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. I calibrate the model to firm microdata for Indonesia, Spain, Chile, and Colombia. I find 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 show that the strength of selection does not covary strongly with output per capita. This suggests that the selection channel is not a quantitatively important explanation for differences in development across countries.

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. I thank Marcos Correa for excellent research assistance. 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. I capture misallocation of inputs across firms with a parsimonious model of distortions. Firms with heterogeneous productivities face distortions that show up as a tax on revenue. The joint distribution of productivities and distortions can feature a covariance between productivity and the revenue wedge. Moreover, firms face a larger revenue tax rate the more they produce, creating decreasing returns to scale and limiting the economy's ability to take advantage of heterogeneous productivities. The model also features cohort exit dynamics: the probability that a firm exits is a function of the firm's productivity, revenue wedge, and age. I solve for the model's steady state distribution of firms, and compute aggregate productivity. 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, 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, the planner 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 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.

Perhaps more important than 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 statistic 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 and wedges). I first lay out the model, and then provide intution 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.

Model Overview

I begin by laying out a model with heterogeneous firms, that differ in their productivity, wedges, and age. There is a continuum of firms corresponding to each cohort, but time (and thus age) is discrete. Firms enter and exit, with an exogenous mass of firms entering each period, and drawing their productivity from an exogenous distribution. The probability of exit depends on age, wedges, 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 exit dynamics. I call this the selection channel.

2.1 Firm's Static Problem

Each firm, indexed by i, has a productivity, z_i, a wedge intercept, $c_{\tau,i}$, and an age, a_i. Firms produce intermediate goods, using a linear production function with labor as their input. Specifically, production is given by:

$$y_i = z_i l_i$$

where y_i is output and l_i is labor.

The intermediates produced by the firms are aggregated by a final good producer, using a CES production function. Normalizing the mass of firms in the economy to one, the final good producer's output is

$$Y = \left(\int_0^1 y_i^{\frac{\theta-1}{\theta}} di \right)^{\frac{\theta}{\theta-1}} = \mathbb{E} \left[y_i^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}$$

The final good producer takes prices as given. This yields the downward sloping firm-level demand curve for each variety:

$$p_i = Y^{\frac{1}{\theta}} \cdot y_i^{\frac{-1}{\theta}}$$

The firm also faces a revenue wedge, which can equivalently be viewed as a revenue tax. The firm's profits are reduced by $\tau_i(y_i)$ 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. Moreover, some firms face a higher tax schedule than others. The tax rate is given by:

$$1 - \tau_i(y_i) = \frac{c_{\tau,i}}{1-\gamma} y_i^{-\gamma}$$

where $c_{\tau,i} > 0$ is the firm-specific wedge intercept, and γ is a parameter, with $0 \leq \gamma < 1$. 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), but is distinct because it is a function of y_i rather than z_i. This generates "correlated distortions," where the most productive firms are the ones who face the highest wedges on the margin, as in Restuccia and Rogerson (2008).

The firm solves the static profit maximization problem:

$$\begin{aligned} \pi_i &= \max_{l_i} (1 - \tau_i(y_i)) p_i y_i - w l_i \\ &= \max_{y_i} (1 - \tau_i(y_i)) p_i y_i - \frac{w}{z_i} y_i \end{aligned}$$

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

$$\log y_i = \mathcal{E} \cdot \left[\log z_i + \log c_{\tau,i} + \frac{1}{\theta} \log Y + \log \left(\frac{1 - \gamma - \frac{1}{\theta}}{1-\gamma} \right) - \log w \right]$$

where $\mathcal{E} := \frac{\theta}{\gamma\theta+1}$ is the elasticity of output with respect to productivity.

Note the important role played by both the wedge slope parameter γ and by the CES parameter θ. Since production is linear, we need something else to ensure that firms face decreasing returns to scale, in order to ensure a bounded solution to the firm's problem. The CES aggregator creates decreasing returns to scale through downward sloping firm-level demand, and the progressive revenue tax also creates decreasing returns to scale. These mechanisms differ, however, in that the presence of a CES aggregator does not necessarily imply that there is misallocation, while the progressive revenue tax will definitely distort the economy away from the efficient allocation.

2.2 Entry and Exit Dynamics

The youngest cohort ($a = 0$) is populated by an exogenous mass of entrants, who draw their productivity and wedge intercepts from an exogenous distribution. Productivity and wedge intercepts are jointly log-normally distributed. That is:

$$(\log z_i, \log c_{\tau,i}) \mid a = 0 \sim N(\mu_0, \Sigma_0)$$

At the end of each period, 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

$$\log(1 - \Pr(\text{exit} \mid z, c_\tau, a)) = \min(0, c_{\text{exit},a} + \delta_{z,a} \log z + \delta_{\tau,a} \log c_\tau)$$

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

This functional form is convenient, because it preserves log-normality, setting aside the issue of truncation at $\Pr(\text{exit} \mid z, c_\tau, a) = 0$. Suppose that $(\log z_i, \log c_{\tau,i}) \mid a \sim N(\mu_a, \Sigma_a)$. Ignoring the issue that the exit rate cannot go below zero, the distribution of productivity and wedge intercepts for surviving firms will also be log-normal, with the same variance-covariance matrix Σ_a. The new means of $(\log z_i, \log c_{\tau,i})$ will be given by:

$$\mathbb{E}[\log z_i \mid \text{Survivor}, a] = \mathbb{E}[\log z_i \mid a] + \delta_{z,a}\text{var}(\log z_i) + \delta_{\tau,a}\text{cov}(\log z_i, \log c_{\tau,i})$$
$$\mathbb{E}[\log c_{\tau,i} \mid \text{Survivor}, a] = \mathbb{E}[\log c_{\tau,i} \mid a] + \delta_{z,a}\text{cov}(\log z_i, \log c_{\tau,i}) + \delta_{\tau,a}\text{var}(\log c_{\tau,i})$$

Focusing on the first line, the mean of log productivity for survivors depends on the initial mean, plus $\delta_{z,a}\text{var}(\log z_i)$, which reflects direct selection on productivity, plus $\delta_{\tau,a}\text{cov}(\log z_i, \log c_{\tau,i})$, which reflects indirect selection, via the covariance between productivity and the wedge intercept. We can rewrite the direct and indirect selection effects as $\text{var}(\log z_i) \cdot [\delta_{z,a} + \delta_{\tau,a}\text{cov}(\log z_i, \log c_{\tau,i}) / \text{var}(\log z_i)]$, and note that $\delta_{z,a} + \delta_{\tau,a}\text{cov}(\log z_i, \log c_{\tau,i}) / \text{var}(\log z_i)$ is equal to the coefficient from a regression of survival on log productivity: the first term is the direct effect of log productivity on survival, and the second term is an omitted variable bias term.

Microfoundations for Exit

There are many 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.

Because there are so many potential microfoundations, I do not take a stand on why survival depends on productivity, wedges and age, but instead I merely note it as a robust feature of the data, and model it in a reduced form way. My model of the firm's static problem allows me to compute the effect of selective exit on aggregate productivity, while remaining agnostic about the underlying model for the firm's exit decision. My results are robust to various microfoundations of exit, as long as (i) the firm's probability of exit depends only on its age, current productivity, and current wedge, and (ii) those microfoundations do not alter the firm's static problem. The former could be violated if, for example, firm's receive signals about their future productivity that are not captured by their current state variables, and decide to exit based on those signals. The latter could be violated if, for example, the firm can invest in productivity-enhancing technology, and decides whether to do so based on its expectation of how future exit rates depend on its own productivity. Despite these potential limitations, I argue that this reduced form is appealing in the breadth of microfoundations that it can handle, and is a useful step towards better understanding the selection channel.

2.3 Productivity Dynamics

In the baseline version of the model, I assume that an individual firm's productivity and wedge intercept does not change over time: for a given firm, $(z_{it}, c_{\tau,it}) = (z_{i0}, c_{\tau,i0})\,\forall t$. To properly solve for the dynamics of productivity and wedges, we need to simulate numerically, since the exit function does not preserve log-normality due to truncation at zero exit. However, we can use the log-normal approximation from before to get a sense of how productivity and wedges evolve across cohorts. Using the results from above, we have:

$$\begin{aligned}\mu_a &\approx \mu_0 + \sum_{s=0}^{a-1}\left[\delta_{z,s}\text{var}\left(\log z_i \mid s\right) + \delta_{\tau,s}\text{cov}\left(\log z_i, \log c_{\tau,i} \mid s\right)\right] \\ &\approx \mu_0 + \sum_{s=0}^{a-1}\left[\delta_{z,s}\text{var}\left(\log z_i \mid a=0\right) + \delta_{\tau,s}\text{cov}\left(\log z_i, \log c_{\tau,i} \mid a=0\right)\right]\end{aligned}$$

In an in-progress extension to the model, I add in Markov dynamics, based on a Gaussian VAR(1). These dynamics induce mean reversion, which attenuates 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.

2.4 Aggregate Productivity

To solve for aggregate productivity, we compute output of the final good, Y, and divide by labor demand. Note, from earlier, that we have $y_i = z_i^{\mathcal{E}} \cdot c_{\tau,i}^{\mathcal{E}} \cdot C$, where C is a constant that does not vary across firms. Similarly, we have $l_i = y_i / z_i = z_i^{\mathcal{E}-1} \cdot c_{\tau,i}^{\mathcal{E}} \cdot C$, where C is the same constant as before. Computing aggregate productivity, we have:

$$\begin{aligned}
\log Z &:= \log Y - \log \mathbb{E}\left[l_i\right] \\
&= \frac{\theta}{\theta-1} \log \mathbb{E}\left[y_i^{\frac{\theta-1}{\theta}}\right] - \log \mathbb{E}\left[l\right] \\
&= \frac{\theta}{\theta-1} \log \mathbb{E}\left[\left(z_i^{\mathcal{E}} \cdot c_{\tau,i}^{\mathcal{E}} \cdot C\right)^{\frac{\theta-1}{\theta}}\right] - \log \mathbb{E}\left[z_i^{\mathcal{E}-1} \cdot c_{\tau,i}^{\mathcal{E}} \cdot C\right] \\
&= \frac{\theta}{\theta-1} \log \mathbb{E}\left[\left(z_i^{\mathcal{E}} \cdot c_{\tau,i}^{\mathcal{E}}\right)^{\frac{\theta-1}{\theta}}\right] - \log \mathbb{E}\left[z_i^{\mathcal{E}-1} \cdot c_{\tau,i}^{\mathcal{E}}\right]
\end{aligned}$$

The fact that firm-level production is constant returns to scale, as is the CES aggregator, means that the constant term C falls out. Aggregate productivity depends solely on the joint distribution of productivities and wedge intercepts. Thus, we do not need to solve for the equilibrium wage in order to compute aggregate productivity.

In general, there will be no convenient expression for aggregate productivity: we need to simulate the steady-state distribution and solve numerically. However, to gain intuition, we will consider the case where $(\log z_i, \log c_{\tau,i})$ is multivariate normal. In this case, we have:

$$\log Z = \underbrace{\log \mathbb{E}\left[z_i\right]}_{\text{Average Productivity}} + \underbrace{\operatorname{cov}\left(\log l_i, \log z_i\right)}_{\text{Allocation of Production}} - \underbrace{\frac{1}{2} \cdot \frac{1}{\theta} \operatorname{var}\left(\log y_i\right)}_{\text{Preference for Variety}}$$

In the lognormal case, aggregate productivity depends on the unweighted average productivity, on an "allocation of production" term which reflects the fact that the labor-weighted average productivity will be higher than the unweighted average, and on a "preference for variety" term, which reflects the fact that under CES aggregation, consumers (represented by the final good producer) prefer to consume a bit of each variety, rather than focusing their consumption on a subset of varieties (i.e. consumers have convex preferences). We can express each of these terms as a function solely of the joint distribution of $\log z_i$ and $\log c_{\tau,i}$:

$$\begin{aligned} \text{Average Productivity} &= \mathbb{E}\left[\log z_i\right] + \frac{1}{2}\text{var}\left(\log z_i\right) \\ \text{Allocation of Production} &= \left(\mathcal{E}-1\right)\text{var}\left(\log z_i\right) + \mathcal{E}\text{cov}\left(\log z_i, \log c_{\tau,i}\right) \\ \text{Preference for Variety} &= -\frac{1}{2}\cdot\frac{1}{\theta}\mathcal{E}^2\cdot\left[\text{var}\left(\log z_i\right) + \text{var}\left(\log c_{\tau,i}\right) + 2\text{cov}\left(\log z_i, \log c_{\tau,i}\right)\right] \end{aligned}$$

Note that when $\gamma = 0$ (e.g. as in Hsieh and Klenow, 2009), the elasticity of output with respect to productivity is simply $\mathcal{E} = \theta$, which simplifies the expressions and makes the $\text{cov}\left(\log z_i, \log c_{\tau,i}\right)$ term cancel out.[1] Conversely, in the linear aggregation case, where $\theta \to \infty$, we have that $\frac{1}{\theta} \to 0$, so the preference for variety term disappears, and $\mathcal{E} \to \frac{1}{\gamma}$.

This expression reveals that aggregate productivity will depend importantly on average productivity: if every firm's productivity doubles, then aggregate productivity will also double. Moreover, the average wedge does not matter: only the dispersion of wedges matters.[2] Productivity will depend, however, on the variance of log productivity, the variance of the log wedge intercept, and their covariance. Aggregate productivity will be increasing in $\text{var}\left(\log z_i\right)$ if $\theta \geq \frac{1+\sqrt{5}}{2}$. Aggregate productivity will be decreasing in $\text{var}\left(\log c_{\tau,i}\right)$ for any finite θ, and will be increasing in $\text{cov}\left(\log z_i, \log c_{\tau,i}\right)$ if $\gamma > 0$.

Selective exit will obviously affect average productivity. But it may also affect higher order moments: for example, selective exit may narrow the distribution of productivity by cutting off the left tail. If aggregate productivity is increasing in the variance of log productivity, then this will be a countervailing force against the increase in mean productivity. The strength of this countervailing force will depend on parameters: for example, low values of γ and high values of θ will increase the importance of $\text{var}\left(\log z_i\right)$ for aggregate productivity. We will explore this particular idea (that allocation and selection are substitutes) further in Section 5.

2.5 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 and wedges conditional on cohort. More formally:

[1] In this environment, static misallocation is the difference between aggregate productivity under actual wedges and counterfactual aggregate productivity if the wedges were all set to zero. If $\gamma = 0$, we recover the Hsieh and Klenow (2009) result that misallocation is simply $\theta \cdot \text{var}\left(\log \tau_i\right)$. If $\gamma > 0$, then misallocation will be more complicated, and will also depend on the covariance of wedges and productivity.

[2] This is a standard result in models of misallocation with uniform returns to scale: if all firms are affected by a uniform wedge, then that will cause wages to fall, but it will not create misallocation of resources across firms (conditional on the aggregate supply of labor). In the constant returns to scale case, this also means that there will be no effect on aggregate productivity.

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

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

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 next provide a back-of-the-envelope approximation, which shows that the selection channel will be approximately equal to the difference in mean log productivity between survivors and exiters.

2.6 Back-of-the-Envelope Approximation

The basic insight of the model 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. Mirroring our earlier notation, let μ denote average log productivity, and μ^* denote average log productivity in the economy where exit is random conditional on age. The approximate selection channel is $\mu - \mu^*$.

Our derivation will rely on a few key equations. First, we have the following identity:

$$\mu = \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}}$$

where μ_{Exiters} is the average log productivity of exiting firms, $\mu_{\text{Survivors}}$ is the average log productivity of surviving firms, and ρ is the exit rate. This identity represents the fact that the average log productivity in a given period will be a weighted average of the log productivity of firms who will survive to the next period, and the log productivity of firms who will not survive.

Second, we have the following steady-state condition:

$$\begin{aligned} \mu_t &= \mu_{t+1} \\ \Longrightarrow\ \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} &= \rho\mu_{\text{Entrants}} + (1-\rho)\,\mu_{\text{Survivors}} \\ \Longrightarrow\ \mu_{\text{Exiters}} &= \mu_{\text{Entrants}} \end{aligned}$$

where μ_{Entrants} is the average log productivity of new entrants. This steady state condition relies on no population growth and no change in firm productivity over time. Under these conditions, in steady state, the productivity of the entrants must exactly equal the productivity of the exiting firms that they replace, in order to maintain the same distribution of productivity.

Finally, we have an equation for the counterfactual average log productivity under no selection:

$$\mu^* = \mu_{\text{Entrants}}$$

This is a result of no productivity growth over time. Since the productivity of a given firm is the same in all periods it is alive, the no-selection counterfactual will simply feature the same productivity distribution as the initial entrants.

Combining these three equations, we can derive our back of the envelope:

$$\begin{aligned}
\text{Selection Channel} &\approx \mu - \mu^* \\
&= \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} - \mu_{\text{Entrants}} \\
&= \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} - \mu_{\text{Exiters}} \\
&= (1-\rho)\,(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}) \\
&\approx (\mu_{\text{Survivors}} - \mu_{\text{Exiters}}) \\
&=: -\beta_{\text{exit},z}
\end{aligned}$$

where the second to last line comes from assuming that ρ is small.

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. I will refer to this quantity as $-\beta_{\text{exit},z}$, because it is the (negative) coefficient from a regression of log productivity on an indicator for exit.

There are three main reasons why effect on aggregate productivity will differ from this back-of-the-envelope approximation. First, the back-of-the-envelope focuses on the effect on the mean of log productivity, which is not the same thing as aggregate productivity. Aside from the fact that the mean of the log is not the log of the mean, 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. The full model addresses this, by layering a model of static allocation on top of the model of entry and exit dynamics.

Second, the addition of firm-level productivity dynamics will affect the back-of-the-envelope, both by altering the steady state condition and potentially by changing the equation for μ^*. In the appendix, I show how to incorporate firm dynamics into the back-of-the-envelope approximation. Mean reverting firm dynamics will tend to attenuate the selection channel: if there is mean reversion, then selection on today's productivity will not necessarily imply selection on tomorrow's productivity, and so the selection channel will be weaker than $\beta_{\text{exit},z}$ suggests. In contrast, life-cycle productivity growth (firms enter at a low productivity but become more productive as they age) will make the selection channel stronger than suggested by $(\mu_{\text{Survivors}} - \mu_{\text{Exiters}})$. Intuitively, life-cycle productivity growth means that the gap between exiters and survivors will understate the gap between survivors and entrants, and thus $\beta_{\text{exit},z}$ will understate the true size of the selection channel.

Finally, population growth will attenuate the selection channel. In the appendix, I show how to incorporate population growth into the back-of-the-envelope approximation. Population growth changes the steady-state condition: instead of entrants replacing exiters, there are now more entrants than exiters, and entrant productivity is thus a weighted average of the productivity of exiters and the productivity of survivors. Since the selection channel depends on the difference between $\mu_{\text{Survivors}}$ and μ_{Entrants}, this will make the selection channel weaker than implied by $\beta_{\text{exit},z}$.

In the next two sections, I will calibrate the full model to the data, and compare the strength of the selection channel to this back-of-the-envelope approximation. To preview the results: I will find that the back-of-the-envelope is in fact a very good approximation to the true selection channel, within the empirically relevant range of parameters. This will then allow me to extend my analysis further, to countries for which I cannot fully calibrate the model, but can compute the difference in mean log productivity between survivors and exiters. This cross-country analysis allows me to take a development accounting perspective, and answer the question of whether the selection channel is an important contributor to cross-country differences in output per capita.

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.[3] 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

[3]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
$\log(\text{Value Added TFP})$	0	1.107	-1.199	-0.107	1.401	341305
$\log(\text{Workers})$	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.

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, would be $(1-\rho)/\rho \approx 12$ years.[4] 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:

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

[4] 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.

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 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

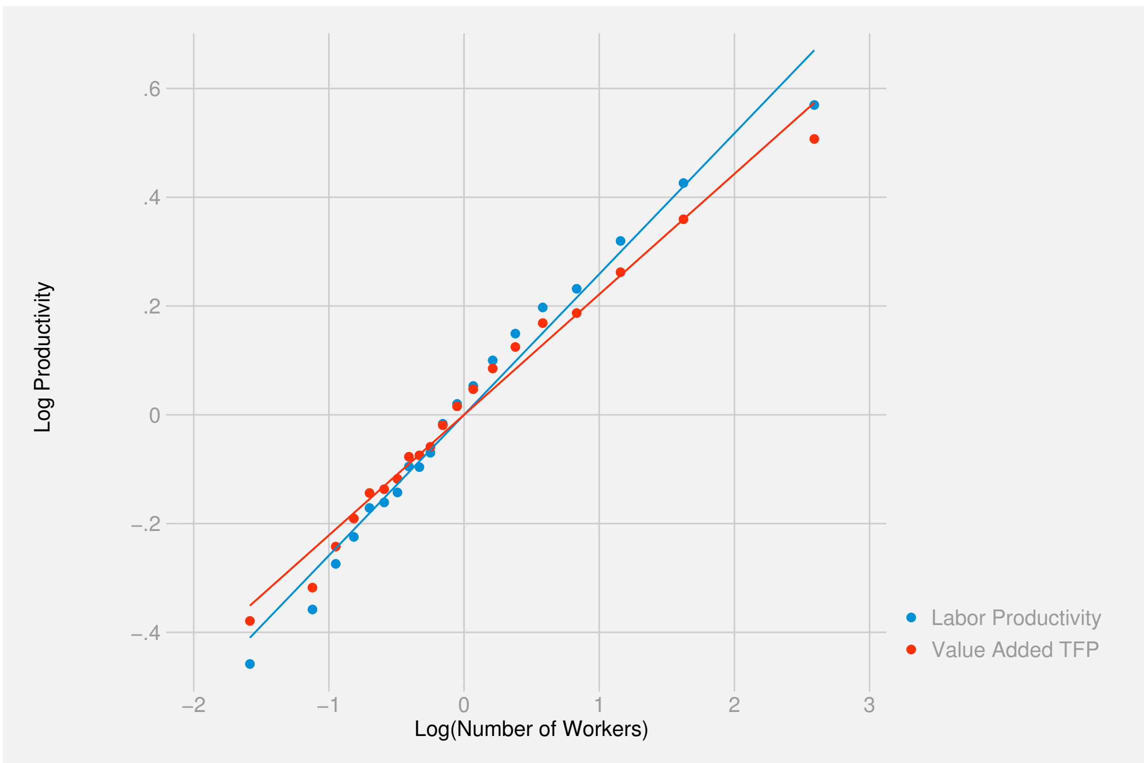
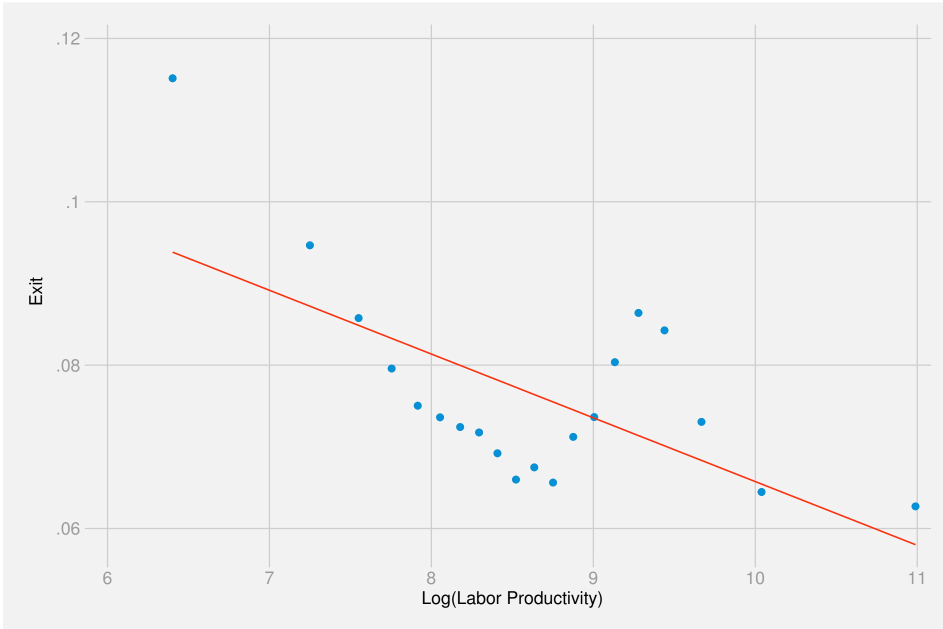
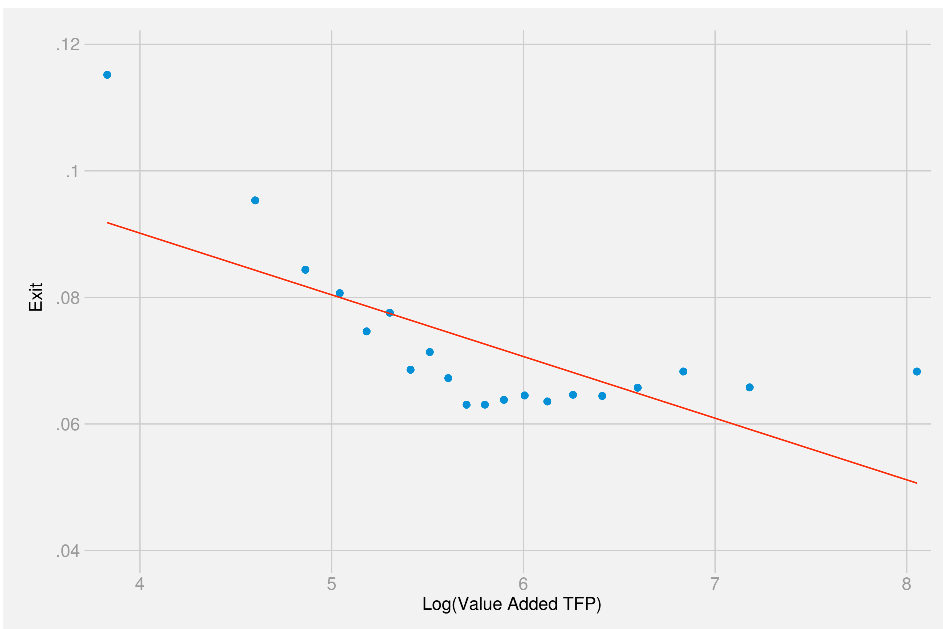


Figure 1: Productivity vs. Number of Workers for Indonesia (Controlling for Industry-Year Fixed Effects)

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.

measures of productivity and under different specifications. My main specification is

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

where $\log(\text{Productivity}_{it})$ 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 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).

Finally, the fifth column shows β_{exit} based on $\log z$ using a CES parameter $\theta = 3$. Here, I define $\log z$ as $\frac{\theta}{\theta-1}\log y - \log l$, where y is value added and l is number of workers. This estimate is meaningfully different from the others: it is roughly three times the magnitude of the first column, which corresponded to $\theta \to \infty$. Using lower values of θ corresponds to an increase in the magnitude of β_{exit}; this is because $\log z = (\log y - \log l) + \frac{1}{\theta-1}\log y$, and so the regression with finite θ will yield the regression of log value added per worker on exit, plus $\frac{1}{\theta-1}$ times the regression of log value added on exit. Since large firms are less likely to exit, this will yield larger (more negative)

	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$	$\log(\text{VA TFP})$	$\log(\text{Rev TFP})$	$\log z$, $\theta = 3$
Exit	-0.123	-0.161	-0.139	-0.185	-0.378
	(0.006)	(0.006)	(0.007)	(0.007)	(0.009)
Observations	604151	606569	325516	325910	604151
Industry-Year FE	Yes	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 $\frac{\text{Revenue}}{l^{\frac{2}{3}} k^{\frac{1}{3}}}$. I define $\log z$ as $\frac{\theta}{\theta-1} \log y - \log l$, where y is value added and l is number of workers.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.

estimates of β_{exit}.[5] This will be the main implication of the choice of the CES parameter for the size of the selection channel: lower values of θ will result in larger measurements for β_{exit}.

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, at least for large values of θ. 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 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.

[5] Larger values of θ will yield smaller magnitudes of β_{exit}. To the extent that $\theta = 3$ is a lower bound value for the CES parameter, as argued in Hsieh and Klenow 2009, this estimate of β_{exit} will be an upper bound on the true magnitude of β_{exit}.

	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$
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.

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, and in the sixth column I measure $\log z$ for $\theta = 3$. 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, and in the sixth column I measure $\log z$ for $\theta = 3$. 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, and in the sixth column I measure $\log z$ for $\theta = 3$. In each case except the last, the estimate does not change substantially (the last estimate is substantially larger than the previous ones). Moreover, for large θ, 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.[6] These estimates also differ slightly from my estimates because they

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

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.[7] 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.[8] 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 Calibration

I now turn to calibration of the model. I base the main calibration on the Indonesian data, since the long panel allows me to estimate the age profiles of the exit rate and of the selectivity of exit. I then provide alternate calibrations for Chile, Colombia, and Spain, re-estimating key parameters but not age profiles. 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 $\beta_{\text{exit},z,a}$ to refer to the difference in mean log productivity, by age, between exiters and survivors, and the notation $\beta_{\text{exit},\tau,a}$ to similarly refer to the difference in the wedge intercept, $\log c_{\tau,i}$. I use the notation ρ_a to refer to exit rates by age. As long as the other

[7] 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.

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

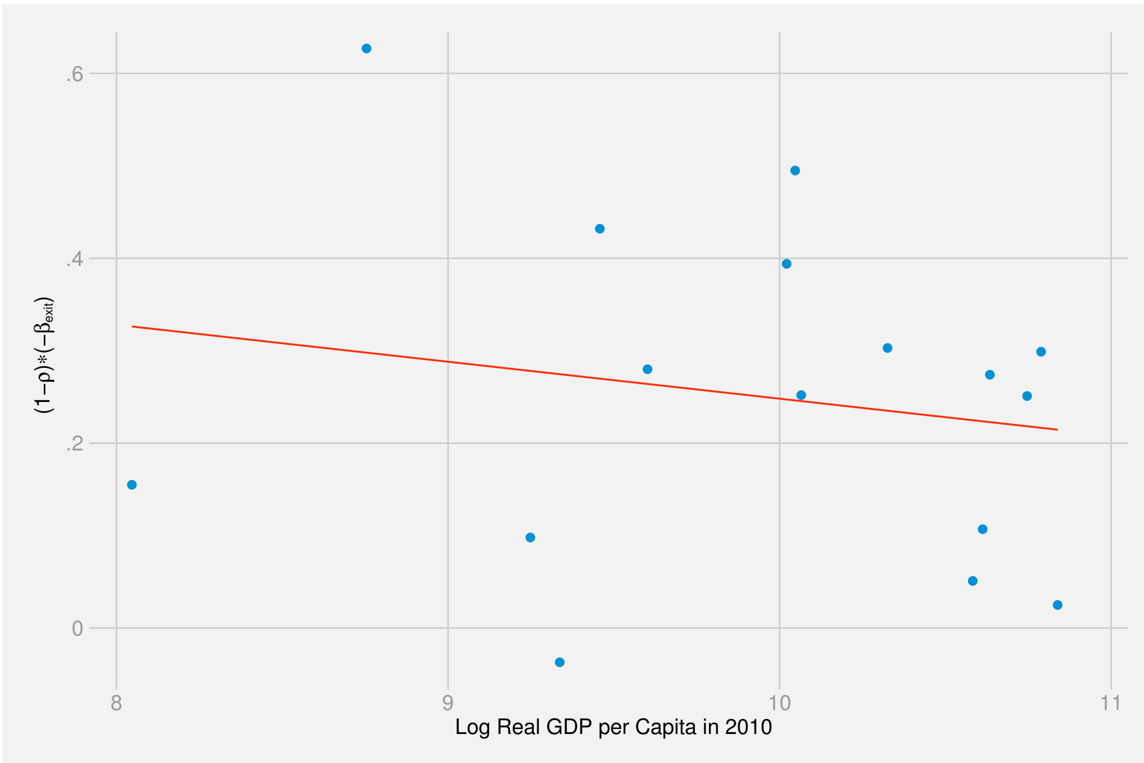


Figure 3: β_{exit} vs. GDP per Capita

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 (Calibration 1)	Estimate (Calibration 2)
θ	Assumed	∞	3
γ	Entrant Dist./Assumed	0.906	0
$\text{var}(\log z_i \mid a = 0)$	Entrant Distribution	1.014	1.618
$\text{var}(\log c_{\tau,i} \mid a = 0)$	Entrant Distribution	0.685	1.014
$\text{cov}(\log z_i, \log c_{\tau,i} \mid a = 0)$	Entrant Dist./Assumed	0	-1.595
ρ_a	Exit Rates	See Table 5	See Table 5
$\beta_{\text{exit},z,a}$	Exiter vs. Survivor Gap	See Table 5	See Table 5
$\beta_{\text{exit},\tau,a}$	Exiter vs. Survivor Gap	See Table 5	See Table 5

Table 4: Main Calibration

Notes: This table shows the main calibration of the model for Indonesia. The CES parameter, θ, is set by assumption, depending on the calibration. The remaining parameters are estimated to match a particular target in the data. The parameter matrix Σ_0, representing the distribution of log productivity among entrants, is set to match the distribution of log labor productivity for entrants, demeaned with industry-year fixed effects. The parameter γ, representing the inverse elasticity of slope of wedges with respect to output, is set in Calibration 1 to generate a zero covariance term in Σ_0, while in Calibration 2 it is set to zero by assumption. The age profile of $\beta_{\text{exit},z,a}$ is calibrated to match the productivity gap between exiters and survivors, by cohort. The age profile of $\beta_{\text{exit},\tau,a}$ is calibrated to match the gap in wedge intercepts 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 $\beta_{\text{exit},z,a}$, $\beta_{\text{exit},\tau,a}$, 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.

parameters are known, knowing $(\beta_{\text{exit},z,a}, \beta_{\text{exit},\tau,a}, \rho_a)$ will pin down the parameters $\delta_{z,a}, \delta_{\tau,a}, c_{\text{exit},a}$.

4.1 Distribution of Productivity and Wedge Intercepts for Entrants

To calibrate the distribution of $\log z_i$ and $\log c_{\tau,i}$ for entrants, I must first generate these variables in the data. To do this, I note that with some manipulations, I can obtain the following formula for $\log z_i$ as a function only of data on $p_i y_i$ (which I measure as value added in nominal terms) and l_i (which I measure as number of workers):

$$\log z_i = \frac{\theta}{\theta - 1} \log (p_i y_i) - \log l_i - \frac{1}{\theta - 1} \log Y$$

This formula allows me to identify $\log z_i$ up to a constant, given a particular value of θ. Since the mean of log productivity (and of the log wedge intercept) has no effect on my results anyway, I normalize both to have mean zero.

I also can compute $\log c_{\tau,i}$, up to a constant, as a function of observable data on $p_i y_i$ and l_i.

We have:

$$\begin{aligned}
\log c_{\tau,i} &= \frac{1}{\mathcal{E}}\log y_i - \log z_i - \left[\frac{1}{\theta}\log Y + \log\left(\frac{1-\gamma-\frac{1}{\theta}}{1-\gamma}\right) - \log w\right] \\
&= \frac{1}{\mathcal{E}}\log y_i - \log y_i + \log l_i + C \\
&= \left(\frac{1}{\mathcal{E}} - 1\right)\left(\frac{\theta}{\theta-1}\log\left(p_i y_i\right)\right) + \log l_i + C
\end{aligned}$$

where C is constant across firms (and thus not important), and again we have $\mathcal{E} := \frac{\theta}{\gamma\theta+1}$ as the elasticity of output with respect to productivity.

Note that for $\gamma = 0$, our formula for $\log c_{\tau,i}$ simplifies to $\log c_{\tau,i} = -\log\left(p_i y_i\right) + \log l_i + C$. This mirrors the result of Hsieh and Klenow (2009): in a CES model with $\gamma = 0$ and constant returns to scale, measured differences in revenue per worker reflect differences in wedges, rather than differences in underlying productivity.

To compute productivity, I need to take a stand on the CES parameter θ, and to compute the wedge intercept, I need to take a stand on both θ and the wedge slope parameter γ. I estimate results under two calibrations. In the first calibration, I let $\theta \to \infty$, and I set γ to make the covariance of $\log z_i$ and $\log c_{\tau,i}$ equal to zero for entrants. Under this calibration, productivity is simply value-added per worker. In the second calibration, I set $\theta = 3$ and $\gamma = 0$. This is the calibration of Hsieh and Klenow (2009), and these paramaters are common in the literature on misallocation.

To compute the variance-covariance martrix for $\log z_i$ and $\log c_{\tau,i}$ for entrants, I demean the log labor productivity and log wedge intercept of entrants using industry-year fixed effects, and then compute the variance-covariance matrix of the demeaned data for entrants. The mean is mechanically zero for both calibrations. For the $\theta \to \infty, \text{cov}\left(\log z_i, \log c_{\tau,i}\right) = 0$ calibration, the variance of $\log z_i$ is 1.014, the variance of $\log c_{\tau,i}$ is 0.685, and $\gamma = 0.906$. For the $\theta = 3, \gamma = 0$ calibration, the variance of $\log z_i$ is 1.618, the variance of $\log c_{\tau,i}$ is 1.014, and their covariance is -1.595.

4.2 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, $\alpha \cdot \exp\left(\beta a\right)$. 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:

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

measuring firm age as before and excluding firms that are present at the start of the data in

Selective Exit

	Calibration 1	Calibration 2
α_z	-0.165	-0.381
	(0.02)	(0.03)
β_z	-0.067	-0.015
	(0.02)	(0.008)
α_τ	-0.259	0.165
	(0.008)	(0.02)
β_τ	0.029	-0.067
	(0.003)	(0.02)

Gompertz-Makeham

α	0.064
	(0.002)
β	-0.130
	(0.008)
λ	0.047
	(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 described in the text. 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 $\rho_a = \alpha \cdot \exp(\beta a) + \lambda$, 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.

1975. I report the results in Table 5. I also plot the fitted estimates compared to the raw data in Figure 4. 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.[9]

4.3 Selective Exit

Next, I calibrate $\beta_{\text{exit},z,a}$, 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, under Calibration 1, 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 $\beta_{\text{exit},z,a}$, 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

[9] 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.

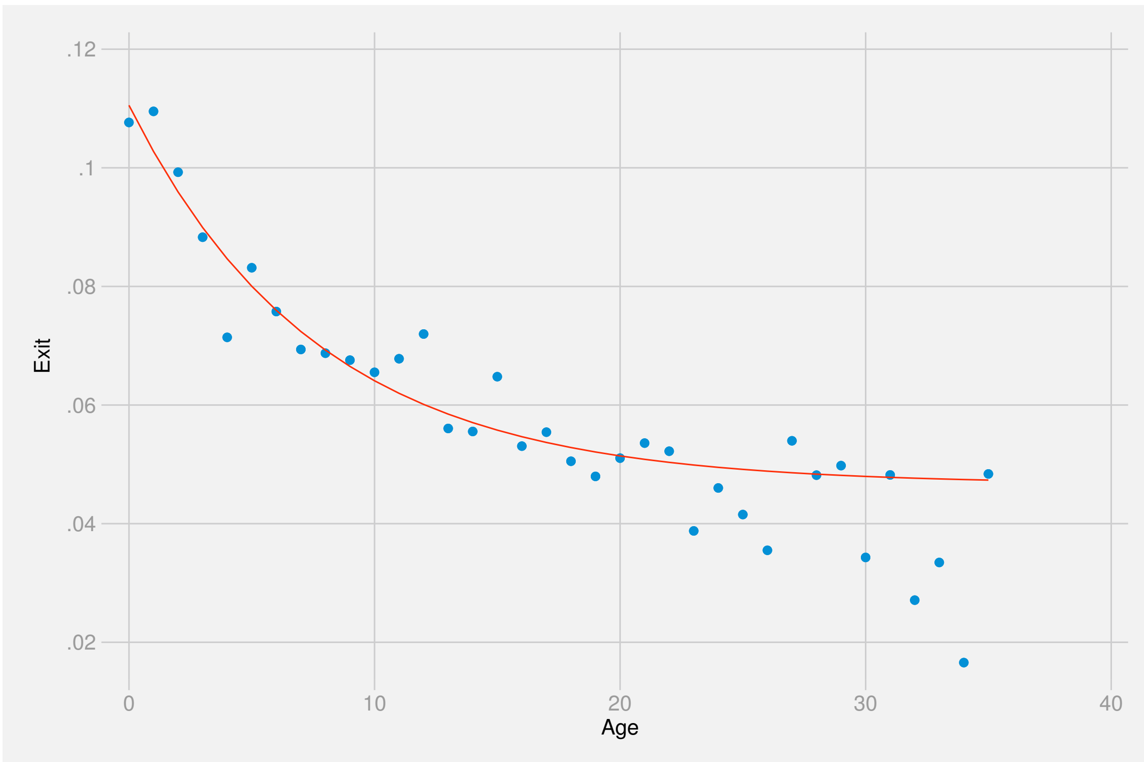


Figure 4: Gompertz-Makeham Fitted Estimates of Exit Rates by Age

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.

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 do this for both calibrations, and I show the results for Calibration 1 in Figure 5a, 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 a model to the estimated coefficients. I let $\beta_{\text{exit},z,a}$ match the data exactly for $a = 0$ and $a = 1$, and then I fit the following model to the estimated coefficients for $a \geq 2$:

$$\beta_{\text{exit},z,a} = \alpha_z \cdot \exp\left(\beta_z a\right)$$

I estimate the model using weighted non-linear 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 estimates to obtain a fitted value of $\beta_{\text{exit},a}$ for $a \geq 2$. The estimates imply that selection is always in the same direction (less productive firms are more likely to exit), and also that future rounds of selection weaken as firms age. In Figure 5b, 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.

I also perform the same analysis for $\beta_{\text{exit},\tau,a}$. I follow the analogous procedure to get the estimated coefficients, as well as the fitted estimates. I model the fitted estimates as $\beta_{\text{exit},\tau,a} = \alpha_\tau \cdot \exp\left(\beta_\tau a\right)$. The results are in Table 5. The estimates for Calibration 1 imply that selection diverges as the firm gets very old: this is not possible, and is likely a result of over-extrapolation. When estimating the model under Calibration 1, I thus stop the growth of $\beta_{\text{exit},\tau,a}$ at $a = 20$, and then set it to zero for $a > 70$.[10]

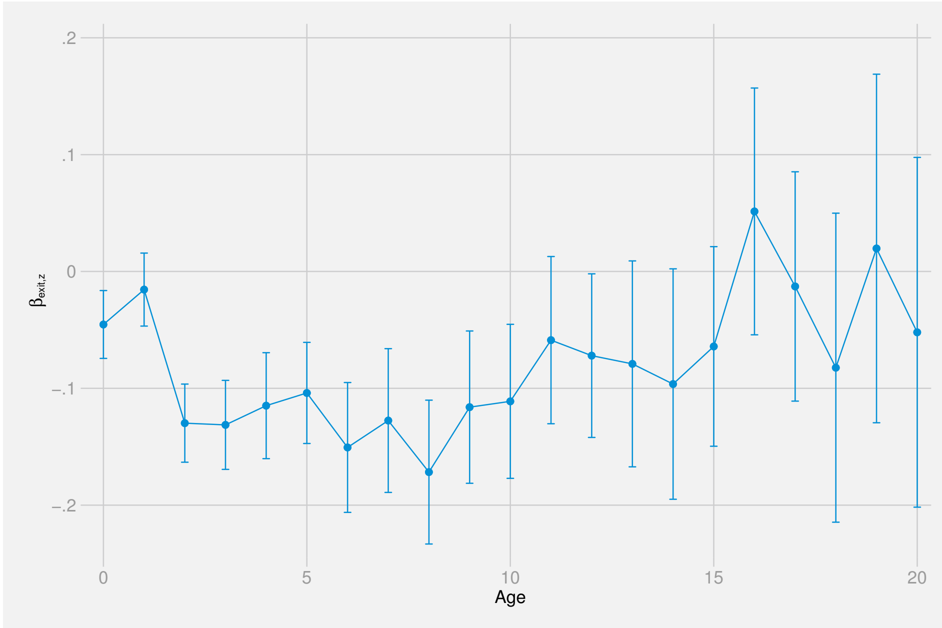
4.4 Recalibration for Chile, Colombia, and Spain

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.[11]

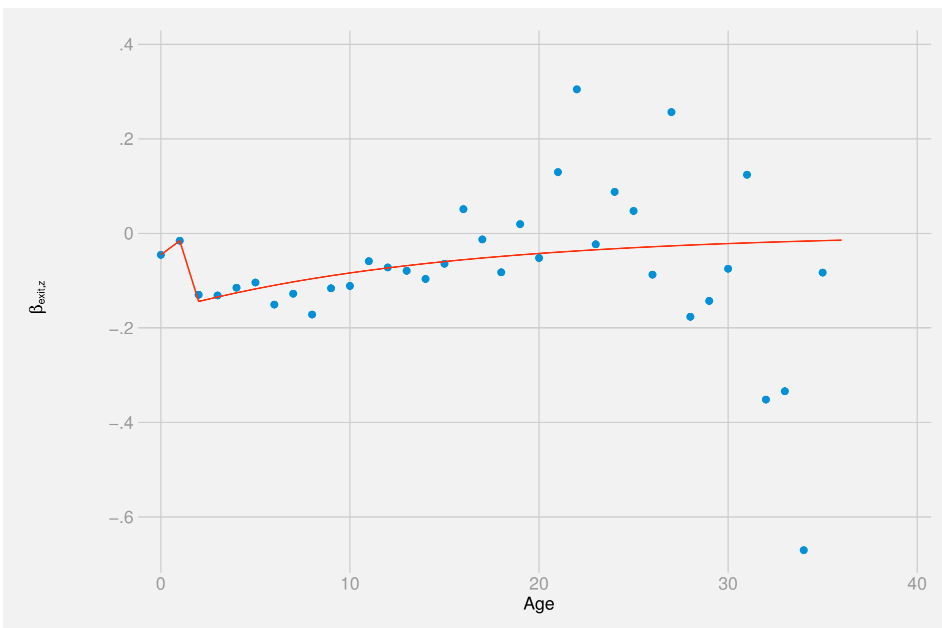
To recalibrate γ and Σ_0, I follow the same procedure for Chile and Colombia as I do for Indonesia. However, because I do not have industry codes in the Spanish data, I absorb year fixed effects but do not absorb industry-year fixed effects. I summarize the results for γ, under Calibration 1, in Table 6. Note that, for Spain, the γ called for in Calibration 1 is actually greater

[10]This is necessary to solve the model: otherwise, the distribution of $c_{\tau,i}$ is eventually too narrow to deliver the desired combination of exit rate ρ_a and selection on the wedge intercept, $\beta_{\text{exit},\tau,a}$.

[11]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.



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



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

Figure 5: Estimates of $\beta_{\text{exit},a}$

Notes: This figure shows estimates of $\beta_{\text{exit},z,a}$, the age profile of the gap in log productivity between exiters and survivors, in the Indonesian data. I estimate $\beta_{\text{exit},z,a}$ 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 then fit the raw estimates for $a \geq 2$ to the model $\beta_{\text{exit},z,a} = \alpha_z \cdot \exp(\beta_z a)$, estimating with weighted least squares and using the inverse variance of the estimate as the weight. I allow the fitted estimates to match the data exactly for $a = 0$ and $a = 1$. The results are depicted in the second panel, which shows the raw estimates as blue dots, and the fitted model as a red line.

	γ	$(\theta \to \infty)$ $\beta_{\text{exit},z}$	$\beta_{\text{exit},\tau}$	$(\theta = 3, \gamma = 0)$ $\beta_{\text{exit},z}$	$\beta_{\text{exit},\tau}$	ρ
Indonesia	0.906	-0.123	-0.338	-0.378	0.123	7.6%
Chile	0.823	-0.502	-0.289	-0.982	0.502	6.3%
Colombia	0.848	-0.342	-0.487	-0.830	0.342	10.9%
Spain	1.010	-0.390	-0.325	-0.744	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.

than 1: this implies that more productive firms in the Spanish data actually employ less labor than less productive firms, at least among entrants.

In addition to reporting γ, which I use in the recalibration, I also report $\beta_{\text{exit,z}}$, $\beta_{\text{exit},\tau}$, and the exit rate, ρ. Note, given the discussion earlier in this section, that $\beta_{\text{exit},\tau}$ under Calibration 2 is the same as $-\beta_{\text{exit},z}$ under Calibration 1. Knowing $\beta_{\text{exit,z}}$ 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 focus on the mapping between $\beta_{\text{exit},z}$ 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 fully calibrated model (for each calibration). I then recalibrate the model using estimated parameters for Chile, Colombia, and Spain, allowing me to compare the strength of the selection channel in each of these countries. I find that the back-of-the-envelope approximation is fairly accurate, and the accuracy of the back-of-the-envelope does not seem to vary across country calibrations. Given this, I conclude with a broader analysis 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. The calibration section offers a profile of exit rates and selective exit gaps, $(\rho_a, \beta_{\text{exit},z,a}, \beta_{\text{exit},\tau,a})$. I use these estimates, along with the baseline distribution of productivity at age a, to iteratively back out $(c_{\text{exit},a}, \delta_{z,a}, \delta_{\tau,a})$. Sometimes, the values of $\beta_{\text{exit},z,a}$

or $\beta_{\text{exit},\tau,a}$ are too large to be easily generated by the underlying distribution of z or τ: in these cases, I "winsorize" the value of $\beta_{\text{exit},z,a}$ or $\beta_{\text{exit},\tau,a}$ to a feasible value.[12]

I re-solve the model for different levels of selection, by multiplying the profile of $\beta_{\text{exit},z,a}$ by a scalar. I also solve a version of the model where $\beta_{\text{exit},\tau,a}$ is equal to zero for all ages. For each run of the model, I compute the implied $\beta_{\text{exit,z}}$, i.e. what a regression of log productivity on exit would yield if run on data generated by the model. I also compute aggregate productivity for each model run. Since one of the model runs I solve sets both $\beta_{\text{exit},z,a}$ and $\beta_{\text{exit},\tau,a}$ equal to zero, I know aggregate productivity under the no-selection counterfactual, and thus can compute the strength of the selection channel for each calibration of the model. Thus, I am able to trace out the implied relationship between $\beta_{\text{exit},z}$, 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 $\beta_{\text{exit},z}$ 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 well. As a result, the empirical finding that $\beta_{\text{exit},z}$ 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. $\beta_{\text{exit,z}}$

I begin, in Figure 6, by comparing the selection channel to $\beta_{\text{exit,z}}$. To keep everything in a log scale, I use $\log(Z)$ as the vertical axis; this results in roughly linear relationships. I normalize aggregate productivity so that $Z = 1$ for the "no-selection" calibration where $\beta_{\text{exit},z,a} = \beta_{\text{exit},\tau,a} = 0$. This ensures that the level of $\log Z$ also corresponds to the strength of the selection channel. I also include vertical dashed lines, which mark the preferred estimates of $\beta_{\text{exit,z}}$ for Indonesia and Spain, and facilitate cross-country comparisons for each model. I provide results for Calibration 1 and Calibration 2: the vertical dashed lines correspond to the estimates of $\beta_{\text{exit,z}}$ under the figure's calibration of θ.

I plot three versions of the relationship between β_{exit} and the selection channel. The diagonal dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit}} = $ Selection Channel. The thin orange line is the model with no selection on wedges, $\beta_{\text{exit},\tau,a} = 0$. This version of the model

[12] Specifically, I compute the mean, 10th percentile, and 90th percentile of the distribution of $\log z$ for cohort a. If $\beta_{\text{exit},z,a}$ is less than the 10th percentile minus the mean, then I replace it with the 10th percentile minus the mean. If it is greater than the 90th percentile minus the mean, then I replace it with the 90th percentile minus the mean. Since the exit rate is typically below 10%, this ensures that it is feasible to achieve $\beta_{\text{exit},z,a}$. Eventually, the underlying distribution may get so tight that it cannot easily accomodate any selection. If the 10th percentile is greater than the mean, or if the 90th percentile is less than the mean, then I simply set $\beta_{\text{exit},z,a} = 0$. I follow the same procedure for τ.

shows that selection on wedges is of little importance compared to selection on productivity, justifying the focus on $\beta_{\text{exit},z}$. The thick teal line is the main model. The main model is very close to the back-of-the-envelope in Calibration 1, and somewhat below the back-of-the-envelope in Calibration 2 (in fact, being below the back-of-the-envelope for Calibration 2 roughly cancels out the fact that $\beta_{\text{exit},z}$ tends to be larger for Calibration 2).

Broadly, 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 relatively little loss, be reduced to an easily measured statistic, even in a model with productivity dispersion, misallocation, and changing exit and selection dynamics over the lifecycle.

5.2 Cross-Country Comparison

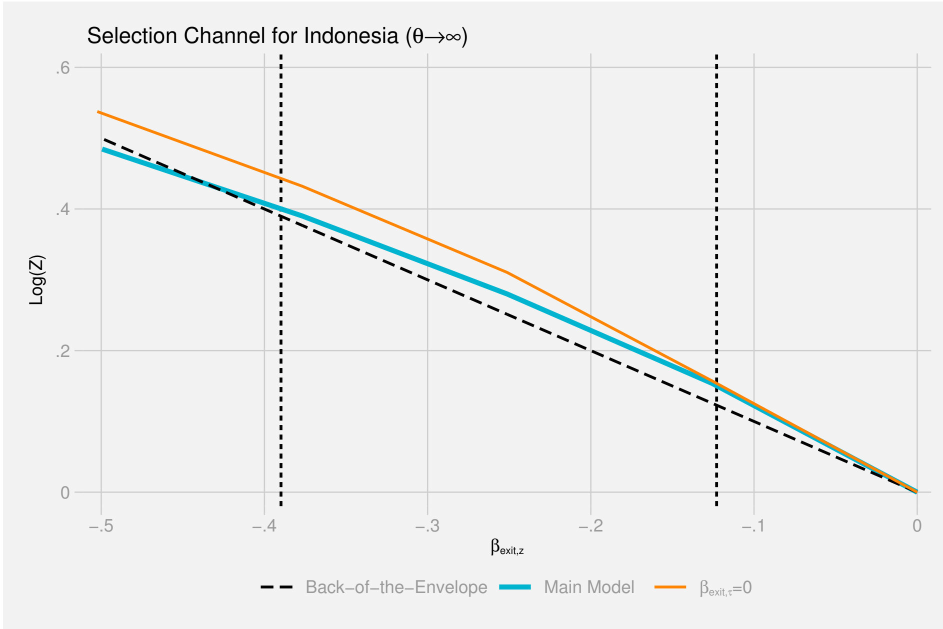
The strength of the selection channel in a given country will depend on the estimated $\beta_{\text{exit},z}$ in that country, and on the mapping from $\beta_{\text{exit},z}$ to the selection channel, given the other relevant parameters for that country. To examine how the selection channel varies across countries, I will show how the model's mapping between $\beta_{\text{exit},z}$ and the selection channel changes when we recalibrate γ and Σ_0 to match the Spanish data, the Chilean data, and the Colombian data.[13]

I show the results of the recalibration in Figure 7: the thick teal line is the model calibrated to Indonesian data, while the orange lines in different shades show the recalibrations for other countries. There is little difference between the lines; in fact, in Calibration 2, the lines lie essentially 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.

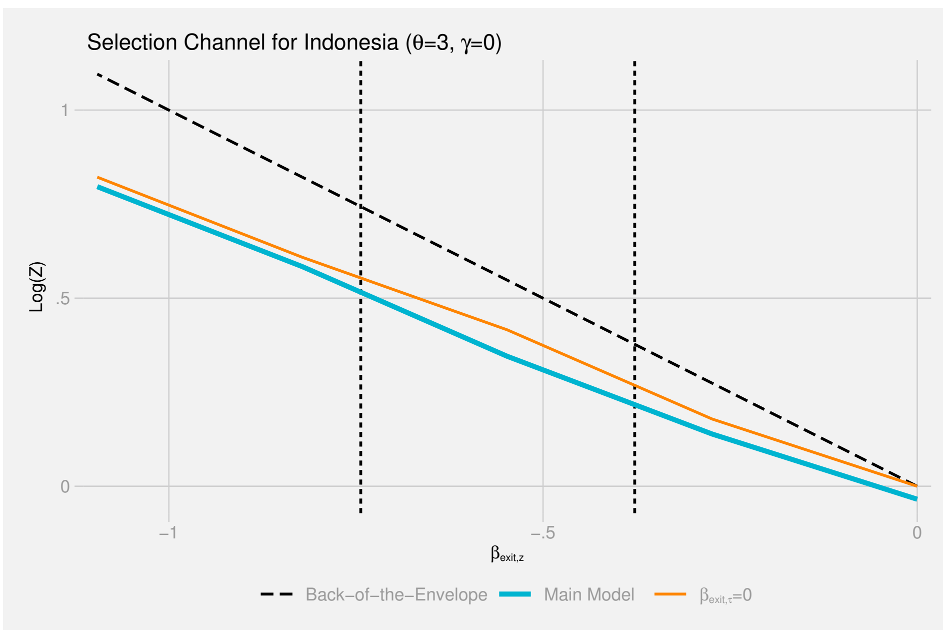
Indonesia has a lower $\beta_{\text{exit},z}$ than Spain. If Indonesia raised its level of selection to that of Spain, based on Calibration 1, its aggregate productivity would increase by roughly 30%, using either the calibrated model or the back-of-the-envelope approximation. The selection channel plays 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 (in fact, selection seems to be stronger in Chile than Spain). 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 in β_{exit} are also conclusions about cross-country differences in the selection channel. The selection channel is an important component of

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



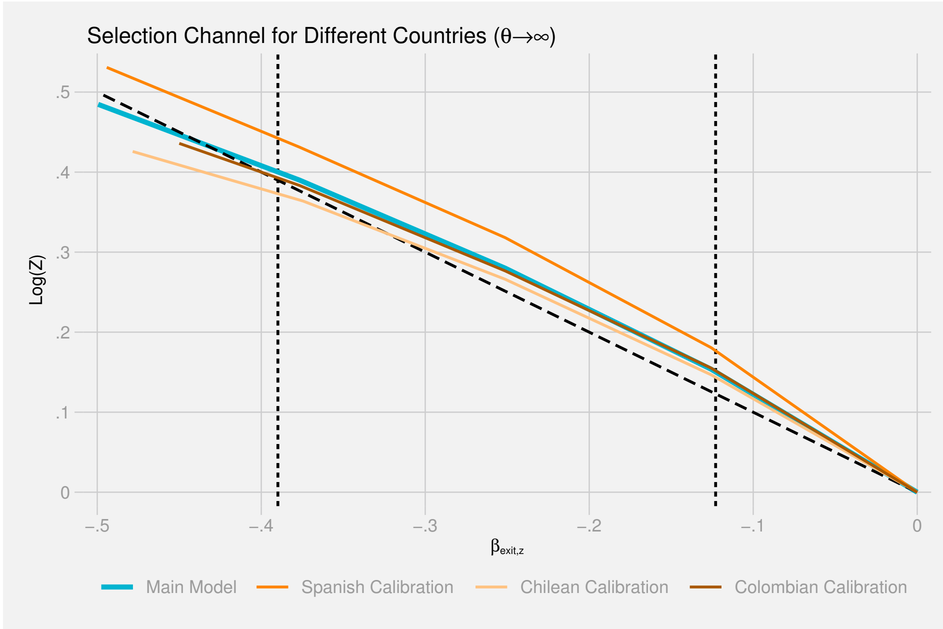
(a) Calibration 1



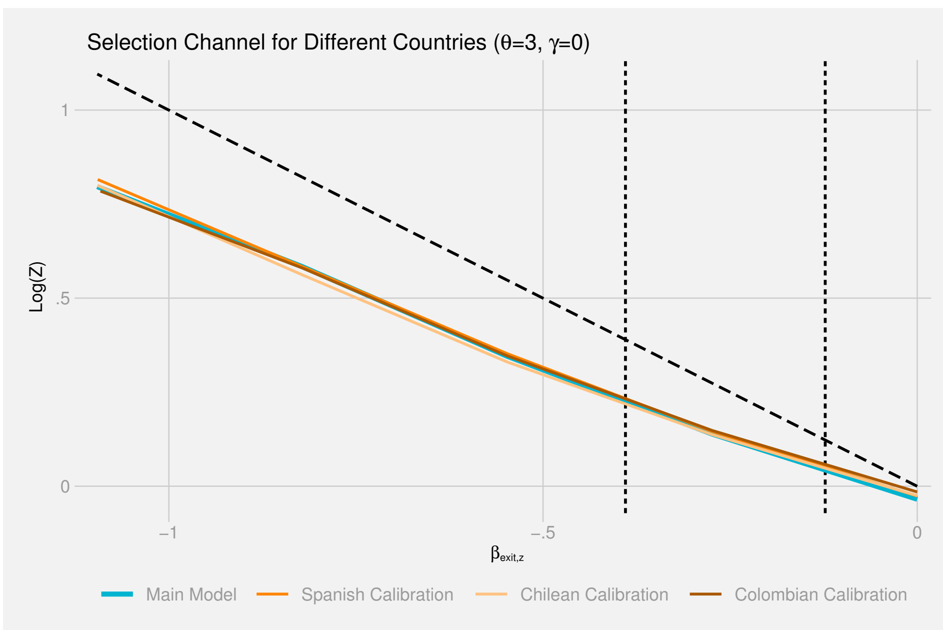
(b) Calibration 2

Figure 6: Selection Channel vs. β_{exit}

Notes: This figure shows the relationship between aggregate productivity and $\beta_{\text{exit},z}$. Aggregate productivity is normalized so that $Z = 1$ under the no-selection benchmark, so $\log(Z)$ also equals the selection channel. The first panel shows results for Calibration 1, and the second panel shows results for Calibration 2. The diagonal dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit},z} = \text{Selection Channel}$. The thin orange line is the model with no selection on wedges, $\beta_{\text{exit},\tau,a} = 0$. The thick teal line is the main model, with $\beta_{\text{exit},\tau,a}$ determined by the data. Each point on the orange and teal lines represents a solution to the model, with $\beta_{\text{exit},z,a}$, scaled proportionally up or down. For each run of the model, I compute aggregate productivity and also compute what the estimated $\beta_{\text{exit},z}$ would be in data generated from the solved model. I also include vertical dashed lines, which mark the preferred estimates of $\beta_{\text{exit},z}$ for Indonesia and Spain under the panel's calibration of θ, to facilitate cross-country comparisons for each model.



(a) Calibration 1



(b) Calibration 2

Figure 7: Selection Channel vs. β_{exit} (Indonesia vs. Spain, Chile, and Colombia)

Notes: This figure shows the relationship between aggregate productivity and $\beta_{\text{exit},z}$, with Σ_0 and γ calibrated to different countries. Aggregate productivity is normalized so that $Z = 1$ under the no-selection benchmark for each country, so $\log(Z)$ also equals the selection channel. The first panel shows results for Calibration 1, and the second panel shows results for Calibration 2. The diagonal dashed line is the back-of-the-envelope approximation, in which $-\beta_{\text{exit},z} =$ Selection Channel. The thick teal line is the main model, calibrated to the Indonesian data. The thin orange lines in different shades represent recalibrations to different countries: Spain, Chile, and Colombia. Each point on the orange and teal lines represents a solution to the model, with $\beta_{\text{exit},z,a}$, scaled proportionally up or down. For each run of the model, I compute aggregate productivity and also compute what the estimated $\beta_{\text{exit},z}$ would be in data generated from the solved model. I also include vertical dashed lines, which mark the preferred estimates of $\beta_{\text{exit},z}$ for Indonesia and Spain under the panel's calibration of θ, to facilitate cross-country comparisons for each model.

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:

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

where the coefficient λ 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:

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

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:

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

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

What λ should we use when implementing this? A natural choice is $\lambda = 1$: 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 that $\Delta \log Y = \Delta \log A + \frac{2}{3}\Delta \log L + \frac{1}{3}\Delta \log K$: 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

	$\lambda = 1$	$\lambda = 1.5$
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 $R^2 = \frac{\lambda^2 \cdot \text{Var}\left(\beta_{\text{exit},i}\right) + 2\lambda \cdot \text{Cov}\left(\beta_{\text{exit},i}, \varepsilon_i\right)}{\text{Var}(\log(\text{GDP per capita}_i))}$, and I also compute an alternative formula, $R^2_{\text{Alternative}} = \frac{\lambda^2 \cdot \text{Var}\left(\beta_{\text{exit},i}\right) + \lambda \cdot \text{Cov}\left(\beta_{\text{exit},i}, \varepsilon_i\right)}{\text{Var}(\log(\text{GDP per capita}_i))}$. The formula depends on λ, which governs the relationship between $-\beta_{\text{exit}}$ and the selection channel. The first column shows results for $\lambda = 1$, which corresponds to the direct effect of the selection channel on aggregate productivity, and the second column shows results for $\lambda = 1.5$, 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.

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 $\lambda = 1.5$.

I estimate R^2 and $R^2_{\text{Alternative}}$ for both $\lambda = 1$ and $\lambda = 1.5$, using the data from Figure A.10. I use the non-parametric bootstrap to construct bias-corrected confidence intervals.[14] 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 $\lambda = 1$) to 11.1% (R^2 for $\lambda = 1.5$).

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

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

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, 7, 8, 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

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. 2.2

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, 2.4, 1, 5, 4.1

_ 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

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

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

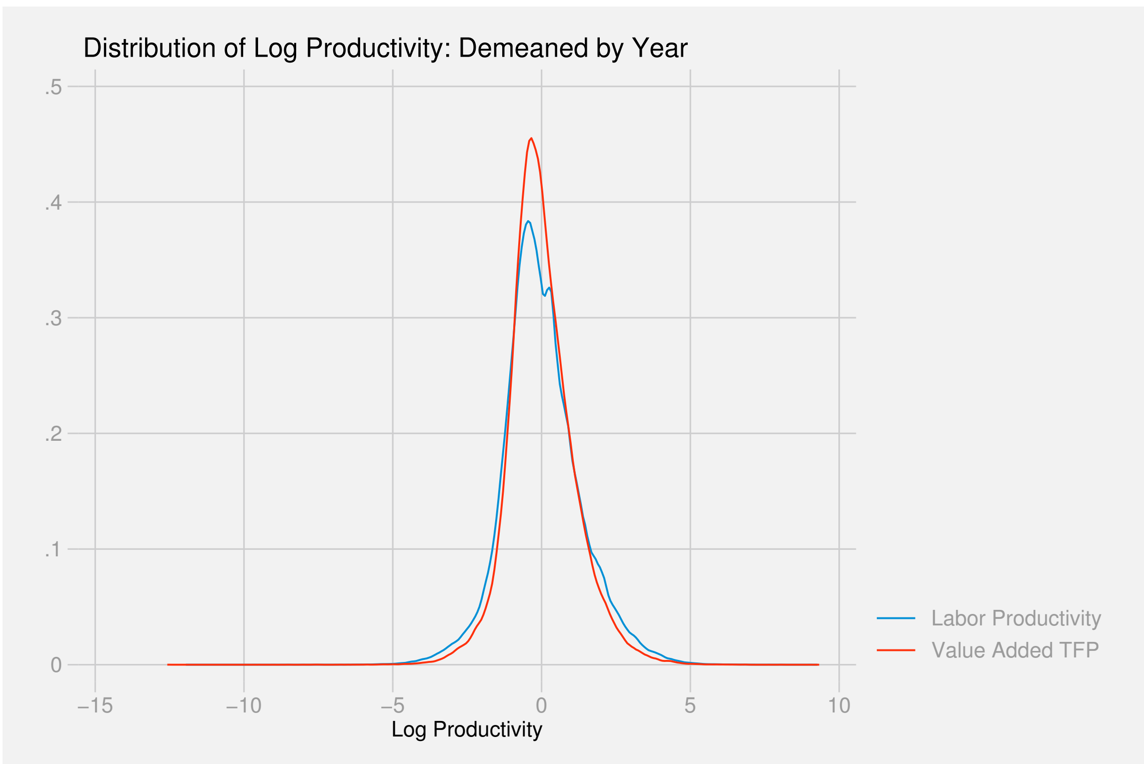


Figure A.1: Distribution of Productivity for Indonesia (Demeaned by Year)
Notes: This figure shows kernel density plots of the distribution of log productivity, using the Indonesian data. The data are residualized on industry-year fixed effects before plotting the density. I define labor productivity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

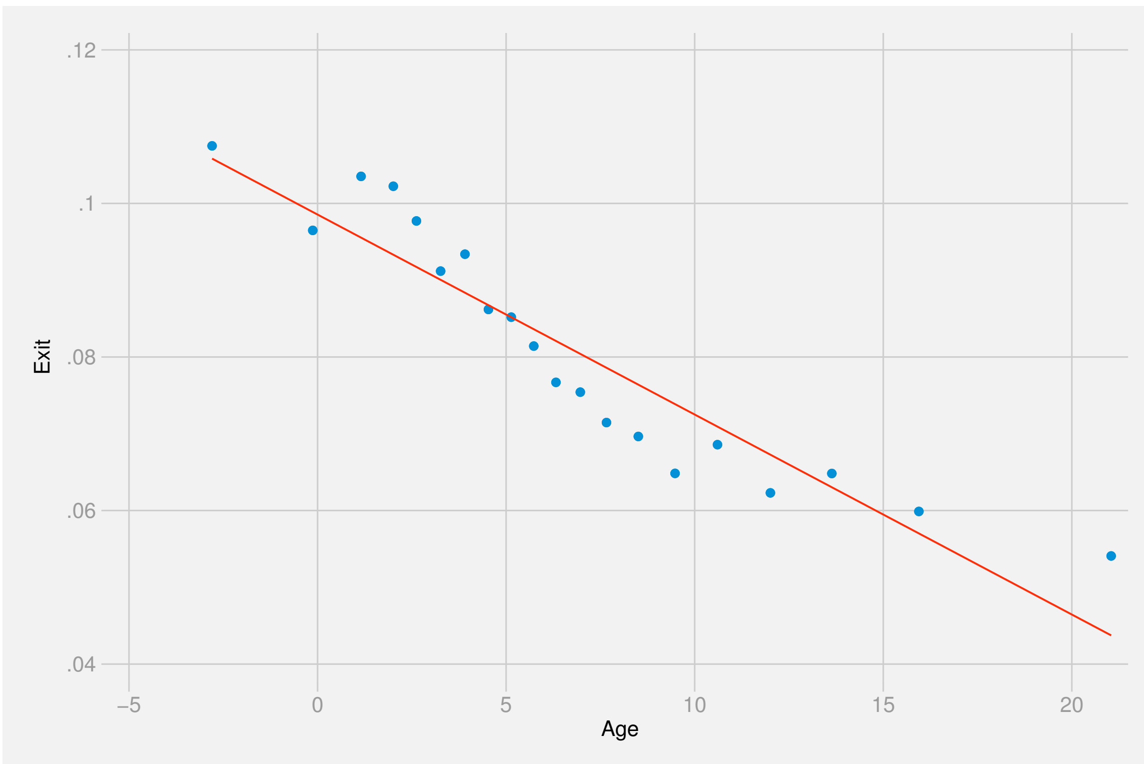


Figure A.2: Exit vs. Age for Indonesia (Controlling for Industry-Year Fixed Effects)

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.

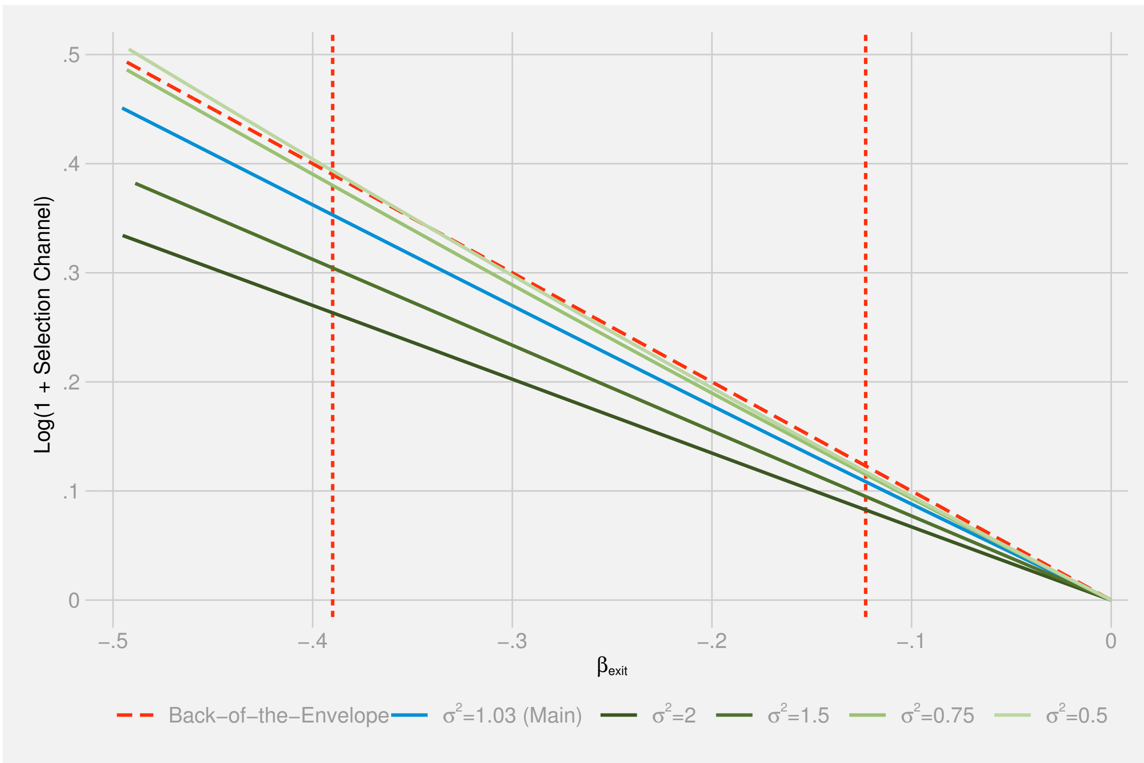


Figure A.3: Selection Channel vs. β_{exit} (Varying σ^2)

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	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log$ (Value Added TFP)	$\log$ (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	$\log$ (Workers)	Exit	$\log$ (Workers)
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.00781	0.221		
	(0.000345)	(0.00361)		
$\log$ (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	$\log$ (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
$\log(\text{Workers})$	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	$\log(\text{Workers})$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$	$\log z$, $\theta =$ 3
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.0372 (0.00122)	0.313 (0.0108)				
Exit			-0.502 (0.016)	-0.406 (0.023)	-0.418 (0.014)	-0.982 (0.026)
Observations	76071	81326	76071	50884	75115	76071
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
$\geq$ 20 Workers	No	No	No	Yes	No	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, fifth, and sixth 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. The sixth column uses $\log z$, where I define $\log z$ as $\frac{\theta}{\theta-1}\log y - \log l$, where y is value added and l is number of workers.

	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
$\log(\text{Workers})$	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	$\log(\text{Workers})$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$	$\log z$, $\theta = 3$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.0534 (0.00161)	0.373 (0.0138)				
Exit			-0.342 (0.010)	-0.330 (0.016)	-0.358 (0.010)	-0.830 (0.017)
Observations	95521	102817	95521	58539	95533	95521
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
≥ 20 Workers	No	No	No	Yes	No	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, fifth, and sixth columns 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. The sixth column uses $\log z$, where I define $\log z$ as $\frac{\theta}{\theta-1} \log y - \log l$, where y is value added and l is number of workers.

	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 (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 (Workers)	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	$\log\left(\frac{\text{Revenue}}{\text{Worker}}\right)$	$\log z,\ \theta = 3$
$\log\left(\frac{\text{Value Added}}{\text{Worker}}\right)$	-0.0281	0.0786				
	(0.000169)	(0.00154)				
Exit			-0.390	-0.342	-0.400	-0.744
			(0.002)	(0.007)	(0.002)	(0.004)
Observations	3741407	3741407	3741407	406019	4378371	3741407
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
$\geq$ 20 Workers	No	No	No	Yes	No	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, fifth, and sixth columns 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. The sixth column uses $\log z$, where I define $\log z$ as $\frac{\theta}{\theta-1}\log y - \log l$, where y is value added and l is number of workers.

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 Extensions to the Back-of-the-Envelope Approximation

In this appendix, I extend the back-of-the-envelope approximation in a variety of directions. I show how to incorporate population growth, mean-reverting firm-level productivity dynamics, and life-cycle growth in firm productivity. In each case, I will focus on the limit as the exit rate, ρ, grows small. This corresponds to a continuous time limit of the model: the instantaneous exit rate is zero, but the rate at which the exit rate goes to zero is proportional to other quantities, such as the rate of growth in population or productivity, or the rate of mean reversion.

B.1 Baseline Proof

We begin by reviewing the proof of the back-of-the-envelope in the case without growth or mean reversion. This will be useful because it lets us introduce an appropriate notion of the limit as the exit rate grows small, and because it sets the framework for the later modifications to the benchmark version.

We will use the same notation as in the main text. Let μ denote the mean of log productivity, μ_{Exiters} denote the mean of log productivity among exiting firms, and similarly for $\mu_{\text{Survivors}}$ and μ_{Entrants}. We will use μ^* to denote mean log productivity under the no-selection counterfactual, and the approximate selection channel will be $\mu - \mu^*$. Finally, the exit rate will be ρ. To appropriately define the limit as ρ goes to zero, we will define $\rho = 1 - e^{-\tilde{\rho}\Delta}$, which will go to zero as $\Delta \to 0$. This will not be important in our baseline proof, since nothing else will go to zero as Δ shrinks, but it will matter for the extensions.

Our derivation will rely on three equations:

$$\begin{aligned}
\text{Identity: } & \mu = \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} \\
\text{Steady State: } & \mu_{\text{Entrants}} = \mu_{\text{Exiters}} \\
\text{Counterfactual: } & \mu^* = \mu_{\text{Entrants}}
\end{aligned}$$

The first of these is an identity, and will always hold. The second is a steady-state condition: since firm productivities do not change, and the population of firms is constant, each exiter must be replaced by an entrant with the same productivity (our extensions will alter this steady state equation). The final one is an expression for the counterfactual mean of log productivity under no selection. Since firms do not experience any change in their log productivity on average, the counterfactual mean is simply the mean of the entrants: this will be altered in our extension that allows for life-cycle productivity growth.

Our proof then proceeds as in the main text:

$$\begin{aligned}
\text{Selection Channel} &\approx \mu - \mu^* \\
&= \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} - \mu_{\text{Entrants}} \\
&= (1-\rho)\left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right) \\
&\to \left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right)
\end{aligned}$$

B.2 Extension: Population Growth

We now consider an extension with population growth. Suppose that each entering cohort is $(1+g)$ times larger than the previous cohort. So that we can take the limit as $\Delta \to 0$, we will define $1 + g = e^{\tilde{g}\Delta}$.

Adding in population growth will not affect the identity equation, nor will it affect the equation for the counterfactual. However, it will affect the steady-state equation. Since the population is growing each period, we have:

$$\begin{aligned}
(1+g)\, N_t &= N_{t+1} \\
\implies (1+g) &= (1-\rho) + \text{Entry Rate} \\
\implies \text{Entry Rate} &= \rho + g
\end{aligned}$$

This lets us derive the new steady state condition:

$$\begin{aligned}
\rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} &= (\rho+g)\,\mu_{\text{Entrants}} + (1-\rho-g)\,\mu_{\text{Survivors}} \\
\implies \rho\mu_{\text{Exiters}} &= (\rho+g)\,\mu_{\text{Entrants}} - g\mu_{\text{Survivors}} \\
\implies \mu_{\text{Entrants}} &= \mu_{\text{Exiters}} + \frac{g}{\rho}\left(\mu_{\text{Survivors}} - \mu_{\text{Entrants}}\right)
\end{aligned}$$

Intuitively, population growth means that entrants are now replacing a mix of exiters and survivors, and so their average log productivity must be higher than that of exiters.

We can then use this expression to derive our modified back-of-the-envelope:

$$\begin{aligned}
\text{Selection Channel} &\approx \mu - \mu^* \\
&= \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} - \mu_{\text{Entrants}} \\
&= (1-\rho)\left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right) - \frac{g}{\rho}\left(\mu_{\text{Survivors}} - \mu_{\text{Entrants}}\right) \\
&= (1-\rho)\left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right) - \frac{g}{\rho}\left(\mu_{\text{Survivors}} - \mu^*\right) \\
\implies \text{Selection Channel} &\approx \frac{(1-\rho)}{1+\frac{g}{\rho}}\left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right) \\
&\to \frac{1}{1+\frac{g}{\rho}}\left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right)
\end{aligned}$$

Population growth thus weakens the selection channel. To verify that $\frac{g}{\rho}$ converges to a positive constant as $\Delta \to 0$, we can use L'Hospital's Rule:

$$\begin{aligned}
\frac{g}{\rho} &= \frac{e^{\tilde{g}\Delta} - 1}{1 - e^{-\tilde{\rho}\Delta}} \\
&\to \frac{\tilde{g} \cdot e^{\tilde{g}\Delta}}{-\tilde{\rho} \cdot \left(-e^{-\tilde{\rho}\Delta}\right)} \\
&\to \frac{\tilde{g}}{\tilde{\rho}}
\end{aligned}$$

B.3 Extension: Life-cycle Productivity Growth

We now consider an extension with firm-level productivity growth over the life-cycle. We now assume zero population growth, but firm productivity will grow proportionally by $(1+g)$ each period. Once again, we will define $1+g=e^{\tilde{g}\Delta}$, which will ensure that $\frac{g}{\rho}$ converges to a positive constant as $\Delta \to 0$.

The identity equation will remain the same, but both the steady-state equation and the counterfactual equation will change. We begin with the counterfactual equation:

$$\begin{aligned}
\mu^* &= \left[\sum_{t=0}^{\infty} \left(\left(1-\rho\right)\left(1+g\right)\right)^t \cdot \mu_{\text{Entrants}}\right] \cdot \left[\sum_{t=0}^{\infty} \left(1-\rho\right)^t\right]^{-1} \\
&= \mu_{\text{Entrants}} \cdot \frac{1}{1-\left(1-\rho\right)\left(1+g\right)} \cdot \rho \\
&= \mu_{\text{Entrants}} \cdot \frac{\rho}{\rho - g + \rho g} \\
&= \mu_{\text{Entrants}} \cdot \frac{1}{1-\frac{g}{\rho}+g} \\
&\to \mu_{\text{Entrants}} \cdot \frac{1}{1-\frac{g}{\rho}}
\end{aligned}$$

The other key change is a change to the steady-state equation. We now have to account for productivity growth. We have:

$$\begin{aligned}
\rho\mu_{\text{Exiters}} + \left(1-\rho\right)\mu_{\text{Survivors}} &= \rho\mu_{\text{Entrants}} + \left(1-\rho\right)\left(1+g\right)\mu_{\text{Survivors}} \\
\implies \rho\mu_{\text{Exiters}} &= \rho\mu_{\text{Entrants}} + \left(1-\rho\right) g\mu_{\text{Survivors}} \\
\implies \rho\mu_{\text{Entrants}} &= \rho\mu_{\text{Exiters}} - \left(1-\rho\right) g\mu_{\text{Survivors}} \\
\implies \mu_{\text{Entrants}} &= \mu_{\text{Exiters}} - \frac{\left(1-\rho\right)}{\rho} g\mu_{\text{Survivors}} \\
&\to \mu_{\text{Exiters}} - \frac{g}{\rho}\mu_{\text{Survivors}}
\end{aligned}$$

Intuitively, if there is productivity growth, then the entrants have to be less productive than the exiters, because the survivors will be even more productive tomorrow.

Let us now compute the approximate selection channel, in the case where the exit rate is small. We have:

$$\begin{aligned}
\text{Selection Channel} &\approx \mu - \mu^* \\
&= \rho\mu_{\text{Exiters}} + \left(1-\rho\right)\mu_{\text{Survivors}} - \mu_{\text{Entrants}} \cdot \frac{1}{1-\frac{g}{\rho}} \\
&\to \mu_{\text{Survivors}} - \mu_{\text{Entrants}} \cdot \frac{1}{1-\frac{g}{\rho}} \\
&\to \mu_{\text{Survivors}} - \left(\mu_{\text{Exiters}} - \frac{g}{\rho}\mu_{\text{Survivors}}\right) \cdot \frac{1}{1-\frac{g}{\rho}} \\
&= \frac{1}{1-\frac{g}{\rho}} \cdot \left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right)
\end{aligned}$$

Thus, life-cycle productivity growth strengthens the selection channel.

Both population growth and life-cycle productivity growth affect the selection channel through a multiplier that depends on the ratio of the growth rate to the exit rate: population growth multiplies the selection channel by $\frac{1}{1+\frac{g}{\rho}}$, while life-cycle productivity growth multiplies the selection channel by $\frac{1}{1-\frac{g}{\rho}}$. Intuitively, population growth shifts the mass of production towards younger firms, relative to the no growth benchmark, while life-cycle productivity growth makes older firms more important. Since older firms are more selected than younger firms, it makes sense that population growth would weaken the selection channel, while life-cycle productivity growth strengthens it.

B.4 Extension: Mean Reversion

We now consider an extension with mean-reverting productivity dynamics. Assume there is no population growth, and that, in the random exit world, there is no trend in average productivity.

Instead, we have AR(1) productivity dynamics:

$$\mathbb{E}\left[\log z_{it}\right] = (1-\alpha)\,\mu_{\text{Entrants}} + \alpha \log z_{i,t-1}$$

where $\alpha = e^{-(1-\tilde{\alpha})\Delta}$ (this form reflects that persistence will decay over time, but that when α is higher, the persistence is slower). Note also that, using L'Hospital's rule, we have $\frac{(1-\alpha)}{\rho} = \frac{1-e^{-(1-\tilde{\alpha})\Delta}}{1-e^{-\tilde{\rho}\Delta}} \to \frac{1-\tilde{\alpha}e^{-(1-\tilde{\alpha})\Delta}}{\tilde{\rho}e^{-\tilde{\rho}\Delta}} \to \frac{1-\tilde{\alpha}}{\tilde{\rho}}$. Note also that this derivation bakes in the assumption that $\alpha > 0$. If $\alpha = 0$, then we instead have $\frac{(1-\alpha)}{\rho} = \frac{1}{\rho} \to \infty$.

The identity equation will stay the same as in the baseline. The counterfactual equation will also stay the same: which firm is most productive will change over time, but in a world without selection, the mean will always be μ_{Entrants}.

The steady-state equation will change. We have:

$$\begin{aligned}
\rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} &= \rho\mu_{\text{Entrants}} + (1-\rho)\left((1-\alpha)\,\mu_{\text{Entrants}} + \alpha\mu_{\text{Survivors}}\right) \\
\implies \rho\mu_{\text{Exiters}} &= \rho\mu_{\text{Entrants}} + (1-\rho)\,(1-\alpha)\left(\mu_{\text{Entrants}} - \mu_{\text{Survivors}}\right) \\
\implies \rho\mu_{\text{Entrants}} &= \rho\mu_{\text{Exiters}} - (1-\rho)\,(1-\alpha)\left(\mu_{\text{Entrants}} - \mu_{\text{Survivors}}\right) \\
\implies \mu_{\text{Entrants}} &= \mu_{\text{Exiters}} - \frac{1-\rho}{\rho}\,(1-\alpha)\left(\mu_{\text{Entrants}} - \mu_{\text{Survivors}}\right) \\
&\to \mu_{\text{Exiters}} - \frac{1-\alpha}{\rho}\left(\mu_{\text{Entrants}} - \mu_{\text{Survivors}}\right)
\end{aligned}$$

Intuitively, because there is mean reversion, the new entrants will be more productive than the exiters, since they are being balanced by the survivors in $t+1$, who are not as productive as the survivors in time t.

We can now compute the approximate selection channel, in the case where the exit rate is small. We have:

$$
\begin{aligned}
\text{Selection Channel} &\approx \mu - \mu^* \\
&= \rho\mu_{\text{Exiters}} + (1-\rho)\,\mu_{\text{Survivors}} - \mu_{\text{Entrants}} \\
&\to \mu_{\text{Survivors}} - \mu_{\text{Exiters}} + \frac{1-\alpha}{\rho}\left(\mu_{\text{Entrants}} - \mu_{\text{Survivors}}\right) \\
&= \mu_{\text{Survivors}} - \mu_{\text{Exiters}} - \frac{1-\alpha}{\rho}\left(\mu_{\text{Survivors}} - \mu^*\right) \\
\implies \text{Selection Channel} &\approx \frac{1}{1+\frac{1-\alpha}{\rho}} \cdot \left(\mu_{\text{Survivors}} - \mu_{\text{Exiters}}\right)
\end{aligned}
$$

Thus, mean reversion will tend to weaken the strength of the selection channel. When $\frac{1-\alpha}{\rho}$ is zero, that means we have full persistence ($\alpha = 1$). In that case, the selection channel is the same as in the baseline. When $\frac{1-\alpha}{\rho}$ is higher, then we have attenuation of the selection channel. When $\alpha = 0$, then $\frac{1-\alpha}{\rho} \to \infty$ as $\rho \to 0$, and the selection channel will be zero.