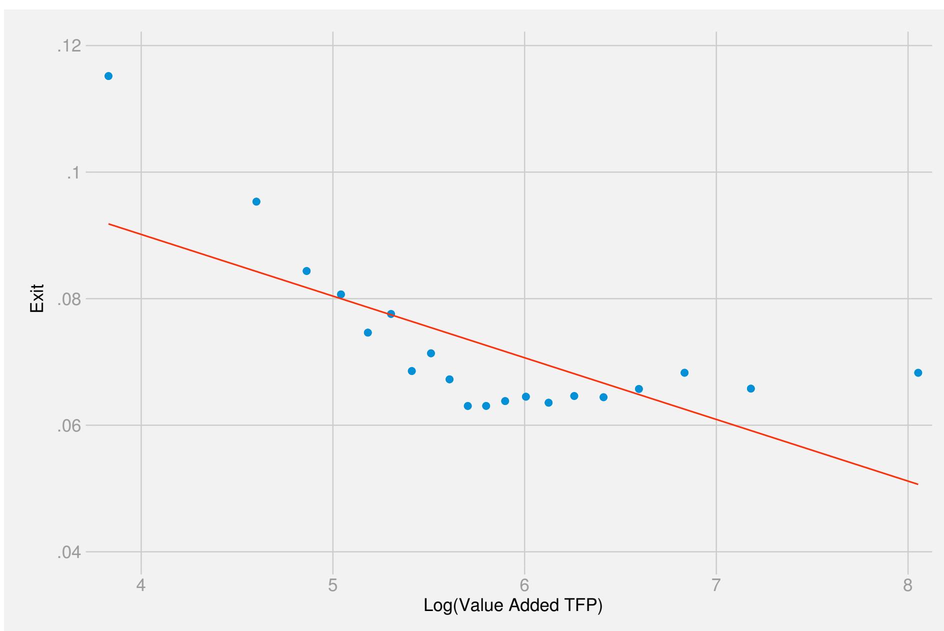


Figure 1: Productivity vs. Number of Workers for Indonesia (Controlling for Industry-Year Fixed Effects)

Notes: This figure shows a binned scatter plot of measures of log productivity on the log number of workers, using the Indonesian data. The data are residualized on industry-year fixed effects before forming bins and plotting. I define labor productivity as value-added per worker. I define Value Added TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}} k^{\frac{1}{3}}}$.



(a) Labor Productivity



(b) Total Factor Productivity

Figure 2: Exit vs. Productivity (Controlling for Industry-Year Fixed Effects)

Notes: These figures show binned scatter plots of exit on measures of log productivity, using the Indonesian data. The data are residualized on industry-year fixed effects before forming bins and plotting. I define labor productivity as value-added per worker. I define Value Added TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year of the sample.

appears in the sample (I set this indicator to be missing in the last year of the data, because for that year I cannot tell which establishments will exit and which will survive).

In the model, the probably of exit is a function solely of productivity and age. In reality, this is a reduced form for a process that may depend on many characteristics. If other variables are correlated with productivity, and the probability of exit is directly affected by those other variables, we may wonder whether we should control for those omitted variables when estimating β_{exit}. In other words, are we interested in correlation or causation?

For the purposes of calibrating the model, we are in fact interested in the correlation of log productivity and exit, rather than necessarily the causal effect of productivity on exit. The selective exit of firms depends on many characteristics, many of which will be correlated with productivity. The selection channel reflects these many correlations. If, for example, the correlation of productivity and exit is a result of the fact that well-managed firms are more productive and less likely to exit, then we would still want to include that as part of the selection channel. The difference in log productivity between survivors and exiters serves as a sufficient statistic to capture these correlations.

In Table 2, I show estimates of β_{exit} using different measures of productivity. Each estimate in the table controls for industry-year fixed effects, and standard errors are clustered by firm. The first column shows my preferred estimate, which uses (log) labor productivity, measured as value added per worker. Exiting firms are 12.3% less productive than stayers. The second column uses revenue per worker, the third column uses value added TFP, and the fourth column uses revenue TFP (which is defined the same as value-added TFP, but uses revenue as the numerator instead of value added). The magnitude of β_{exit} is somewhat larger under these alternative measures: β_{exit} ranges from -12.3% (the preferred estimate) to -18.5% (under revenue-based TFP).

In Table 3, I show estimates of β_{exit} under different specifications. The first column again uses the preferred specification. The second column uses year effects instead of industry-year effects. This changes the results substantially: β_{exit} more than doubles, to 25.6%. Finally, the third column uses industry-year fixed effects, but restricts the sample by dropping certain years (1984, 1987, 1990, and 2000) in which there appear to be anomalously elevated exit rates (the reason for these exit anomalies is not well documented, but seems to be driven by changes in the sampling frame). Dropping these years does not change the results meaningfully.

Overall, these results instill confidence that β_{exit} is somewhere in the ballpark of -12% for Indonesia. However, measurement choices can matter somewhat, especially the decision of whether or not to include industry-year fixed effects. The model does not include a concept of industries, and thus is not very helpful in deciding which specification is more appropriate. However, I favor the specification with industry-year fixed effects for two reasons. First is the simple reason of convention: previous analyses of exit, as well as previous models of misallocation, typically focus on within-industry differences. Second, the specification without industry-year fixed effects is capturing cross-sector selection: firms in less productive industries are more likely to exit. In

	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$	$\log(\text{VA TFP})$	$\log(\text{Rev TFP})$
Exit	-0.123	-0.161	-0.139	-0.185
	(0.006)	(0.006)	(0.007)	(0.007)
Observations	604151	609804	325516	325910
Industry-Year FE	Yes	Yes	Yes	Yes

Table 2: Estimates of β_{exit} for Indonesia: Different Productivity Measures

Notes: This table shows estimates of β_{exit} for the Indonesian data. I define Value Added (VA) TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$. I define Revenue (Rev) TFP as $\frac{\text{Revenue}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$. The number of observations is substantially smaller for TFP measures, because data on capital is missing for many firms and years. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$
Exit	-0.123	-0.256	-0.107
	(0.006)	(0.007)	(0.006)
Observations	604151	604488	548065
Year FE	Yes	Yes	Yes
Industry-Year FE	Yes	No	Yes
Restricted Years	No	No	Yes

Table 3: Estimates of β_{exit} for Indonesia: Different Specifications

Notes: This table shows estimates of β_{exit} for the Indonesian data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year. The first column shows the baseline specification. The second column uses year fixed effects instead of industry-year fixed effects. The third column returns to the baseline specification, but drops the years (1984, 1987, 1990, and 2000) in which there appear to be anomalously elevated exit rates.

the thought experiment underlying the selection channel, less productive firms become relatively more or less likely to exit. If cross-industry differences were not controlled for, then this thought experiment would really be about reallocation across sectors, which is not the intended focus of my paper.

3.3 Data for Chile, Colombia, and Spain

To complement the empirical analysis for Indonesia, I use data from other countries to study exit dynamics. This allows me to conduct cross-country comparisons, and will later facilitate a calibration of the model for those countries, and a quantification of the selection channel in different countries. Here, I describe the data that I use for Chile, Colombia, and Spain.

3.3.1 Data for Chile and Colombia

The data for Chile and Colombia come from censuses of manufacturing plants. The Chilean data covers all firms with more than ten employees, from 1979-1996. The Colombian data covers all establishments with more than ten employees, from 1981-1991.

I use the definition for value added favored by each manufacturing survey. The surveys provide the number of workers, allowing me to construct value added per worker. For revenue, I take the sum of finished goods sold, finished goods shipped to other establishments, and the change in inventory of finished goods (thus, revenue represents the output of finished goods at the plant level, multiplied by the price of those finished goods). I measure capital using deflated capital, constructed following Demirer (2020). I construct firm age and exit in the same manner as for the Indonesian data.

3.3.2 Spanish Data

I analyze data on non-financial firms in Spain. I download the Orbis data set from Bureau van Dijk, via Wharton Research Data Services (WRDS). The data are provided to Bureau van Dijk via information providers, and are originally collected as administrative data by the local government. Spanish firms are required by regulation to provide this information, and so the data set is fairly comprehensive.

I focus on the years 2009 to 2016, because for these years I can measure exit rates accurately. For most companies in my data, the most recent year available is 2018. I thus cannot measure exit accurately in 2018, and I also cannot measure it well in 2017, as a number of firms had not yet reported their 2018 data when WRDS received the Orbis data (this would lead to many spurious exits in 2017). Orbis keeps the ten most recent years of data for a given firm. This creates a selection bias for earlier years: I thus drop years before 2009 (for years before 2009, any firm that is observed in the data is a firm that must not have survived to 2018). After subsetting to 2009-2016, exit rates are fairly stable across years, and the estimate of β_{exit} also appears to be stable across years.

3.4 Cross-Country Comparison

Chile

I show summary statistics for Chile in Appendix Table A.4, computed in the same way as the earlier summary statistics for Indonesia. The exit rate is 6.3%, which is similar to the 7.6% exit rate in Indonesia. The average age is low, but this is because the lower number of years means that I cannot observe age for many Chilean firms. I thus focus on analyses that do not rely on age. The distribution of log productivity has a somewhat lower standard deviation than does the Indonesian data: 0.874 rather than 1.107. The firms in the Chilean data are somewhat smaller than those in the Indonesian data, however, this is driven by the lower size cutoff in the Chilean data.

I show productivity regressions for Chile in Appendix Table A.5. More productive firms employ more labor, and are less likely to exit. The effect of productivity on exit is about five times as strong as in Indonesia. As a result, even though the variance of productivity is somewhat lower, the estimated β_{exit} is substantially larger, at -50% (compared to -12% in Indonesia). In the fourth column, I subset to firms with at least 20 employees, to improve comparability with the Indonesian data. In the fifth column, I measure productivity as revenue per worker. The estimates provide a consistent picture: selection is much stronger in Chile than in Indonesia.

Colombia

I show summary statistics for Colombia in Appendix Table A.6. The exit rate is 10.9%, which is moderately higher than 7.6% exit rate in Indonesia. As in the Chilean data, the average age is artificially low, due to the shorter panel. Similarly, the firms in the Colombian data are somewhat smaller than those in the Indonesian data, driven by the lower size cutoff. Log labor productivity is less dispersed than in Indonesia, with a standard deviation of 0.781.

I show productivity regressions for Colombia in Appendix Table A.7. More productive firms are larger and less likely to exit. The estimated β_{exit} is -34% (compared to -12% in Indonesia). In the fourth column, I subset to firms with at least 20 employees, to improve comparability with the Indonesian data. In the fifth column, I measure productivity as revenue per worker. Selection is substantially stronger in Colombia than in Indonesia.

Spain

The Spanish data enable a comparison with a developed economy. In Appendix Table A.8, I show summary statistics for Spain, computed in the same way as the earlier summary statistics for Indonesia. The exit rate is 8%, which is similar to the 7.6% exit rate in Indonesia. The average age is quite low, because the Spanish panel is quite short and so I cannot observe age for most Spanish firms. I only observe productivity for a subset of Spanish firms. For those firms for which I observe productivity, the distribution of log productivity has a somewhat lower standard deviation than does the Indonesian data: 0.868 rather than 1.107. The firms in the Spanish data are smaller than those in the Indonesian data: this is because the Spanish data do not have the same firm size restrictions.

I show productivity regressions for Spain in Appendix Table A.9. I do not have industry codes for the Spanish data, so I include year fixed effects but cannot include industry-year effects. More productive firms employ more labor, although the effect is somewhat weaker than in Indonesia. The effect of productivity on exit, however, is substantially stronger than in Indonesia. As a result, even though the variance of productivity is somewhat smaller, the estimated β_{exit} is substantially larger, at -39% (compared to -12% in Indonesia). In the fourth column, I subset to firms with at least 20 employees, as in the Indonesian data. In the fifth column, I measure productivity as revenue per worker. In each case, the estimate does not change substantially. Moreover, even the smallest estimate of β_{exit} for Spain is substantially larger than the largest estimate for Indonesia.

Many Countries

To conclude this section, I extend the analysis of β_{exit} to many countries. To do this, I rely on the work of Bartelsman et al. (2009). In their handbook chapter, they provide summary statistics from their effort to provide harmonized firm census data from many countries. While they do not provide an estimate of β_{exit}, they do provide an estimate of the difference in log productivity between incumbents and exiters, which is equal to $(1-\rho)\times(-\beta_{\text{exit}})$, where ρ is the exit rate. This will in general be very similar to β_{exit}, and also serves as a good back-of-the-envelope approximation to the selection channel.[4] These estimates also differ slightly from my estimates because they measure productivity as revenue per worker rather than value added, and because they use a three-year time period rather than a one year period (they measure exit as equal to one if the firm exits some time in the next three years). I make my estimates comparable to theirs by using log revenue per worker as my outcome and using a three-year definition of exit, and then computing $-\beta_{\text{exit}}$ and multiplying by $1-\rho$.

The results are in Appendix Table A.10. I first list the estimates from Bartelsman et al. (2009), taken from column 3 of Table 1.9 of their chapter. I then list my own estimates for Indonesia and Spain, as well as my own estimates for Colombia and Chile.[5] In the second column I include real GDP per capita in 2010, from the World Bank's World Development Indicators. I use Germany's GDP per capita for West Germany. The WDI do not have data for Taiwan.

The main pattern to notice is that there is no pattern: there is variation in the strength of selection across countries, but it does not appear to be strongly correlated with GDP. I confirm this visually in Figure 3, which compares the estimated $(1-\rho)\times(-\beta_{\text{exit}})$ to log real GDP per capita.[6] Moreover, despite the small sample size, a regression rules out strong correlations: a regression of the estimated $(1-\rho)\times(-\beta_{\text{exit}})$ on log real GDP per capita yields a coefficient of -0.04, with a standard error of 0.07. The upper bound of the 95% confidence interval is 0.10, which would suggest that a country whose log real GDP per capita was 100 log points higher would, at most, have a selection channel that is 10 percentage points higher. I discuss these results further in the Section 5, but they suggest that even if the selection channel is an important contributor to aggregate productivity, it is unlikely to explain much of cross-country differences in GDP per capita.

[4] In fact, the derivation of the back-of-the-envelope approximation makes the assumption $(1-\rho)\,\beta_{\text{exit}} \approx \beta_{\text{exit}}$.

[5] My own estimates for Chile are quite similar to those of Bartelsman et al. (2009). The estimates differ, however, for Colombia: my estimates suggest $(1-\rho)\times(-\beta_{\text{exit}})$ in Colombia is in the "middle of the pack" at 33%, while Bartelsman et al. (2009) find it has the strongest selection in their data, at 63%. Whatever the source of this discrepancy, it does not affect the broader message: differences in selection across countries cannot explain substantial differences in GDP per capita.

[6] In this figure, I use the Bartelsman et al. (2009) estimates for Chile and Colombia, rather than those from my own analysis.

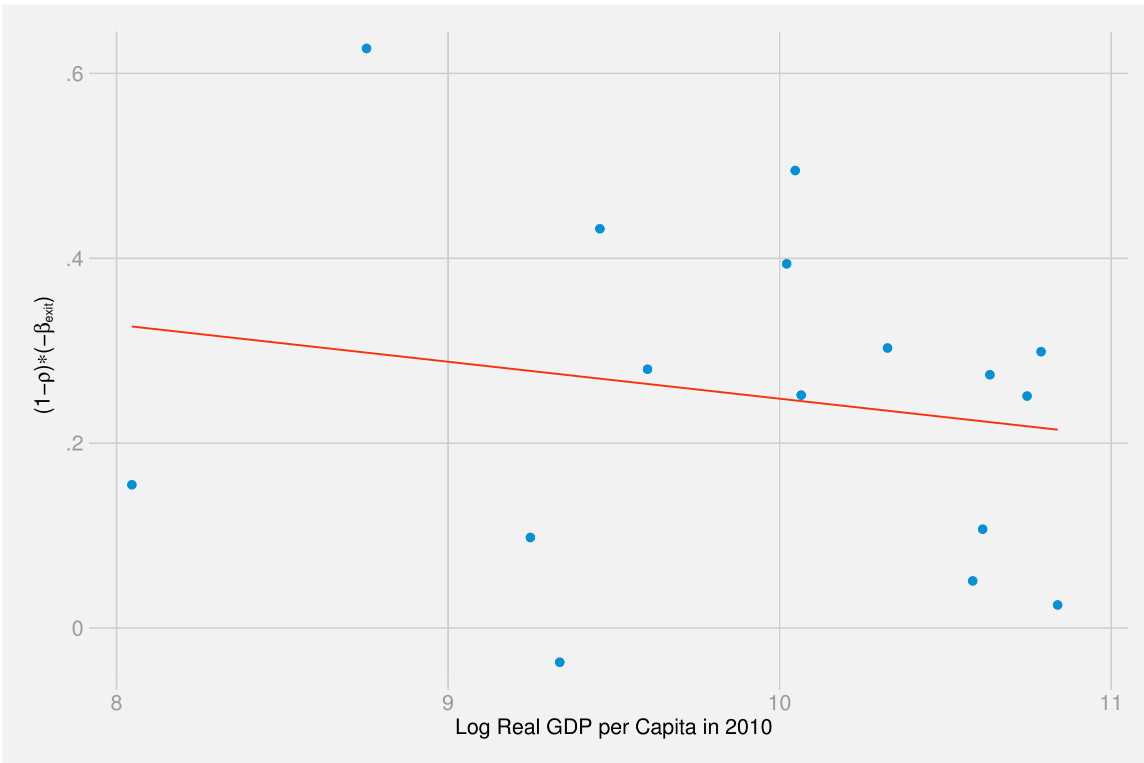


Figure 3: β_{exit} vs. GDP per Capita

Notes: This figure shows estimates, by country, of $(1-\rho)\cdot(-\beta_{\text{exit}})$, or one minus the exit rate times the difference in log productivity between stayers and exiters, plotted against each country's log real GDP per capita in 2010. Data on real GDP per capita come from the World Bank's World Development Indicators. I use Germany's GDP per capita for West Germany. The WDI do not contain GDP data for Taiwan. I draw estimates of $(1-\rho)\cdot(-\beta_{\text{exit}})$ from column 3 of Table 1.9 in Bartelsman et al. (2009). I supplement their estimates with my own estimates for Indonesia and Spain; I use the Bartelsman et al. (2009) estimates for Colombia and Chile instead of my own. Bartelsman et al. (2009) use log revenue per worker as their outcome variable, and define exit as equal to one if the firm exits any time over the next three years: I use the same definitions for my analysis of Indonesia and Spain to maintain consistency. The red line shows the best-fit line from a univariate regression of $(1-\rho)\cdot(-\beta_{\text{exit}})$ on log real GDP per capita in 2010.

Parameter	Data Target	Estimate
(μ, σ^2)	Distribution of Initial Productivity	$(0, 1.03)$
γ	Inverse Elasticity of Output to Productivity	0.819
$\beta_{\text{exit},a}$	Productivity Gap of Exiters vs. Survivors	See Table 5
ρ_a	Exit Rates	See Table 5

Table 4: Main Calibration

Notes: This table shows the main calibration of the model for Indonesia. Each set of parameters is estimated to match a particular target in the data. The paramaters (μ, σ^2), representing the distribution of log productivity among entrants, match the distribution of log labor productivity for entrants, demeaned with industry-year fixed effects (the mean, μ, is thus zero as a normalization). The parameter γ, representing the inverse elasticity of output to productivity, is calibrated to match a regression of log value added on log labor productivity, controlling for industry-year fixed effects. The age profile of $\beta_{\text{exit},a}$ is calibrated to match the productivity gap between exiters and survivors, by cohort. The age profile of exit rates, ρ_a, is also calibrated to match the age profile of exit rates. The age profiles of $\beta_{\text{exit},a}$ and ρ_a are each assumed to follow a particular functional form: the estimated parameters of these functional forms are in Table 5. See text for further details on the calibration.

4 Calibration

I now turn to calibration of the model. I base the main calibration on the Indonesian data, although I provide alternate calibrations for Chile, Colombia, and Spain. To organize this exercise, Table 4 shows the relevant parameters, their target in the data, and the estimated value. For the parameters that vary by age, I refer to the relevant table. I use the notation $\beta_{\text{exit},a}$ to refer to the difference in mean log productivity, by age, between exiters and survivors. In the log-normal model (without truncation), this is equal to $\frac{\delta_a \sigma^2}{\rho_a}$. Thus, $\beta_{\text{exit},a}$ pins down δ_a, as long as the other parameters are known. In Appendix B, I discuss how this is a function of parameters in the log-normal model with truncation.

4.1 Productivity Distribution for Entrants

To calibrate the productivity distribution for entrants, I demean the log labor productivity of entrants using industry-year fixed effects, and then compute the mean and standard deviation of the demeaned effects. The mean is mechanically zero, and the standard deviation is 1.03. I set the upper bound on productivity to be three standard deviations above the mean: this leaves 99.9% of the distribution below the upper bound.

4.2 Inverse Elasticity of Output to Productivity

To calibrate γ, the inverse elasticity of output to productivity, I run a regression. In the model, there is a one-to-one correspondence between log output and log productivity, and so it is not obvious which should be the regressor and which should be the outcome. Since there is no residual in the model, the two regressions should give equivalent results, which is not true in the data.

In Appendix C, I show that, if there is a residual in the model, the regression of log output on log productivity is correct, and the inverse of the regression coefficient will recover γ.[7] Under appropriate assumptions, the addition of a residual will not affect the computation of aggregate productivity (see Appendix C for details). Intuitively, the benefits to allocation, $\left(\frac{1}{\gamma}-1\right)\sigma^2$, reflect the covariance of log productivity and log inputs; multiplying the coefficient from a regression on log productivity by the variance of log productivity will correctly recover this covariance.

I thus run the following regression

$$\log\left(q_{it}\right)=\alpha_{st}+\frac{1}{\gamma}\cdot\log\left(z_{it}\right)+\varepsilon_{it}$$

where q_{it} is the output of firm i in year t, α_{st} is an industry-year fixed effect, z_{it} is measured as value added per worker, and the regression coefficient on log z_{it} is equal to $\frac{1}{\gamma}$. Regressing log output on log labor productivity, controlling for industry-year fixed effects, yields $\gamma = 0.819$. This implies benefits to allocation of roughly $\left(\frac{1}{\gamma}-1\right)\sigma^2 \approx 0.22 * 1.03^2$, or a roughly 23% increase in aggregate productivity relative to a benchmark where all firms receive the same inputs.

4.3 Selective Exit

Next, I calibrate $\beta_{\text{exit},a}$, which is the difference in mean log productivity between exiters and survivors, by age. To calibrate this, I regress log labor productivity on an indicator that is equal to one in the period in which the firm exits, controlling for industry-year fixed effects. In the full sample, this regression yields $\beta_{\text{exit}} = -0.123$ with a standard error of 0.0055 (See Table 2 in the previous section). However, for the full model, we want to estimate $\beta_{\text{exit},a}$, which can vary by age.

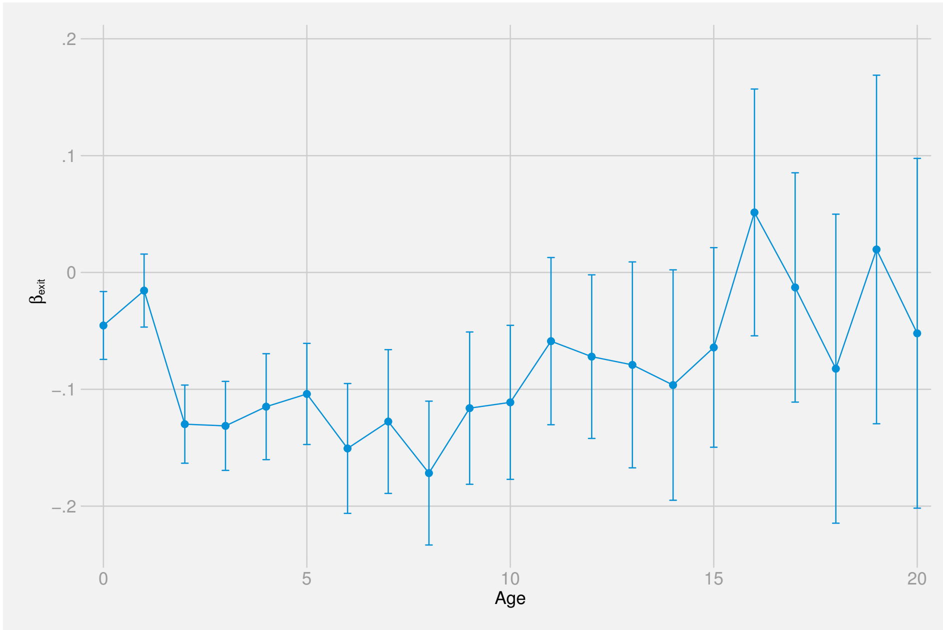
I calibrate $\beta_{\text{exit},a}$ with a two step procedure. First, I estimate $\beta_{\text{exit},a}$ by running the above regression separately for each firm age I observe in the data, where I measure firm age as the number of years since the first time the firm appears in the data, excluding firms that are present in 1975 (the first year of the data). This provides me with estimates of $\beta_{\text{exit},a}$from age 0 to 35. I show the results in Figure 4a, for firms up to age 20.

These results are informative, but they are noisy at higher ages, and they do not extend past age 35. To smooth the results and extend them to all ages, I fit the following linear model to the coefficients:

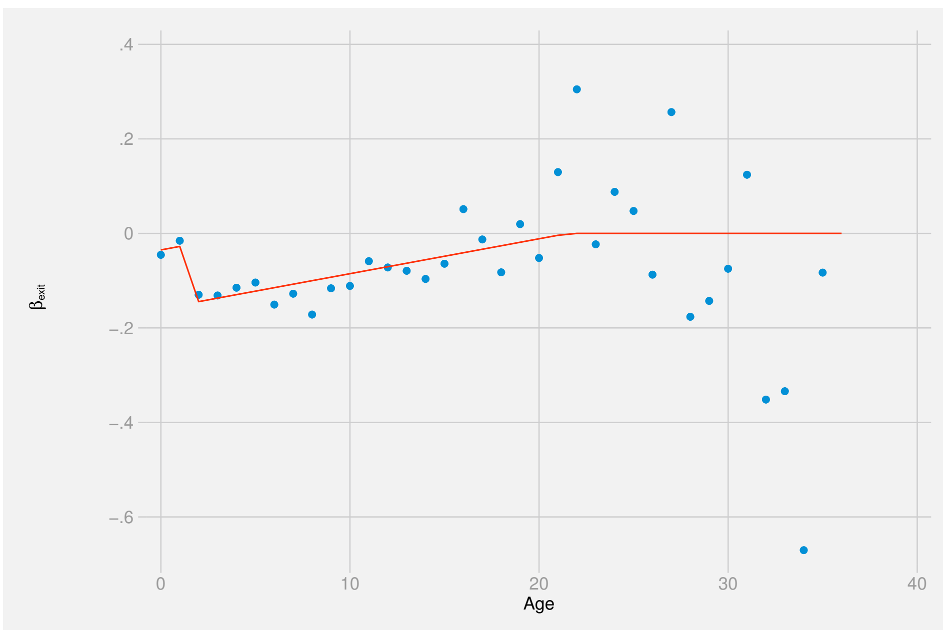
$$\beta_{\text{exit},a}=\theta_0+\theta_{\mathbf{1}(a\leq 1)}\times\mathbf{1}\left(a\leq 1\right)+\theta_a\times a$$

I estimate the model using weighted least squares, using the inverse squared standard error of each coefficient as its weight (i.e. precision weights). I show the results in Table 5. To obtain the final set of $\beta_{\text{exit},a}$, I use the estimated θs to obtain a fitted value of $\beta_{\text{exit},a}$, and then replace that value with zero if the fitted value is positive. This ensures that the exiting firms of given cohort are never more productive than the surviving firms, and also implies that after a certain year,

[7] This is also the route taken by prior literature (Hsieh and Klenow, 2014; Bento and Restuccia, 2017).



(a) Raw Estimates of $\beta_{\text{exit},a}$



(b) Fitted Estimates of $\beta_{\text{exit},a}$

Figure 4: Estimates of $\beta_{\text{exit},a}$

Notes: This figure shows estimates of $\beta_{\text{exit},a}$, the age profile of the gap in log productivity between exiters and survivors, in the Indonesian data. I estimate $\beta_{\text{exit},a}$ by regressing $\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$ on an indicator for exit separately for each age, controlling for industry-year fixed effects. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year of the sample. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. These raw estimates are plotted in the first panel, with 95% confidence intervals (clustering by firm), up to age 20. I fit the raw estimates to the model $\beta_{\text{exit},a} = \theta_0 + \theta_{\mathbf{1}(a\leq 1)} \times \mathbf{1}\left(a \leq 1\right) + \theta_a \times a$, estimating with weighted least squares and using the inverse variance of the estimate as the weight. The results are depicted in the second panel, which shows the raw estimates as blue dots, and the fitted model as a red line. For the fitted line, I assume that $\beta_{\text{exit},a}$ stays at zero rather than ever going positive.

selection stops. In Figure 4b, I compare the fitted values to the raw estimates for the full sample. Note that the later estimates are quite noisy. Nonetheless, the functional form I fit appears to capture the age path of $\beta_{\text{exit},a}$ well.

4.4 Firm Mortality

Next, I turn to calibrate firm mortality by age. I model firm mortality rates using the Gompertz-Makeham law of mortality, which states that mortality rates are the sum of a constant component, λ, and an exponential component, $\alpha \cdot \exp(\beta a)$. The Gompertz-Makeham law is commonly used to model human mortality rates, which tend to rise with age. I use non-linear least squares to estimate the model:

$$\rho_a = \alpha \cdot \exp(\beta a) + \lambda$$

measuring firm age as before and excluding firms that are present at the start of the data in 1975. I report the results in Table 5. I also plot the fitted estimates compared to the raw data in Figure 5. The Gompertz-Makeham functional form fits the data extremely well. Firm mortality is declining in age, asymptoting to a λ of 4.7%. This declining mortality rate also implies that the average firm age is greater than $\frac{1-\rho}{\rho}$: after passing through the high exit early years, firms have lower mortality rates and thus can expect to live a while.[8]

4.5 Recalibration for Chile, Colombia, and Spain

To understand the importance of the selection channel in other countries, I recalibrate the parameters to fit the data for Chile, Colombia, and Spain. Because of the limited number of years available for each of the non-Indonesian data sets, I do not attempt to recalibrate the age profile of exit rate and selection, ρ_a and $\beta_{\text{exit},a}$. However, note that the unconditional exit rate, ρ, is similar in Indonesia, Spain, and Chile, and is only moderately higher in Colombia.[9]

To recalibrate γ and σ^2, I follow the same procedure for Chile and Colombia as I do for Indonesia. However, because I do not have industry codes in the Spanish data, I absorb year fixed effects but do not absorb industry-year fixed effects. The recalibration yields $\gamma = 0.927$ and $\sigma^2 = 1.04$ for Spain, $\gamma = 0.762$ and $\sigma^2 = 0.74$ for Chile, and $\gamma = 0.728$ and $\sigma^2 = 0.61$ for Colombia. I summarize the results in Table 6. In addition to σ^2 and γ, which I use in the recalibration, I also report β_{exit} and the exit rate, ρ. Knowing β_{exit} is particularly useful because it helps us know how much to scale up or down the strength of selection for each country: in Section 5 I will explicitly

[8] Although the Gompertz-Makeham law fits mortality well for both humans and firms, the exponential component is declining in age for firms, whereas for humans it rises with age. Because of this, average firm age (ignoring population growth) is higher than life expectancy, whereas the opposite is true for humans. This also means that the Gompertz-Makeham law fits firm mortality even better than it fits human mortality: for humans, the law must eventually break down because it would predict mortality rates above one for the very old.

[9] Much of the Colombian data comes from the aftermath of the 1982 crisis that affected many Latin American economies, which may explain much of the elevated exit rate.

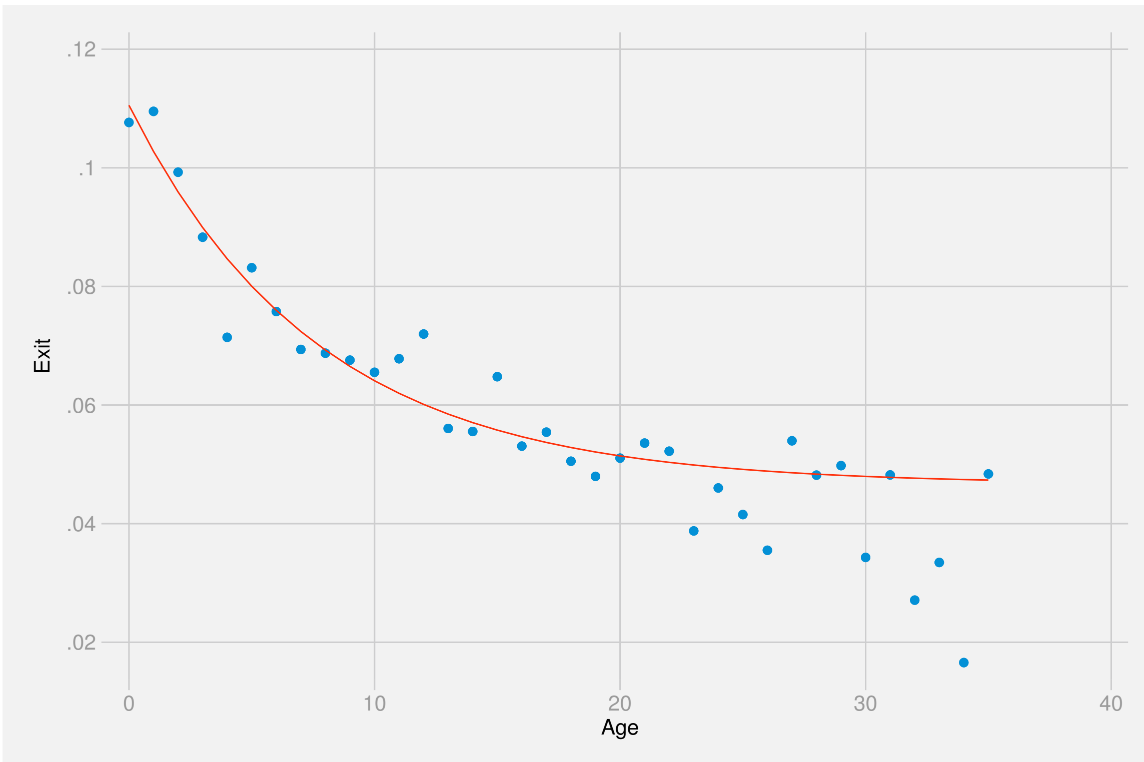


Figure 5: Gompertz-Makeham Fitted Estimates of Exit Rates by Age

Notes: This figure shows exit rates by age in the Indonesian data, compared to the fitted values from the estimated Gompertz-Makeham model. I use non-linear least squares to estimate the model $\rho_a = \alpha \cdot \exp(\beta a) + \lambda$, where a is age and ρ_a is the exit rate at age a. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year of the sample. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. The blue dots show the average exit rates by age, while the red line shows the fitted values from the Gompertz-Makeham model.

Selective Exit		Gompertz-Makeham	
θ_0	-0.159	α	0.064
	(0.011)		(0.002)
$\theta_{\mathbf{1}(a\leq 1)}$	0.124	β	-0.130
	(0.013)		(0.008)
θ_a	0.007	λ	0.047
	(0.001)		(0.002)

Table 5: Parameter Estimates for Selective Exit and Mortality by Age

Notes: This table shows estimates for the age profile of $\beta_{\text{exit},a}$ and ρ_a in the Indonesian data. I model $\beta_{\text{exit},a}$, the age profile of the gap in log productivity between exiters and survivors using a parametric functional form. I estimate $\beta_{\text{exit},a}$ by regressing $\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$ on an indicator for exit separately for each age, controlling for industry-year fixed effects. I then fit the raw estimates to the model $\beta_{\text{exit},a} = \theta_0 + \theta_{\mathbf{1}(a\leq 1)} \times \mathbf{1}\left(a \leq 1\right) + \theta_a \times a$, estimating with weighted least squares and using the inverse variance of the estimate as the weight. I report the estimated parameters, with standard errors in parentheses, in the first column. I model ρ_a, the exit rate by age, using the Gompertz-Makeham law of mortality. I use non-linear least squares to estimate the model $\rho_a = \alpha \cdot \exp\left(\beta a\right) + \lambda$, where a is age and ρ_a is the exit rate at age a. I report the estimated parameters, with standard errors in parentheses (clustered by firm), in the second column. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year of the sample. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data.

focus on the mapping between β_{exit} and the selection channel. Although I will not recalibrate the profile of exit rates by age, it is reassuring to note that the average exit rate is not very different across countries.

5 Results

With the calibrated parameters in hand, I now examine the selection channel quantitatively using the model. I begin by showing results for Indonesia, comparing the back-of-the-envelope to the simplified model (homogeneous parameters across ages) and then to the fully calibrated main model (with heterogeneous parameters by age). I next examine how the model results depend on key parameters, namely γ (the inverse elasticity of output to productivity) and σ^2 (the variance of log productivity for entrants). Finally, I turn to a quantitative comparison of the selection channel in Indonesia, Chile, Colombia, and Spain, and conclude with a broader discussion of the role of the selection channel in explaining cross-country differences in output per capita.

To solve the model, I use the calibrated parameters described in Section 4, and I assume that all remaining firms die after reaching age 200. To solve the model with the truncated log-normal distribution of productivity requires finding a fixed point, such that the shift in μ_a yields the desired $\beta_{\text{exit},a}$; I describe the details of this procedure in Appendix B. I then re-solve the model for different levels of selection, by multiplying the profile of $\beta_{\text{exit},a}$ by a scalar. For each run of the model, I compute the implied β_{exit}, i.e. what a regression of log productivity on exit would

	σ^2	γ	β_{exit}	ρ
Indonesia	1.03	0.819	-0.123	7.6%
Chile	0.74	0.762	-0.502	6.3%
Colombia	0.61	0.728	-0.342	10.9%
Spain	1.04	0.927	-0.390	8.0%

Table 6: Alternative Calibrations for Multiple Countries

Notes: This table compares the main calibration, based on the Indonesian data, to alternative estimates for Chile, Colombia, and Spain. The first column shows the estimates of σ^2, the variance of log productivity. The second column shows estimates of γ, the inverse elasticity of output to productivity, calibrated to match a regression of log value added on log labor productivity. The third column shows β_{exit}, gap in log productivity between exiters and survivors. Each of the first three columns is estimated after controlling for industry-year fixed effects, except for the results for Spain, which only control for year fixed effects. The fourth column shows ρ, the average exit rate in the sample. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year of the sample.

yield if run on data generated by the model. I also compute the implied selection channel for each model run. Thus, I am able to trace out the implied relationship between β_{exit}, a coefficient that is measurable in the data, and the selection channel, an equilibrium object for which we need the model.

The figures in this section focus on plotting this relationship between $\beta_{\text{exit},a}$ and the selection channel, and showing how it varies under different parametrizations of the model. Broadly, I find that β_{exit} serves as a useful approximation to the true selection channel, and the back-of-the-envelope approximation performs extremely well. As a result, the empirical finding that $\beta_{\text{exit},a}$ does not seem to covary with output per capita also implies that the strength of the selection channel is not, on average, substantially higher or lower in developed vs. developing countries. Thus, while my findings suggest that the selection channel is an important component of productivity, it is not an important explanation of why poor countries are poor.

5.1 Selection Channel vs. β_{exit}

I begin, in Figure 6, by comparing the selection channel to β_{exit}. To keep everything in a log scale, I use $\log(1 + \text{Selection Channel})$ as the vertical axis; this results in roughly linear relationships. I also include vertical dashed lines, which mark the preferred estimates of β_{exit} for Indonesia and Spain, and facilitate cross-country comparisons for each model. I plot three versions of the relationship between β_{exit} and the selection channel. The (diagonal) red dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit}} = \log(1 + \text{Selection Channel})$. The green line is the homogeneous model, in which $\beta_{\text{exit},a}$ and ρ_a are the same for all ages; this results in a relationship that is very close to the back-of-the-envelope. The blue line is the main model: it shows the calibrated model with $\beta_{\text{exit},a}$ and ρ_a varying across ages. The main model is somewhat above the back-of-the-envelope and the homogeneous model, implying that more accurately modeling the

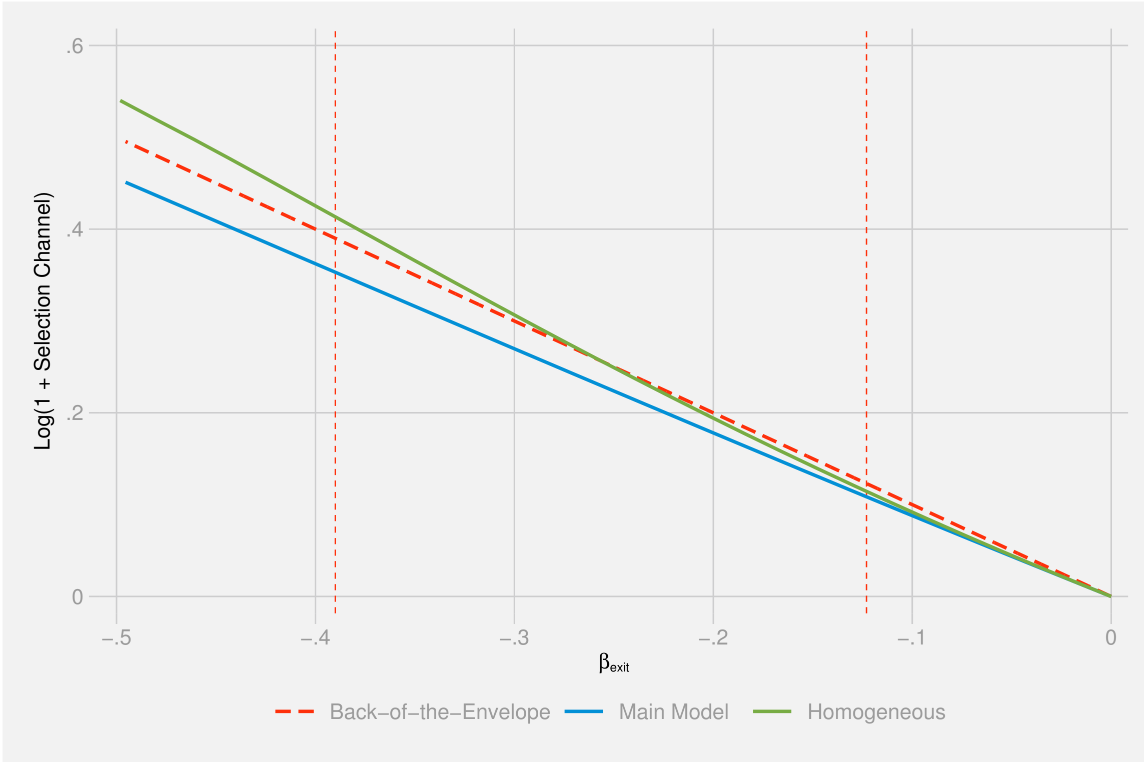


Figure 6: Selection Channel vs. β_{exit}

Notes: This figure shows the relationship between the selection channel and β_{exit}. The (diagonal) red dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit}} = \log(1 + \text{Selection Channel})$. The green line is the homogeneous model, in which $\beta_{\text{exit},a}$ and ρ_a are the same for all ages. The blue line is the main model, with $\beta_{\text{exit},a}$ and ρ_a varying across ages. Each point on the blue and green lines represents a solution to the model, run with exit rates by age held constant and with selection by age, $\beta_{\text{exit},a}$, scaled proportionally up or down. For each run of the model, I compute the selection channel and also compute what the estimated β_{exit} would be in data generated from the solved model. I also include vertical dashed lines, which mark the preferred estimates of β_{exit} for Indonesia and Spain, to facilitate cross-country comparisons for each model.

timing of exit and selection modestly increases the selection channel.

However, the back-of-the-envelope still appears to be a good approximation to the selection channel in the main model. This suggests that a fairly complex question, "How does selective exit affect aggregate productivity?" can, with little loss, be reduced to a relatively easily measured statistic, even in a model with productivity dispersion, misallocation, and changing exit and selection dynamics over the lifecycle.

Comparative Statics: Selection vs. Allocation

In an economy with heterogeneous firms, a hypothetical planner will wish to allocate less resources to less productive firms. This can be accomplished in two ways. The planner can *starve* the unproductive firms by allocating them less labor. Or, the planner can *kill* the unproductive firms, by having them exit the market. The former channel, which I refer to as allocation, is the intensive

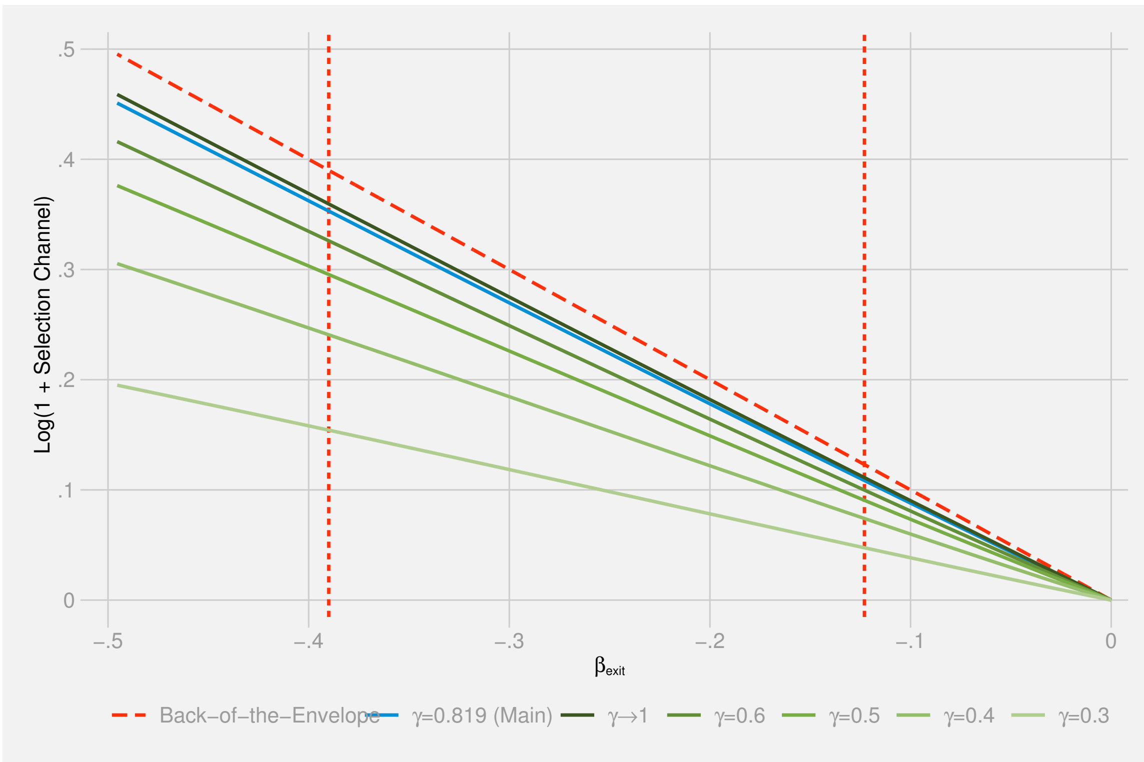


Figure 7: Selection Channel vs. β_{exit} (Varying γ)

Notes: This figure shows the relationship between the selection channel and β_{exit}, for different levels of γ. The (diagonal) red dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit}} = \log(1 + \text{Selection Channel})$. The blue line is the main model, with $\gamma = 0.819$. The green lines represent different choices of γ, with darker lines corresponding to higher values of γ. Each point on the blue and green lines represents a solution to the model, run with exit rates by age held constant and with selection by age, $\beta_{\text{exit},a}$, scaled proportionally up or down. For each run of the model, I compute the selection channel and also compute what the estimated β_{exit} would be in data generated from the solved model. I also include vertical dashed lines, which mark the preferred estimates of β_{exit} for Indonesia and Spain, to facilitate cross-country comparisons for each calibration of the model.

margin channel traditionally studied by the misallocation literature. The latter is the selection channel, which is distinguished in my model as being a dynamic, extensive margin phenomenon.

This dichotomy between killing or starving the less productive firms provides intuition suggesting that selection and allocation may be substitutes: if the planner implements an efficient static allocation by starving the inefficient firms, then she does not need to kill them. In the extreme, where all labor is allocated to the most efficient firm, it does not matter if the planner has killed off the inefficient firms, since they are not receiving labor anyway. On the other hand, if the static allocation of inputs is such that productive and unproductive firms get similar levels of inputs, then the selection of which firms are around to receive inputs will matter a great deal.

Another perspective on the interaction between selection and allocation is to consider how selective exit affects the variance of log productivity. Under the log-normal distribution, aggregate productivity is given by $\log Z = \mu + \frac{1}{2}\sigma^2 + \left(\frac{1}{\gamma} - 1\right)\sigma^2$. While this only holds exactly under log-normality, it is a useful approximation, and reveals that the importance of allocation, represented by $\left(\frac{1}{\gamma} - 1\right)\sigma^2$, depends on the variance of log productivity. Selection will have two effects on the variance of log productivity. It raises the variance of the average log productivity across cohorts, through accumulating selection effects over time. However, for a truncated log-normal distribution, it decreases the within-cohort variance of log productivity, as the distribution compresses against the upper bound.[10] Selection will be a subsitute for allocation if the latter channel dominates.

I explore the interaction between selection and allocation by varying γ, the parameter which governs the covariance between log productivity and log labor. I show the results of this exercise in Figure 7. The vertical and diagonal dashed lines are as before. The blue line again represents the main calibration of the model. The green lines represent different choices of γ: darker lines correspond to higher choices of γ. At the calibrated parameters, allocation and selection are substitutes: as γ falls and the allocation of inputs improves, the selection channel shifts down.

To give some sense of magnitudes, it can be useful to benchmark against the misallocation literature. While estimates have varied (in fact, there is some disagreement about whether misallocation is an important source of cross-country differences in TFP), I will use the estimates of Hsieh and Klenow (2009) for exposition. They find that reducing misallocation to US levels would raise Indian manufacturing TFP by 40-60%. In my model, that increase in TFP would be (roughly) equivalent to moving from $\gamma = 0.819$ to $\gamma = 0.6$.[11] This change in γ does not severely change the relationship between β_{exit} and the selection channel. Thus, for values of γ that align with the literature on misallocation, the back-of-the-envelope approximation remains fairly accurate.

[10] Readers may recognize these two effects as representing the two components of the law of total variance: the total variance of log productivity is the variance of the conditional expectation across cohorts, plus the within-cohort variance.

[11] This can be seen most straightforwardly via the formula for aggregate productivity under log-normality, in which the benefits of allocation are given by $\left(\frac{1}{\gamma} - 1\right)\sigma^2$. For $\sigma^2 = 1.03$, moving from $\gamma = 0.819$ to $\gamma = 0.6$ leads a productivity improvement of 46 log points. Moving to $\gamma = 0.5$ would yield an improvement of 80 log points, overshooting the Hsieh and Klenow estimates.

Comparative Statics: Changing Entrant Heterogeneity

I also explore how changing the variance of entrants' log productivity, σ^2, interacts with the selection channel. Whereas varying γ represents changes in static allocative efficiency, changes in σ^2 represent changes in fundamental dispersion. I show the effects of changing σ^2 in Appendix Figure A.3. As before, the main model is in blue, and alternative calibrations, with different levels of σ^2, are shown in corresponding shades of green (darker green represents higher σ^2). Note that while I change σ^2, I do not change the upper bound on log productivity: it is calibrated to be 3σ above μ_0, based on the original calibrated value of σ.

Higher values of σ^2 reduce the strength of the selection channel. This can again be understood through the competing effects of selection on the total variance of log productivity: selection raises the variance of log productivity by creating variation across cohorts, but lowers log-productivity within-cohort due to truncation. With a lower σ^2, truncation is less binding, and so the latter effect becomes relatively less important.

5.2 Cross-Country Comparison

The strength of the selection channel in a given country will depend on the estimated β_{exit} in that country, and on the mapping from β_{exit} to the selection channel, given the other relevant parameters for that country. To examine how the selection channel varies across countries, I will show how the model's mapping between β_{exit} and the selection channel changes when we recalibrate γ and σ^2 to match the Spanish data, the Chilean data, and the Colombian data.[12]

I show the results of the recalibration in Figure 8: the blue line is the model calibrated to Indonesian data, while the green lines in different shades show the recalibrations for other countries. There is little difference between the lines; in fact, the Spanish and Indonesia calibrations lie almost on top of each other. The upshot is that what matters for the selection channel is β_{exit}; other parameters, to the extent that they differ across countries, are of secondary importance for the selection channel.

Indonesia has a lower β_{exit} than Spain. If Indonesia raised its level of selection to that of Spain, its aggregate productivity would increase by roughly 30%, using either the calibrated model or the back-of-the-envelope approximation. The selection channel plays an important role in determining aggregate productivity.

Yet this does not imply that the selection channel can meaningfully explain cross-country differences in output. Colombia is poorer than Chile, which is poorer than Spain, yet the selection channel is similar in each of these countries. As we saw in Section 3, estimates of β_{exit} do not strongly covary with output per capita, and the variance of β_{exit} is much smaller than the variance of log GDP per capita. The results of this section confirm that the back-of-the-envelope approximation is fairly accurate, and thus our earlier conclusions about cross-country differences

[12]The parameters for each country's recalibration are summarized in Table 6.

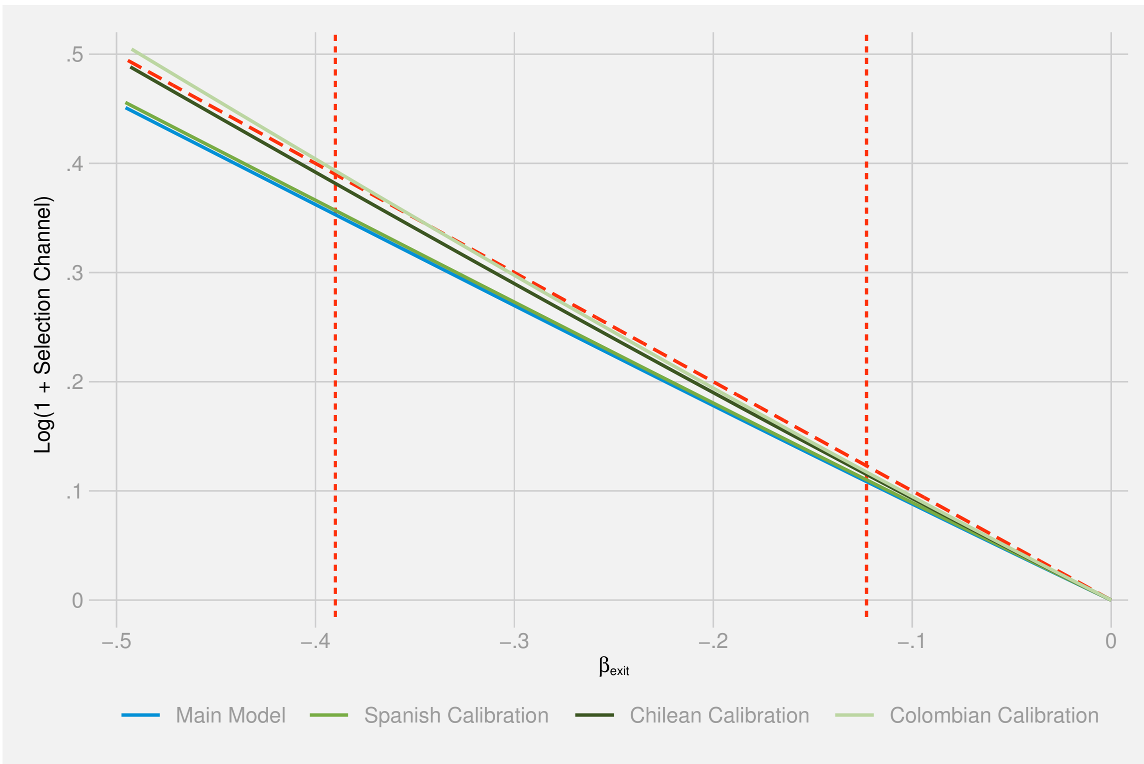


Figure 8: Selection Channel vs. β_{exit} (Indonesia vs. Spain, Chile, and Colombia)

Notes: This figure shows the relationship between the selection channel and β_{exit}, with γ and σ^2 calibrated to different countries. The (diagonal) red dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit}} = \log(1 + \text{Selection Channel})$. The blue line is the main model, calibrated to Indonesia. The green lines in different shades represent recalibrations to different countries: Spain, Chile, and Colombia. Each point on the blue and green lines represents a solution to the model, run with exit rates by age held constant and with selection by age, $\beta_{\text{exit},a}$, scaled proportionally up or down. For each run of the model, I compute the selection channel and also compute what the estimated β_{exit} would be in data generated from the solved model. I also include vertical dashed lines, which mark the preferred estimates of β_{exit} for Indonesia and Spain, to facilitate cross-country comparisons for each calibration of the model.

in β_{exit} are also conclusions about cross-country differences in the selection channel. The selection channel is an important component of productivity, but it does not strongly covary with log GDP per capita, and is not an important channel for explaining cross-country differences in output per capita.

How Much of the Variance in Log Output per Capita Can Be Explained by the Selection Channel?

To argue more formally that the selection channel is not important for cross-country differences in output per capita, I focus on computing the R^2 of the selection channel. To do this, consider the following equation determining log output per capita in country i:

$$\log\left(\text{GDP per capita}_i\right) = \underbrace{\theta \cdot \beta_{\text{exit},i}}_{\text{Selection Channel}} + \varepsilon_i$$

where the coefficient θ determines the mapping from $\beta_{\text{exit},i}$ into the selection channel, and where ε_i represents other factors that affect a country's log output per capita. I define the R^2 of the selection channel as the share of the variance that would go away if the selection channel did not vary across countries. We can derive the formula for the R^2 by decomposing the variance of log GDP per capita:

$$\begin{aligned}\text{Var}\left(\log\left(\text{GDP per capita}_i\right)\right) &= \theta^2 \cdot \text{Var}\left(\beta_{\text{exit},i}\right) + 2\theta \cdot \text{Cov}\left(\beta_{\text{exit},i}, \varepsilon_i\right) + \text{Var}\left(\varepsilon_i\right) \\ \implies R^2 &= \frac{\theta^2 \cdot \text{Var}\left(\beta_{\text{exit},i}\right) + 2\theta \cdot \text{Cov}\left(\beta_{\text{exit},i}, \varepsilon_i\right)}{\text{Var}\left(\log\left(\text{GDP per capita}_i\right)\right)}\end{aligned}$$

This formula for the R^2 tells us what share of the variance in log output per capita would disappear if the selection channel were the same in all countries. One disadvantage of this formula is that, due to the covariance term, it is not additive across different factors that might explain GDP per capita (see discussion in Hsieh and Klenow (2010)). A popular alternative formula instead attributes half the covariance to each factor, yielding:

$$R^2_{\text{Alternative}} = \frac{\theta^2 \cdot \text{Var}\left(\beta_{\text{exit},i}\right) + \theta \cdot \text{Cov}\left(\beta_{\text{exit},i}, \varepsilon_i\right)}{\text{Var}\left(\log\left(\text{GDP per capita}_i\right)\right)}$$

which is also convenient because this is also the regression coefficient from an OLS regression of $\theta\beta_{\text{exit},i}$ on $\log\left(\text{GDP per capita}_i\right)$.

What θ should we use when implementing this? A natural choice is $\theta = 1$: we have seen that the effect of selection on aggregate productivity is very well approximated by β_{exit}. However, since we have log *output* per capita on the right hand side, we may also wish to account for the effect of differences in aggregate productivity on capital. Suppose that $\frac{K}{Y}$ is constant, as implied by a Solow model with constant savings rate and depreciation. Then, $\Delta \log Y = \Delta \log K$. Suppose also

	$\theta = 1$	$\theta = 1.5$
R^2	-12.9%	-23.1%
95% Confidence Interval	(-51.1%, 9.2%)	(-82.5%, 11.1%)
$R^2_{\text{Alternative}}$	-4.0%	-6.0%
95% Confidence Interval	(-21.5%, 6.5%)	(-32.2%, 10.0%)
Observations	16	16

Table 7: R^2 of Selection Channel Across Countries

Notes: This table shows estimates of the R^2 of the selection channel in explaining cross-country differences in output. The estimates rely on the same data as Figure 3. I compute R^2 as $R^2 = \frac{\theta^2 \cdot \text{Var}(\beta_{\text{exit},i}) + 2\theta \cdot \text{Cov}(\beta_{\text{exit},i}, \varepsilon_i)}{\text{Var}(\log(\text{GDP per capita}_i))}$, and I also compute an alternative formula, $R^2_{\text{Alternative}} = \frac{\theta^2 \cdot \text{Var}(\beta_{\text{exit},i}) + \theta \cdot \text{Cov}(\beta_{\text{exit},i}, \varepsilon_i)}{\text{Var}(\log(\text{GDP per capita}_i))}$. The formula depends on θ, which governs the relationship between $-\beta_{\text{exit}}$ and the selection channel. The first column shows results for $\theta = 1$, which corresponds to the direct effect of the selection channel on aggregate productivity, and the second column shows results for $\theta = 1.5$, which incorporates the indirect effect on output through capital accumulation. I compute point estimates using the sample analogs of each formula, and compute bias-corrected confidence intervals using the non-parametric bootstrap.

that $\Delta \log Y = \Delta \log A + \frac{2}{3}\Delta \log L + \frac{1}{3}\Delta \log L$: this is true globally for a Cobb-Douglas production function $Y = AK^{\frac{1}{3}}L^{\frac{2}{3}}$, and is true locally for a constant returns to scale aggregate production function with the appropriate factor shares. If labor supply is constant, then it follows that $\Delta \log Y = \frac{3}{2} \cdot \Delta \log A$. This would suggest a choice of $\theta = 1.5$.

I estimate R^2 and $R^2_{\text{Alternative}}$ for both $\theta = 1$ and $\theta = 1.5$, using the data from Figure A.10. I use the non-parametric bootstrap to construct bias-corrected confidence intervals.[13] I report the results in Table 7.

The point estimates for R^2 are consistently negative: since β_{exit} is negatively correlated with development in sample, the gap between rich and poor countries would be larger if not for the selection channel. Because I only have data for 16 countries, the confidence intervals are wide. Nevertheless, I can rule out large R^2 for the selection channel. Across estimates, the high end of the 95% confidence interval ranges from 6.5% ($R^2_{\text{Alternative}}$ with $\theta = 1$) to 11.1% (R^2 for $\theta = 1.5$).

The selection channel does not appear able to explain large differences in output per capita across countries. The variance of β_{exit} is too small, and the covariance with other factors that affect development is negative. Due to the small sample size, I cannot rule out modest explanatory power for the selection channel, but I can rule out a sizable role for the selection channel.

6 Conclusion

In this paper, I have examined the effect of selective exit of firms on aggregate productivity. A back-of-the-envelope approximation tells us that the selection channel is roughly equal to the difference in mean log productivity between surviving and exiting firms. The full model, calibrated

[13] I construct confidence intervals using 10,000 bootstrap replications.

to microdata from Indonesia, Chile, Colombia, and Spain, verifies that this back-of-the-envelope approximation is highly accurate.

The estimates of β_{exit} suggest that the selection channel is an important component of aggregate productivity. I find that Indonesia's aggregate productivity is 12% higher thanks to the selection channel, relative to a benchmark of random exit. If Indonesia's exit became as selective as Spain, it could increase its aggregate productivity by roughly 30%.

Yet the selection channel does not explain a meaningful component of differences in output per capita across countries. In a cross-country comparison, the strength of selection is not strongly correlated with development. This makes sense given the relative size of cross-country differences in output per capita: gaps in output per capita are much larger than cross-country differences in the selection channel. The selection channel is important, but it is not a primary explanation for lack of development in poor countries.

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, Reallocation, and Growth," *American Economic Review*, nov 2018, *108* (11), 3450–3491. 1

Asturias, Jose, Sewon Hur, Timothy J. Kehoe, and Kim J. Ruhl, "Firm Entry and Exit and Aggregate Growth," *American Economic Journal: Macroeconomics*, jan 2023, *15* (1), 48–105. 1

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Measuring and Analyzing Cross-Country Differences in Firm Dynamics," in "Producer Dynamics," University of Chicago Press, 2009, pp. 15–80. 1, 3.4, 5, 6, 3, A.10

Bento, Pedro and Diego Restuccia, "Misallocation, Establishment Size, and Productivity," *American Economic Journal: Macroeconomics*, July 2017, *9* (3), 267–303. 1, 2.1, 7

Castillo-Martinez, Laura, "Sudden Stops, Productivity, andthe Exchange Rate," Technical Report 2020. 1

Cooley, Thomas F and Vincenzo Quadrini, "Financial Markets and Firm Dynamics," *American Economic Review*, dec 2001, *91* (5), 1286–1310. 1

Demirer, Mert, "Production Function Estimation with Factor-Augmenting Technology: An Application to Markups," December 2020. 3.3.1

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, February 2008, *98* (1), 394–425. 2.3

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, sep 1992, *60* (5), 1127. 1

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448. 1, 5.1

_ and _ , "Development Accounting," *American Economic Journal: Macroeconomics*, January 2010, *2* (1), 207–223. 5.2

_ and _ , "The Life Cycle of Plants in India and Mexico," *The Quarterly Journal of Economics*, May 2014, *129* (3), 1035–1084. 1, 2.1, 7

Jaef, Roberto N. Fattal, "Entry and Exit, Multiproduct Firms, and Allocative Distortions," *American Economic Journal: Macroeconomics*, apr 2018, *10* (2), 86–112. 1

Jovanovic, Boyan, "Selection and the Evolution of Industry," *Econometrica*, may 1982, *50* (3), 649. 1

Kalouptsidi, Myrto, Paul T. Scott, and Eduardo Souza-Rodrigues, "Identification of counterfactuals in dynamic discrete choice models," *Quantitative Economics*, 2021, *12* (2), 351–403. 2.5

Kochen, Federico, "Finance Over the Life Cycle of Firms," Technical Report 2022. 1

Luttmer, E. G. J., "Selection, Growth, and the Size Distribution of Firms," *The Quarterly Journal of Economics*, aug 2007, *122* (3), 1103–1144. 1

Nakamura, Emi and Jón Steinsson, "Identification in Macroeconomics," *Journal of Economic Perspectives*, aug 2018, *32* (3), 59–86. 1

Peters, Michael, "Heterogeneous Markups, Growth, and Endogenous Misallocation," *Econometrica*, 2020, *88* (5), 2037–2073. 1

_ **and Fabrizio Zilibotti**, "Creative Destruction, Distance to Frontier, and Economic Development," Technical Report oct 2021. 1

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, October 2008, *11* (4), 707–720. 1, 2.1

Yang, Mu-Jeung, "Micro-Level Misallocation and Selection," *American Economic Journal: Macroeconomics*, oct 2021, *13* (4), 341–368. 1

A Additional Figures and Tables

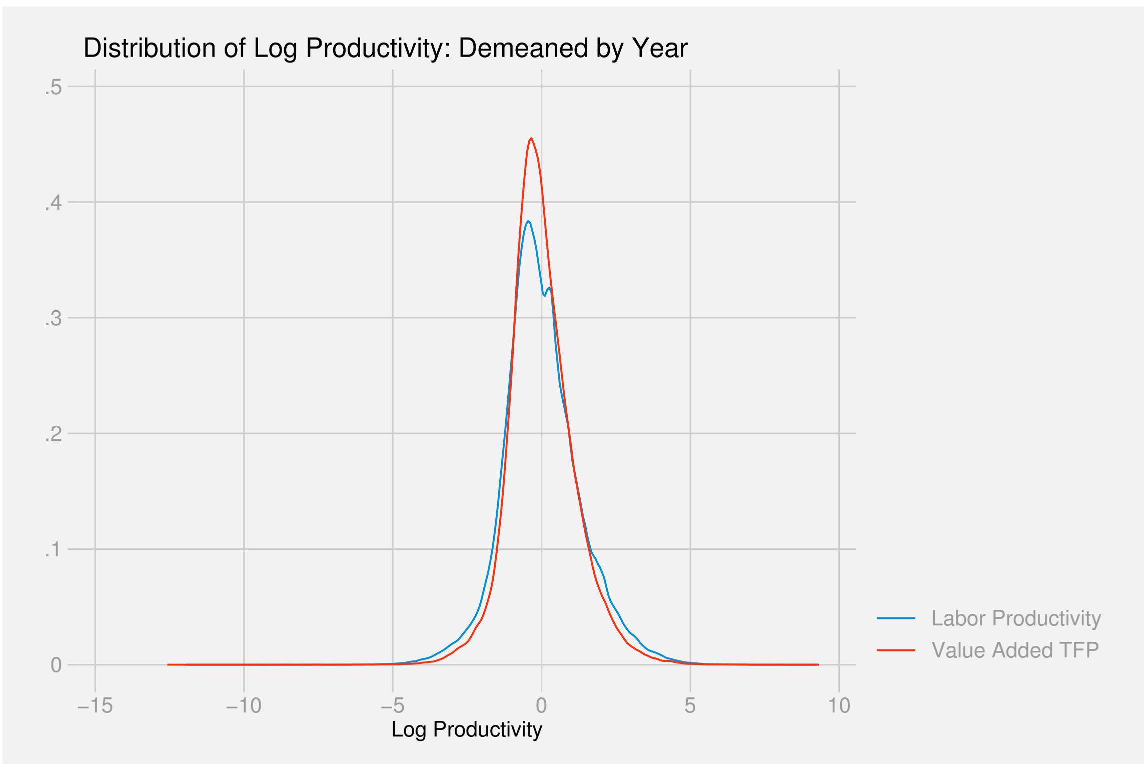


Figure A.1: Distribution of Productivity for Indonesia (Demeaned by Year)

Notes: This figure shows kernel density plots of the distribution of log productivity, using the Indonesian data. The data are residualized on industry-year fixed effects before plotting the density. I define labor productivity as value-added per worker. I define Value Added TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}} k^{\frac{1}{3}}}$.

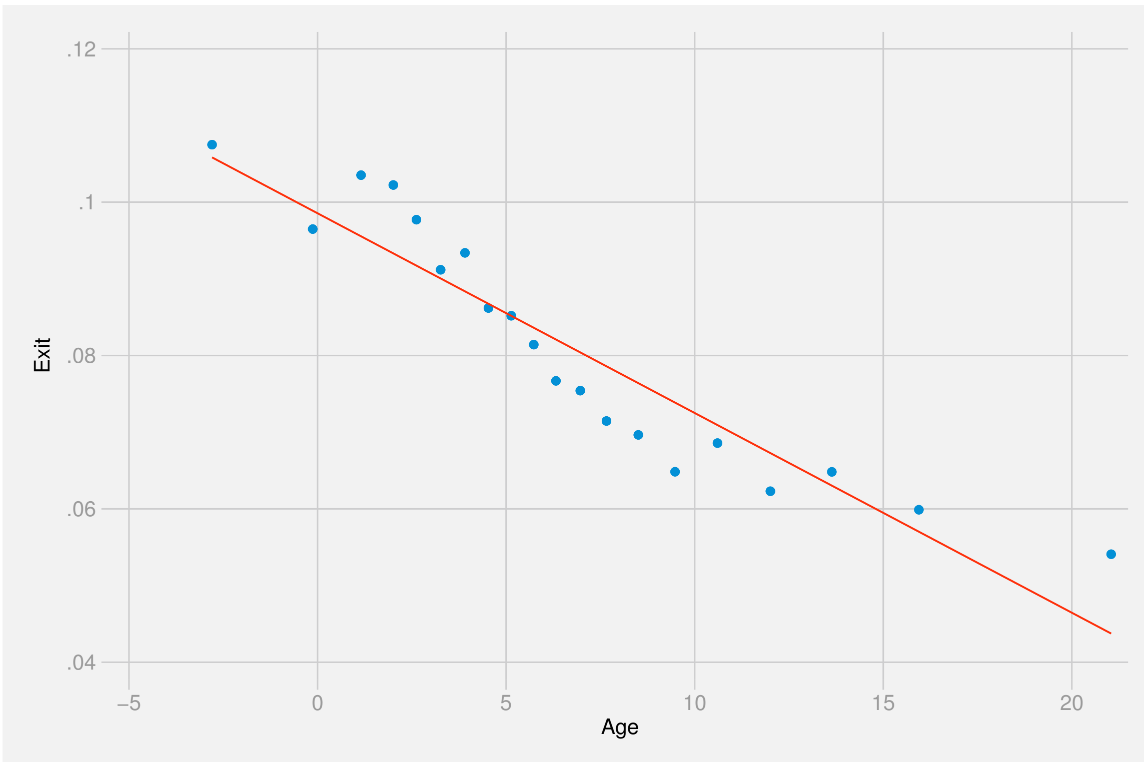


Figure A.2: Exit vs. Age for Indonesia (Controlling for Industry-Year Fixed Effects)

Notes: This figure shows a binned scatter plots of exit on age, using the Indonesian data. The data are residualized on industry-year fixed effects before forming bins and plotting. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year of the sample.

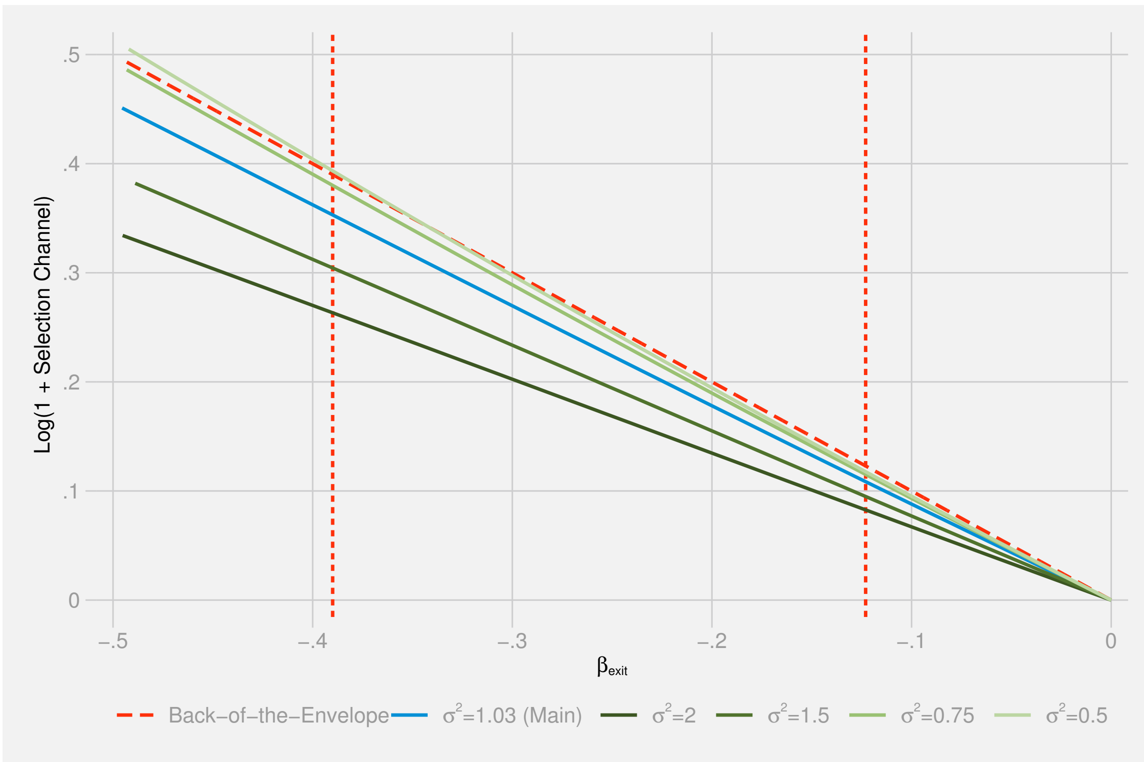


Figure A.3: Selection Channel vs. β_{exit} (Varying σ^2)

Notes: This figure shows the relationship between the selection channel and β_{exit}, for different levels of σ^2. The (diagonal) red dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit}} = \log(1 + \text{Selection Channel})$. The blue line is the main model, with $\sigma^2 = 1.03$. The green lines represent different choices of σ^2, with darker lines corresponding to higher values of σ^2. Each point on the blue and green lines represents a solution to the model, run with exit rates by age held constant and with selection by age, $\beta_{\text{exit},a}$, scaled proportionally up or down. For each run of the model, I compute the selection channel and also compute what the estimated β_{exit} would be in data generated from the solved model. I also include vertical dashed lines, which mark the preferred estimates of β_{exit} for Indonesia and Spain, to facilitate cross-country comparisons for each calibration of the model.

	Exit	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log(\text{Value Added TFP})$	$\log(\text{Workers})$
Age	-0.00260	0.00411	0.00272	0.0369
	(0.0000602)	(0.000573)	(0.000592)	(0.000843)
Observations	514725	531957	311612	534383
Industry-Year FE	Yes	Yes	Yes	Yes

Table A.1: Regressions on Age for Indonesia

Notes: This table shows regressions of exit, log productivity, and the number of workers on age for the Indonesian data. All regressions control for industry-year fixed effects. I define Value Added (VA) TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$. The number of observations is substantially smaller for TFP measures, because data on capital is missing for many firms and years. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	Exit	$\log(\text{Workers})$	Exit	$\log(\text{Workers})$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.00781	0.221		
	(0.000345)	(0.00361)		
$\log(\text{Value Added TFP})$			-0.00975	0.234
			(0.000497)	(0.00399)
Observations	604151	627485	325516	340824
Industry-Year FE	Yes	Yes	Yes	Yes

Table A.2: Regressions on Productivity for Indonesia

Notes: This table shows regressions of exit and the number of workers on log productivity for the Indonesian data. All regressions control for industry-year fixed effects. I define Value Added (VA) TFP as $\frac{\text{Value Added}}{l^{\frac{2}{3}}k^{\frac{1}{3}}}$. The number of observations is substantially smaller for TFP measures, because data on capital is missing for many firms and years. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	Exit	$\log(\text{Workers})$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.00650	0.202
	(0.000374)	(0.00342)
Age	-0.00259	0.0358
	(0.0000607)	(0.000798)
Observations	509484	531957
Industry-Year FE	Yes	Yes

Table A.3: Regressions on Productivity and Age for Indonesia

Notes: This table shows regressions of exit and the number of workers on log productivity and age for the Indonesian data. All regressions control for industry-year fixed effects. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	Mean	Standard Deviation	10th Percentile	Median	90th Percentile	Observations
Exit	0.063	0.243	0.000	0.000	0.000	76226
Age	3.915	3.633	0.000	3.000	9.000	27292
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	0.000	0.874	-1.034	0.036	0.999	81326
$\log\left(\text{Workers}\right)$	3.578	1.001	2.485	3.332	5.056	81516

Table A.4: Summary Statistics for Chile

Notes: This table shows summary statistics for the Chilean data. I demean log productivity by year, thus the mean is mechanically equal to zero. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	Exit	$\log\left(\text{Workers}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.0372	0.313			
	(0.00122)	(0.0108)			
Exit			-0.502	-0.406	-0.418
			(0.0158)	(0.0232)	(0.0144)
Observations	76071	81326	76071	50884	75115
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
$\geq$ 20 Workers	No	No	No	Yes	No

Table A.5: Regressions on Productivity for Chile

Notes: This table shows regressions of exit and the number of workers on log productivity, as well as estimates of β_{exit}, for the Chilean data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year. All regressions control for industry-year fixed effects. The third, fourth, and fifth column provide estimates of β_{exit}. The estimate in the third column is the baseline estimate. The fourth column subsets to firms with twenty or more workers in order to improve comparability with the Indonesian data. The fifth column uses log revenue per worker as the outcome instead of log value added per worker.

	Mean	**Standard Deviation**	**10th Percentile**	**Median**	**90th Percentile**	**Observations**
Exit	0.109	0.312	0.000	0.000	1.000	95612
Age	3.027	2.955	0.000	2.000	7.000	48540
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	0.000	0.781	-0.813	-0.045	0.913	102817
log (Workers)	3.448	1.113	2.303	3.258	5.004	102844

Table A.6: Summary Statistics for Colombia

Notes: This table shows summary statistics for the Colombian data. I demean log productivity by year, thus the mean is mechanically equal to zero. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	Exit	log (Workers)	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.0534 (0.00161)	0.373 (0.0138)			
Exit			-0.342 (0.00987)	-0.330 (0.0160)	-0.358 (0.0102)
Observations	95521	102817	95521	58539	95533
Industry-Year FE	Yes	Yes	Yes	Yes	Yes
$\geq$ 20 Workers	No	No	No	Yes	No

Table A.7: Regressions on Productivity for Colombia

Notes: This table shows regressions of exit and the number of workers on log productivity, as well as estimates of β_{exit}, for the Colombian data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year. All regressions control for industry-year fixed effects. The third, fourth, and fifth column provide estimates of β_{exit}. The estimate in the third column is the baseline estimate. The fourth column subsets to firms with twenty or more workers in order to improve comparability with the Indonesian data. The fifth column uses log revenue per worker as the outcome instead of log value added per worker.

	Mean	Standard Deviation	10th Percentile	Median	90th Percentile	Observations
Exit	0.080	0.271	0	0	0	6313709
Age	1.746	1.683	0	1	4	1425250
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	0	0.868	-0.925	-0.006	0.968	3741407
$\log(\text{Workers})$	1.374	1.191	0.000	1.099	2.944	4421497

Table A.8: Summary Statistics for Spain

Notes: This table shows summary statistics for the Spanish data. I demean log productivity by year, thus the mean is mechanically equal to zero. I compute age as the number of years since the firm first entered the sample, with age being missing for firms that appear in the sample in the first year of the data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year.

	Exit	$\log(\text{Workers})$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.0281	0.0786			
	(0.000169)	(0.00154)			
Exit			-0.390	-.342	-0.400
			(0.002)	(0.007)	(0.002)
Observations	3741407	3741407	3741407	406019	4378371
Year FE	Yes	Yes	Yes	Yes	Yes
$\geq$ 20 Workers	No	No	No	Yes	No

Table A.9: Regressions on Productivity for Spain

Notes: This table shows regressions of exit and the number of workers on log productivity, as well as estimates of β_{exit}, for the Spanish data. I compute exit as being equal to one in the last year that the firm ever appears in the sample: for that reason, exit is missing in the last year. All regressions control for year fixed effects, but do not contain industry-year fixed effects. The third, fourth, and fifth column provide estimates of β_{exit}. The estimate in the third column is the baseline estimate. The fourth column subsets to firms with twenty or more workers in order to improve comparability with the Indonesian data. The fifth column uses log revenue per worker as the outcome instead of log value added per worker.

Country	$(1-\rho) \times (-\beta_{\text{exit}})$	Real GDP per Capita (2010, USD)
Bartelsman et al. (2009)		
Argentina	0.098	10,385
Chile	0.432	12,808
Colombia	0.627	6,336
Estonia	0.28	14,790
Finland	0.251	46,459
France	0.107	40,638
Korea	0.495	23,087
Latvia	-0.037	11,348
Netherlands	0.025	50,950
Portugal	0.394	22,498
Slovenia	0.252	23,509
Taiwan	0.264	—
U.K.	0.051	39,435
U.S.	0.299	48,467
West Germany	0.274	41,531
Own Analysis		
Indonesia	0.155	3,122
Spain	0.303	30,502
Chile	0.380	12,808
Colombia	0.329	6,336

Table A.10: Cross-Country Comparison of β_{exit}

Notes: This table shows estimates, by country, of $(1-\rho) \cdot (-\beta_{\text{exit}})$, or one minus the exit rate times the difference in log productivity between stayers and exiters, plotted against each country's log real GDP per capita in 2010. Data on real GDP per capita come from the World Bank's World Development Indicators. I use Germany's GDP per capita for West Germany. The WDI do not contain GDP data for Taiwan. I draw estimates of $(1-\rho) \cdot (-\beta_{\text{exit}})$ from column 3 of Table 1.9 in Bartelsman et al. (2009). I supplement their estimates with my own estimates for Indonesia, Spain, Chile, and Colombia. Bartelsman et al. (2009) use log revenue per worker as their outcome variable, and define exit as equal to one if the firm exits any time over the next three years: I use the same definitions for my analysis of Indonesia, Spain, Chile, and Colombia to maintain consistency.

B Complications for the Truncated Lognormal Distribution

When I implement the calibrated model, I use a truncated log-normal distribution for productivity, rather than a log-normal distribution. In this appendix, I work out the issues that arise due to using this truncated distribution. I begin by showing that my results on aggregate productivity for the log-normal distribution go through, but with some correction terms. I show that, as was true in the log-normal case, the wage falls out, and so it is not necessary to solve for the equilibrium wage in order to determine aggregate productivity. I then show how the truncation of the productivity distribution affects the mapping from $\beta_{\text{exit},a}$ to δ_a, and how I solve this issue numerically using a contraction mapping.

B.1 Aggregate Productivity with Truncated Log-Normal Productivity

Let $\bar{z}$ be the upper bound on productivity. We then have:

$$\log z \mid a \sim N\left(\mu_a, \sigma^2\right)$$

$$\mathbb{E}\left[z \mid a, z < \bar{z}\right] = \exp\left(\mu_a + \frac{1}{2}\sigma^2\right) \cdot \frac{\Phi\left(\frac{\log \bar{z} - \mu_a - \sigma^2}{\sigma}\right)}{\Phi\left(\frac{\log \bar{z} - \mu_a}{\sigma}\right)}$$

Recall that output and labor are linear in logs:

$$\log q = \frac{1}{\gamma}\log z + \frac{1}{\gamma}\log\left(\frac{c_\tau}{w}\right)$$
$$\log l = \frac{1-\gamma}{\gamma}\log z + \frac{1}{\gamma}\log\left(\frac{c_\tau}{w}\right)$$

We thus have

$$\log \mathbb{E}\left[q \mid a\right] = \frac{\mu_a}{\gamma} + \frac{1}{2\gamma^2}\sigma^2 + \frac{1}{\gamma}\log\left(\frac{c_\tau}{w}\right) + \log\Phi\left(\frac{\log \bar{z} - \mu_a - \frac{1}{\gamma}\sigma^2}{\sigma}\right) - \log\Phi\left(\frac{\log \bar{z} - \mu_a}{\sigma}\right)$$
$$\log \mathbb{E}\left[l \mid a\right] = \frac{\mu_a}{\gamma}\left(1-\gamma\right) + \frac{\left(1-\gamma\right)^2}{2\gamma^2}\sigma^2 + \frac{1}{\gamma}\log\left(\frac{c_\tau}{w}\right) + \log\Phi\left(\frac{\log \bar{z} - \mu_a - \frac{1-\gamma}{\gamma}\sigma^2}{\sigma}\right) - \log\Phi\left(\frac{\log \bar{z} - \mu_a}{\sigma}\right)$$

Note that, as in the case without the truncation, the wage appears only as a $\frac{1}{\gamma}\log\left(\frac{c_\tau}{w}\right)$ term in the expressions for the log expected output and log expected labor (this is also the only place that c_τ appears). This multiplier term will cancel out when computing a cohort's aggregate productivity, and similarly will fall out when we compute the labor share of each cohort.

This yields the formula for cohort level productivity:

$$\log Z := \log\frac{\mathbb{E}\left[q\right]}{\mathbb{E}\left[l\right]} = \underbrace{\mu_a + \frac{1}{2}\sigma^2}_{\log \mathbb{E}[z]} + \underbrace{\left(\frac{1}{\gamma} - 1\right)\sigma^2}_{\text{Benefits of Allocation}} + \underbrace{\log\Phi\left(\frac{\log \bar{z} - \mu_a - \frac{1}{\gamma}\sigma^2}{\sigma}\right) - \log\Phi\left(\frac{\log \bar{z} - \mu_a - \frac{1-\gamma}{\gamma}\sigma^2}{\sigma}\right)}_{\text{Truncation Correction Terms}}$$

which is very similar to the original formula for aggregate productivity (as noted earlier, this formula does not depend on c_τ or on w). The formula for the cohort-level labor share is calculated similarly. I use the two formulas together to calculate aggregate productivity.

B.2 Selection with Truncated Log-Normal Productivity

With log-normal productivity, the mapping from $\beta_{\text{exit},a}$ to δ_a was straightforward: $\rho_a\left(-\beta_{\text{exit}}\right)$ is the difference in the mean of log productivity between the full population and the survivors, so $\delta_a\sigma^2 = \rho_a\left(-\beta_{\text{exit},a}\right)$. With truncated log-normal productivity, we still have that $\rho_a\left(-\beta_{\text{exit}}\right)$ is the shift in log productivity. However, we must select the appropriate shifter of μ to achieve this difference after truncation.

Let $\log z \sim N\left(\mu, \sigma^2, \bar{z}\right)$, where $\bar{z}$ is the upper bound, and let $\log z \mid \text{survival} \sim N\left(\mu + \theta, \sigma^2, \bar{z}\right)$. Note that $\mathbb{E}\left[\log z\right] = \mu - \sigma \cdot \frac{\phi\left(\frac{\bar{z}-\mu}{\sigma}\right)}{\Phi\left(\frac{\bar{z}-\mu}{\sigma}\right)}$. We then have:

$$\mathbb{E}\left[\log z \mid \text{survival}\right] - \mathbb{E}\left[\log z\right] = \theta + \sigma\left(\frac{\phi\left(\frac{\bar{z}-\mu}{\sigma}\right)}{\Phi\left(\frac{\bar{z}-\mu}{\sigma}\right)} - \frac{\phi\left(\frac{\bar{z}-\mu-\theta}{\sigma}\right)}{\Phi\left(\frac{\bar{z}-\mu-\theta}{\sigma}\right)}\right)$$

We can solve this with the contraction mapping (our initial guess of θ can be zero, which will initially yield the "naive" $\theta = \mathbb{E}\left[\log z \mid \text{survival}\right] - \mathbb{E}\left[\log z\right]$):

$$\theta' = \underbrace{\mathbb{E}\left[\log z \mid \text{survival}\right] - \mathbb{E}\left[\log z\right]}_{-\rho_a \cdot \beta_{\text{exit},a}} - \sigma\left(\frac{\phi\left(\frac{\bar{z}-\mu}{\sigma}\right)}{\Phi\left(\frac{\bar{z}-\mu}{\sigma}\right)} - \frac{\phi\left(\frac{\bar{z}-\mu-\theta}{\sigma}\right)}{\Phi\left(\frac{\bar{z}-\mu-\theta}{\sigma}\right)}\right)$$

This can be proven to be a contraction mapping because the derivative of the right hand side with respect to θ is bounded between zero and one.[14] The contraction will be especially fast when $\frac{\bar{z}-\mu-\theta}{\sigma}$ is large, since the derivative of the right hand side will be very close to zero.

[14] For my purposes, it would be sufficient to simply find the fixed point of this equation, but the fact that there is a contraction makes computation fast.

C Details on Estimating γ

How should we measure γ? We use γ in two places: to compute within-cohort aggregate productivity, and to compute the share of labor employed by each cohort. The model does not allow for any variation in labor except for the variation induced by differences in productivity, and so it is not obvious how to account for this in the estimation of γ, which governs the firm size-productivity relationship. In this appendix, I will show that γ should be computed based on a regression of $\log q$ on $\log z$. I show that this regression will yield the right parameter in each of these cases. I will also clarify what assumptions from the model are necessary to ensure that this method of estimating γ gives us correct results.

C.1 Aggregate Productivity

In the aggregate productivity case, assume $(\log z, \log l)$ are multivariate normal. Then, we have:

$$
\begin{aligned}
\log Z \mid a &:= \log \frac{\mathbb{E}\left[zl \mid a\right]}{\mathbb{E}\left[l \mid a\right]} \\
&= \left(\mathbb{E}\left[\log z \mid a\right] + \mathbb{E}\left[\log l \mid a\right] + \frac{1}{2}\mathrm{Var}\left(\log z + \log l \mid a\right)\right) \\
&\quad - \left(\mathbb{E}\left[\log l \mid a\right] + \frac{1}{2}\mathrm{Var}\left(\log l \mid a\right)\right) \\
&= \mathbb{E}\left[\log z \mid a\right] + \frac{1}{2}\left(\mathrm{Var}\left(\log z \mid a\right) + 2 \cdot \mathrm{Cov}\left(\log z, \log l \mid a\right)\right) \\
&\quad + \frac{1}{2}\mathrm{Var}\left(\log l \mid a\right) - \frac{1}{2}\mathrm{Var}\left(\log l \mid a\right) \\
&= \underbrace{\mu_a + \frac{1}{2}\sigma^2}_{\log \mathbb{E}[z|a]} + \underbrace{\left(\frac{1}{\gamma} - 1\right)\sigma^2}_{\mathrm{Cov}(\log z, \log l)}
\end{aligned}
$$

This derivation did not use anything from the model except for the joint lognormality of $\log z$ and $\log l$. The benefits of allocation term, which is $\left(\frac{1}{\gamma} - 1\right)\sigma^2$ in the model is simply the covariance of $\log z$ and $\log l$. Thus, for aggregate productivity, we want the covariance of log productivity and log labor, and so we want to have $\left(\frac{1}{\gamma} - 1\right) = \frac{\mathrm{Cov}(\log z, \log l)}{\sigma^2}$, or equivalently $\frac{1}{\gamma} = \frac{\mathrm{Cov}(\log z, \log q)}{\sigma^2}$. This is exactly what we get from a regression of $\log q$ on $\log z$.

C.2 Labor Share of Each Cohort

Now we move to the share of labor employed by each cohort. We again assume that $(\log z, \log l)$ are multivariate normal. Labor employed is:

$$
\begin{aligned}
\log \mathbb{E}\left[l \mid a\right] &= \mathbb{E}\left[\log l \mid a\right] + \frac{1}{2}\mathrm{Var}\left(\log l \mid a\right) \\
&= \mathbb{E}\left[\mathbb{E}\left[\log l \mid z, a\right] \mid a\right] + \frac{1}{2}\mathrm{Var}\left(\mathbb{E}\left[\log l \mid z, a\right] \mid a\right) + \frac{1}{2}\mathbb{E}\left[\mathrm{Var}\left(\log l \mid z, a\right) \mid a\right]
\end{aligned}
$$

Assume $\mathbb{E}[\log l \mid z, a] = \mathbb{E}[\log l \mid z]$. Also assume $\frac{(1-\gamma)}{\gamma}$ is obtained from a regression of $\log l$ on $\log z$, and $\frac{1}{\gamma}\mathbb{E}\left[\log\left(\frac{c_\tau}{w}\right)\right]$ is the intercept from that regression. We then have:

$$\log \mathbb{E}[l \mid a] = \mu_a \frac{(1-\gamma)}{\gamma} + \frac{1}{\gamma}\mathbb{E}\left[\log\left(\frac{c_\tau}{w}\right)\right] + \frac{(1-\gamma)^2}{2\gamma^2}\sigma^2 + \frac{1}{2}\mathbb{E}[\mathrm{Var}(\log l \mid z, a) \mid a]$$

If we assume that $\mathbb{E}[\mathrm{Var}(\log l \mid z, a) \mid a] = \mathrm{Var}(\log l \mid z)$, then the last term will be the same across cohorts, and so we still have that the share of labor employed by the cohort is

$$\frac{N_a \cdot \mathbb{E}[l \mid a]}{\sum_s N_s \cdot \mathbb{E}[l \mid s]} = \frac{N_a \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_a\right)}{\sum_s N_s \cdot \exp\left(\frac{1-\gamma}{\gamma}\mu_s\right)}$$

Thus, measuring $\frac{1}{\gamma}$ as the coefficient from a regression of $\log q$ on $\log z$ will give us the correct labor shares for each cohort.

This rests, again, on the assumption of joint log-normality of z and l. But it now also rests on two additional assumptions: $\mathbb{E}[\log l \mid z, a] = \mathbb{E}[\log l \mid z]$ and $\mathbb{E}[\mathrm{Var}(\log l \mid z, a) \mid a] = \mathrm{Var}(\log l \mid z)$. These assumptions require that the slope and residual variance of the relationship between log labor and log productivity are the same across cohorts. Under these three assumptions, the model results remain unchanged, and measuring $\frac{1}{\gamma}$ with the regression of $\log q$ on $\log z$ ensures the correct result.