Appendices for Online Publication Only

A Data

We begin with annual ORBIS data on firm financials in Spain from 2005-2019. Since our interest is in firms as a legal concept rather than on, say, physical locations or lines of work, we restrict our sample to consolidated financial statements.

Our revenue measure is the ORBIS variable `opre`, i.e., operating revenue or turnover measured at the firm-year level. This variable, in logs, residualized with respect to 4-digit NAICS industry and year effects, is our baseline revenue measure referred to as y in the text.

We trim our panel of residualized revenue y data at the 0.1% and 99.9% thresholds. We also lightly clean our sample in some other ways to guard against the possibility that observed exit or entry might be driven by missing information in a specific year. We first form candidate indicators for firm entry and exit events based on that firm's data availability in our historical ORBIS panel dataset. If for a given year the firm has data for at least one of the most populated variables (e.g. employment and payroll) but not revenue, then the firm is dropped altogether. Second, we ensure that data "holes" do not generate spurious entry or exit by verifying that the firm is not ever present in the dataset before (after) the candidate entry (exit) year, with the "after" window extending for a buffer of four years.[17]

The benchmark ORBIS sample we construct following the guidelines above results in a panel dataset in Spain with a total of 5,157,769 firm-years for 1,032,098 firms in the 2005-2014 period. In the analysis that follows in this appendix, we provide statistics from various alternative datasets we consider as part of robustness checks to our baseline empirical approach. We also introduce various ancillary empirical results referenced throughout the main text.

A.1 Empirical Robustness Checks

Table A.1 reports moments of residualized log revenue y and revenue growth Δy from our baseline Spanish ORBIS dataset on over one million firms for over five million

[17] For this reason, our effective sample period never goes beyond 2014, the end date quoted in the text. This restriction allows us to verify that a firm does not show up again between 2015 and 2019, since 2019 is the formal end of the ORBIS historical dataset in the ORBIS vintage we used.

Table A.1: Empirical Moments under Alternative Datasets

	Revenue y			Revenue Growth Δy		
	Std dev	Skewness	Kurtosis	Std dev	Skewness	Kurtosis
Baseline	1.548	0.025	4.196	0.656	-0.312	29.212
Before 2009	1.515	0.007	4.182	0.692	-0.071	29.920
After 2009	1.583	0.025	4.224	0.652	-0.482	26.974
Mfg	1.561	0.056	3.862	0.480	-0.829	41.572
Non-Mfg	1.546	0.021	4.250	0.679	-0.285	27.767
Unconsolidated only	1.527	-0.031	4.125	0.655	-0.321	29.160
No M&A	1.545	0.028	4.214	0.655	-0.320	28.937
Year Effects Only	1.699	0.131	3.756	0.660	-0.415	28.765
No Trimming	1.641	-0.068	5.088	0.774	-0.774	35.153
1% Trimming	1.416	-0.008	3.336	0.586	-0.201	21.646
Remove Firm Age	1.496	-0.063	4.356	0.664	-0.593	25.882
Italy	1.764	-0.660	6.070	0.956	-0.165	33.889
Portugal	1.514	-0.122	5.077	0.736	0.347	28.683
France	1.379	-0.152	6.483	0.604	0.637	72.380
Norway	1.705	-0.233	4.274	0.681	-0.153	26.698

Notes: This table reports moments of firm revenue y and revenue growth Δy under alternative empirical approaches. The Baseline moments in the top row represent our benchmark ORBIS sample of just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain. In this case, revenue levels y represent log firm revenue demeaned by sector and year, while revenue growth Δy is the first difference of revenue levels. In subsequent rows, we report moments from datasets constructed from ORBIS using different nations, subsamples, time periods, or data treatment approaches.

firm years together with analogous moments for a range of robustness checks and alternative samples described in the main text in Section 7.

A.1.1 An Extended Parametric Model

In a robustness check we consider an extended parametric AR(1) model. First, we recover observed profitability z_{it} for firm i in year t from residualized revenue y_{it} by inverting the labor optimality condition according to (14). In our extended model, we decompose profitability as

$$z_{it} = \mu_i e^{\log \hat{z}_{it} + \nu_{it}}. \tag{18}$$

Table A.2: Extended Parametric AR(1) Model

Panel A: Extended AR(1) Parameters	Symbol	Value
Autocorrelation	ρ	0.9522
Persistent shock variance	σ_ε^2	0.0189
Transitory shock variance	σ_ν^2	0.0150
Pareto fixed effects lower bound	μ_{min}	0.7941
Pareto fixed effects shape	μ_{shape}	4.4176
Panel B: Extended AR(1) Moments	**Data**	**Model**
Autocorrelation, log z	0.9071	0.9071
Variance , Δ log z	0.0480	0.0480
Variance, log z	0.2692	0.2692
Top 1% share, z	0.0408	0.0408
Mean, z	1.1441	1.1441
Panel C: Model Predictive Accuracy	**RMSE**	**LPS**
Nonparametric	1.000	-3.25
Extended AR(1)	1.021	-3.6
Benchmark AR(1)	1.033	-3.7

Notes: Panels A and B reports simulated method of moment estimates and fit for our extended parametric AR(1) model (18). The estimates were computed using a simulated panel of identical size to our benchmark ORBIS data for just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain. As in our baseline model calibration, we recover profitability z from log firm revenue demeaned by sector and year using the labor optimality condition (14). Panel C reports a battery of predictive accuracy tests for $\log z$ – relative root mean squared errors (RMSE) and log predictive scores (LPS) — for our nonparametric model from Section 2.2, our benchmark AR(1), and the extended AR(1) from Panels A and B.

In the equation above μ_i is a firm fixed effect cross-sectionally distributed Pareto with scale parameter μ_{min} and shape parameter μ_{shape}. Inside the exponent, $\log \hat{z}_{it} = \rho \log z_{it-1} + \varepsilon_{it}$ is a Gaussian AR(1) component with $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, and $\nu_{it} \sim N(0, \sigma_\nu^2)$ is an iid transitory shock.

We estimate the model (18) with an exactly identified simulated method of moments strategy. While the identification is joint, roughly speaking, the autocorrelation of profitability disciplines ρ, the variance of both profitability growth and levels discipline shock innovations σ_ε^2 and σ_ν^2, and finally the mean and top 1% share of profitability discipline the Pareto scale and shape parameters μ_{min} and μ_{mean}. Panels A and B in Table A.2 report our point estimates and targeted moments, revealing an exact fit as well as high estimated persistence, conditional volatility close to evenly split between persistent and transitory sources, and a nontrivial distribution of firm heterogeneity.

Panel C of A.2 subjects three models – the empirical or nonparametric model defined in Section 2.2, our benchmark calibrated parametric AR(1), and the extended AR(1) model – to a battery of predictive accuracy tests for log profitability. In the second column we report the root mean squared error (RMSE) of the mean one-year predictions implied by each model, normalizing the nonparametric model's RMSE to 1. This statistic measures the point forecast accuracy of each model, with higher values indicating a poorer performance. In the column "LPS" we report the log predictive score, a measure of a model's predictive accuracy over the full distribution of one-year ahead profitability in which higher values indicate more accurate prediction. Under either measure, the performances of both the baseline and extended parametric models are poor relative to the nonparametric model, although the extended model does improve on the benchmark AR(1)'s performance meaningfully. Quantitatively, the extended model closes only around a third of the accuracy gap with the nonparametric model measured using mean forecast RMSE's and around a quarter of the accuracy gap measured using the broader LPS measure.

A.2 Predicting Market Value with Lifetime Revenue

To examine the predictive content of our lifetime revenue measure $W(y)$ defined in Section 2, over and above current revenue y, we restrict our baseline Spanish ORBIS dataset to a subset of only publicly listed firms. For this subsample, we observe realized market value. We see in Table A.3's regression results in columns 1-3 that contemporaneous revenue is highly correlated with a firm's market value. Yet, our constructed firm lifetime revenue variable is a better predictor of a firm's market value. In particular, once lifetime revenue is included, contemporaneous revenue ceases to be statistically significant. We view these results as validating the empirical relevance of our lifetime revenue measure.

A.3 Industry Clustering and Exit Rates

We develop a framework linking firm exit to our notion of observed lifetime revenue $W(y)$ developed in Section 2. Recall that the stationary distribution $H(y)$ of current revenue y implies a stationary distribution $H(W)$ of lifetime revenue W. Similarly,

Table A.3: Lifetime Revenue, Current Revenue, and Market Value

	Market Value$_{it}$			
	(1)	(2)	(3)	(4)
Revenue$_{it}$	0.284***	0.141***	0.141***	-0.057
	(0.029)	(0.018)	(0.018)	(0.036)
Lifetime Revenue$_{it}$				0.362***
				(0.076)
Fixed Effects	-	Industry	Industry Year	Industry Year
Firm-Years	4273	4273	4273	4273

Notes: The table reports OLS estimates of market value, in logs, for firm i in year t on log revenue and log lifetime revenue. Industry refers to four-digit industry codes. The sample is drawn from the subset of publicly listed firms within our baseline Spanish ORBIS dataset spanning 2005-2014 for both listed and unlisted firms. Unconditionally, the correlation of log revenue and market value is 0.24, and the correlation of log lifetime revenue and log market value is 0.27. Standard errors are clustered at the firm level. Significance is indicated as * = 10% level, ** = 5% level, and *** = 1% level.

the revenue exit hazard $\mathbb{P}(\text{Exit}|y)$ implies a lifetime revenue exit hazard $\mathbb{P}(\text{Exit}|W)$. We rewrite the exit rate as

$$\mathbb{P}(\text{Exit}) = \int \mathbb{P}(\text{Exit}|W)dH(W).$$

In this *purely statistical* model, the new exit rate predicted in partial equilibrium after a windfall increase ϵ in lifetime revenue W is given by

$$\mathbb{P}(\text{Exit}) = \int \mathbb{P}(\text{Exit}|W+\epsilon)dH(W).$$

Thus, the sensitivity of exit to this windfall revenue increase can then be computed as the distributional "clustering statistic" $\mathcal{C}$ given by

$$\mathcal{C} = -\frac{\partial \mathbb{P}(\text{Exit})}{\partial \epsilon}|_{\epsilon=0} = -\int \frac{\partial \mathbb{P}(\text{Exit}|W)}{\partial W}dH(W). \tag{19}$$

Table A.4: Clustering across Sectors

Clustering Statistic $\mathcal{C}$	NAICS Sector
0.091	Construction, 23
0.0543	Real Estate, 53
0.0492	Professional Technical Services, 54
0.0484	Retail Trade, 44
0.0453	Retail Trade, 45
0.0447	Information, 51
0.0401	Manufacturing, 33
0.0399	Wholesale Trade, 42
0.0399	Arts & Entertainment, 71
0.0392	Administrative Support Services, 56
0.0389	Accommodation and Food Services, 72
0.0376	Manufacturing, 32
0.0371	Educational Services, 61
0.0369	Other Services, 81
0.0351	Manufacturing, 31
0.0324	Transportation and Warehousing, 48
0.0288	Finance and Insurance, 52
0.0236	Health Care and Social Assistance, 62

Notes: This table reports the value of the clustering statistic $\mathcal{C}$ defined in (19) at the 2-digit NAICS sector level. The underlying data is our benchmark ORBIS data for just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain.

Intuitively, $\mathcal{C}$ is simply a weighted average of the slope of exit hazard. We use $\mathcal{C}$ as a measure of clustering simply because the statistic captures the coincidence of high distributional density with steep exit hazards. In other words, a lifetime revenue distribution with a higher value of $\mathcal{C}$ has higher distributional weight and is "clustered" in regions where exit is more marginal.

We compute the clustering statistic $\mathcal{C}_s$ using the nonparametric lifetime revenue distributions for each 2-digit NAICS sector s in our baseline Spanish ORBIS dataset. Table A.4 reports the bunching statistics by sector, which vary widely. For instance, construction, real estate, professional services and retail trade are characterized by larger clustering statistics than health care, transportation or manufacturing.

We then use more disaggregated industry classifications for 4-digit NAICS industries j within 2-digit sector s for year t in our data to estimate versions of the following

Table A.5: Clustering and Exit

	Exit Rate$_{jt}$			
	(1)	(2)	(3)	(4)
Δ Revenue$_{jt}$	-0.045***	-0.046***	-0.050***	-0.039***
	(0.008)	(0.012)	(0.011)	(0.011)
Δ Revenue$_{jt}$ $\times$ Clustering$_s$	-0.011*	-0.013*	-0.014*	
	(0.006)	(0.006)	(0.007)	
Clustering$_s$	0.428**	0.413**		
	(0.211)	(0.209)		
Δ Revenue$_{jt}$ $\times$ I(Highly Clustered$_s$)				-0.052**
				(0.022)
Fixed Effects	-	Year	Year, Sector	Year, Sector
Industry-Years	1584	1584	1584	1584
Years	2006-13	2006-13	2006-13	2006-13

Notes: The table reports OLS estimates from (20) of 4-digit NAICS industry j exit rates in year t on industry j's revenue growth in year t and standardized clustering statistics $\mathcal{C}_s$ for 2-digit NAICS sector s containing j. "Highly Clustered" sectors are those with clustering in the top quartile across sectors. Standard errors are clustered at the 4-digit industry j level. Significance is indicated as * = 10% level, ** = 5% level, and *** = 1% level. The underlying data is our benchmark ORBIS data for just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain.

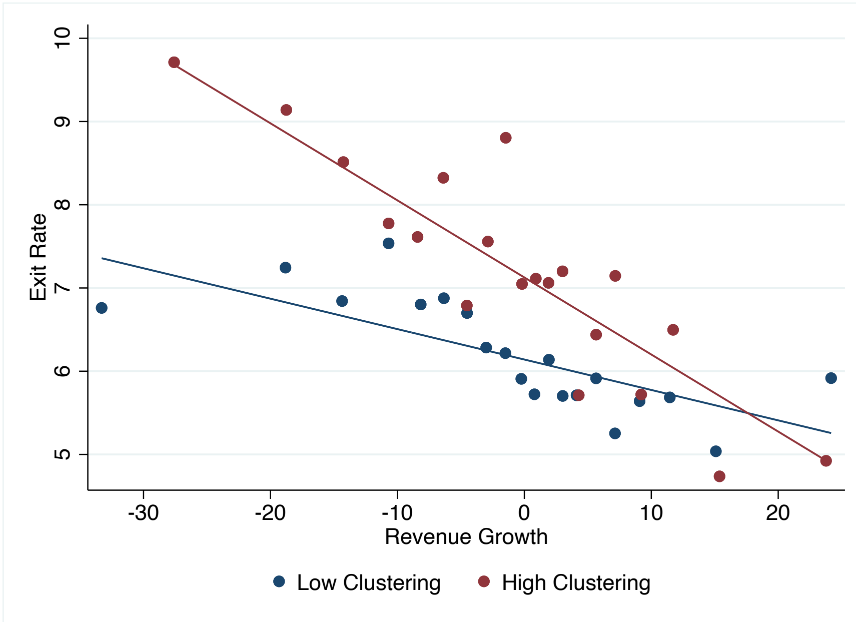
specification

$$\mathbb{P}(\text{Exit})_{jt} = \alpha + \beta\Delta\text{Revenue}_{jt} + \gamma\Delta\text{Revenue}_{jt} \times \mathcal{C}_{s(j)} + \delta\mathcal{C}_{s(j)} + \varepsilon_{jt}. \tag{20}$$

Above, $\mathbb{P}(\text{Exit})_{jt}$ is the exit rate of industry j in year t, $\Delta\text{Revenue}_{jt}$ is the industry j growth rate of revenue in year t, and $\mathcal{C}_{s(j)}$ is the clustering statistic for sector s containing industry j. Note that our model-based intuition predicts $\gamma < 0$ if more clustering is linked to higher exit sensitivity.[18] Table A.5 presents estimates of (20).

[18] Our maintained assumption is that the degree of clustering at the 4-digit level is relatively homogeneous within a given 2-digit sector and stable over our sample period. We rely on this assumption since at the 4-digit level, with too few observations in each cell, the resulting $\mathcal{C}_s$ statistics

Figure A.1: Clustering, Revenue Growth, and Exit Rates



Notes: The binscatter plots exit rates on the vertical axis against revenue growth rates on the horizontal axis. Both variables are measured at the industry (4-digit NAICS) by year level, with a total of 198 industries and 1980 industry-years in total. In red, with associated line of best fit, the plot is based on observations from "high clustering" 2-digit NAICS sectors in which the clustering statistic $\mathcal{C}$ from (19) is in the top quartile across sectors, while the blue observations and line plot data from other "low clustering" sectors with $\mathcal{C}$ below the top quartile. The underlying data is our benchmark ORBIS data for just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain.

Column 1 shows that high revenue growth at the 4-digit industry level is associated with lower exit and that – via the estimated interaction term – this negative association is stronger in sectors with a higher degree of clustering, consistent with our model's intuition. Columns 2 and 3 show the robustness of this pattern to the inclusion of year fixed effects as well as fixed effects for 2-digit sector s. In both columns, the interaction term continues to be negative and statistically significant at the 10% level. In column 4, we replace the linear interaction term with a categorical approach. We define a highly clustered sector as a sector with a clustering statistic

are too noisy. Also, note that output in our model is stationary while, naturally, output exhibits positive growth in the data. So (20) links the exit rate to the transformed stationary *growth rate* of sectoral revenue rather than its level. This transformation allows the empirical test to be consistent with the interpretation of the model.

$\mathcal{C}_s$ in the upper quartile of the distribution of $\mathcal{C}_s$ across sectors. The interaction term is significant at the 5% level, emphasizing again that clustering is indeed statistically linked to the dynamics of industry exit rates. Figure A.1 presents the same fact from column 4, with heterogenous sensitivities in high vs low clustering sectors, using a simple binscatter plot. So, to summarize, Table A.5 shows that industries with more clustered lifetime revenue distributions exhibit higher exit rate sensitivity to changes in revenue growth, intuitively consistent with our key model mechanism.

A.4 Predicting Firm Exit with Profit versus Revenue

Table A.6: Predicting Firm Outcomes with Revenue vs Profits

	Regressor	
	Revenue	Profit Margin
Regressand	(1)	(2)
Exit	-0.021***	-0.001***
	[0.001]	[0.002]
R^2	0.037	0.027
Employment Growth	0.021***	0.0011***
	[0.001]	[0.0013]
R^2	0.022	0.018

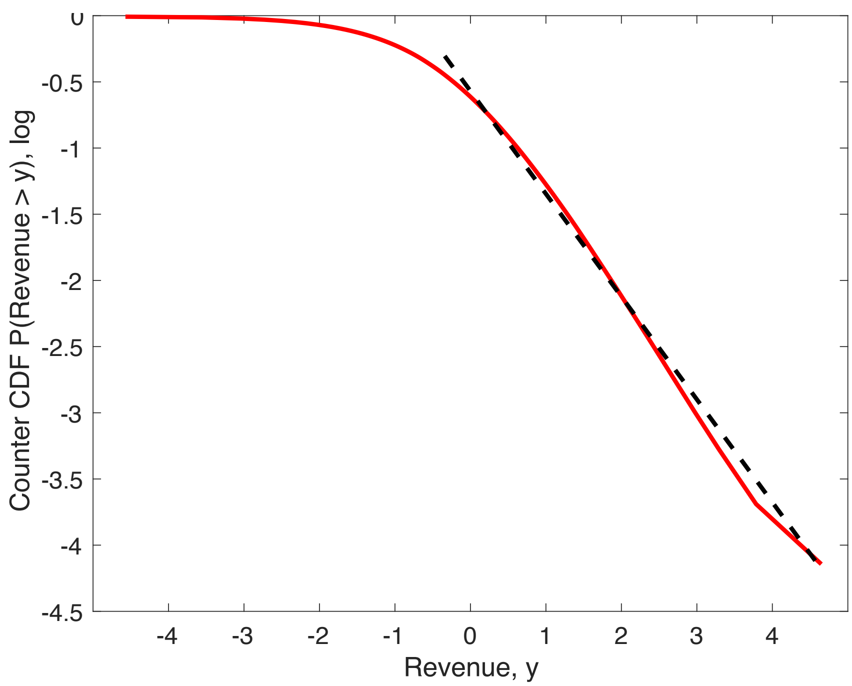
Notes: The table reports results from a serious of predictive regressions of firm exit (top panel) or firm employment growth (bottom panel) in a given year on the firm's revenue, in logs, or profit margin, the ratio of earnings before interest and taxes to revenue, measured in the previous year. Year and industry fixed effects are included in all specifications. The sample is our benchmark ORBIS data for just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain. p-values, based on clustering at the industry level, are reported in square brackets. Significance is indicated as * = 10% level, ** = 5% level, and *** = 1% level.

Our empirical analysis centers on firm-level revenue. Table A.6 reports the results of a set of predictive regressions demonstrating that our revenue variable is a better predictor of both exit and employment growth at the firm level than a natural alternative measure of firm profits.

A.5 A Power Law Tail in Firm Revenue

Fat-tailed cross-sectional size distributions are ubiquitous in many economic contexts and can in principle be detected by a telltale linear relationship between log size and the log counter-CDF of a distribution (Gabaix, 2016). In Figure A.2, we see that a linear relationship of this sort matches the shape of the right tail of our baseline sample's stationary distribution of revenue $H(y)$.

Figure A.2: A Power Law Tail in Firm Revenue



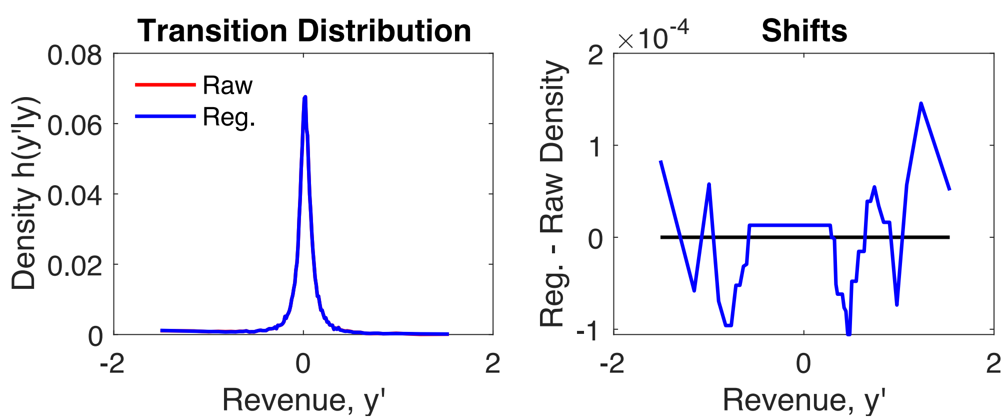
Notes: The solid red line in the figure plots the stationary distribution of revenue $H(y)$ computed from our baseline ORBIS sample of just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain. The horizontal axis is revenue y, in logs, and the vertical axis is the log counter CDF of the revenue distribution. The dotted line is the line of best fit estimated on the upper half of our revenue distribution with a slope coefficient of -0.77.

B Model

In this appendix we provide further information on our solution and calibration of the quantitative model. We start with our approach to (very lightly) regularizing the raw nonparametric empirical objects from Section 2 to satisfy standard assumptions for firm dynamics models. We then discuss the numerical techniques we employ while solving and calibrating both the nonparametric and parametric models. Finally, we present details on our quantitative model robustness checks and recalibrations in a set of summary tables.

B.1 Regularizing the Raw Data

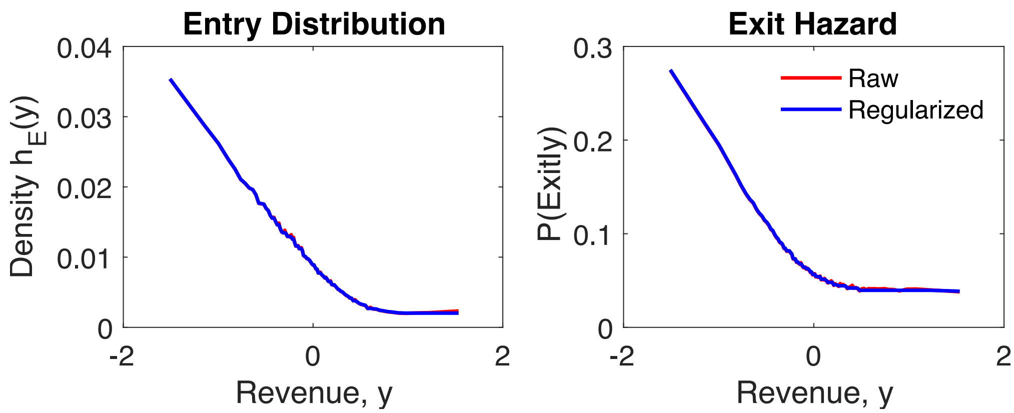
Figure B.1: Regularized vs Raw Transition Distribution



Notes: The left panel of the figure plots raw (in red) and regularized (in blue) transition densities $h(y'|y)$ for next year's revenue y' conditional upon median revenue y in the current year. The right panel plots the regularization shifts or the difference between the regularized and raw densities. The horizontal axis in each figure is next year's revenue y', in logs. The underlying data is drawn from our baseline ORBIS sample of just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain.

In order to embed $H(y'|y)$, $H_E(y)$, and $P(\text{Exit}|y)$ into a canonical heterogeneous firms model, each needs to be adjusted in order to satisfy some standard technical or regularity assumptions. To ensure monotonicity of firm value functions and a well behaved value function iteration algorithm, we first require that the transition distribution H exhibits persistence via first-order stochastic dominance: for two states

Figure B.2: Regularized vs Raw Entry and Exit Patterns



Notes: The left panel of the figure plots the entry density $h_E(y)$, and the right panel plots the exit hazard $\mathbb{P}(\text{Exit}|y)$. The horizontal axis in each figure is the current year's revenue y, in logs. In both panels, the raw object is presented in red, and the regularized object is presented in blue. The underlying data is drawn from our baseline ORBIS sample of just over 5 million firm-years for over one million firms covering the 2005-2014 period in Spain.

such that $y_2 \geq y_1$, we require that $H(y'|y_2) \leq H(y'|y_1)$ for all y'. Second, to ensure that the fixed cost distribution $G(\phi_F)$ has a nondecreasing CDF, i.e., to ensure that the fixed cost distribution is in fact a distribution, we require that the exit hazard $\mathbb{P}(\text{Exit}|y)$ is nonincreasing in y. We do not technically face a need to regularize the entry distribution, but given its overall declining shape in the raw data we also impose that the entry density, not just the exit hazard, is nonincreasing as well.

To impose these regularizations, we design and employ a simple procedure with the intention of making only minimal modifications to the raw data. First, we note that our assumptions of first-order stochastic dominance and a downward sloping hazard can be written as a large set of inequality restrictions which must be satisfied by each value of our extracted distributions and exit hazards. Our procedure then operates as follows. First, we initialize a regularized object, either a transition matrix or an exit hazard, to the raw data equivalent. Second, we compute the existing "gaps" in each of our inequality restrictions. We then distribute weight proportional to the size of this gap to the remaining entries in the corresponding distribution or hazard. Third, we recompute the inequality gaps or errors in our regularized empirical objects, ending the procedure if the gaps are absent but restarting if they are not. An example may help build intuition for our procedure. If an exit hazard is slightly less than downward sloping in the raw data due to apparent noise and a bump upwards in observed exit rates for a given revenue bin, we simply take a small portion of the exit rate in that bin and distribute it elsewhere along the hazard, repeating this approach for all points iteratively until the resulting hazard is downward sloping as a whole.

Helpfully, this regularization procedure turns out to impose only extremely light modifications to the raw data, i.e., the raw data is already very close to satisfying these regularity conditions absent apparent statistical noise. To illustrate this for our baseline ORBIS sample, Figure B.1 compares the raw and regularized transition density for revenue y' conditional upon median revenue y in the current year. The left panel plots the two resulting densities, which are virtually identical to the naked eye. The right panel plots the "shift" or difference between the regularized and raw densities, which remains trivial across revenue levels. Figure B.2 compares the raw and regularized entry distribution and exit hazards which are, again, virtually indistinguishable. In both figures, where applicable, the raw data is presented in red and the regularized objects are plotted in blue.

B.2 Solving the Empirical Nonparametric Model

Note that the static optimality condition for the input n in equation (14) and the residualized log revenue grid y_i (indexing our partition of the revenue space into N_y equally weighted intervals) together imply a quantile-based grid for profitability shocks z_i, $i = 1, ..., N_z$, where $N_z = N_y$ and $\log z_i = (1 - \alpha)y_i$ for all i. Similarly, the empirical objects $H(y'|y)$, $H_E(y)$, and $\mathbb{P}(Exit|y)$ imply an incumbent profitability transition $F(z'|z)$, an entry distribution $F_E(z)$, and an exit hazard $\mathbb{P}(\text{Exit}|z)$ on the profitability grid z_i.

We assume that exit occurs for the highest profitability firms in our sample for only exogenous reasons, i.e., that $\delta = \mathbb{P}(\text{Exit}|z_{N_z})$. In our baseline ORBIS sample in Spain, the resulting exogenous exit rate is $\delta = 3.9\%$. The remaining parameters to be calibrated in our nonparametric model include only the labor share α, the household's rate of time preference β, the fixed labor supply $\bar{N}$, and the sunk entry cost ϕ_E. Given a parameterization of the model, i.e., a list of these parameters, we solve the model with an outer loop-inner loop approach as follows.

1. **Outer Loop on GE Objects** Guess values for the wage W and the entry

mass M_E, and fix a GE tolerance $\epsilon^{GE} > 0$.

(a) **Inner Loop on Firm Value Function** Initialize $k = 0$, guess a value function $V^{(k)}(z)$, and fix a value function error tolerance $\epsilon^V > 0$.

i. Compute the implied continuation values ${\phi_F^*}^{(k)}(z)$ via equation (2) and using $V^{(k)}(z)$.

ii. Infer the distribution $G^{(k)}(\phi_F)$ of fixed cost shocks ϕ_F consistent with ${\phi_F^*}^{(k)}(z)$, $V^{(k)}(z)$, and the empirical exit hazard by using the mapping

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

iii. Compute an updated value function $V^{(k+1)}(z)$ via the Bellman equation

$$V^{(k+1)}(z) = \left\{ \begin{array}{c} \max_n (zn^\alpha - Wz) \\ -\int_0^{{\phi_F^*}^{(k)}(z)} \phi_F dG(\phi_F) \end{array} + \beta(1-\delta) \int V^{(k)}(z') dF(z'|z) \right\}.$$

iv. If the error in the Bellman equation $\max_z |V^{(k+1)}(z) - V^{(k+1)}(z)|$ is smaller than ϵ^V, then the firm value function $V(z) = V^{(k)}(z)$, continuation values $\phi_F^*(z) = {\phi_F^*}^{(k)}(z)$, and the fixed cost distribution $G(\phi_F) = G^{(k)}(\phi_F)$ are computed. Otherwise, set $k = k+1$ and return to step (1(a)i).

(b) **Inner Loop on Firm Distribution** Initialize $k = 0$, guess an operating distribution $F_O^{(k)}(z)$ for firms, guess a mass $M_O^{(k)}$ of operating firms, and fix a tolerance $\epsilon^F > 0$ for distributional convergence.

i. Compute the implied mass of operating firms $M_O^{(k+1)}$ via

$$M_O^{(k+1)} = (1-\delta) M_O^{(k)} \int G(\phi_F^*(z)) dF_O^{(k)}(z) + M_E.$$

ii. Compute the implied distribution of operating firms $F_O^{(k+1)}(z)$ via

$$F_O^{(k+1)}(z') = (1-\delta) \frac{M_O^{(k)}}{M_O^{(k+1)}} \int G(\phi_F^*(z)) F(z'|z) dF_O^{(k)}(z) + \frac{M_E}{M_O^{(k+1)}} F_E(z').$$

iii. If the errors in the operating mass update $|M_O^{(k+1)} - M_O^{(k)}|$ and distributional update $\max_z |F_O^{(k+1)}(z) - F_O^{(k)}(z)|$ are both less than ϵ^F, then the operating mass $M_O = M_O^{(k)}$ and operating distribution $F_O(z) = F_O^{(k)}(z)$ are computed. Otherwise, set $k = k + 1$ and return to step (1(b)i).

2. Compute the implied value to entry V_E via

$$V_E = \int V(z) dF_E(z).$$

3. Compute the implied labor demand N via

$$N = M_O \int n^*(z) dF_O(z),$$

   where $n^*(z)$ is optimal static labor demand for an individual firm with profitability z.

4. If the error in the free entry condition $|V_E - \phi_E|$ and the error in the labor market clearing condition $|N - \bar{N}|$ are both less than the GE tolerance ϵ^{GE}, then the model is solved. Otherwise, update your guesses for the wage and entry mass and return to step (1).

When the algorithm above is complete, the nonparametric version of our model is solved in a manner not only consistent with general equilibrium but also, by construction, with the observed revenue transitions, the entry distribution, and the exit hazard measured nonparametrically.

A few additional technical details are useful. We implement all of the calculations above continuously, linearly interpolating value functions, fixed cost distributions, operating distributions, and continuation values on the grid z_i. Where integration is required, we use Simpson quadrature with densities $f_O(z)$, $f_E(z)$, and $f(z'|z)$ consistent with linear interpolation of the CDFs $F_O(z)$, $F_E(z)$, and $F(z'|z)$ in a manner which preserves the empirical weight on equal-mass intervals containing the revenue quantiles y_i. Because the free entry condition is separable from the entry mass M_E, we first employ bisection on the aggregate wage W to ensure that the free entry condition is satisfied, then we update M_E so that (3) is exactly satisfied. In our baseline, we employ $N_y = N_z = 101$ grid points or quantiles, and on a 2017 iMac

Pro model solution takes around a minute or two in MATLAB without requiring aggressive parallelization.

B.3 Solving the AR(1)/Parametric Model

In our AR(1) or parametric model version, the parameters to be calibrated include the labor share α, the household's rate of time preference β, the fixed labor supply $\bar{N}$, the sunk entry cost ϕ_E, the upper bound $\bar{\phi}_F$ of the fixed cost distribution $G(\phi_F) = U(0, \bar{\phi}_F)$, the persistence of the lognormal AR(1) profitability process ρ, the conditional variance of the lognormal AR(1) profitability process σ^2, and the mean of the lognormal entry distribution μ_E. The exogenous exit hazard δ is carried over identically from our nonparametric model solution as described above. Given a parameterization of the model, i.e., a list of these parameters, we solve the model with an outer loop-inner loop approach as follows.

1. **Outer Loop on GE Objects** Guess values for the wage W and the entry mass M_E, and fix a GE tolerance $\epsilon^{GE} > 0$.

   (a) **Inner Loop on Firm Value Function** Initialize $k = 0$, guess a value function $V^{(k)}(z)$, and fix a value function error tolerance $\epsilon^V > 0$.

      i. Compute an updated value function $V^{(k+1)}(z)$ via the Bellman equation

$$V^{(k+1)}(z) = \left\{ \begin{array}{c} \max_n (zn^\alpha - Wz) \\ -\int_0^{\phi_F^{*,(k)}(z)} \phi_F dG(\phi_F) \end{array} + \beta(1-\delta) \int V^{(k)}(z') dF(z'|z) \right\}.$$

      ii. If the error in the Bellman equation $\max_z |V^{(k+1)}(z) - V^{(k+1)}(z)|$ is smaller than ϵ^V, then the firm value function $V(z) = V^{(k)}(z)$ is computed. Otherwise, set $k = k + 1$ and return to step (1(a)i).

   (b) **Inner Loop on Firm Distribution** Initialize $k = 0$, guess an operating distribution $F_O^{(k)}(z)$ for firms, guess a mass $M_O^{(k)}$ of operating firms, and fix a tolerance $\epsilon^F > 0$ for distributional convergence.

      i. Compute the implied mass of operating firms $M_O^{(k+1)}$ via

$$M_O^{(k+1)} = (1-\delta) M_O^{(k)} \int G(\phi_F^*(z)) dF_O^{(k)}(z) + M_E.$$

ii. Compute the implied distribution of operating firms $F_O^{(k+1)}(z)$ via

$$F_O^{(k+1)}(z') = (1-\delta)\frac{M_O^{(k)}}{M_O^{(k+1)}}\int G(\phi_F^*(z))F(z'|z)dF_O^{(k)}(z) + \frac{M_E}{M_O^{(k+1)}}F_E(z').$$

iii. If the errors in the operating mass update $|M_O^{(k+1)} - M_O^{(k)}|$ and distributional update $\max_z |F_O^{(k+1)}(z) - F_O^{(k)}(z)|$ are both less than ϵ^F, then the operating mass $M_O = M_O^{(k)}$ and operating distribution $F_O(z) = F_O^{(k)}(z)$ are computed. Otherwise, set $k = k + 1$ and return to step (1(b)i).

2. Compute the implied value to entry V_E via

$$V_E = \int V(z)dF_E(z).$$

3. Compute the implied labor demand N via

$$N = M_O \int n^*(z)dF_O(z),$$

where $n^*(z)$ is optimal static labor demand for an individual firm with profitability z.

4. If the error in the free entry condition $|V_E - \phi_E|$ and the error in the labor market clearing condition $|N - \bar{N}|$ are both less than the GE tolerance ϵ^{GE}, then the model is solved. Otherwise, update your guesses for the wage and entry mass and return to step (1).

Note that unlike in the empirical or nonparametric version of the model, the fixed cost distribution $G(\phi_F) = U(0, \bar{\phi}_F)$ is predetermined. Also note that the entry and transition distributions $F_E(z)$ and $F(z'|z)$ are parametric, following conventional lognormal processes converted to a uniform profitability grid as in Tauchen (1986). Just as in the nonparametric solution of the model, however, we continue to solve the model continuously, storing value functions via linear interpolation, computing integrals via Simpson quadrature, and evaluating entry, operating, and transition distributions using linear interpolation of the CDFs $F_E(z)$, $F_O(z)$, and $F(z'|z)$. In our baseline, we again employ $N_z = N_y = 101$ points for our interpolation procedures, and model solution takes around a minute or two on a 2017 iMac Pro in MATLAB without aggressive parallelization.

B.4 Calibrating the Model

There are multiple model parameters which we fix or calibrate externally before engaging in a moment-matching exercise, as outlined in Section 5.1. We set $\alpha = 2/3$ to generate a conventional labor share of 2/3, we set $\beta = 1/1.04$ to be consistent with a conventional 4% real interest rate and an annual solution of the model, and we set $\bar{N}$ to be equal to the aggregate employment rate (resulting in $\bar{N} = 0.5974$ in our baseline Spanish sample and comparable values for our other samples). We also set the exogenous exit hazard δ based on the observed exit rate of the largest firms in our empirical sample, resulting in $\delta = 3.9\%$ for our baseline Spanish sample and comparable values for our other samples. Each of the versions of our model, nonparametric and parametric, is solved holding these externally calibrated parameters fixed.

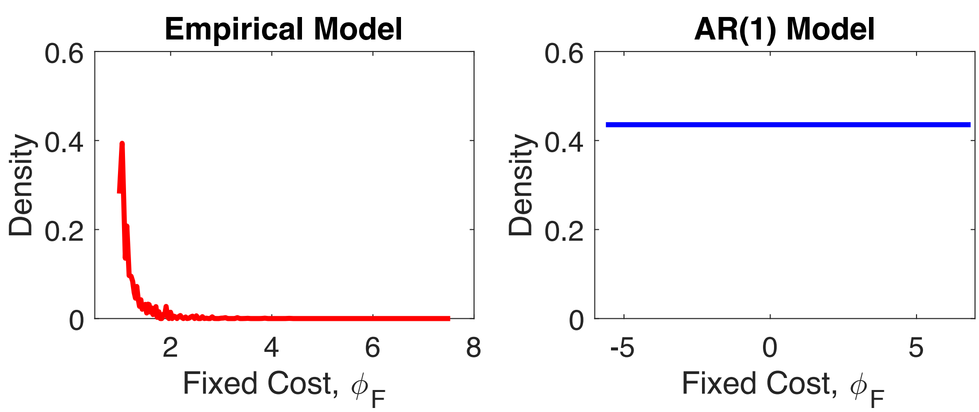
Nonparametric Calibration With the externally calibrated parameters listed above fixed, only the sunk entry cost ϕ_E must be calibrated for the nonparametric model. We choose the value of ϕ_E to match the observed average number of employees per firm. The number of employees per firm declines in the wage W, which adjusts to satisfy the free entry condition as the parameter ϕ_E is shifted.

Parametric Calibration With the externally calibrated parameters above fixed, we must still fix the values of the lognormal AR(1) profitability process (ρ, σ^2), the mean of the lognormal entry distribution μ_E, the upper bound $\bar{\phi}_F$ of the fixed cost distribution $G(\phi_F) = U(0, \bar{\phi}_F)$, as well as the sunk entry cost ϕ_E. Following convention in the parametric firm dynamics literature, we first set ρ to the autocorrelation of the profitability process $\log z$ inferred from our observed revenue series y, and we set σ^2 to match the observed variance of $\log z$.

Then, with ρ and σ^2 fixed, we choose the remaining three parameters $(\mu_E, \bar{\phi}_F, \phi_E)$ to jointly match three moments. As in the nonparametric model, we match (i) the observed average number of employees per firm. We also match (ii) the observed exit rate $\mathbb{P}(\text{Exit})$ which naturally moves with the fixed cost upper bound $\bar{\phi}_F$. Finally, we match (iii) the mean difference between log revenue for entering and operating firms, which naturally moves with the mean of the entry distribution μ_E. One might wonder why we did not target moments (ii) nor (iii) in our nonparametric model solution. But the nonparametric model matches both of these moments by construction, since both moments are implied by the combination of incumbent revenue transitions, exit hazards, and the entry distribution which are fully matched in the nonparametric model.

Fixed Cost Distributions The nonparametric and parametric calibration techniques yield fixed cost distributions $G(\phi_F)$, which we plot in Figure B.3. In the nonparametric case, our procedure yields a distribution with high density at low fixed cost realizations: this is required in order to match the strongly declining exit hazard found empirically and plotted in Figure B.2.

Figure B.3: Calibrated Fixed Cost Distributions



Notes: The left panel plots the density $g(\phi_F)$ of fixed costs recovered in our baseline nonparametric quantitative model analysis, while the right panel plots the same density for our baseline calibrated AR(1) quantitative model. The horizontal axis is the fixed cost shock ϕ_F, in logs, while the vertical axis is the density $g(\phi_F)$.

B.5 Robustness Checks

Section 7 overviews a large number of quantitative model robustness checks. For each check, we redo the calibration process summarized above for an alternative sample or model assumption, resulting in the recalibrated values list in Appendix Table B.1. The associated counterfactual implications of a subsidy to operating firms, in the nonparametric vs AR(1) cases, are available in Appendix Table B.2.

Table B.1: Alternative Model Calibrations

	Empirical Case	Parametric AR(1) Case				
	ϕ_E	ρ	σ	μ_E	$\bar{\phi}_F$	ϕ_E
Panel A: Alternative Model Assumptions						
Endogenous Labor Supply	22.9	0.94	0.19	-0.44	2.30	5.18
Higher $\alpha = 0.75$	16.8	0.94	0.14	-0.33	1.96	4.42
Lower $\alpha = 0.60$	28.4	0.94	0.23	-0.53	2.50	5.66
Panel B: Alternative Datasets						
Before 2009	16.7	0.93	0.19	-0.52	2.46	4.02
After 2009	25.1	0.95	0.19	-0.46	1.96	5.48
Manufacturing Only	25.2	0.98	0.13	-0.48	2.11	4.25
Non-Manufacturing Only	21.5	0.93	0.20	-0.44	2.25	4.99
Unconsolidated Accounts	16.9	0.94	0.19	-0.44	2.15	4.79
Excluding M&A	22.7	0.94	0.19	-0.36	2.55	4.97
Year Effects Only	14.3	0.96	0.19	-0.45	2.17	4.84
No Trimming	129.7	0.91	0.24	-0.33	3.12	7.65
Trimming at 1% and 99%	9.39	0.94	0.18	-0.41	1.88	4.00
Remove Firm Age	64.7	0.92	0.21	0.11	2.40	12.1
Italy	28.0	0.91	0.24	-0.65	3.13	5.08
Portugal	13.4	0.95	0.17	-0.84	1.57	2.18
France	22.7	0.96	0.15	-0.65	1.62	3.67
Norway	24.8	0.95	0.19	-0.85	1.23	3.33
Baseline	22.9	0.94	0.19	-0.43	2.30	5.18

Notes: This table reports calibrated parameters for each of our model robustness checks and alternative datasets, for both the empirical and parametric AR(1) model versions. Note that in the extension with endogenous labor supply, we also calibrate $\omega = 1.51$ (empirical case) and $\omega = 1.39$ (parametric case) to match the Spanish employment rate.

Table B.2: Relative Subsidy Impacts in Our Empirical vs AR(1) Models

	Exit Rate	Output
Panel A: Alternative Model Assumptions		
Endogenous Labor Supply	3.0660	0.8125
Higher $\alpha = 0.75$	3.0728	0.6872
Lower $\alpha = 0.60$	3.0765	0.6859
Panel B: Alternative Datasets		
Before 2009	2.5787	0.7095
After 2009	2.4338	0.6176
Manufacturing Only	3.6506	0.4211
Non-Manufacturing Only	2.7832	0.7108
Unconsolidated Accounts	3.1550	0.7109
Excluding M&A	3.0750	0.7161
Year Effects Only	2.8076	0.7015
No Trimming	3.3206	0.7883
Trimming at 1% and 99%	2.6755	0.7461
Remove Firm Age	2.9401	0.3934
Italy	3.9723	0.7040
Portugal	5.8445	0.7905
France	1.8992	0.5814
Norway	2.8302	0.6560
Baseline	3.0701	0.6845

Notes: This table reports relative changes at the aggregate level from a fixed cost subsidy equal to 5% of pre-subsidy output in our calibrated empirical nonparametric versus the parametric AR(1) model. Panel A reports results under various alternative model assumptions, while Panel B considers calibrations based on alternative ORBIS datasets. For each experiment indicated in the first column, we first calculate the change in the aggregate exit rate, in percentage points, and aggregate output, in percent, relative to the no-subsidy values for both the nonparametric and AR(1) models. We then report the ratio of the nonparametric to the AR(1) model's changes. The second column reports this ratio for the exit rate, while the third column reports this ratio for output.