The Empirical Distribution of Firm Dynamics and Its Macro Implications*

Nir Jaimovich
University of California, San Diego, *NBER*, and *CEPR*
nijaimovich@ucsd.edu

Stephen J. Terry
Boston University and NBER
stephent@bu.edu

Nicolas Vincent
HEC Montréal
nicolas.vincent@hec.ca

June 2023

Abstract

Heterogeneous firm models are ubiquitous in modern macroeconomics. We revisit a central feature of these models: the idiosyncratic shock process faced by firms. Using a large representative firm-level dataset, we document nonparametrically that the common assumption, a Gaussian AR(1) shock process, is at odds in important ways with observed fat-tailed firm dynamics. We embed these findings within a standard quantitative general equilibrium heterogeneous firm dynamics model and show that the nature of firm-level shocks has a sizable quantitative effect on the economy's responsiveness to aggregate shifts.

Keywords: firm dynamics, nonparametric shocks, selection, subsidy policy

*We would like to thank Mathieu Potvin for excellent research assistance on the project. Financial support from the Fondation HEC Montréal and the Social Sciences and Humanities Research Council of Canada (Vincent) is gratefully acknowledged. We also thank participants at numerous seminars and conferences for valuable feedback. We use historical ORBIS data accessed through an NBER consortium agreement with Moody's Bureau van Dijk.

"The overall effect on aggregate employment seems ambiguous, depending on the stochastic structure of firm-level shocks. This being the case, evidence on the firm-level stochastic environment is necessary." Hopenhayn and Rogerson (1993)

1 Introduction

Heterogeneous firm models are at the center of modern macroeconomics and are nowadays routinely used both for positive and normative analysis. Since the seminal contribution of Hopenhayn (1992), this type of framework has increasingly been employed in macroeconomics to study, among many other topics, the contributions of labor frictions to aggregate outcomes (Hopenhayn and Rogerson, 1993); the cyclical implications of firm entry and exit (Bilbiie et al., 2012; Clementi and Palazzo, 2016; Lee and Mukoyama, 2018); the decline in business dynamism (Decker et al., 2016, 2020; Karahan et al., 2022); the role of firm heterogeneity in shaping aggregate investment dynamics (Khan and Thomas, 2008, 2013; Winberry, 2021); the propagation of financial frictions (Moll, 2014; Midrigan and Xu, 2014; Ottonello and Winberry, 2020); uncertainty shocks (Bloom et al., 2018); and the drivers and consequences of resource misallocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Bento and Restuccia, 2017; Kehrig and Vincent, 2020). This broad family of models has also been highly influential in the trade literature (Melitz, 2003).

In this paper, we revisit standard assumptions regarding the driving force in these models, namely the idiosyncratic shocks faced by firms. The broad importance of these shocks was recognized early on by Hopenhayn and Rogerson (1993), as evidenced by the quote at the top of this page. Specifically, we show that the common parametric assumption made in the literature for idiosyncratic shocks – a Gaussian AR(1) – leads to firm-level distributions and dynamics that differ in important ways from those observed empirically. Furthermore, we demonstrate that these differences have a sizable impact on aggregate predictions when analyzing the response of quantitative models to a range of simply policy experiments. In what follows, we describe these findings in more detail.

We start by constructing a large firm-level panel dataset from the historical ORBIS database. Relative to common alternative US datasets, ORBIS proves attractive by providing, for a set of countries, a representative size distribution – due to its inclusion of both private and publicly listed firms – and information on a broad range of firm-level outcomes. For our baseline country, Spain, our sample covers just over one million firms that were active at some point during the 2005-2014 period. We also show that our results are robust to instead using other countries with broad coverage in ORBIS.

Next, we exploit this large dataset to nonparametrically characterize revenue dynamics for incumbent firms, as well as firm entry and exit patterns. We compare our empirical estimates to similar objects implied by the widely used Gaussian AR(1) model. Broadly speaking, we find that the *stationary distribution* of firm revenue is not strikingly different in the two cases, except for the well-known presence of a fat or power law tail in the empirical cross section. The *dynamics*, however, are very different. For instance, a firm initially in the center of the revenue distribution has a much higher probability of placidly remaining near the median *or* dramatically moving to the tails than implied under a Gaussian AR(1). Conversely, conditional on starting in the tails of the distribution, firms face a greater probability of returning towards the center than in the Gaussian AR(1). These differences result in a fat-tailed or leptokurtic distribution of revenue *growth* in the data, incompatible with the standard Gaussian AR(1) assumption.

What are the economic implications of these statistical findings? In any model, a firm's expected continuation value is ultimately the object that shapes its optimal decisions, such as whether to exit or how much to hire and invest. But firm value is, of course, not typically observable in representative datasets like ours, for which private firms represent the vast majority of businesses. Hence, for the remainder of our empirical analysis we define and construct a related concept of firm-level "lifetime revenue," i.e., the present discounted value of expected revenue. We empirically measure two versions of lifetime revenue based on the firm dynamics implied by the nonparametric empirical and parametric Gaussian AR(1) models. We show that the implied firm-level lifetime revenue distributions differ substantially between these two cases. In particular, lifetime revenue is more clustered at lower revenue states where the likelihood of exit is high, i.e., the probability density function is much higher in these regions, for the nonparametric version. The reason is intuitive: the higher probability of moving away from the tails and, conversely, remaining near the center of the revenue distribution creates a pronounced compression of the distribution of firm lifetime prospects under the nonparametric empirical specification.

Our empirical findings have potential implications for almost any model featuring firm optimization in the face of idiosyncratic shocks. As a first step, we demonstrate analytically, within a simple pared-down firm dynamics model with entry and exit, that clustered distributions like those observed empirically cause a higher sensitivity of the firm exit rate to shifts in the economy. We then generalize this intuition by building and analyzing a canonical quantitative general equilibrium firm dynamics model with heterogeneous firms. Our particular model is intentionally standard and based on Hopenhayn (1992) and Hopenhayn and Rogerson (1993), featuring firm-level profitability shocks, fixed costs of operation, endogenous exit, and sunk entry costs. We consider and solve both nonparametric and parametric versions of the model. For the nonparametric version, we develop a novel solution and calibration technique allowing us to perfectly match our key empirical objects: (i) the revenue transition distribution for incumbents; (ii) the exit hazard; and (iii) the relative size distribution of entrants. For the parametric version, we follow the literature in calibrating a Gaussian AR(1) shock process as well as other parameters to match a number of straightforward empirical moments. Aside from the modeling of the idiosyncratic shock processes, the two models are identical and therefore directly comparable.

Motivated by some of the earliest theoretical experiments in the firm dynamics literature in Hopenhayn (1992), we then compare the impacts of two types of simple policies in both models: a fixed subsidy to each operating firm and a subsidy to entrants. In both cases, the exit rate responds more in the nonparametric case because firm value is more clustered at lower revenue levels, where exit is both more likely and more sensitive. In other words, the intuition from our empirical analysis of lifetime revenue and from our simple model carries over directly to the quantitative model. Yet, we show that the implications for the response of aggregate output to the subsidies is more subtle and depends on the nature of the policy under consideration: the response in the nonparametric model is significantly *smaller* response for the fixed operating subsidy but *larger* for the entry subsidy. As we explain next, the reason is tied to the role of selection.

Consider first the subsidy to operating firms, which directly increases firms' values and lowers the overall exit rate. As a result, selection worsens, leading to more low-profitability firms surviving. Due to a stronger decline of the exit rate, this negative selection effect is more pronounced in the nonparametric model, dampening the aggregate increase in output relative to the AR(1) version. On the other hand, the entry subsidy generates an *increase* in the exit rate. As a result, selection improves, boosting output more in the nonparametric case.

With our benchmark results in hand, we then undertake multiple robustness checks and extensions. Empirically, we construct alternative ORBIS datasets based on other countries, time periods, sectoral subsamples, and data treatment approaches, including an investigation of the role played by firm age. In every case, we continue to uncover fat-tailed revenue dynamics inconsistent with a Gaussian AR(1). In addition, we repeat our model-based quantitative analysis and show that under each robustness check, we continue to uncover dramatically higher sensitivity of exit rates in our nonparametric model. The same findings hold when we consider various alternative model assumptions and parameterizations, or rely on an extended parametric statistical model. Finally, we document a correlation between clustered distributions and higher exit rate sensitivity at the industry level that is in line with the predictions of the model.

In summary, we not only show that the common use of the Gaussian AR(1) model in heterogeneous firms models is at odds with observed fat-tailed revenue dynamics, but we also argue that these differences matter for quantitative macroeconomic analysis.

We view this paper as closely linked to three main strands of the literature. First, the paper naturally relates to theoretical and empirical work on firm dynamics (Dunne et al., 1989; Hopenhayn, 1992; Davis and Haltiwanger, 1992; Kehrig, 2015; Clementi and Palazzo, 2016; Karahan et al., 2022). This literature exploits firm heterogeneity to rationalize stylized facts about firm dynamics and draw macro conclusions. Our paper contributes both new facts and new quantitative implications. Second, we contribute to existing work contrasting empirical and "conventional" distributions (Midrigan, 2011; Carvalho and Grassi, 2019; Forneron, 2020; Guvenen et al., 2021; Sterk et al., 2021; Guvenen et al., 2023; Boar et al., 2023; Barro and Ursúa, 2012), showing that the common parametric assumptions used in heterogeneous agent models are poor approximations of reality. Third, our work relates to the allocative implications of policy and shocks in the presence of firm heterogeneity (Hopenhayn and Rogerson, 1993; Guner et al., 2008; Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Davies and Eckel, 2010; Gourio and Miao, 2010; Asker et al., 2014; Garicano et al., 2016; Catherine et al., 2018; Kehrig and Vincent, 2020; Ottonello and Winberry, 2020; Bils et al., 2021; Sraer and Thesmar, 2021). We show that our empirical findings alter policy impacts in a quantitatively significant manner.

The rest of the paper is organized as follows. Section 2 introduces our data and facts. Section 3 analyzes our simple model. Section 4 builds a canonical quantitative general equilibrium firm dynamics model. Section 5 employs separate nonparametric and parametric approaches to solve and calibrate the model. Section 6 analyzes simple experiments within each version of the model that highlight the economic implications of the nonparametric calibration. Section 7 discusses empirical and quantitative robustness checks. Section 8 concludes. Online appendices provide further details on our empirical analysis (Appendix A) and quantitative analysis (Appendix B).

2 Data

In this section, we introduce our representative firm microdata and present our nonparametric approach to extracting several key empirical objects. With this framework in hand, we then describe empirical firm revenue dynamics.

2.1 ORBIS Data

Measuring firm outcomes, even key variables such as revenue, is notoriously difficult in the US. Accessible sources such as Compustat tend to limit their scope to publicly listed – large, nonrepresentative – firms, while government data sources such as those within the US Census often lack financial outcomes, or also suffer from unrepresentativeness (Crouzet and Mehrotra, 2020). To overcome these challenges, we broaden our focus and turn to Moody's, formerly Bureau van Dijk's, historical ORBIS dataset. ORBIS is drawn mostly from government business registers and contains many firm-level outcomes for both private and publicly listed companies at yearly frequency, for many countries. Coverage and representativeness vary greatly across countries, and researchers working with ORBIS data must also be mindful of differentiating the commercial from the historical ORBIS datasets, with varying sample selection criteria. Despite these subtleties, Kalemli-Ozcan et al. (2022) and Bajgar et al. (2020) demonstrate in detail that for multiple countries, mostly located in Western Europe, the historical ORBIS data yields a representative sample matching administrative measures – such as those aggregated in the COMPNET database – of the firm size distribution for revenue or employment.

Our baseline sample consists of just over one million private and public Spanish firms that are active at some point over the years 2005-2014, for a total of around five million firm-year observations. Appendix Table A.1 presents some summary statistics on this sample. Although our Spanish ORBIS data is a useful benchmark, we show in Section 7 that our results are robust to instead using data from Italy, Portugal, France, and Norway, nations for which ORBIS data is also representative.

2.2 Measuring Three Key Empirical Objects

ORBIS includes many economic and financial variables, but we focus on revenue given its role in canonical firm dynamics models as an outcome linked to both shock processes and a firm's production choices. Revenue is also, helpfully, one of the most populated outcomes in ORBIS across countries. Given our focus on idiosyncratic patterns, we analyze log firm revenue residualized with respect to both sector and year effects, sometimes succinctly referring to this measure as "revenue" below. Omitting subscripts, we denote this outcome y for a given firm year and let y' indicate the following year's outcome at the same firm. We also construct indicators for firm entry and exit, a task made easier by ORBIS' firm panel structure. We separate firms into the potentially overlapping categories of "incumbents" including all those operating in a given year; "entrants" including only first-year incumbents; and "continuing" firms which operate in future year(s). With this dataset in hand, we nonparametrically measure three objects.

1. The transition distribution, i.e., the distribution of next-year's revenue conditional upon current revenue for continuing incumbent firms, denoted $H(y'|y)$
2. The distribution of revenue for entrant firms, denoted $H_E(y)$
3. The exit hazard for incumbent firms, denoted $\mathbb{P}(\text{Exit}|y)$

Our extensive sample allows us to perform straightforward nonparametric estimation. First, we discretize firm revenue into 101 equally weighted intervals. Next, we estimate three objects: the matrix $H(y'|y)$ describing incumbent dynamics, obtained using transitions of firm revenue across intervals for continuing incumbents; the vector $H_E(y)$, defined as the distribution of entrants across revenue intervals; and the vector $\mathbb{P}(\text{Exit}|y)$, capturing the exit rates of incumbent firms across revenue intervals. Below, we will interchangeably refer to these objects as "nonparametric" or "empirical," and all of the facts we lay out below are functions of these three items.

For comparison, we also consider a Gaussian AR(1) parametric model for revenue $y' = \rho_y y + \sigma_y \varepsilon$, where $\varepsilon \sim N(0,1)$. We estimate ρ and σ to match the autocorrelation and unconditional variance of revenue in our data.[1] This parametric model implies a transition distribution $H_{AR(1)}(y'|y)$, different from the empirical distribution $H(y'|y)$. Our analysis below contrasts nonparametric empirical facts and those implied by the parametric Gaussian AR(1) case.

Appendix A presents more detailed information on our sample construction, data treatment, and estimation approaches. Section 7 demonstrates our results' robustness to a range of alternative data treatment choices and sample splits. Notably, our robustness exercises also confirm that our results are not driven by firm age, a variable not directly incorporated in our baseline analysis.

2.3 Facts

This section lays out some key stylized facts. Where relevant, red lines indicate our empirical nonparametric estimates while blue lines indicate outcomes implied by a parametric Gaussian AR(1).

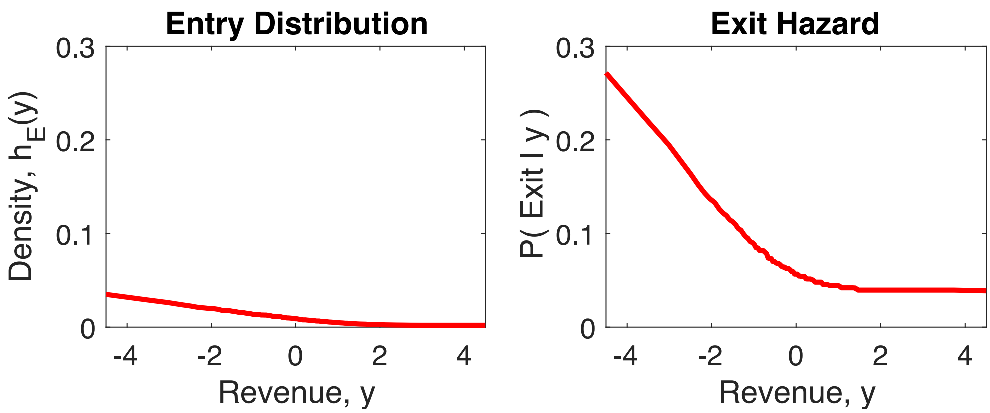
Entry and Exit Patterns Figure 1 plots the entry density $h_E(y)$ (left panel) and exit hazard $\mathbb{P}(\text{Exit}|y)$ (right panel). Both objects are convex and downward sloping in revenue y, although the entry distribution is somewhat flatter. The downward slope of the exit hazard aligns with the predictions of canonical firm dynamics models.[2]

Revenue Dynamics The top row of Figure 2 plots two measures of revenue dynamics for continuing incumbents. The top left panel plots the densities from our empirical transition distribution $H(y'|y)$ and the parametric AR(1) transition distribution $H_{AR(1)}(y'|y)$, conditional upon a starting level of revenue y equal to the median. Specifically, conditional on being initially at its median value, the probability of log revenue remaining around the median is much higher empirically than

[1] For our baseline Spanish (log) revenue dataset, we find $\hat{\rho}_y = 0.94$ and $\hat{\sigma}_y = 0.57$ for continuing incumbent firms.

[2] The sharp downward slope of the exit hazard in revenue provides another justification for our focus on revenues to discipline our firm dynamics model. In fact, in Appendix Table A.6 we also show that firm revenue contains more explanatory power for firm exit than firm profits.

Figure 1: Firm Entry and Exit Patterns in the Data



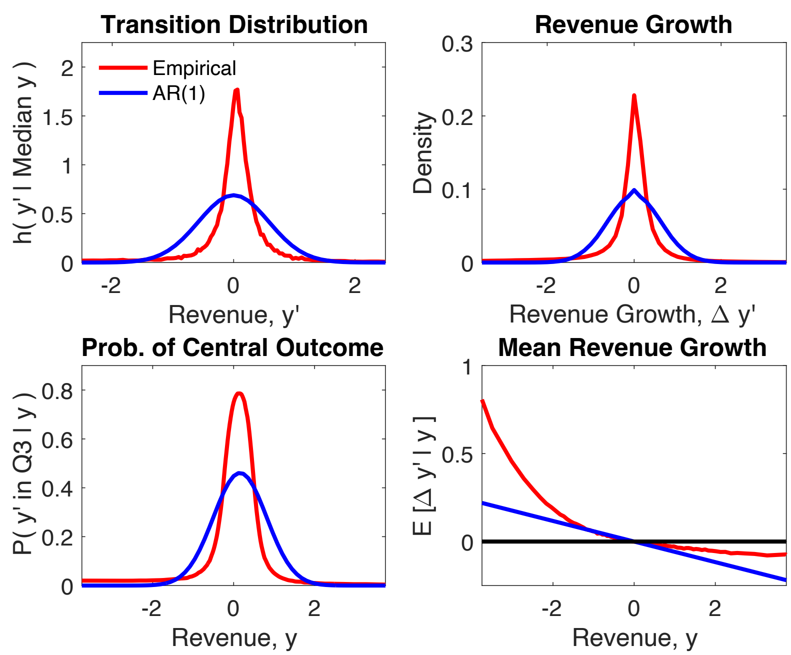
Notes: The figure plots the estimated entry density $h_E(y)$ (left panel) as well as the estimated exit hazard $\mathbb{P}(\text{Exit}|y)$ (right panel) as a function of y, i.e., log revenue residualized by sector and year. Both objects are estimated nonparametrically using our baseline Spanish ORBIS sample covering around five million firm years for around one million firms over the 2005-2014 period.

under the AR(1) specification. Yet, the empirical distribution also features a higher (though low) likelihood of moving to the tails from the median.

The top right panel plots unconditional distributions of revenue growth. The empirical (parametric) revenue growth distribution features a standard deviation, skewness, and kurtosis of 0.65 (0.59), -0.31 (0), and 29.21 (3) respectively. Hence, although the dispersion and skewness are roughly similar, the empirical revenue dynamics are distinctly leptokurtic or "fat-tailed." Intuitively, firms are overwhelmingly more likely to experience very small, but also sometimes very large, yearly revenue growth rates relative to those implied by the standard Gaussian AR(1).

The bottom row of Figure 2 provides insight into the revenue mobility of incumbents, especially mobility from the tails. Specifically, the bottom left panel plots the probability that a firm's revenue next year lies in the 3rd quintile, i.e., the center of the distribution, while the bottom right panel plots mean revenue growth over the next year. For the empirical and parametric Gaussian AR(1) versions, outcomes are plotted conditional upon firm revenue today. We can see in the bottom left panel that smaller firms are empirically more likely to grow towards the center of the distribution than an AR(1) would imply. This pattern is echoed in the high conditional mean of revenue growth rates for such firms in the bottom right panel. Despite the

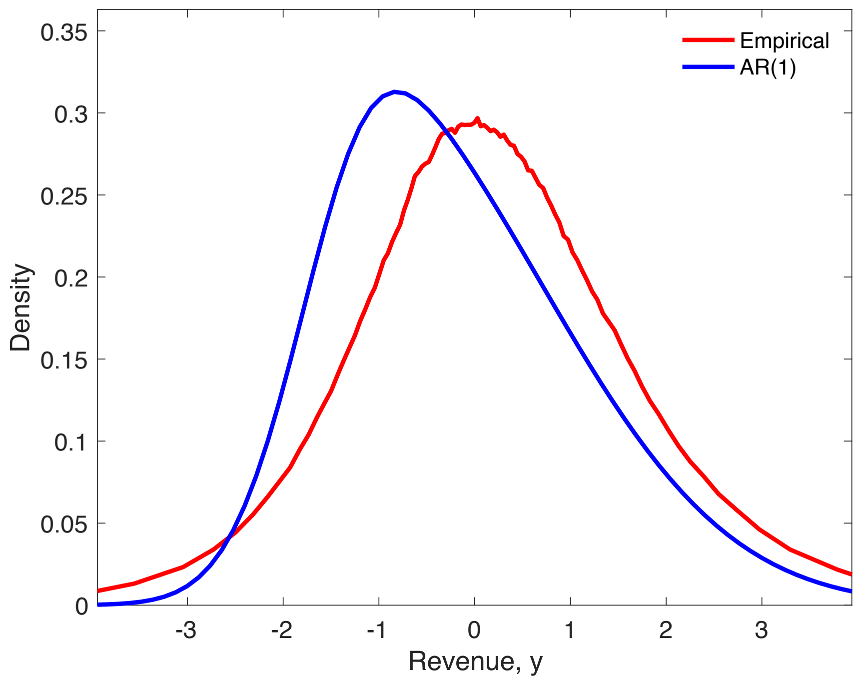
Figure 2: Revenue Dynamics



Notes: The top left panel of the figure plots the distribution of next year's firm revenue y' conditional upon median revenue y in the current year. The top right panel plots the stationary distribution of yearly revenue growth $\Delta y'$. The bottom left panel plots the probability that next year's firm revenue y' lies within the central or 3rd quintile, conditional upon revenue y in the current year. The bottom right panel plots mean revenue growth $\Delta y'$ over the next year, conditional upon revenue y in the current year. Here, y is log revenue residualized by sector and year from our baseline Spanish ORBIS sample covering around five million firm years for around one million firms over the 2005-2014 period. In each panel, the red line is computed from the empirical nonparametric estimates $H(y'|y)$ while the blue line reflects the transition distribution $H_{AR(1)}(y'|y)$ implied by the parametric AR(1) case.

fact that AR(1) implied transitions at the right tail are more aligned with those extracted from the data, significant differences remain. In particular, the conditional mean of revenue growth is linear in revenue in the AR(1) case while clearly nonlinear in the data. In summary, these patterns reveal that firms are more empirically likely to grow quickly towards, and then remain within, the center of the distribution than implied by the parametric model.

Figure 3: Stationary Revenue Distributions in the Data



Notes: The figure plots the stationary distribution of firm revenue y. Here, y is log revenue residualized by sector and year from our baseline Spanish ORBIS sample covering around five million firm years for around one million firms over the 2005-2014 period. The red line is computed from our empirical nonparametric estimates, while the blue line reflects the transition distribution implied by the parametric AR(1) case.

Stationary Revenue Distribution Next, we compute the empirical nonparametric stationary distribution $H(y)$ of firm revenue satisfying

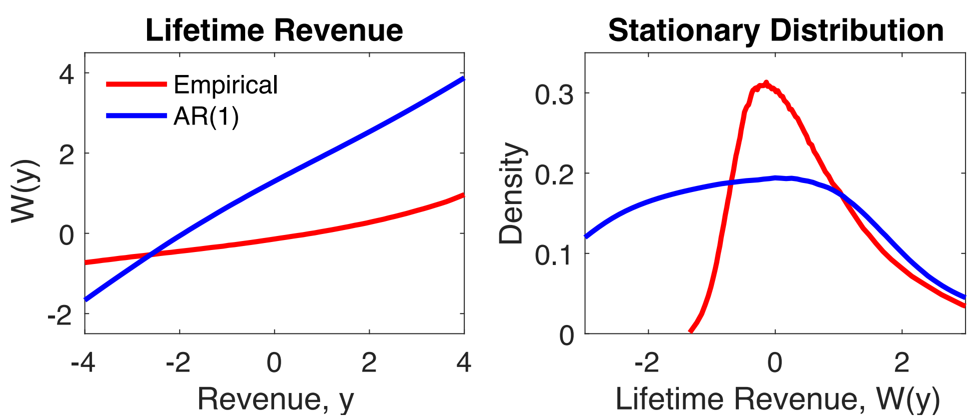
$$H(y') = \int H(y'|y)\,(1 - \mathbb{P}(\text{Exit}|y))\,dH(y) + \mathbb{P}(\text{Exit})H_E(y'),$$

where $H(y'|y)$, $\mathbb{P}(\text{Exit}|y)$, and $H_E(y')$ match their data counterparts.[3] For comparison, we also compute an otherwise identical parametric stationary distribution of revenue $H_{AR(1)}(y)$ by simply replacing the empirical transition distribution $H(y'|y)$ with its counterpart $H_{AR(1)}(y'|y)$. Figure 3 plots the densities associated with these two unconditional distributions, which do not appear dramatically different to the

[3]Note that for the last term we rely on the fact that in a stationary steady state $\mathbb{P}(\text{Exit}) = \mathbb{P}(\text{Entry}) = \int \mathbb{P}(\text{Exit}|y)dH(y)$.

naked eye. Of course the two distributions do differ in meaningful ways.[4] In particular, outside the plotted range the empirical distribution $H(y)$ exhibits the well known fat tail consistent with a power law which no Gaussian model can match, a fact we document in Appendix Figure A.2. But, overall, the figure draws a contrast between the reasonably high similarity of the empirical and parametric cross-sectional distributions of firm revenue *levels* versus their markedly divergent implications for firm revenue *dynamics* documented above.

Figure 4: Lifetime Revenue



Notes: The figure plots lifetime revenue as a function of current revenue (left panel) as well the stationary distribution of lifetime revenue (right panel). Here, y is log revenue residualized by sector and year, and for ease of reference we present lifetime revenue $W(y)$ in logs and subsequently demeaned. Our empirical estimates come from our baseline Spanish ORBIS sample covering around five million firm years for around one million firms over the 2005-2014 period. In each panel, the red line is computed from our empirical nonparametric estimates while the blue line reflects the transition distribution implied by the parametric AR(1) case.

Firm "Lifetime Revenue" If the documented differences in revenue dynamics between the empirical and parametric models do not generate large apparent differences in the stationary distribution of revenue *levels*, do they still matter? The answer is yes. Revenue dynamics impact expected firm *lifetime* outcomes, such as firm value. In canonical firm dynamics models, firm value – the expected present discounted value of payouts – is not only the key decision-relevant measure for firm entry and exit, it

[4] The standard deviation, skewness, and kurtosis of the empirical (parametric) stationary distributions for revenue levels y are roughly comparable at 1.77 (1.36), 0.01 (0.60), and 7.54 (3.28).

also shapes the firm's choice along many dimensions, such as hiring and investment.

Directly obtaining information on firm value for a representative set of firms is, unfortunately, impossible. Since the vast majority of firms are not publicly listed, their market value is not observable. Instead, our approach is to construct a novel proxy of firm value, which we refer to as "lifetime revenue:" the expected present discounted value of future firm revenue. As we show below, this object can be computed from any dataset that includes information on firm revenue and exit patterns for both listed and unlisted firms, suchs as ORBIS. Note that the measure relies on *revenue* instead of *payouts*, because information on payouts is often missing or of dubious quality in representative firm-level datasets. Yet, in most widely used quantitative firm dynamics models, payouts in firm value equations are highly correlated with and dominated quantitatively by firm revenue.[5] Moreover, we later show in Appendix A.2 that for the small subsample of publicly listed firms, our measure of lifetime revenue is a good predictor of observed market value.

Lifetime revenue $W(y)$ can be easily computed as a function of current firm revenue y, using only the estimated objects already in hand, via the Bellman equation

$$W(y) = e^y + \left(\frac{1 - \mathbb{P}(\text{Exit}|y)}{R}\right) \int W(y')dH(y'|y).$$

We choose $R > 1$ to deliver a conventional constant yearly real interest rate of 4%. Besides this assumption, lifetime revenue $W(y)$ is otherwise purely a function of our estimated empirical transition distribution $H(y'|y)$ and exit hazard $\mathbb{P}(\text{Exit}|y)$. We also compute an analogous lifetime revenue object $W_{AR(1)}(y)$ using the same Bellman equation but replacing the transition distribution $H(y'|y)$ with its parametric AR(1) counterpart $H_{AR(1)}(y'|y)$. Finally, relying on the stationary revenue distributions $H(y)$ and $H_{AR(1)}(y)$ computed above, we immediately obtain stationary distributions of lifetime revenue $H(W)$ and $H_{AR(1)}(W)$ for the empirical and nonparametric cases.

In the left panel of Figure 4, we plot the lifetime revenue constructs $W(y)$ and $W_{AR(1)}(y)$ as a function of current revenue y. The two functions are strikingly differ-

[5] In particular, note that in most quantitative applications of models in the Hopenhayn (1992) tradition, firm payouts can be divided into two terms. The first term is proportional to firm revenue, and the second term reflects transitory adjustments based on flow factors such as investment, financial frictions, or adjustment costs. The first term tends to be meaningfully larger in magnitude and more persistent than the second, driving a high correlation between revenue outcomes over a firm's lifetime and its underlying difficult-to-measure expected firm payouts.

ent, with the empirical version being much flatter than its AR(1) counterpart. This finding is a key insight from our paper with important implications for the predictions of dynamic heterogeneous firm models, as we show in later sections.

To understand the intuition behind this result, recall two facts that we documented earlier. First, revenue dynamics in the data are fat-tailed or leptokurtic, a property highlighted in Figure 2. In particular, for the median firm, small revenue transitions are by far the most common outcomes. Second, firms face a higher likelihood of transitioning out of the tails of the size distribution than what is implied by the AR(1) process (see the bottom row of Figure 2). That is, movement towards the center is more prevalent than generally assumed, particularly from low-y states. Together, these two facts about *dynamics* generate an implied lifetime revenue function that is much less sensitive to current revenue under the non-parametric empirical case than under the parametric AR(1). In other words, firms with different current revenue levels of y have lifetime revenue values that are not as different from one another as the AR(1) model suggests. As a result, the stationary distribution of lifetime revenue, shown in the right panel of Figure 4, is therefore much less dispersed and exhibits more "clustering" or higher densities at low levels in the data than under the parametric AR(1).[6]

In the remainder of the paper, we draw out the implications of this key insight for firm dynamics models. Specifically, we show that the degree of clustering for firm values links directly to the sensitivity of overall firm exit to changes in the economic environment, i.e., that these empirical facts directly discipline and change the quantitative aggregate implications of workhorse firm dynamics models.

3 Simple Model

In the previous section, we showed that the cross-sectional firm size distribution in the data is roughly similar to the one implied by the parametric Gaussian AR(1) case. Yet, firm revenue dynamics differ markedly, generating important differences in expected lifetime outcomes. In this section, we first analyze the implications for firm dynamics of this divergence in a simple and transparent analytical framework. We then proceed in Section 4 to a quantitative general equilibrium heterogeneous firms

[6] The standard deviation, skewness, and kurtosis of the empirical (parametric) log lifetime revenue distributions in the right panel of Figure 4 are 1.17 (1.85), 1.67 (0.20), and 6.66 (2.53), respectively.

model.

Time is discrete. Each of a unit mass of existing firms chooses at the start of period $t = 0$ whether to exit or to continue operating. Continuing commits a firm to operate forever, from $t = 0$ onwards, while a firm exiting immediately receives an outside option of 0. Firms are risk neutral and discount the future at the constant exogenous rate $R > 1$. Each firm observes its own current exogenous profitability state z in period 0 before choosing to continue or exit. A firm's net payoff in any period equals its profitability z plus an exogenous constant $\mu\left(\frac{R-1}{R}\right)$. At the start of period 0, the cross-sectional firm profitability distribution is exogenously given by $z \sim N(0, \sigma_z^2)$. That is, for all cases discussed below, we assume an identical cross-sectional distribution of z at $t = 0$. However, we examine three distinct cases for the *dynamics* of z and their implications for the time-0 distribution of the firm values V.

Permanent z In this case, firm profitability z is permanent and fixed, so that a firm's value is

$$V_{perm}(z) = \mu + z + \frac{1}{R}z + \frac{1}{R^2} + ... = \mu + z\left(\frac{R}{R-1}\right).$$

Hence, the distribution of firm values at the start of period 0 is

$$V_{perm} \sim N\left(\mu, \sigma_z^2\left(\frac{R}{R-1}\right)^2\right) = N(\mu, \sigma_{perm}^2).$$

Persistent z In this case, each firm's profitability follows an independent Gaussian AR(1) with persistence satisfying $0 < \rho < 1$ and shock variance $\sigma^2 = (1-\rho^2)\sigma_z^2$. A firm's expected value is

$$V_{pers}(z) = \mu + z + \frac{1}{R}\rho z + \frac{1}{R^2}\rho^2 z + ... = \mu + z\left(\frac{R}{R-\rho}\right).$$

The distribution of firm values at the start of period 0 is therefore

$$V_{pers} \sim N\left(\mu, \sigma_z^2\left(\frac{R}{R-\rho}\right)^2\right) = N(\mu, \sigma_{pers}^2).$$

Transitory z In this case, firm profitability $z \sim N(0, \sigma_z^2)$ is iid across time and firms. A firm's expected value is

$$V_{iid}(z) = \mu + z + \frac{1}{R}\mathbb{E}(z) + \frac{1}{R^2}\mathbb{E}(z) + ... = \mu + z.$$

As a result, the firm value distribution at the start of period 0 is

$$V_{iid} \sim N\left(\mu, \sigma_z^2\right) = N(\mu, \sigma_{iid}^2).$$

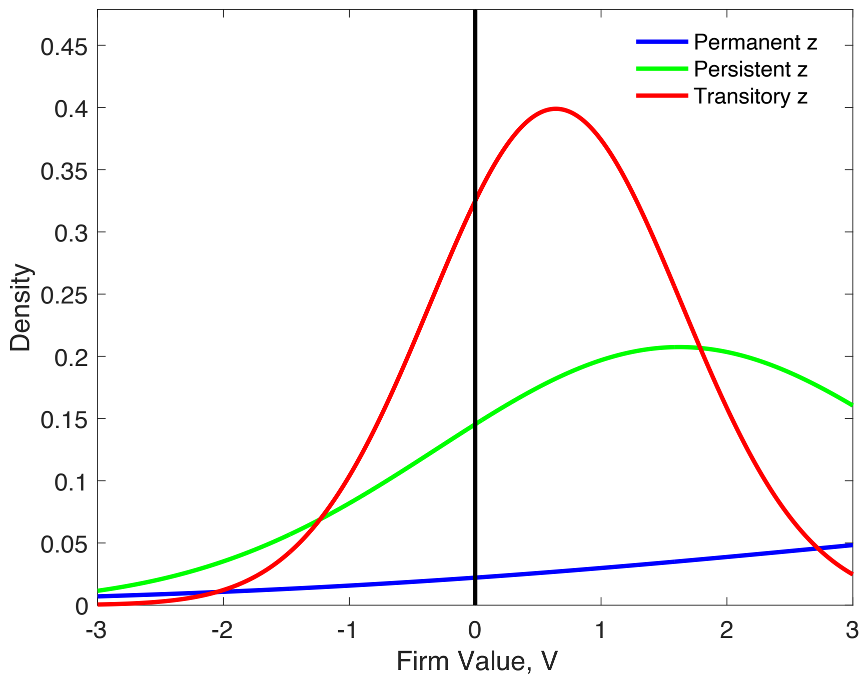
Note that since $0 < \rho < 1 < R$, we can rank the variances of the underlying firm value distributions across cases as $\sigma_{iid}^2 < \sigma_{pers}^2 < \sigma_{perm}^2$. Intuitively, faster mean reversion at the firm level generates a more compressed distribution of firm value. Hence, although all cases exhibit an identical cross-sectional distribution of profitability or size z at $t = 0$, the divergent dynamics of profitability imply different distributions of firms' decision-relevant object: firm value.

For an illustrative parameterization, Figure 5 plots the firm value distribution under each scenario. Firms with negative value below the plotted threshold level at 0 choose to exit in period 0. To allow comparison across cases, we choose the mean payoff parameter μ in each case to guarantee an identical exit rate.[7] Firm value dispersion varies widely across the three cases, despite identical cross-sectional size distributions, with higher mean reversion generating more compressed firm value distributions in the transitory and persistent cases relative to the permanent case.

But do these distributional differences matter for firm dynamics, and in particular exit rates? To answer this question, first note that the exit rate is simply $\mathbb{P}(\text{Exit}) = F(0)$, where $F(V)$ is the firm value CDF. Next, consider the implementation of a one-time subsidy $s > 0$ paid to all firms. This policy naturally shifts the firm value distribution to the right and implies a new, lower exit rate $\mathbb{P}(\text{Exit}|s) = F(0 - s)$. Specifically, the local sensitivity of the exit rate to the subsidy is $\frac{\partial \mathbb{P}(\text{Exit})}{\partial s}|_{s=0} = -f(0)$, where $f(V) = F'(V)$. Consequently, when the distribution of firm value is more clustered, with a higher density of firms at the exit threshold $V = 0$, the exit rate is more responsive to the subsidy. Such clustering varies widely across the cases in Figure 5.

[7] We target 10%, a round value, in Figure 5. We pick μ_i such that $\Phi\left(\frac{z_i^* - \mu_i}{\sigma_i}\right) = p_{exit}$, where z_i^* is such that $V_i(z_i^*) = 0$, σ_i is the standard deviation of the potential firm value distribution as defined in the text for each case i, and Φ is the standard normal CDF.

Figure 5: Firm Value Distributions in the Simple Model



Notes: The figure plots the distribution of firm value at the start of period 0 in the simple model. For this illustrative parameterization we choose $\rho = 0.4$, $\sigma_z^2 = 1$, and $R = 1/1.2$. The figure plots the firm value distributions for the permanent z case (blue line), the persistent z case (green line), and the transitory or iid z case (red line) together with the exit threshold of 0 (black vertical line). We also normalize μ separately for each case to generate an exit rate of 10%.

As we showed in Section 2, the rich firm revenue dynamics observed in the data generate an empirical distribution of lifetime revenue outcomes that is more clustered, with higher densities at low levels, than the one implied by the standard Gaussian AR(1) process ubiquitous in the literature. The basic intuition from our simple model suggests that such clustering should cause the aggregate exit rate to be more sensitive to changes in the economy. To formalize and quantify this intuition, we now turn our attention to a quantitative general equilibrium model of firm dynamics.

4 Quantitative Model

Our quantitative general equilibrium firm dynamics model is in the spirit of Hopenhayn (1992), Hopenhayn and Rogerson (1993) and the literature spawned by their seminal work. Time is discrete, and the economy exhibits stationarity with no aggregate uncertainty.

4.1 Operating Firms and Exit

Two types of firms in the economy, incumbents and entrants, form a mass M_O of operating firms in any given period. Each operating firm produces a homogenous numeraire good in the amount $y = zn^{\alpha}$. Firms hire undifferentiated labor n at a competitive wage W. Production exhibits decreasing returns to scale with $0 < \alpha < 1$ and is shifted by an exogenous idiosyncratic profitability shock $z > 0$ following a first-order Markov chain with transition distribution $F(z'|z)$.[8] An operating firm's dynamic problem is summarized by its value function

$$V(z) = \max_{n} \left[zn^{\alpha} - Wn + \mathbb{E}_{\phi_F} \max \left\{ 0, -\phi_F + \frac{1}{R}(1-\delta) \int V(z') dF(z'|z) \right\} \right]. \quad (1)$$

At the start of a period, each operating firm solves a static profit maximization problem for labor input n, which is the sole input used in production. Then, each operating firm receives an iid fixed production cost draw ϕ_F which is denominated in output units and drawn from an exogenous distribution $G(\phi_F)$. Operating firms must pay the fixed cost ϕ_F in order to continue to produce in the future, although a firm can alternatively choose to exercise an option to endogenously exit with limited liability and outside option 0. If an operating firm chooses not to exit, and avoids an iid exogenous death shock with probability satisfying $0 < \delta < 1$, then the firm transits as an incumbent to the next period. An operating firm therefore trivially

[8] Importantly, we note that the decreasing returns to scale production function specification is isomorphic to a monopolistic competition framework with love of variety; our results are not specific to the formulation chosen. So we prefer to refer to z as "profitability" instead of the narrower term "productivity." In order to match the revenue dynamics we do *not* need to take a stand on whether the driving shocks are supply or demand shocks; the revenue function in such a model is given by

$$Revenue = z^{\nu-1} \daleth^{\nu} \text{Aggregates},$$

where z and $\daleth$ respectively denote idiosyncratic productivity shocks and demand shocks.

chooses to endogenously exit if and only if its fixed cost ϕ_F exceeds a threshold level, or equivalently a continuation value, given by

$$\phi_F^*(z) = \frac{1}{R}(1-\delta)\int V(z')dF(z'|z). \tag{2}$$

4.2 Entry

A mass of potential entrants in the economy considers whether to enter at the start of each period. Entry requires that a firm pay an exogenous constant sunk cost $\phi_E > 0$, denominated in output units, in order to obtain an initial profitability shock draw z from the exogenous distribution $F_E(z)$. After receiving an initial profitability draw z, each entrant firm joins the mass of currently operating firms in the current period. Free entry implies that the sunk cost ϕ_E must weakly exceed the value to entry

$$\phi_E \geq \int V(z)dF_E(z), \tag{3}$$

with equality whenever the mass M_E of entry is greater than zero.

4.3 Households

The economy is populated by a measure one of identical households. Households consume the numeraire good and supply labor inelastically in the exogenous amount $\bar{N} > 0$. In addition to labor income, households also receive dividends from operating firms. The household problem reflects an optimal choice of the level of consumption, C, to maximize welfare given by the discounted sum of log utility payoffs. The simple dynamic problem is represented by

$$S = \max_C \{\log(C) + \beta S'\} \tag{4}$$

where the time discount rate satisfies $0 < \beta < 1$ and a standard budget constraint holds. As usual, household intertemporal optimization in a stationary steady state implies that the real interest rate is proportional to the household time discount rate

$$\beta = \frac{1}{R}. \tag{5}$$

4.4 Timing

To summarize, the timing of the model within each period is as follows:

1. New entrant firms pay entry costs.
2. Incumbent firms and new entrants receive their idiosyncratic profitability draws z, drawn from $F_E(z)$ for entrants and according to the transition distribution $F(z|z_{-1})$ for incumbents with previous profitability z_{-1}.
3. Operating firms, i.e., both incumbents and entrants, produce output $y = zn^{\alpha}$ by combining z and labor n hired at the prevailing wage W.
4. Operating firms draw an iid fixed cost $\phi_F \sim G(\phi_F)$.
5. Operating firms form expectations of their continuation value $\phi_F^*(z)$, choosing whether to exit endogenously or remain in operation for next period. Operating firms that choose to remain pay the fixed cost ϕ_F.
6. Households receive firm profits and labor income and then consume.
7. A fraction δ of operating firms exogenously exits.
8. Surviving operating firms transition to the next period as incumbents.

4.5 Stationarity and Aggregates

Stationarity requires that the distribution $F_O(z)$ of operating firms is stable across periods according to the mapping

$$M_O F_O(z) = (1-\delta) M_O \int G(\phi_F^*(z_{-1})) F(z|z_{-1}) dF_O(z_{-1}) + M_E F_E(z), \tag{6}$$

which implicitly defines the distribution $F_O(z)$ but also implies proportionality of the operating and entrant masses according to

$$M_O \mathbb{P}(\text{Exit}) = M_E.^{9} \tag{7}$$

[9] Take $z \to \infty$ in Equation (6) to obtain $M_O = M_O(1-\delta) \int G(\phi_F^*(z_{-1})) dF_O(z_{-1}) + M_E$. Since $\int (1-\delta) G(\phi_F^*(z_{-1})) dF_O(z_{-1}) = \int \left[1 - \mathbb{P}(\text{Exit}|z_{-1})\right] dF_O(z_{-1}) = 1 - \mathbb{P}(\text{Exit})$, we immediately obtain $M_O \mathbb{P}(\text{Exit}) = M_E$.

Aggregates in the economy can be written as a function of the stationary distribution. Output Y, total fixed costs Φ_F, and total sunk costs Φ_E satisfy the equations

$$Y = M_O \int y(z, W) dF_O(z) \tag{8}$$

$$\Phi_F = M_O \int \int_{\{\phi_F \leq \phi_F^*(z)\}} \phi_F dG(\phi_F) dF_O(z) \tag{9}$$

$$\Phi_E = M_E \phi_E \tag{10}$$

which together imply the level of consumption C via the aggregate resource constraint

$$Y = C + \Phi_F + \Phi_E. \tag{11}$$

Total labor demand N is given by

$$N = M_O \int n(z, W) dF_O(z), \tag{12}$$

which, of course, must equal exogenous labor supply if markets clear

$$N = \bar{N}. \tag{13}$$

4.6 General Equilibrium

Stationary general equilibrium in this economy is a value function $V(z)$, exit thresholds $\phi_F^*(z)$, a stationary distribution $F_O(z)$ of operating firms, an operating mass M_O, an entrant mass M_E, aggregate output Y, aggregate fixed operating costs Φ_F, aggregate sunk entry costs Φ_E, aggregate consumption C, aggregate labor demand N, a wage W, and an interest rate R such that operating firms' optimal value satisfies (1), exit thresholds are optimal according to (2), the stationary distribution replicates itself according to (6), the operating mass is proportional to entry via (7), free entry holds in (3), the aggregate production and resource constraints in (8), (9), (10), and (11) hold, the labor market clears with demand in (12) equal to supply via (13), and household intertemporal optimality holds in (5).

5 Calibration and Solution

In this section, we lay out our approach to calibrating and solving the quantitative general equilibrium framework described in Section 4. We consider two versions of the same model. The first case is based on our empirical nonparametric estimates from Section 2, while the second case employs the standard Gaussian AR(1) parametric assumptions adopted in the literature.

5.1 Calibration

We calibrate the model at annual frequency. As an initial step, we first externally calibrate four parameters shared by both versions of our model. In particular, we choose $\alpha = 0.67$ to match a labor share of two thirds. The value $\beta = 1/1.04$ is picked to generate a yearly net real interest rate $R - 1$ of 4%. We also normalize $\bar{N}$ to the mean employment rate of 59.7% in Spain during our sample period. And, finally, we set the exogenous exit rate δ equal to the 3.9% exit rate observed among the largest firms in our empirical sample, another normalization.

In both versions of our model, we discipline our calibration of profitability shocks z using empirical evidence on firm revenue. In fact, by inverting a firm's static labor demand from the optimization problem in (1), we obtain and employ the simple formula

$$\log z = (1 - \alpha)y + \text{Constant}, \tag{14}$$

which allows us to obtain z directly, up to a normalizing aggregate constant, from observed firm revenue y.

Nonparametric Empirical Calibration In Section 2, we nonparametrically estimated three key objects: the empirical revenue transition distribution for continuing incumbents $H(y'|y)$, the entrant revenue distribution $H_E(y)$, and the revenue exit hazard $\mathbb{P}(\text{Exit}|y)$. Inverting firm profitability z from revenue y via equation (14) directly yields equivalent empirical estimates as functions of z, which we label $H(z'|z)$, $H_E(z)$, and $\mathbb{P}(\text{Exit}|z)$.

Helpfully, the profitability transition and entrant profitability distributions are primitive exogenous objects. Hence, to empirically calibrate the nonparametric version of our model we simply set $F(z'|z) = H(z'|z)$ and $F_E(z) = H_E(z)$, i.e., we can

Table 1: Model Calibration

	Value	Empirical Target
Panel A: Nonparametric Case		
Profitability transition, $F(z'\|z)$	-	$H(z'\|z)$
Entrant distribution, $F_E(z)$	-	$H_E(z)$
Fixed cost distribution, $G(\phi_F)$	-	$\mathbb{P}(\text{Exit}\|z)$
Sunk entry cost, ϕ_E	22.9	Employees per firm, 12.3
Panel B: Parametric AR(1) Case		
Profitability persistence, ρ	0.94	Profitability autocorr., 0.94
Profitability volatility, σ	0.19	Profitability st. dev., 0.56
Entrant profitability mean, μ_E	-0.43	Mean entrant vs operating $\log z$, -0.36
Fixed cost support, $\bar{\phi}_F$	2.30	Exit rate $\mathbb{P}(\text{Exit})$, 6.9%
Sunk entry cost, ϕ_E	5.18	Employees per firm, 12.3

Notes: Panel A of the table lists internally calibrated model objects, their calibrated values where relevant, and the associated empirical targets for the nonparametric version of the model, while Panel B reports the same information for the parametric version of the model. All empirical targets come from our baseline Spanish ORBIS sample covering around five million firm years for around one million firms over the 2005-2014 period. In Panel A, dash placeholders are used to denote three distributions pinned down nonparametrically, as discussed in the main text, to exactly match the indicated empirical targets.

directly choose the model distributions to exactly replicate their empirical equivalents. Calibrating the exogenous fixed cost distribution, by contrast, requires more care, since exit is endogenous in the model. To do so, we exploit the theoretical identity

$$\mathbb{P}(\text{Exit}|z) = 1 - (1-\delta)G(\phi_F^*(z)) \tag{15}$$

linking the endogenous, but observable, exit hazard $\mathbb{P}(\text{Exit}|z)$ to the exogenous, but unknown, fixed cost distribution $G(\phi_F)$. We observe that the exit thresholds $\phi_F^*(z)$ can be determined straightforwardly as a function of the firm value function $V(z)$ through equation (2). Therefore, taking the value function $V(z)$ and the exit thresholds $\phi_F^*(z)$ as given, the identity in equation (15) directly and nonparametrically implies a unique fixed cost distribution $G(\phi_F)$ which is precisely consistent with the observed exit hazard $\mathbb{P}(\text{Exit}|z)$.

Of course, we do not observe $V(z)$ ex ante. However, our solution algorithm for the nonparametric version of the model, summarized below, employs conventional dynamic programming or value function iteration to solve the Bellman equation (1). Within each step of this iteration, we employ our ongoing updated guesses for the value function $V(z)$, and hence continuation values $\phi_F^*(z)$, to compute ongoing updated guesses for the fixed cost distribution $G(\phi_F)$. Convergence of $V(z)$ then delivers convergence of the fixed cost distribution $G(\phi_F)$.

Taken as a whole, our approach to calibration of this version of the model delivers nonparametric distributions $F(z'|z)$, $F_E(z)$, and $G(\phi_F)$ that allow us to perfectly replicate both our empirical estimates of $H(z'|z)$, $H_E(z)$, and $\mathbb{P}(\text{Exit}|z)$, as well as their revenue-indexed versions $H(y'|y)$, $H_E(y)$, and $\mathbb{P}(\text{Exit}|y)$. We emphasize that since the empirical results in Section 2 were computed as functions of these empirical targets, our calibrated model also, *by construction*, matches all of the nonparametric results we presented in that section, including observed revenue dynamics, revenue mobility, the stationary distributions of current and lifetime revenue, exit hazards, and exit rates.

One parameter, the sunk entry cost ϕ_E, remains to be calibrated. Note that higher levels of the sunk cost ϕ_E cause, via the free entry condition (3), an increase in mean entry values, requiring lower equilibrium wages W and driving up mean employment per firm. We therefore calibrate ϕ_E, jointly with the distributions above, to exactly match the mean ratio of employees per firm in our baseline Spanish dataset. Panel

A of Table 1 summarizes the results of our internal calibration of the nonparametric version of the model.

Parametric AR(1) Calibration The calibration of our parametric model is more conventional and relies on various distributional assumptions. We assume that the transition distribution for profitability z, $F(z'|z)$, is governed by an exogenous Gaussian AR(1) process

$$\log z' = \rho \log z + \sigma \varepsilon', \quad \varepsilon \sim N(0,1).$$

where persistence and volatility satisfy $0 < \rho < 1$ and $\sigma > 0$. For the entrant distribution $F_E(z)$, we also assume that entrants' log profitability $\log z \sim N(\mu_E, \sigma^2)$ is drawn from a Gaussian distribution with mean μ_E. We further assume that the distribution of iid fixed costs $G(\phi_F)$ is uniformly distributed between 0 and an upper bound $\bar{\phi}_F > 0$, so $G(\phi_F) = U(0, \bar{\phi}_F)$.

Our distributional assumptions, together with the sunk entry cost parameter ϕ_E, imply a total of five internally calibrated parameters. Panel B of Table 1 lists the parameters, the resulting calibrated values, and the empirical targets which we exactly match in our calibrated parametric model through a joint procedure. Following standard practice, ρ and σ are disciplined by the observed autocorrelation and variance of profitability z in our sample. The mean difference between the log profitability of entrants versus operating firms varies directly with the entrant profitability mean μ_E. Higher values of the upper bound $\bar{\phi}_F$ for the distribution of the fixed cost generate higher mean fixed costs and, therefore, higher mean exit rates in the model. Finally, as in the nonparametric case, we target the mean number of employees per firm in order to help identify the sunk cost of entry ϕ_E. The results of our calibration in Table 1 are unsurprising, with high persistence of profitability of $\rho = 0.94$, moderately high conditional volatility of $\sigma \approx 20\%$ annually, and a meaningful reduction of $\mu_E = -43\%$ in entrant profitability relative to all operating firms.

Note that both the nonparametric and parametric AR(1) versions of the model offer exact fits to their empirical targets. But the parametric model, as is conventional in the firm dynamics literature, only matches a narrower, selected set of moments implied by the empirical profitability distributions and exit hazards. By contrast, the nonparametrically calibrated model offers an exact fit to all of these distributions at all points of the support. Consequently, the nonparametric version of our model also matches by construction each moment targeted by the parametric version, while simultaneously exploiting far more information from our empirical dataset.

5.2 Solution

To numerically solve both versions of the model, we employ conventional dynamic programming methods, i.e., value function iteration. We approximately solve the key operating firm Bellman equation (1) over a continuous state space for profitability z in an "inner loop." We embed this firm-level solution inside a general equilibrium "outer loop" over the wage W and entry mass M_E in order to satisfy the free entry and labor market clearing conditions (3) and (13). At a high level, this approach is quite standard within the quantitative firm dynamics literature.

Successfully solving the nonparametric version of our model in a manner fully consistent with our empirical targets requires a few novel ingredients beyond the standard approach, however. First, we lightly regularize the empirical transition distribution $H(z'|z)$, imposing that the z process exhibits persistence in a first-order stochastic dominance sense. Second, we also regularize the exit hazard $\mathbb{P}(\text{Exit}|z)$, imposing that the hazard is nonincreasing in profitability z. The first condition improves the stability of our value function iteration algorithm, while the second condition ensures that our recovered fixed cost distribution $G(\phi_F)$ is in fact nondecreasing. Fortunately, as we document in Appendix Section B.1, the raw empirical objects quite nearly satisfy both conditions, resulting in only extremely light adjustments in practice. Finally, as mentioned in our calibration discussion above, we must nonparametrically recover updated guesses for the fixed cost distribution $G(\phi_F)$ which are consistent with observed exit hazards within each step of our value function iteration algorithm. We defer further technical information on our solution techniques for both versions of the model to a detailed discussion in Appendix B.

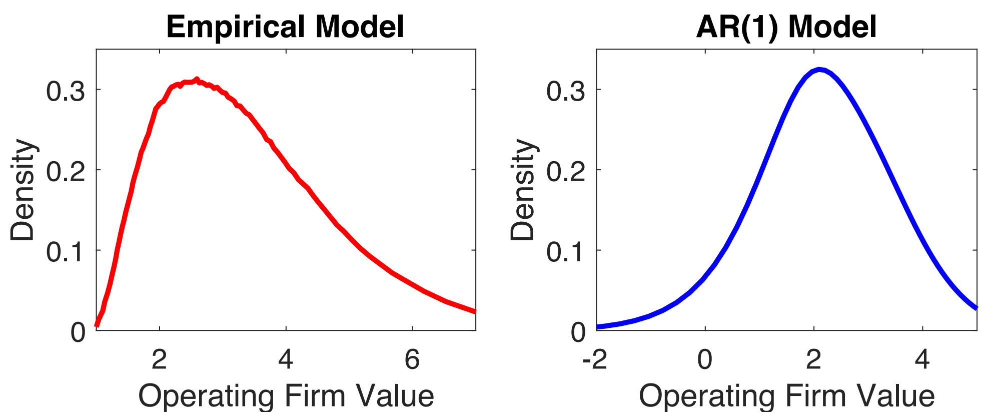
6 Inspecting the Mechanism: Empirical vs AR(1)

Our goal in this section is to assess the quantitative relevance of our empirical findings for the predictions of macroeconomic models featuring heterogeneous firms. To do so, we consider the impact of two simple experiments: (i) a subsidy s_F for all operating firms, and (ii) a subsidy s_E to entrant firms. These experiments conveniently mirror the changes in the fixed operating cost and the sunk entry cost theoretically studied in Hopenhayn (1992). In each experiment, we compute and analyze the quantitative response of macroeconomic aggregates, taking into account general equilibrium. Note that in this model with perfectly competitive output and labor markets, our focus is on descriptively analyzing the impacts of each experiment rather than on normative questions.

6.1 The Model-Implied Distribution of Firm Values

Using our simple analytical model, we argued earlier that the clustering of the distribution at low firm values, where exit occurs, causes a higher sensitivity of the exit rate to economic changes.

Figure 6: Empirical vs AR(1) Firm Value Distributions



Notes: The figure plots the stationary distribution of operating firm continuation values ϕ_F^*, in logs, in the calibrated nonparametric (left panel) and parametric AR(1) (right panel) models.

Figure 6 plots the stationary distribution of operating firm value in our quantitative model in both the calibrated empirical (left panel) and parametric AR(1) (right panel) cases. Recall that the empirical model matches, by construction, the more clustered distribution of lifetime revenue in the data in Figure 4. We see from Figure 6 that our earlier intuition about lifetime revenue carries over to the underlying firm value functions.[10] The distribution of firm value is indeed more clustered at the low

[10]In addition to the fact that Figure 4 and Figure 6 plot different objects conceptually, i.e., lifetime revenue vs firm value distributions, one additional technical detail differentiates the two figures. The

end, with higher densities towards the left of the distribution where exit is most likely.

To demonstrate the quantitative relevance of the firm values distribution, we proceed with our two simple experiments.

6.2 Subsidy to Operating Firms

In the first experiment, the subsidy s_F to operating firms effectively lowers their net fixed operating costs. We distribute the subsidy, denominated in numeraire output, to all operating firms each period, financing the subsidy through lump sum taxes on households. Figure 7 plots the response of various aggregates to the subsidy, for both the empirical nonparametric (red lines) and parametric AR(1) (blue lines) models. We find that while the response of the exit rate is significantly more pronounced under the nonparametric case, the response of output is smaller. In what follows, we provide an overview of the economic forces behind these responses.

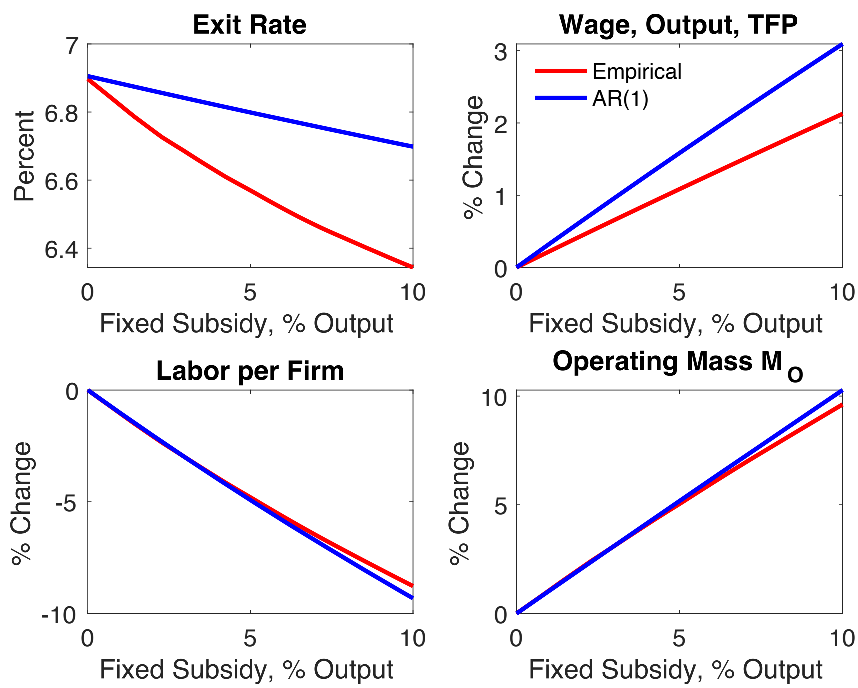
Exit and Selection The subsidy mechanically raises the value of every operating firm, which leads to a decline in the exit rate in both models (top left panel). Our intuition from Figure 6 suggests that the exit rate should be more sensitive to the subsidy under the empirical model, since firm values are more clustered where exit is more likely. Figure 7 confirms that this intuition holds in our full quantitative model. Specifically, with a subsidy of 5% of output, the exit rate decline in the empirical model is three times as large as in the parametric AR(1) model. The fall in exit in turn triggers a negative selection effect, as lower-z firms now survive with higher probability. Given the higher decline in the exit rate, this negative selection effect is more pronounced in the empirical version, partly counteracting the direct increase in mean firm value due to the subsidy.

Wages The equilibrium wage W in the top right panel of Figure 7 rises to offset the increase in the expected value of entry due to the subsidy, ensuring that the free entry condition (3) continues to hold.[11] While this logic holds in both models, the

empirical vs AR(1) lifetime revenue distributions in Figure 4 both rely on the empirical exit hazard and entry distributions, while varying only the incumbent revenue transition distributions. In Figure 6, with the parametric model's structural exit hazards and entry distributions already defined and in hand, we also vary the entry and exit patterns.

[11] Recall that firms enter based on an expected continuation value: only after entry do they learn their profitability level, produce and then choose whether to exit. For this reason, there is no

Figure 7: Impact of the Fixed Operating Subsidy s_F



Notes: Each panel in the figure plots an aggregate outcome as the subsidy is increased from zero for the calibrated nonparametric empirical (red lines) and parametric AR(1) (blue lines) models. For comparability, the horizontal axis in each panel is the aggregate size of the subsidy as a percentage of zero-subsidy aggregate output in each economy. The vertical axes plot either the levels of the outcomes or, where natural, percent changes from the zero-subsidy level.

size of the response is different. Because a stronger decline in the exit rate in the empirical model generates a stronger negative selection effect, average firm value after entry rises less as a direct result of the subsidy. As a result, the equilibrium wage increase required by the free entry condition is smaller in the empirical model.

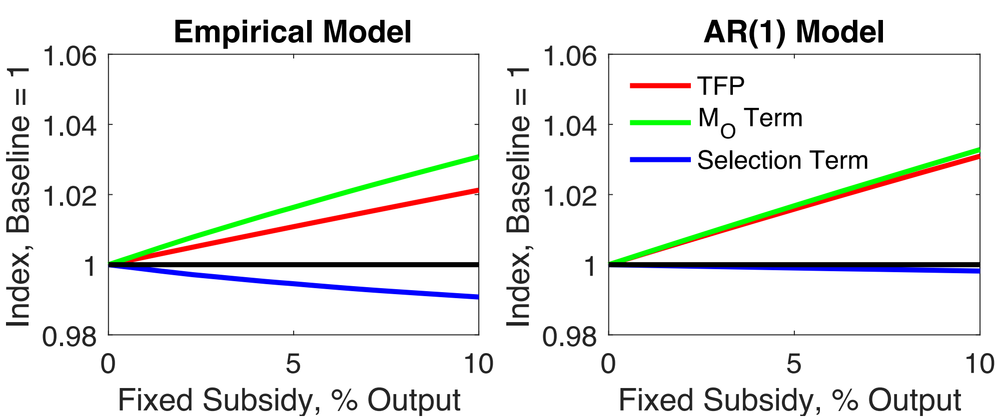
Labor and the Mass of Firms Recall that total labor in the economy is in fixed supply $\bar{N}$. The labor market clearing condition

$$N = M_O \int n(z, W) dF_O(z) = \bar{N} \tag{16}$$

selection through entry, and thus the sunk entry costs must always equal the average firm value across the entrant profitability distribution in the free entry condition (3).

equalizes total demand N to this fixed supply. But, as a result, the mass of operating firms, M_O, must move inversely to the average labor per firm, i.e., $\int n(z, W) dF_O(z)$. Next, note that average labor per firm is pinned down by two margins: the size of operating firms $n(z, W)$, which is governed by the wage; and the selection of operating firms through the distribution $F_O(z)$. The operating subsidy acts on both. Higher wages reduce firm size $n(z, W)$ for a given z (a stronger force in the AR(1) model), while negative selection in $F_O(z)$ lowers the mean of z across firms (a stronger force in the empirical model). Because of the offsetting strengths of these two channels, the overall decline in average labor per firm in the bottom left panel of Figure 7 turns out to be very similar across in the empirical and AR(1) models. Now, since labor per firm declines, the mass of operating firms M_O must rise to restore labor market clearing. But because the decline in average labor per firm is comparable across models, the rise in the mass of operating firms in the bottom right panel of Figure 7 is also similar.

Figure 8: Decomposed TFP under the Fixed Operating Subsidy s_F



Notes: The figure plots each component of the TFP decomposition in (17) as the subsidy is increased from zero for the calibrated nonparametric empirical (left panel) and parametric AR(1) (right panel) models. For comparability, the horizontal axis in each panel is the aggregate size of the subsidy as a percentage of zero-subsidy aggregate output in each economy. The vertical axes index each component to one in the zero-subsidy baseline. The black line depicts the no-change line.

TFP and Output In this economy with fixed labor supply, the aggregate levels of the wage W, output Y, and measured $TFP = Y/N^\alpha$ are proportional to one another.

As a result, the percentage changes in output and TFP under the subsidy exactly match those plotted in the top right panel of Figure 7 for the wage. We conclude that output and TFP in this economy respond more in the AR(1) model than in the empirical version. Specifically, we note that with a subsidy of 5% of output, the percent change in these outcomes is only two thirds as strong in the empirical model as in the parametric AR(1). Our discussion above suggested that this difference is the result of a strong negative selection force generated by the sharp decline in exit in the empirical economy. To highlight this point, we decompose measured aggregate TFP in this economy into two margins

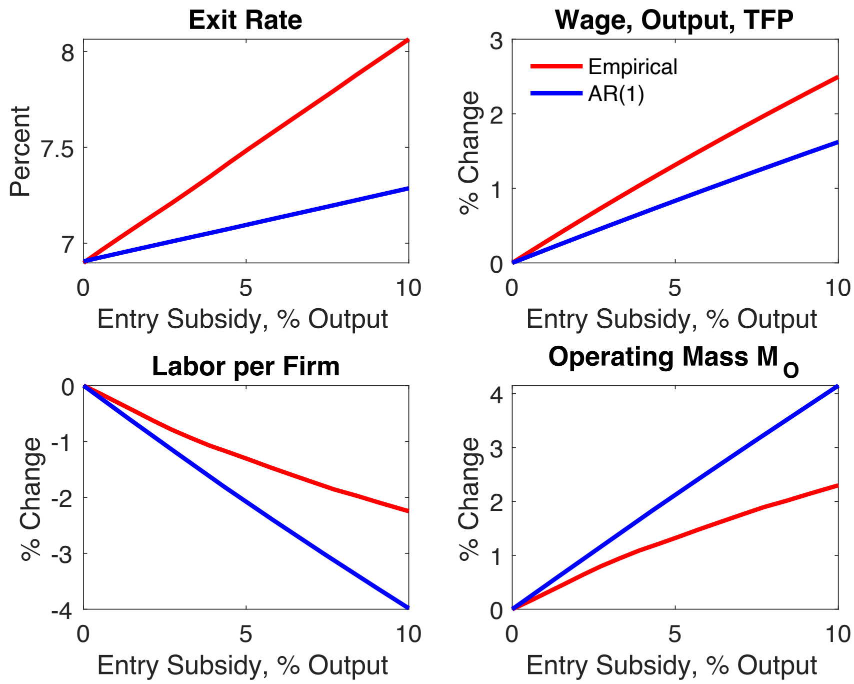
$$TFP = \underbrace{M_O^{1-\alpha}}_{\text{Operating Mass of Firms}} \underbrace{\left(\int z^{\frac{1}{1-\alpha}} dF_O(z)\right)^{1-\alpha}}_{\text{Selection}}. \tag{17}$$

The first term increases with the operating mass of firms through a standard extensive margin effect under decreasing returns. The second term is a geometric mean of operating profitability, i.e., a measure of firm selection. Figure 8 plots the respective contributions of these two components from equation (17) to the change in TFP in the empirical (left panel) and AR(1) (right panel) models. Our decomposition confirms that the negative contribution from selection is indeed more pronounced in the empirical model, due to the larger fall in the exit rate. This selection margin entirely explains the more muted response of TFP (and output) relative to the AR(1) case.

6.3 Subsidy to Entrants

In our next experiment, the subsidy s_E is given only to entrants, lowering their net entry costs. We again finance this output-denominated subsidy with a simple lump sum tax on households. As in the case of the operating subsidy, we find that the exit rate is more sensitive in the empirical model. This time, however, we note that the response of output is significantly larger than under the AR(1) specification. As in our first experiment, we conclude that a selection effect driven by shifts in the exit rate is key to understanding differences in the response across our two model versions.

Figure 9: Impact of the Entry Subsidy s_E

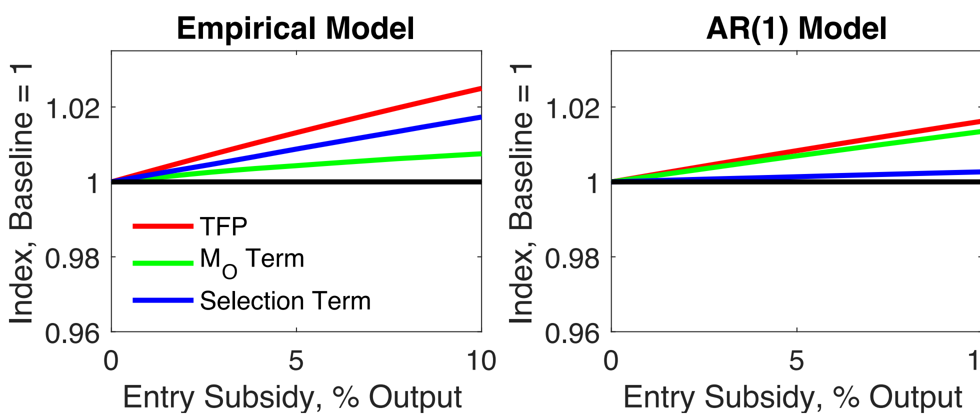


Notes: Each panel in the figure plots an aggregate outcome as the subsidy is increased from zero for the calibrated nonparametric empirical (red lines) and parametric AR(1) (blue lines) models. For comparability, the horizontal axis in each panel is the aggregate size of the subsidy as a percentage of zero-subsidy aggregate output in each economy. The vertical axes plot either the levels of the outcomes or, where natural, percent changes from the zero-subsidy level.

Wages, Exit, and Selection Our entry subsidy lowers the cost of entry to $\phi_E - s_E$ in the free entry condition (3). In order to restore equilibrium, the wage must increase in both versions of our model to reduce the post-entry expected value of operating. Wage increases trigger exit among low-profitability firms and hence generate a positive selection effect. However, due to the shape of the value distributions in Figure 6 with more clustering of firm value at low levels, we see in the top left panel of Figure 9 that the exit rate increases more sharply in the empirical versus the AR(1) model. Specifically, with a subsidy of 5% of output, the exit rate increase in the empirical model is three times as large as in the parametric AR(1) model. To offset the resultingly stronger selection effect, the wage rises by more in the empirical model in the top right panel of Figure 9.

Labor and the Mass of Firms A higher wage following the subsidy directly drives down labor demand $n(z, W)$ at individual firms, conditional upon profitability z. Yet the selection channel, driving more low-z firms to exit following the subsidy, has the opposite effect through indirect changes in the distribution of operating firms. Ultimately, the net effect on labor per firm is ambiguous. In the bottom left of Figure 9 we see that the direct effect dominates, since average labor demand per operating firm falls in both models. But the nonparametric model features a stonger indirect selection effect, generating a smaller overall fall in labor per firm in this case. Finally, recall from the labor market clearing condition in equation (16) that labor per firm and the mass of operating firms must move inversely due to the fixed total labor supply. As a result, we see in the bottom right panel of Figure 9 that the shift in the mass of operating firms M_O is also smaller in the empirical than in the AR(1) model.

Figure 10: Decomposed TFP under the Entry Subsidy s_E



Notes: The figure plots each component of the TFP decomposition in (17) as the subsidy is increased from zero for the calibrated nonparametric empirical (left panel) and parametric AR(1) (right panel) models. For comparability, the horizontal axis in each panel is the aggregate size of the subsidy as a percentage of zero-subsidy aggregate output in each economy. The vertical axes index each component to one in the zero-subsidy baseline.

TFP Recall that in this economy, measured TFP and aggregate output are both proportional to the wage and therefore rise following the subsidy (top right panel of Figure 9). In Figure 10, we rely on equation (17) to decompose the rise in TFP into separate contributions from the operating mass and selection effects for the empirical (left panel) and AR(1) (right panel) models. The sharper response of exit in the empirical model generates a stronger positive selection effect. The operating mass of firms, however, rises more in the parametric case. On net, the selection channel induced by exit is stronger, underlying the larger impact of the subsidy on TFP, and output, in our empirical nonparametric model. Specifically, with a subsidy of 5% of output, the increase in TFP and output in the empirical model is one and a half times as large as in the parametric AR(1) model.

6.4 Taking Stock

Our analysis across both subsidy experiments highlights the quantitative importance of the shape of firm value distribution in driving aggregate responses in a canonical general equilibrium model of firm dynamics. Specifically, our nonparametric model matching the more pronounced empirical clustering of the distribution of lifetime revenue at lower levels where exit is more likely to occur in Figure 4 also features a more clustered distribution of underlying firm value in Figure 6 relative to the standard AR(1) case. As a result, exit rates and hence selection shift more strongly in our nonparametric model, causing a large difference in the quantitative predictions of the two models. We conclude that embedding a shock process that adequately matches the rich distributional dynamics found in the data proves to be crucial for quantitative work with this class of models.

7 Discussion, Extensions, and Robustness

In this section, we discuss a range of additional robustness checks and extensions to our empirical and quantitative analysis. We frame our discussion around a number of natural and sensible questions. In each case, our empirical approach or quantitative conclusions prove to be robust.

Do country, data treatment, time period, or industry composition drive our results? Our analysis naturally involves many decisions regarding data treatment and sample construction. Do these choices drive our conclusions? We investigate whether these choices drive our conclusions using a combination of empirical but also quantitative model robustness checks.

Focusing first on the empirical moments that are key to our findings, Appendix Table A.1 reports a number of statistics on revenue and revenue growth moments in our baseline ORBIS dataset as well as in a number of robustness checks. We first consider the role of the sample period by splitting our dataset into pre- and post-2009 samples. Second, we divide our broad representative dataset into manufacturing and non-manufacturing subsamples. Third, we use unconsolidated firm-level accounts instead of the consolidated statements from our baseline. Fourth, we exclude firm-years with reported M&A activity. Fifth, we demean log revenue by year only, rather than our baseline year and sector demeaning. Sixth, we consider different treatments of outliers relative to our baseline baseline trimming of 0.1% of revenue outliers. Seventh, we consider data from Italy, Portugal, France, and Norway instead of our Spain baseline. To judge the results in Appendix Table A.1, recall a key fact from Section 2: revenue growth in Figure 2 is leptokurtic or fat-tailed. Appendix Table A.1 reveals extremely high baseline revenue growth kurtosis of about 30, compared to exactly 3 in any Gaussian case. Uniformly, we find fat-tailed revenue growth in all of our robustness checks.

Next, for each of these alternative data samples, we perform a full recalibration of our empirical nonparametric and parametric AR(1) models. Appendix Table B.1 lists the recalibrated parameters for all the quantitative model robustness checks. We recompute the changes in the aggregate exit rate and output induced by an operating subsidy s_F totaling 5% of pre-subsidy output. Appendix Table B.2 reports the ratio of these responses in the empirical vs parametric AR(1) models. Our baseline nonparametric model's exit rate response is 3 times as strong as the one in the parametric model, driving a negative selection effect which dampens the output response to only around two-thirds that of the parametric model. The same overall pattern is observed in all of our robustness checks.

Do our exact model assumptions drive our results? Our quantitative model is purposefully conventional within the Hopenhayn (1992) tradition, but we explore our results' robustness to multiple alternative assumptions. First, while we fix aggregate labor in our baseline, as a robustness check we instead consider the case of an endogenous labor supply.[12] Second, the parameter α, which plays an important

[12]The extension is straightforward. We replace household log consumption preferences $\log C$ with log-linear utility $\log C - \omega N$ for some $\omega > 0$. Then, we replace the labor market clearing condi-

role in the TFP decomposition (17), has multiple interpretations. A literal view links α to the labor share, rationalizing our baseline external calibration $\alpha = 2/3$. But a revenue function view of our production technology under imperfect competition links α to production and demand elasticities. We therefore entertain values of α of 0.6 and 0.75. After performing model recalibrations and counterfactual analyses for each scenario, Appendix Table B.2 reports the relative impacts of an operating subsidy in the nonparametric vs parametric models. Our conclusions are little changed from baseline.

Does firm age drive our results? Our baseline analysis, like much work following Hopenhayn (1992), features no separate role for firm age conditional upon size in predicting growth or exit. Yet, many papers rationalize related evidence, recently documented authoritatively by Sterk et al. (2021), with mechanisms such as learning (Jovanovic, 1982; Arkolakis et al., 2018), demand accumulation (Foster et al., 2008; Gourio and Rudanko, 2014; Moreira, 2018), or financial frictions (Moll, 2014). This rich firm age literature is complementary to, but quite distinct from, our analysis contrasting nonparametric versus parametric approaches. Nevertheless, we observe firm demographic data in ORBIS and can calculate firm age. We therefore conduct another robustness check by residualizing revenue against firm age – with a full set of age indicators denominated in years – in addition to our baseline demeaning by sector and fiscal year. The moments for this alternative dataset, shown in Appendix Table A.1, reveal that revenue growth remains strongly fat-tailed or leptokurtic, i.e., features that are incompatible with a Gaussian AR(1). Finally, Appendix Table B.2 reports that once the model is calibrated, solved and simulated based on the dataset controlling for age, the relative impact of the operating subsidy across the two models is in line with that in our baseline for firm exit and in fact stronger in the case of firm output.

Do firm heterogeneity and noise drive our results? Following the standard specification adopted in the literature, we naturally contrast our nonparametric empirics to the predictions from a Gaussian AR(1) parametric benchmark. One might

tion with the household intratemporal optimality condition $W = \omega C$. Otherwise, the equilibrium structure remains unchanged. We calibrate $\omega = 1.51$ (empirical case) and $\omega = 1.39$ to match labor supply N to the Spanish employment rate.

however wonder whether a richer extension of our AR(1) that includes permanent firm fixed effects and transitory shocks, ingredients ubiquitous in household incomplete markets analyses, might allow us to match the predictions of our nonparametric model. In Appendix Section A.1.1, we therefore specify and estimate an extended parametric Gaussian AR(1) model augmented with a Pareto distribution of firm fixed effects as well as Gaussian transitory shocks. We then subject all our models to a battery of tests gauging their predictive accuracy for both the mean and full distribution of observed revenue dynamics. Appendix Table A.2 reports our extended model's estimates and shows that the extended model, while still failing to capture the fat-tailed nature of revenue growth observed in the data, does predict firm revenue somewhat better than our benchmark Gaussian AR(1). Ultimately, however, even this richer parametric model remains less accurate for prediction than our nonparametric structure, giving us some assurance that our comparison of parametric versus nonparametric approaches is not unduly driven by our choice of parametric benchmark.

Can we empirically link lifetime revenue and market value? Outcomes summarizing a firm's lifetime prospects such as firm value are not typically available for unlisted private firms. For this reason, in Section 2 we proposed a new measure, lifetime revenue, defined as the expected presented discounted value of firm revenue. This proxy for firm value can be constructed simply using information on revenue and exit alone. We found earlier that the distributions of data-driven lifetime revenues (Figure 4) and model-implied firm values (Figure 6) both display similar clustering, providing support for our proxy. But for the small subset of publicly listed Spanish firms in our sample, we can push further empirically without directly relying on our structural model. Regressions in Appendix Table A.3 reveal that, while both current revenue and lifetime revenue are highly correlated with observed market value, current revenue loses its predictive power for firm value once we account for lifetime revenue. This result confirms that our lifetime revenue measure does in fact capture useful variation in a firm's long-term prospects, as captured by realized market value.

Can we empirically link clustered distributions and exit rate sensitivity? A reader might accept our empirical evidence of fat-tailed revenue dynamics and clustered lifetime revenue outcomes for firms but still harbor two natural objections.

First, we do not provide a formal definition of distributional clustering in our analysis above. Second, we use our quantitative model, inevitably laden with assumptions, rather than a more direct empirical approach to link our intuitive notion of clustering to higher exit rate sensitivity.

In Appendix Section A.3 we push further in both directions. To begin, we develop a reduced-form, purely statistical model that allows us to predict the aggregate exit rate based on the distribution of firm lifetime revenue. Within this framework, we analytically derive the predicted local response of the exit rate to a hypothetical one-time revenue windfall for all firms. This derivative, which is directly computable in our data, has a natural interpretation as a clustering statistic. This clustering statistic is higher when the lifetime revenue distribution has, on average, higher density in regions with steeper exit hazards. Empirically, we compute and report the value of the statistic in Appendix Table A.4 for each two-digit sector within our sample. Clustering varies widely, with particularly high values in sectors including construction and retail trade and particularly low values in sectors including finance and health care. We then exploit this cross-sectoral heterogeneity by running a set of panel regressions, whose results are presented in Appendix Table A.5, demonstrating that exit rates covary more negatively with sales growth at the industry level in the presence of higher clustering. Our quantitative model, of course, does not incorporate sectoral shocks or clustering heterogeneity. Yet we view these empirical results as consistent with our quantitative model's central prediction, which links the high sensitivity of the exit rate in the nonparametric specification to the high level of clustering of the firm value distribution.

8 Conclusion

In this paper, we argue that the standard parametric assumption for firm-level shocks – a Gaussian AR(1) process – used in the heterogeneous firms literature is not realistic. In particular, we find that nonparametrically solving a model consistent with the firm-level revenue dynamics we observe in the data, i.e., fat-tailed growth and high mobility from the tails, has a large impact on the behavior of a canonical firm dynamics model at the macro level. The standard parametric model implies a firm value distribution which is far too dispersed relative to the firm value distribution consistent with empirical firm dynamics. As a result, the empirical, nonparametric model's more clustered value distribution generates substantially higher sensitivity of the exit rate to a set of standard policy experiments. The stronger extensive margin reaction in our nonparametric model drives strong selection effects serving to amplify or dampen the response of aggregate output, depending upon the exact details of the underlying policy. As a result, we conclude that the standard parametric assumptions adopted in the quantitative firm dynamics literature are far from innocuous but instead directly change the macro implications of firm-level mechanisms.

References

Arkolakis, C., T. Papageorgiou, and O. A. Timoshenko (2018). Firm Learning and Growth. *Review of Economic Dynamics 27*, 146–168.

Asker, J., A. Collard-Wexler, and J. De Loecker (2014). Dynamic Inputs and Resource (Mis)Allocation. *Journal of Political Economy 122*(5), 1013–1063.

Bajgar, M., G. Berlingieri, S. Calligaris, C. Criscuolo, and J. Timmis (2020). Coverage and Representativeness of Orbis Data. Working paper.

Barro, R. J. and J. F. Ursúa (2012). Rare Macroeconomic Disasters. *Annual Review of Economics 4*, 83–109.

Bento, P. and D. Restuccia (2017). Misallocation, Establishment Size, and Productivity. *American Economic Journal: Macroeconomics 9*(3), 267–303.

Bilbiie, F. O., F. Ghironi, and M. J. Melitz (2012). Endogenous Entry, Product Variety, and Business Cycles. *Journal of Political Economy 120*(2), 304–345.

Bils, M., P. J. Klenow, and C. Ruane (2021). Misallocation or Mismeasurement? *Journal of Monetary Economics 124*, S39–S56.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018). Really uncertain business cycles. *Econometrica 86*(3), 1031–1065.

Boar, C., D. Gorea, and V. Midrigan (2023). Why are Returns to Private Business Wealth So Dispersed? Working paper.

Carvalho, V. M. and B. Grassi (2019). Large Firm Dynamics and the Business Cycle. *American Economic Review 109*(4), 1375–1425.

Catherine, S., T. Chaney, Z. Huang, D. Sraer, and D. Thesmar (2018). Quantifying Reduced-Form Evidence on Collateral Constraints. Forthcoming, *Journal of Finance*.

Clementi, G. L. and B. Palazzo (2016). Entry, Exit, Firm Dynamics, and Aggregate Fluctuations. *American Economic Journal: Macroeconomics 8*(3), 1–41.

Crouzet, N. and N. R. Mehrotra (2020). Small and Large Firms over the Business Cycle. *American Economic Review 110*(11), 3549–3601.

Davies, R. B. and C. Eckel (2010). Tax Competition for Heterogeneous Firms with Endogenous Entry. *American Economic Journal: Economic Policy 2*(1), 77–102.

Davis, S. J. and J. Haltiwanger (1992). Gross Job Creation, Gross Job Destruction, and Employment Reallocation. *Quarterly Journal of Economics 107*(3), 819–863.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2016). Declining business dynamism: What we know and the way forward. *American Economic Review 106*(5), 203–07.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2020). Changing business dynamism and productivity: Shocks versus responsiveness. *American Economic Review 110*(12), 3952–90.

Dunne, T., M. J. Roberts, and L. Samuelson (1989). The Growth and Failure of US Manufacturing Plants. *Quarterly Journal of Economics 104*(4), 671–698.

Forneron, J.-J. (2020). A Sieve-SMM Estimator for Dynamic Models. Working paper.

Foster, L., J. Haltiwanger, and C. Syverson (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review 98*(1), 394–425.

Gabaix, X. (2016). Power laws in economics: An introduction. *Journal of Economic Perspectives 30*(1), 185–206.

Garicano, L., C. Lelarge, and J. Van Reenen (2016). Firm Size Distortions and the Productivity Distribution: Evidence from France. *American Economic Review 106*, 3439–3479.

Gourio, F. and J. Miao (2010). Firm Heterogeneity and the Long-Run Effects of Dividend Tax Reform. *American Economic Journal: Macroeconomics 2*(1), 131–68.

Gourio, F. and L. Rudanko (2014). Customer Capital. *Review of Economic Studies 81*(3), 1102–1136.

Guner, N., G. Ventura, and Y. Xu (2008). Macroeconomic Implications of Size-dependent Policies. *Review of Economic Dynamics 11*(4), 721–744.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). What Do Data on Millions of US Workers Reveal About Lifecycle Earnings Dynamics? *Econometrica 89*(5), 2303–2339.

Guvenen, F., A. McKay, and C. Ryan (2023). A Tractable Income Process for Business Cycle Analysis. Working paper.

Hopenhayn, H. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica 60*(5), 1127–1150.

Hopenhayn, H. and R. Rogerson (1993). Job Turnover and Policy Evaluation: A General Equilibrium Analysis. *Journal of Political Economy 101*(5), 915–938.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics 124*(4), 1403–1448.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica 50*(3), 649–670.

Kalemli-Ozcan, S., B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas (2022). How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications. Working paper.

Karahan, F., B. Pugsley, and A. Sahin (2022). Demographic Origins of the Startup Deficit. Working paper.

Kehrig, M. (2015). The Cyclical Nature of the Productivity Distribution. Working paper.

Kehrig, M. and N. Vincent (2020). Good Dispersion, Bad Dispersion. Working paper.

Khan, A. and J. K. Thomas (2008). Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics. *Econometrica 76*(2), 395–436.

Khan, A. and J. K. Thomas (2013). Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity. *Journal of Political Economy 121*(6), 1055–1107.

Lee, Y. and T. Mukoyama (2018). A Model of Entry, Exit, and Plant-level Dynamics Over the Business Cycle. *Journal of Economic Dynamics and Control 96*, 1–25.

Melitz, M. J. (2003). The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.
Midrigan, V. (2011). Menu Costs, Multiproduct Firms, and Aggregate Fluctuations. *Econometrica 79*(4), 1139–1180.
Midrigan, V. and D. Y. Xu (2014). Finance and Misallocation: Evidence from Plant-level Data. *American Economic Review 104*(2), 422–58.
Moll, B. (2014). Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation? *American Economic Review 104*(10), 3186–3221.
Moreira, S. (2018). Firm dynamics, persistent effects of entry conditions, and business cycles. Working paper.
Ottonello, P. and T. Winberry (2020). Financial Heterogeneity and the Investment Channel of Monetary Policy. *Econometrica 88*(6), 2473–2502.
Restuccia, D. and R. Rogerson (2008). Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics 11*(4), 707–720.
Sraer, D. and D. Thesmar (2021). A Sufficient Statistics Approach for Aggregating Firm-Level Experiments. Working paper.
Sterk, V., P. Sedláček, and B. Pugsley (2021). The Nature of Firm Growth. *American Economic Review 111*(2), 547–79.
Tauchen, G. (1986). Finite State Markov-chain Approximations to Univariate and Vector Autoregressions. *Economics Letters 20*(2), 177–181.
Winberry, T. (2021). Lumpy Investment, Business Cycles, and Stimulus Policy. *American Economic Review 111*(1), 364–396.