Quantitative Economics 5 (2014), 377–416 1759-7331/20140377

Size-dependent regulations, firm size distribution, and reallocation

François Gourio
Federal Reserve Bank of Chicago, Boston University, and NBER

Nicolas Roys
University of Wisconsin Madison

In France, firms that have 50 employees or more face substantially more regulation than firms that have less than 50. As a result, the size distribution of firms is visibly distorted: there are many firms with exactly 49 employees. We model the regulation as the combination of a sunk cost that must be paid the first time the firm reaches 50 employees and a payroll tax that is paid each period thereafter when the firm operates with more than 50 employees. We estimate the model using indirect inference by fitting the discontinuity of the size distribution. The key finding is that the regulation is equivalent to a combination of a sunk cost approximately equal to about 1 year of an average employee salary and a small payroll tax of 0.04%. Our structural model fits well the discontinuity in the size distribution. Removing the regulation improves labor allocation across firms, leading in steady state to an increase in output per worker slightly less than 0.3%, holding the number of firms fixed. However, if firm entry is elastic, the steady-state gains are an order of magnitude smaller.

Keywords. Firm size distribution, regulation, threshold effect, reallocation.
JEL classification. E23, O1, O40.

1. Introduction

In many countries, small firms face lighter regulation than large firms. Regulation, broadly defined, takes many forms, from hygiene and safety rules to mandatory elections of employee representatives and to larger taxes. The rationale for exempting small firms from some regulations is that the compliance cost is too high relative to their sales. A necessary consequence, however, is that regulations are phased in as the firm grows,

François Gourio: francois.gourio@chi.frb.org
Nicolas Roys: nroys@ssc.wisc.edu
We thank Gadi Barlevy, Nick Bloom, Jeff Campbell, Russell Cooper, Dean Corbae, John Kennan, Claire Lelarge, Theodore Papageorgiou, Marcelo Veracierto, Mark Wright, the two referees, the editor, and conference and seminar participants for useful discussions and comments. We are grateful to Claire Lelarge and Gregory Verriest from INSEE for their help in constructing our data. The views expressed here are those of the authors and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System. A first version of this paper appeared as Chapter 4 of Nicolas Roys' Ph.D. dissertation, submitted in December 2009 at the Paris School of Economics–Paris 1.

DOI: 10.3982/QE338

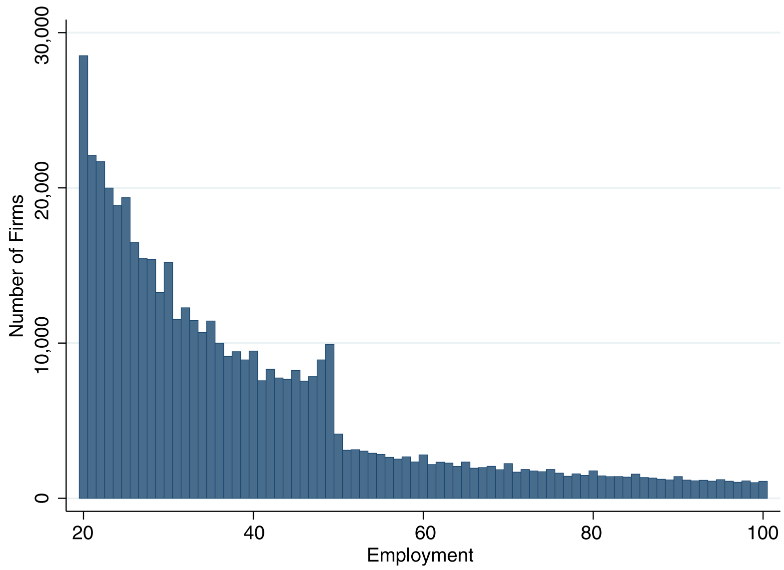

FIGURE 1. Distribution of firm employment between 20 and 100 employees in France.

generating an implicit marginal tax. Because regulations are typically phased in at a few finite points, they are sometimes referred to as "threshold effects": for instance, in the case of France, an important set of regulations applies to firms that have more than 50 employees. As a result, the firm size distribution is distorted: there are few firms that have exactly 50 employees and a large number of firms that have 49 employees. Figure 1 plots the firm size distribution in our French data, illustrating this well known pattern.

These distortions have attracted attention in public policy circles. The common wisdom, as reflected in numerous reports by blue-ribbon panels, is that these regulations are a significant impediment to the growth of small firms and should be suppressed or smoothed out. However, there is little work that formally models these policies to understand and evaluate their effects. In this paper, we evaluate this common wisdom by proposing and estimating a structural model of firm growth that explicitly takes into account the phase-in of the regulation.

The model serves two purposes, positive and normative. On the positive side, a structural model is needed to understand the exact sources of distortion. It is not obvious how the regulations should be modeled, given their scope and complexity (which we discuss in Section 2). Are regulations equivalent to higher per-period costs or to a sunk cost? The puzzle that quickly emerges is, why are there *any* firms at all that have exactly 50 employees given the higher costs? Our intuition is that many of these regulations might be better approximated as a sunk cost (i.e., a one-time investment), since a large fraction of the cost is learning the regulation and adapting to it, and since some regulations might still apply in the future even if the firm operates below 50 employees. The presence of the sunk cost also helps explain why there are some firms that have exactly 50 employees: firms are reluctant to have more than 50 employees the first time that they reach that limit, but the threshold is irrelevant in subsequent periods, since the cost is already paid. On the normative side, what are the potential benefits of removing, or smoothing, the regulation thresholds? The visibly distorted firm distribution suggests that productivity could be increased if firms close to the threshold grow, as labor would be reallocated toward more productive firms. Because we do not believe that we can adequately measure the regulations' benefits, we do not provide a full welfare analysis, but more modestly attempt to measure the steady-state output or employment cost of the regulations.

To address these questions, we estimate a model that incorporates both a sunk cost of complying with the regulation (which captures the cost of learning the regulation for the manager, as well as consulting with lawyers and accountants, and buying equipment required by the regulation, and the fact that some regulations might be "sticky") and a higher per-period cost (which captures the cost that the regulation creates in every period thereafter). Our model can be solved using standard stochastic dynamic optimization techniques (Dixit and Pindyck (1994), Stokey (2008)), and we obtain the cross-sectional distribution in closed form. This is useful when we turn to the estimation, because simulating accurately the highly skewed cross-sectional distribution of firms is challenging. We estimate the model using indirect inference and match the discontinuity in the firm size distribution, which is the key evidence that the regulation matters. We further match the firm size distribution, conditional on having operated above 55 employees in the past. This allows us to separately estimate the sunk cost and the per-period cost. In spite of its parsimony, our model is able to match the distribution (and the conditional distribution) fairly well.

Our main result is that the regulation is equivalent to a combination of a sunk cost of about a year of an average employee wage and a small, but significant, additional payroll tax of 0.04%. We next use our model estimates to infer the social cost of the regulation. Holding the number of firms and total employment constant, we find that output increases by 0.27% in steady state if the regulation is removed. This number captures the misallocation of labor across firms. However, when we allow the number of firms to adjust, we find a smaller effect: output (net of entry costs) rises by only 0.02–0.03% in steady state. This suggests that these regulations may not have large aggregative effects.

The rest of the paper is organized as follows. We first discuss the related literature. Section 2 presents some institutional background, our data, and some reduced-form evidence that motivates our analysis. Section 3 discusses the model. Section 4 covers our estimation method and presents the empirical results. Section 5 uses these estimates to conduct some policy experiments. Section 6 provides some robustness analysis and Section 7 concludes. Additional information is available in a supplementary file on the journal website, http://qeconomics.org/supp/338/code_and_data.zip.

Related literature

Our paper is related to a recent growing literature that studies the effect of misallocation on aggregate productivity. Building on Hopenhayn and Rogerson (1993), Restuccia and Rogerson (2008), and Buera, Kaboski, and Shin (2011) argued that misallocation is an important determinant of aggregate total factor productivity (TFP). Hsieh and

Klenow (2009), and Bartelsman, Haltiwanger, and Scarpetta (2009) presented empirical evidence consistent with higher misallocation in poorer countries with lower TFP. Closely related to our paper, Guner, Ventura, and Xu (2008) and García-Santana and Pijoan-Mas (2011) suggested that size-dependent policies have a large negative impact on total factor productivity since productive firms have less incentive to grow.

In the macroeconomic studies of Restuccia and Rogerson (2008) and Guner, Ventura, and Xu (2008), distortions arise due to implicit "taxes." However, these taxes are not directly measured. The regulations that we discuss are a prime example of these distortions and they very clearly affect the firm distribution, consistent with these studies. While our aim is more modest than these macroeconomic studies, since we focus on one particular distortion, we believe that our focus allows a credible identification of the effect of government regulation on firms outcomes. In particular, we match the distortion in the firm size distribution, which is the prima facie evidence that the size-dependent regulation matters.

While the distortion in the firm size distribution in France is well known (see Cahuc and Kramarz (2004) and the references therein), few papers have studied it in detail. Ceci-Renaud and Chevalier (2011) carefully documented the impact of the various thresholds (10, 20, and 50 employees) on the firm size distribution and on firm dynamics by considering different data sources. We are not aware of any structural modeling that tries to apprehend the costs of the distortion. While finishing this paper, we became aware of a recent working paper (Garicano, Lelarge, and Van Reenen (2013)) that shares some of our goals and approach. Three important differences between our papers are that (i) our model allows the regulation to be a sunk cost rather than a per-period cost, (ii) our estimation method targets the firm size distribution around the threshold, and (iii) our policy experiments assume that the wage adjusts if the regulation is removed. In contrast, their model is static, their estimation method aims at the entire firm size distribution, and their policy experiments assume rigid wages. Hence, while we use similar data, we have different models and estimation methods, and emphasize different policy experiments. We compare our results in more detail in Section 6.3.

2. Motivating evidence

We first describe the institutional background, then we present our data sources, and finally we show some simple reduced-form evidence of the threshold effects.

2.1 *Institutional background*

This section draws heavily from Ceci-Renaud and Chevalier (2011). Labor laws in France as well as various accounting and legal rules make special provisions for firms with more than 10, 11, 20, or 50 employees.

These regulations are not all based on the same definition of "employee." Labor laws, which are likely the most important, are based on the full-time equivalent workforce, computed as an average over the last 12 months. The full-time equivalent workforce takes into account part-time workers, as well as temporary workers, but not trainees or *contrats aidés* (a class of government-subsidized, limited duration contracts, which may be used to hire people who face “special difficulties” in finding employment, such as the very long term unemployed or unskilled youth). Hence, it seems fairly difficult for firms to work around the regulation.

The main additional regulations as the firm reaches 50 employees are the following:

– Possibly mandatory designation of an employee representative.

– A committee for hygiene, safety, and work conditions must be formed and trained.

– A *comité d'entreprise* (works council) must be formed, that must meet at least every other month. This committee, which must have some office space and receives a subsidy equal to 0.2% of the total payroll, has both social objectives (e.g., organizing cultural or sports activities for employees) and an economic role (mostly on an advisory basis).

– A higher payroll tax rate to subsidize training that goes from 0.9% to 1.5% (*formation professionelle*).

– In case of firing more than 9 workers for “economic reasons,” a special legal process must be followed (*plan social*). This process increases dismissal costs and creates legal uncertainty for the firm.

This list is not exhaustive, but clearly one would expect these costs to be significant. Some of these costs are also difficult to model in a tractable manner. In some cases—in particular, the *comité d'entreprise*—the firm is required to fund additional worker benefits. To the extent that the process is reasonably efficient, these rules might simply amount to a substitute form of compensation and have limited effects: the higher benefits may allow firms to attract better workers or to pay them less.

In Section 3, we model these costs as a combination of a per-period cost and a sunk cost that must be paid the first time the firm reaches 50 employees. (The per-period cost is a payroll tax, but as we discuss later, the implications would be similar if it were a per-period fixed cost.) The per-period cost captures the static cost of complying with the regulation each period. The sunk cost captures the investment required to initially comply with the regulation. Some of these investments reflect capital expenditures (e.g., the office space and equipment that must be dedicated to the committees). But the main sunk cost is the time spent by the manager to learn the rules that must be complied with and to learn how to organize the firm to deal with the rules efficiently. These information costs may include consulting with accountants and lawyers. This wasted managerial time is likely a primary cost of the regulation: for small businesses, managerial time is a key scarce input. On the other hand, we suspect that several of these regulations are not very costly once the manager has figured out exactly how to set up processes that comply with the law while minimizing waste. This motivates our modeling of the regulations as sunk costs. Finally, one argument for modeling the cost as sunk is that some regulations may continue to apply if the firm shrinks below 50 employees. For instance, the works council (*comité d'entreprise*) may not be easy to dismantle. In this case, the cost is effectively sunk. However, ultimately it is an empirical question as to what is the best model of the regulation; hence, we design our estimation to allow us to distinguish the two types of costs.

2.2 *Data*

We purchased our data from INSEE, the French statistical institute.[1] INSEE combines administrative (tax) data with statistical surveys to construct the data base SUSE, which has data on employment, total compensation, value added, gross operating surplus, assets, and so forth. All firms with sales over 3.5 million francs (around 530,000 euros) and liable to corporate taxes under the standard regime are included. Moreover, some smaller firms are also included in these data. For our purpose, the 3.5 million threshold implies that almost all firms with more than 30 or so employees are included. Hence, we focus on the threshold at 50 employees, for which our data are essentially exhaustive. We focus on the period 1994–2000.

2.3 *Preliminary data analysis*

Figure 1 plots the distribution of employment, pooling data for the entire period (1994–2000), and truncating at 100 employees. There is clearly a large discontinuity around the thresholds of 50 employees. Many surveys reveal "rounding" of employment, but this figure shows the opposite pattern.

Table 1 reports the size distribution of firms by employment over the range 40–59. There is a clear drop in the number of firms after 49 employees. For example, there are more than three times as many firms with 49 employees as firms with 51 employees.

A common way to summarize firm size distribution is to use power laws. With a power law, the log probability that firm size is greater than x is proportional to x. Formally, $P(\text{size} > x) = Cx^{-\xi}$, where C and ξ are constants. The regulation creates a break in the power law. To illustrate this, Figure 2 displays the results of two estimations. First, we estimate the parameters C and ξ of the power law for firms with more than 100 employees. The power law seems to approximate well the firm size distribution for all but the

Table 1. Distribution of firm employment between 40 and 59 employees.

	Fraction	S.E.	No. Firms		Fraction	S.E.	No. Firms
40	8.42	0.28	9486	50	3.67	0.29	4140
41	6.72	0.29	7575	51	2.75	0.29	3097
42	7.38	0.29	8311	52	2.78	0.29	3130
43	6.88	0.29	7752	53	2.70	0.29	3040
44	6.81	0.29	7666	54	2.57	0.29	2901
45	7.31	0.29	8239	55	2.51	0.29	2826
46	6.70	0.29	7548	56	2.34	0.29	2638
47	6.96	0.29	7841	57	2.24	0.29	2526
48	7.92	0.29	8916	58	2.37	0.29	2670
49	8.80	0.28	9916	59	2.08	0.29	2344

Notes: Fraction is the number of firms for each employment size over the range 40–59, divided by the total number of firms between 40 and 59; S.E. is the associated standard error; and No. Firms is the raw number of firms in each bin.

[1]To obtain our data, researchers can follow the instructions on the INSEE website: http://www.webcommerce.insee.fr/fiche-produit-en.php?id_produit=1540.

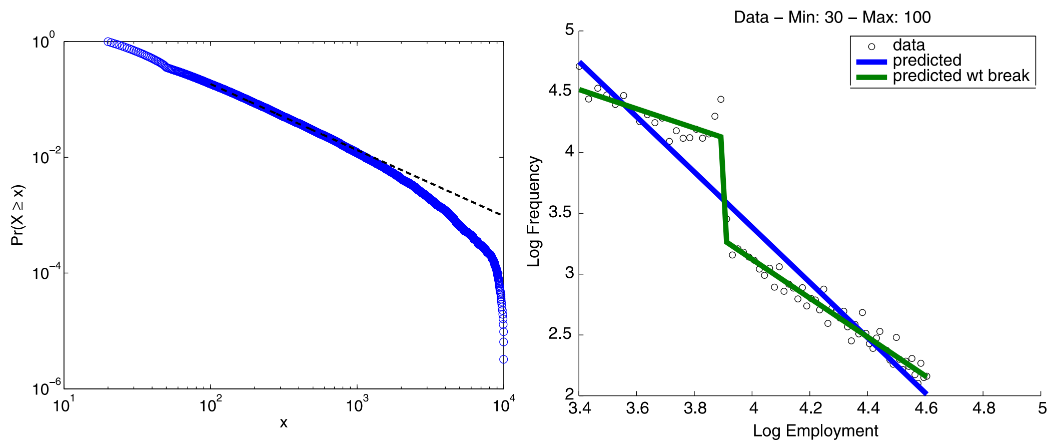

FIGURE 2. Power law estimation. Left: Estimation by maximum likelihood for all the firms with a number of employees greater than 100. Right: Regression of the logged number of firms on the logged number of employees, with and without a structural break at 50. The sample includes only firms with employment level between 30 and 100.

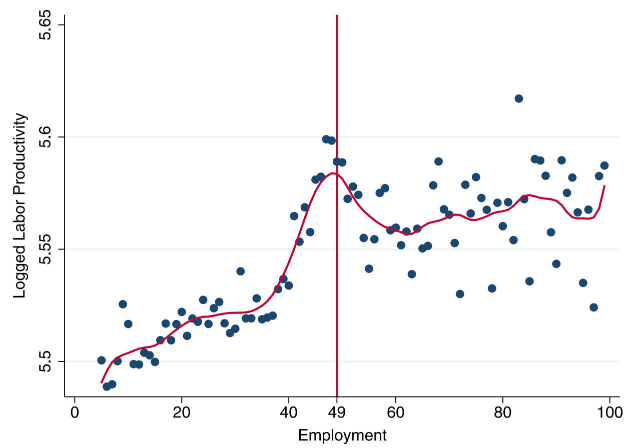

FIGURE 3. Mean logged labor productivity as a function of employment. Each dot represents an employment level. The solid line is a locally weighted regression of logged labor productivity on the employment level with bandwidth 0.18. The vertical line represents a level of employment of 49.

largest firms. This is a well known result (see Axtell (2001) and Di Giovanni, Levchenko, and Ranciere (2011) among others). Second, we run a regression of the log frequency on log size, with or without a structural break (in both slope and intercept) at size 50. The presence of a structural break is clearly visible from this second figure.

Figure 3 plots labor productivity (the ratio of value added to employment) as a function of employment. There are two patterns in this picture: first, labor productivity is on average higher for large firms, as is well known. Second, while there is substantial noise in this figure, a local peak of labor productivity is obtained for 49 employees. This

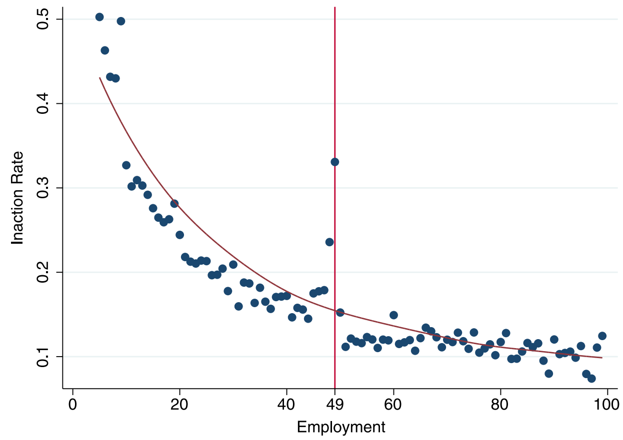

FIGURE 4. Inaction rate: probability that employment stays constant between two consecutive years, as a function of initial employment. Each dot represents a particular employment level. The solid line is a locally weighted regression of the inaction rate on the employment level with bandwidth 0.8. The vertical line represents a level of employment of 49.

is also a natural implication of the regulation: because firms are reluctant to go over the threshold, they hire less labor than they would, generating larger output per worker.

The dynamics of firms around the threshold are also affected. Figure 4 reports the probability that a firm has an employment level that is constant between two periods. Overall, this probability declines with firm size: inaction is more likely for small firms. Yet, this probability increases right before the threshold. The probability of keeping employment constant between two consecutive years is 34% for firms with 49 employees, compared to 17% for firms with 40 employees and 11% for firms with 59 employees. This suggests that the presence of the threshold leads to inaction and slows down employment growth.

To assess the statistical significance of this result, we estimate a probit that characterizes the probability of not adjusting employment. Explanatory variables are a set of dummies variables that indicate whether or not last period employment was $45, \dots, 55$, the growth rate of production, last period employment, and a set of time dummies that capture aggregate shocks. Table 2 reports the coefficients. The probability of inaction increases for firms with a number of employees between 45 and 49. The largest increase is observed for firms of size 49.

As we discussed in the Introduction, it is unclear exactly how to model the cost of the regulation, in particular, whether it is a sunk or a recurrent cost. A simple test is to check if the discontinuity in the firm size distribution is still apparent if one restricts the sample to firms that have already been above the threshold in the past.[2] To reduce the effect of measurement error, Figure 5 compares the firm size distribution in the data with the firm distribution conditional on having been above 55 in the past. While there is still a spike at 50, its size is substantially reduced by conditioning. Whereas in the un-

[2] We thank Theodore Papageorgiou for this suggestion.

Table 2. Probability of inaction around the threshold.

Variable	Coefficient	S.E.
Production growth rate	−0.214	(0.006)
Log of previous period employment	−0.468	(0.002)
Size 45	0.104	(0.023)
Size 46	0.118	(0.024)
Size 47	0.140	(0.024)
Size 48	0.356	(0.021)
Size 49	0.662	(0.018)
Size 50	0.071	(0.035)
Size 51	−0.128	(0.044)
Size 52	−0.057	(0.042)
Size 53	−0.068	(0.043)
Size 54	−0.068	(0.044)
Size 55	−0.029	(0.044)

Notes: This table reports the coefficient estimates of a probit that characterizes the probability of not changing employment. The dependent variable is the inaction rate. Explanatory variables are a set of dummies for last period employment between 45 and 55, the growth rate of production, last period logged employment, and a set of time and industry dummies. Standard errors are given in parentheses.

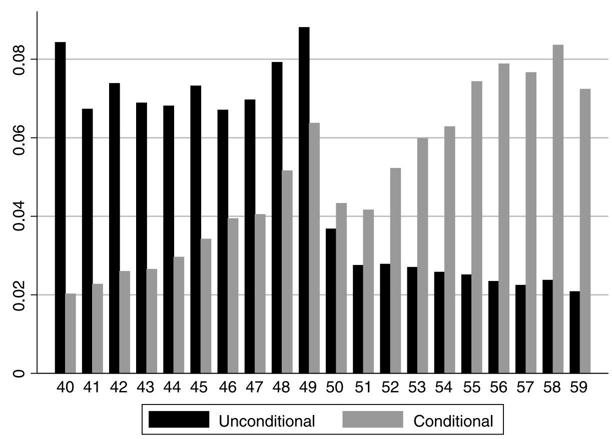

Figure 5. Distribution of employment (between 40 and 59 employees). The figure reports both the unconditional distribution (black) and the distribution conditional on having had more than 55 employees in the past (grey).

conditional distribution, there are 2.40 times more firms with 49 employees than with 50, in the conditional distribution, this ratio is only 1.47. Alternatively, to use a broader measure of discontinuity, there are 2.57 times more firms with employment in the range [47, 49] than in the range [50, 52] in the unconditional distribution, but only 1.13 times in the conditional distribution. This suggests that the sunk cost is a significant element of the regulation. These distributions will serve as target moments in our estimation.

3. Model

In this section, we introduce and solve a simple dynamic model of production and employment based on Lucas (1978). For simplicity, we assume that there is only one threshold. Firms face a regulation that requires them to pay a sunk cost the first time that their employment exceeds the threshold $\underline{n}$ and also face higher per-period costs if they currently have more than $\underline{n}$ employees. Hence, our model incorporates both types of costs, which we separately identify in our estimation.

We start with a partial-equilibrium model, which is the basis of our estimation strategy. Section 5 embeds our model of the firm in a general equilibrium framework to perform some policy experiments.

3.1 *Model assumptions*

Time is continuous and there is no aggregate uncertainty. There is a continuum of firms, which are ex ante homogeneous but differ in their realization of idiosyncratic shocks. Each firm operates a decreasing return to scale, labor-only production function,

$$y = e^z n^\alpha,$$

where $\alpha \in (0, 1)$ and e^z is the exogenous productivity (e denotes the exponential function). We assume that exit is exogenous and occurs at rate λ. We abstract from fixed costs in this problem. Given that exit is exogenous, this is without loss of generality: Fixed costs do not affect the employment decision and we do not use profits data in our estimation.

We assume that log productivity z follows a Brownian motion,

$$dz = \mu\, dt + \sigma\, dW_t.$$

This specification is attractive not only because of its tractability, but because it is consistent with two robust features of the data: (i) firm-level shocks are highly persistent, if not permanent; (ii) the firm size distribution follows a Pareto distribution. As we show below (and as is well known), the geometric Brownian motion dynamics generate a stationary distribution that is Pareto.

We also assume that all firms enter with the same productivity z_0. This simplification has little impact on our results since the moments that we target in our estimation are not sensitive to small firms (which is where the vast majority of entrants are). Employment n can be costlessly adjusted and the wage is w. For simplicity, we assume that n is a continuous choice (i.e., we do not impose indivisibility). If n is greater than $\underline{n}$, a proportional payroll tax τ applies and a fixed cost c_f has to be paid. We assume that the proportional tax applies to all employment, including that below $\underline{n}$, but this is without loss of generality, since we allow the fixed cost c_f to be negative (i.e., the tax could apply only to employment in excess of $\underline{n}$). On top of that, the first time a firm crosses the threshold $\underline{n}$, it has to pay a sunk cost F. As discussed in Section 2, this cost captures the investment necessary to comply with the regulation, including the physical cost of buying equipment, but also the informational costs of learning about the regulation and how to adapt to it.

The presence of the sunk cost makes this a dynamic optimization problem. Let $s \in \{0, 1\}$ denote whether a firm has already paid the sunk cost in the past. The state of the firm is summarized by (z, s).

3.2 *Static subproblem*

We first study the static problem to determine the firm profit function that will enter the dynamic optimization.[3] To find the optimal labor demand and profit of the firm, we first solve the firm's problem conditional on operating below the threshold, then we find the solution conditional on operating above the threshold, and finally we find the overall solution by combining these results.

The current-period profit function for a firm that operates below the threshold is

$$\pi^b(z) = \max_{0 \le n < \underline{n}} \{e^z n^\alpha - wn\}. \tag{1}$$

The superscript b stands for below the threshold. Optimal employment is

$$n^b(z) = \begin{cases} \left(\dfrac{\alpha}{w}\right)^{1/(1-\alpha)} e^{z/(1-\alpha)}, & \text{if } z < \underline{z}, \\ \underline{n}^-, & \text{if } z \ge \underline{z}, \end{cases}$$

where $\underline{z} = \log(\underline{n}^{1-\alpha} \frac{w}{\alpha})$ and $\underline{n}^-$ indicates a value just below $\underline{n}$. Profits are given by the formula

$$\begin{aligned} \pi^b(z) &= e^{z/(1-\alpha)} \left(\frac{\alpha}{w}\right)^{\alpha/(1-\alpha)} (1-\alpha) \quad \text{if } z < \underline{z} \\ &= e^z \underline{n}^\alpha - w\underline{n} \quad \text{if } z \ge \underline{z}. \end{aligned}$$

The current-period profit function for a firm that decides to operate above the threshold, and, hence, to face the regulation, is

$$\pi^a(z) = \max_{n \ge \underline{n}} \{e^z n^\alpha - w(1+\tau)n - c_f\}, \tag{2}$$

where the superscript a stands for above the threshold. The firm operates strictly above the threshold if z is greater than a cutoff value $\overline{z}$, defined as the solution to

$$e^{z/(1-\alpha)} \left(\frac{\alpha}{w(1+\tau)}\right)^{\alpha/(1-\alpha)} (1-\alpha) - c_f = e^z \underline{n}^\alpha - w\underline{n}.$$

It is easy to see that $\overline{z} > \underline{z}$, provided that there is a cost of operating above the threshold: $\tau w \overline{n} + c_f > 0$. We maintain this realistic assumption throughout the paper.

[3] This section thus does not depend on the assumption that z is a Brownian motion.

Summarizing, optimal employment if the firm decides to operate above the threshold is

$$n^a(z) = \begin{cases} \underline{n}, & \text{if } z < \overline{z}, \\ \left(\dfrac{\alpha}{w(1+\tau)}\right)^{1/(1-\alpha)} e^{z/(1-\alpha)}, & \text{if } z \geq \overline{z}. \end{cases} \tag{3}$$

This leads to profits

$$\begin{aligned} \pi^a(z) &= e^{z/(1-\alpha)} \left(\frac{\alpha}{w(1+\tau)}\right)^{\alpha/(1-\alpha)} (1-\alpha) - c_f \quad \text{if } z \geq \overline{z} \\ &= e^z \underline{n}^\alpha - w(1+\tau)\underline{n} - c_f \quad \text{if } z < \overline{z}. \end{aligned}$$

Combining our results, we can now write the firm profit, as a function of the current productivity and state $s \in \{0, 1\}$. Recall that $s = 0$ means that the firm has not paid the sunk cost and, hence, is forced to operate below the threshold, whereas a firm with $s = 1$ can choose to operate either below or above the threshold. Mathematically,

$$\begin{aligned} \pi(z, 0) &= \pi^b(z), \\ \pi(z, 1) &= \max\{\pi^a(z), \pi^b(z)\}. \end{aligned}$$

We can obtain a formula for $\pi(z, 1)$ by noting that (i) if $z < \underline{z}$, then $\pi^b(z) > \pi^a(z)$, since the firm pays lower wages and fixed costs, (ii) for $z > \underline{z}$, the firm will decide to operate above the threshold, and (iii) if $z \in (\underline{z}, \overline{z})$, it is optimal to remain just below the threshold. Hence,

$$\begin{aligned} \pi(z, 1) &= e^{z/(1-\alpha)} \left(\frac{\alpha}{w}\right)^{\alpha/(1-\alpha)} (1-\alpha) \quad \text{for } z < \underline{z} \\ &= e^z \underline{n}^\alpha - w\underline{n} \quad \text{for } \underline{z} \leq z \leq \overline{z} \\ &= e^{z/(1-\alpha)} \left(\frac{\alpha}{w(1+\tau)}\right)^{\alpha/(1-\alpha)} (1-\alpha) - c_f \quad \text{for } z > \overline{z}. \end{aligned}$$

For completeness, we also state the employment demand,

$$\begin{aligned} n(z, 1) &= \left(\frac{\alpha}{w}\right)^{1/(1-\alpha)} e^{z/(1-\alpha)} \quad \text{for } z < \underline{z} \\ &= \underline{n}^- \quad \text{for } \underline{z} \leq z \leq \overline{z} \\ &= \left(\frac{\alpha}{w(1+\tau)}\right)^{1/(1-\alpha)} e^{z/(1-\alpha)} \quad \text{for } z \geq \overline{z} \end{aligned}$$

and $n(z, 0) = n^b(z)$. Overall, firms that have never operated above the threshold $\underline{n}$ are distributed below the threshold or bunched exactly at (more precisely, just below) the threshold, while firms that have operated above $\underline{n}$ in the past will be either below, exactly at, or above the threshold. Both sunk costs and per-period costs lead to bunching at $\underline{n}$.

3.3 *Dynamic optimization*

Given the process for z and the probability of exit λ, the firm's value maximization problem can be written as choosing a stopping time T to cross the threshold. Formally, for a firm that has productivity z today, and that has not yet paid the sunk cost, we have

$$\begin{aligned} V(z,0) = \sup_T E\Bigg[\int_0^T & e^{-(r+\lambda)t}\pi(z_t,0)\,dt \\ & + \left(\int_T^\infty e^{-(r+\lambda)t}\pi(z_t,1)\,dt - Fe^{-(r+\lambda)T}\right)\Bigg]. \end{aligned} \tag{4}$$

We normalized the exit value to zero; since exit is exogenous, this is without loss of generality. Intuitively, the firm will make the switch if its productivity becomes large enough that the benefits from being large overcome the regulation costs. Denote by z^* the cutoff that triggers the firm to pay the sunk cost. Clearly, z^* is greater than $\bar{z}$: given that the evolution of productivity z is uncertain, the firm will delay paying the sunk cost rather than invest as soon as it expects the investment to be just profitable in the present discounted value sense.

The solution of the model can be obtained directly using the results in Stokey (2008), since our model is a special case of the general option exercise problem analyzed in this book. First, we rewrite the problem explicitly as choosing a cutoff z^*, given the current value z:

$$V(z,0) = \sup_{z^* \geq z} E_z\left[\int_0^{T(z^*)} e^{-(r+\lambda)t}\pi(z_t,0)\,dt + e^{-(r+\lambda)T(z^*)}\big(V\big(z^*,1\big) - F\big)\right], \tag{5}$$

with

$$V\big(z^*,1\big) \equiv E_{z^*}\left[\int_0^\infty e^{-(r+\lambda)t}\pi(z_t,1)\,dt\right].$$

In these expressions, E_z denotes the expectation, conditional on $z_0 = z$. The next proposition derives the optimal policy. Denote R_1 and R_2 as the roots of the quadratic $\frac{\sigma^2}{2}R^2 + \mu R - (\lambda + r) = 0$, that is, with $J = \sqrt{\mu^2 + 2(r+\lambda)\sigma^2}$, we have $R_1 = \frac{-\mu - J}{\sigma^2} < 0$ and $R_2 = \frac{-\mu + J}{\sigma^2} > 0$.

Proposition. *The solution to the firm problem* (*equation* (5)) *is* z^*, *the unique value that satisfies*

$$-R_1 \int_{\bar{z}}^{z^*} e^{R_1(z^*-z)}\big[\pi^a(z) - \pi^b(z)\big]\,dz = (r+\lambda)F. \tag{6}$$

See the Appendix for the proof.

The intuition for this proposition is that the firm equates the marginal benefit and marginal cost of waiting to make the investment. The marginal benefit is that the firm avoids paying the cost early, which is attractive given positive discounting and the risk of exit. The left-hand side captures the marginal cost: the firm gives up the increase in profit from operating above the threshold. The integral captures the expected time spent between $\overline{z}$ (at which point it is "statically" profitable to operate above the threshold) and z^*. In the language of Stokey (2008), R_1 discounts the time the process z will spend between $\overline{z}$ and z^*. For given parameters $\{\alpha, \overline{n}, \mu, \sigma, \tau, c_f, F, r, \lambda\}$, this equation allows us to find z^* numerically easily.

An alternative solution method, which does not rely on the results of Stokey (2008), is to write the Hamilton–Jacobi–Bellman equations, value matching and smooth pasting conditions satisfied by the value function. Briefly, we have

$$(r+\lambda)V(z,1) = \pi(z,1) + \mu V_z(z,1) + \frac{\sigma^2}{2}V_{zz}(z,1) \tag{7}$$

for any z and

$$(r+\lambda)V(z,0) = \pi(z,0) + \mu V_z(z,0) + \frac{\sigma^2}{2}V_{zz}(z,0) \tag{8}$$

for $z < z^*$.[4] The boundary conditions are given by value matching,

$$V(z^*,1) = V(z^*,0) - F, \tag{9}$$

and by the smooth pasting condition,

$$V_z(z^*,1) = V_z(z^*,0). \tag{10}$$

Given the expressions of $\pi(z,1)$ and $\pi(z,0)$ found above, it is straightforward to solve this system of differential equation for $V(z,1)$, $V(z,0)$, and z^*. The Appendix provides the algebra.

We conclude this subsection by noting some intuitive comparative statics: higher uncertainty, higher sunk costs, or higher fixed costs all make it optimal to wait longer before crossing the threshold. This is the standard real option effect.

Corollary. *The cutoff z^* is increasing in σ^2, F, τ, c_f, and $\underline{n}$.*

Proof. Differentiation of equation (6) gives the results. □

3.4 *Stationary distribution*

Given our interest in the size distribution, we derive the joint cross-sectional distribution over (z, s) in closed form. Denote the probability density function as $f(z, s)$. Recall that firms enter with $z = z_0$ and then z evolves according to a Brownian motion with parameters (μ, σ). Firms switch from $s = 0$ to $s = 1$ as soon as z reaches z^* and exit upon the realization of a Poisson process with parameter λ. We can write the Kolmogorov forward equation, which reflects the conservation of the total number of firms, net of exit,

$$-\mu\frac{\partial f(z,0)}{\partial z} + \frac{\sigma^2}{2}\frac{\partial^2 f(z,0)}{\partial z^2} = \lambda f(z,0), \tag{11}$$

[4]Note that $\pi(z,0)$ and $\pi(z,1)$ are not C^1 (continuously differentiable): the derivative is discontinuous at $z = \underline{z}$ for $\pi(\cdot,0)$ and $\pi(\cdot,1)$, and at $z = \overline{z}$ for $\pi(\cdot,1)$.

which holds for all $z < z_0$ and all $z \in (z_0, z^*)$. (See Dixit and Pindyck's (1994) Appendix to Chapter 3, for a heuristic derivation, and Chapter 8 for an application similar to our case.) The equation need not hold for $z = z_0$, since there is entry of new firms.

The same equation applies to firms that have made the switch,

$$-\mu \frac{\partial f(z,1)}{\partial z} + \frac{\sigma^2}{2} \frac{\partial^2 f(z,1)}{\partial z^2} = \lambda f(z,1), \tag{12}$$

which holds for all $z \in (-\infty, z^*)$ and for all $z \in (z^*, +\infty)$, but not for $z = z^*$.

Last, we need to state the boundary conditions. The first one is simply the requirement that f is a density, that is,

$$\int_{-\infty}^{+\infty} f(s,1)\, ds + \int_{-\infty}^{+\infty} f(s,0)\, ds = 1.$$

To derive the other boundary conditions, the easiest approach is to approximate the Brownian motion with a discrete random walk, as in Dixit and Pindyck (1994). This yields the conditions

$$f\left(z^*, 0\right) = 0,$$

and $f(\cdot, 0)$ must be continuous at z_0, while $f(\cdot, 1)$ must be continuous at z^*:

$$\lim_{s \to z_0^-} f(s,0) = \lim_{s \to z_0^+} f(s,0),$$
$$\lim_{s \to z_-^*} f(s,1) = \lim_{s \to z_+^*} f(s,1).$$

Finally, a stationarity condition holds for $z = z^*$, reflecting that the number of firms that reach z^* and have $s = 0$ is equal to the number of firms that enter at $s = 1$ with $z = z^*$, and is equal to the number of firms with $s = 1$ that exit in any time period: this leads to

$$-\frac{\sigma^2}{2} f'\left(z^*, 0\right) = \lambda \int_{-\infty}^{\infty} f(s,1)\, ds.$$

Given these boundary equations, solving for the cross-sectional distribution involves some simple algebra, which is relegated to the Appendix. The result is

$$\begin{aligned} f(z,0) &= \frac{\beta_1 \beta_2}{\beta_1 - \beta_2} \left(e^{\beta_2 (z - z_0)} - e^{\beta_1 (z^* - z_0)} e^{\beta_2 (z - z^*)}\right) \quad \text{for } z < z_0 \\ &= \frac{\beta_1 \beta_2}{\beta_1 - \beta_2} \left(e^{\beta_1 (z - z_0)} - e^{\beta_1 (z^* - z_0)} e^{\beta_2 (z - z^*)}\right) \quad \text{for } z^* > z > z_0 \end{aligned}$$

and

$$\begin{aligned} f(z,1) &= \frac{\beta_1 \beta_2}{\beta_1 - \beta_2} e^{\beta_1 (z^* - z_0)} e^{\beta_2 (z - z^*)} \quad \text{for } z < z^* \\ &= \frac{\beta_1 \beta_2}{\beta_1 - \beta_2} e^{\beta_1 (z - z_0)} \quad \text{for } z > z^*. \end{aligned}$$

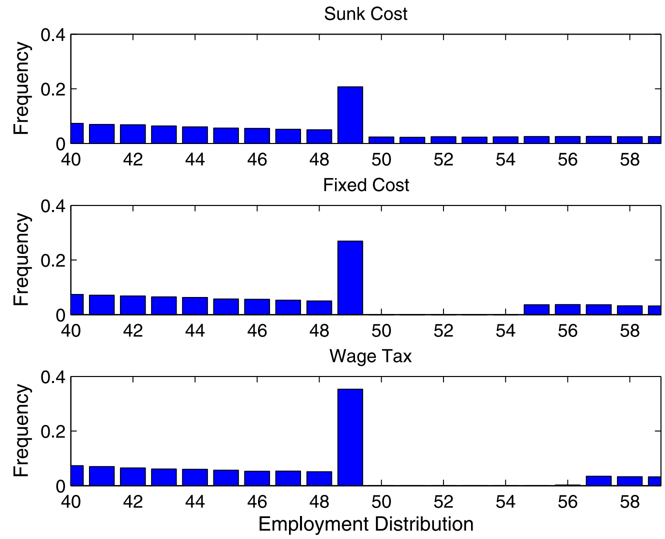

FIGURE 6. Distribution of firm employment between 40 and 59 in the model without measurement error. Top panel: sunk cost only; middle panel: fixed cost per period only; bottom panel: payroll tax only.

This expression implies that z has an exponential distribution in the upper tail. Since log employment and log sales are both proportional to z, employment and sales follow Pareto distributions, and the probability density function (p.d.f.) of employment (or sales) is proportional to n to the power $\beta_1(1-\alpha)-1$.[5]

Figure 6 illustrates some properties of the firm size distribution implied by our model. In the absence of any regulation, this size distribution is Pareto throughout, whereas in our model, it is only Pareto for n large enough. The bottom panel depicts the distribution with a per-period payroll tax. There is a substantial "hole" in the distribution with no firms whatsoever between 50 and 55 employees. This figure presents an empirical challenge, because in the data there are many firms that have an employment level slightly greater than 49. It would be incredible to attribute the presence of all these firms to measurement error. The middle panel shows the distribution if the per-period payroll tax is replaced with a per-period fixed cost. This figure is similar to the bottom panel, which reflects that the two types of per-period costs (fixed cost or payroll tax) lead to the same implications for the employment distribution. Unless one uses data on productivity or profits, it seems extremely difficult to distinguish the two. In our empirical work, we focus on the case of a payroll tax, because one provision of the law explicitly implies higher payroll taxes.

Finally, the top panel shows the impact of a sunk cost on the firm size distribution. The sunk cost model does not suffer from the same deficiency as the per-period cost model: there are no holes in the distribution and, in particular, some firms have exactly

[5]Note that this implies some restrictions on β_1 to ensure that employment be finite. This in turn restricts the parameters μ, λ, and σ^2. Our estimated parameters satisfy these restrictions, so we do not need to impose them in practice.

50 employees. These are firms that crossed the threshold in the past, were subsequently hit by negative productivity shocks, and consequently decided to downsize.

To establish the economic relevance of these regulations, we now turn back to the data and propose a simple structural estimation of our model.

4. Estimation

This section proposes a simple estimation of our model using indirect inference. We take advantage of our closed form solutions, which make calculating model moments computationally easy.

As discussed below, we incorporate classical measurement error in (log) employment; the standard deviation of measurement error is σ_{mrn}. Table 3 lists our parameters. The full set of structural parameters is the vector $\theta = (r, w, \alpha, z_0, \lambda, \mu, \sigma, \tau, F, \sigma_{\text{mrn}})$. We partition this vector into two vectors, that is, $\theta = (\theta_p, \theta_e)$, where $\theta_p = (r, w, \alpha, z_0)$ includes parameters that are set a priori, and $\theta_e = (\lambda, \mu, \sigma, \tau, F, \sigma_{\text{mrn}})$ is the vector of estimated parameters.

Like calibration, indirect inference works by selecting a set of statistics of interest, which the model is asked to reproduce.[6] These statistics are called sample auxiliary parameters $\hat{\Psi}$ (or target moments). For an arbitrary value of θ_e, we use the structural model to generate S statistically independent simulated data sets and compute simulated auxiliary parameters $\Psi^s(\theta_e)$. The parameter estimate $\hat{\theta}_e$ is then derived by searching over the parameter space to find the parameter vector that minimizes the criterion function,

$$\hat{\theta}_e = \arg\min_{\theta_e \in \Theta_e} \left(\hat{\Psi} - \frac{1}{S} \sum_{s=1}^{S} \Psi^s(\theta_e) \right)' W \left(\hat{\Psi} - \frac{1}{S} \sum_{s=1}^{S} \Psi^s(\theta_e) \right),$$

where W is a weighting matrix and Θ_e is the space of estimated parameters. This procedure generates a consistent estimate of θ_e. Following Blundell, Pistaferri, and Preston (2008), we use a diagonal weighting matrix $W = \text{diag}(V^{-1})$, where V is the variance–covariance matrix of the sample auxiliary parameters. This weighting scheme allows for

Table 3. Economic parameters.

Parameters	Definition	
r	Interest rate	Fixed
α	Curvature profit function	Fixed
z_0	TFP level at entry	Fixed
w	Wage	Normalized
λ	Exit probability	Estimated
μ	Drift productivity	Estimated
σ	Std. dev. innovation productivity	Estimated
τ	Payroll tax above $\underline{n}$	Estimated
F	Sunk cost	Estimated
σ_{mrn}	Measurement error	Estimated

[6]See Gourieroux, Monfort, and Renault (1993) for a general discussion of indirect inference.

heteroskedasticity and it has better finite sample properties than the optimal weighting matrix (see Altonji and Segal (1996)). The minimization is performed using the Nelder–Mead simplex algorithm. We used 5000 different starting values to find the global minima. To simulate the model, we draw from the stationary distribution derived in the previous section and use standard approximations to simulate the Brownian motion.

The standard errors are obtained using 500 bootstrap repetitions. In each bootstrap repetition, a new set of data is produced by randomly selecting blocks of observations.[7] We use parametric bootstrap since we draw observations from the fitted model. In the bth bootstrap repetition, auxiliary parameters $\hat{\Psi}^b$ are calculated using the new set of data. An estimator $\hat{\theta}_e^b$ is found by minimizing the weighted distance between the recentered bootstrap auxiliary parameters $(\hat{\Psi}^b - \hat{\Psi})$ and the recentered simulated auxiliary parameters $(\frac{1}{S}\sum_{s=1}^S \Psi^s(\theta_e^b) - \frac{1}{S}\sum_{s=1}^S \Psi^s(\hat{\theta}_e))$:

$$\hat{\theta}_e^b = \arg\min_{\theta_e^b \in \Theta_e} \left((\hat{\Psi}^b - \hat{\Psi}) - \left(\frac{1}{S}\sum_{s=1}^S \Psi^s(\theta_e^b) - \frac{1}{S}\sum_{s=1}^S \Psi^s(\hat{\theta}_e) \right) \right)' \times W \left((\hat{\Psi}^b - \hat{\Psi}) - \left(\frac{1}{S}\sum_{s=1}^S \Psi^s(\theta_e^b) - \frac{1}{S}\sum_{s=1}^S \Psi^s(\hat{\theta}_e) \right) \right).$$

4.1 *Predefined parameters*

Some parameters are not estimated because they are either normalizations or are fairly standard. The wage rate is normalized to 1. We set the real interest rate r to 5%. We assume that α equals 0.66, as in Cooper, Haltiwanger, and Willis (2007). This parameter is a reduced form for the labor share, decreasing returns to scale and the elasticity of demand. There is limited agreement on this parameter; hence, we provide further discussion and comparative statics in Section 6.1. Finally, the parameter z_0 is set so that the average firm has 7.5 employees, as is the case in France.[8]

4.2 *Measurement error*

Our data are based on administrative sources and, hence, are of relatively high quality. Nevertheless, there is likely to be some measurement error in our employment variable. We explicitly introduce measurement error into the simulated moments to mimic the bias these impute into the actual data moments. We do so by multiplying our model employment n_{it} by a measurement error factor mrn_{it} that is independent and identically distributed (i.i.d.) over firm and time, and follows a log-normal distribution with mean $-\frac{1}{2}\sigma_{\mathrm{mrn}}^2$ and standard deviation σ_{mrn}.[9]

[7] See Hall and Horowitz (1996) for more details on the block bootstrap. The sampling is random across firms, but is done in block over the time dimension.

[8] This parameter is largely irrelevant for our estimation, because the moments that we target are not much affected by new firms, as long as they enter with less than 50 employees.

[9] We experimented with a measurement error represented as the difference of two Poisson distributed random variables, which has the advantage that measurement error has to be an integer, but we found very similar results.

Measurement error also helps capture model misspecification, which can take several forms. First, our measure of employment is the arithmetic average of the number of employees at the end of each quarter. This is the relevant measure of employment for some but not all of the regulations. For instance, some regulations are based on employment measured in full-time equivalent and some other regulations apply if there are more than 50 employees in the firm for more than 12 months. Second, measurement error also captures adjustment cost or search frictions that lead to an imperfect control of the size of the workforce.

Last, we transform our model simulated data by rounding to the closest integer.[10]

4.3 *Auxiliary parameters and identification*

Table 5 lists our auxiliary parameters (target moments), which can be divided into three groups. We set to match (i) the average and volatility of growth in employment, and the slope of the power law; (ii) the distribution of employment around the threshold, as approximated by the density of firms between 40 and 46 employees, between 47 and 49 employees, between 50 and 52 employees, and between 53 and 59 employees;[11] (iii) the distribution of employment around the threshold, summarized in the same way, conditional on the firm having had employment above 55 at any point previously in our data. The rationale for the first group of moments (i) is that we want the model to be consistent with key features of firm dynamics. The rationale for the unconditional distribution (ii) is that we want to reproduce well the discontinuity in the firm size distribution, which is the prima facie evidence that the regulation matters. The rationale for the conditional distribution (iii) is that it allows us to distinguish sunk costs from per-period costs. We use a threshold of 55 employees to make this statistic robust to measurement error. Given the normalization of the distributions, we have a total of nine moments, compared to six estimated parameters.

Identification of the model's parameters is achieved by a combination of functional form and distributional assumptions, and is difficult to prove, but the intuition is straightforward. First, the mean employment growth is informative about the drift μ. The variance of employment growth is informative about the variance of productivity shocks σ and the variance of measurement error σ_{mrn}. The slope of the power law is informative regarding the variance of productivity shocks σ, the drift μ, and the exit rate λ. The unconditional distribution is informative regarding the friction parameters τ and/or F and the variance of measurement error σ_{mrn}. Last, the conditional distribution provides an independent source of information on τ, F, and σ_{mrn}.

4.4 *Estimation results*

Table 4 reports the structural parameters estimates together with estimated standard errors. The first row reports the results for the full model, while the second row reports results that assume there are only sunk costs, and the third row assumes that there are only per-period costs. Table 5 evaluates the fit of these three variants of our model.

[10] We set $\underline{n} = 49.01$ so that firms find it costly to be strictly above 49 employees.

[11] The distribution is the number of firms in each bin divided by the length of the bin (7 or 3), and further divided by the total number of firms between 40 and 59.

TABLE 4. Model estimation: parameters.

	λ	μ	σ	σ_{mrn}	F	τ
Full model	0.0301	0.0030	0.0520	0.0211	1.0281	0.0004
	(0.0027)	(0.0004)	(0.0028)	(0.0023)	(0.0758)	(0.0001)
Sunk cost F only	0.0293	0.0029	0.0527	0.0134	1.2908	
	(0.0023)	(0.0004)	(0.0024)	(0.0015)	(0.0531)	
Payroll tax τ only	0.0302	0.0032	0.0504	0.0324		0.0015
	(0.0028)	(0.0004)	(0.0029)	(0.0026)		(0.0001)

Notes: This table reports the parameters obtained using the method of moments estimator, with the associated standard errors, for the full model, the model with only the sunk cost, and the model with only the payroll tax.

TABLE 5. Auxiliary parameters.

	Data	S.E.	Full Model	Sunk Cost F Only	Payroll Tax τ Only
Mean $\Delta \log n$	0.0082	0.0001	0.0081	0.0081	0.0082
Std. dev. $\Delta \log n$	0.156	0.0001	0.156	0.1562	0.1561
Power law coefficient	2.2522	0.0056	2.2754	2.2493	2.2863
Density of firms in each bin, unconditional					
40–46	0.0718	0.0008	0.0654	0.0653	0.0666
47–49	0.0790	0.0008	0.0849	0.0852	0.0783
50–52	0.0307	0.0005	0.0335	0.0352	0.0341
53–59	0.0240	0.0005	0.0267	0.0260	0.0281
Density of firms in each bin, conditional					
40–46	0.0284	0.0016	0.0284	0.0293	0.0306
47–49	0.0519	0.0021	0.0599	0.0442	0.0818
50–52	0.0457	0.002	0.0429	0.0563	0.0464
53–59	0.0726	0.0024	0.0704	0.0705	0.0573
Residuals			224.51	274.09	457.18

Notes: This table reports the target moments, evaluated at the estimated parameter values, and the minimized criterion.

Overall, our data are consistent with a regulation that acts like a sunk cost of slightly more than 1 year of a worker's wage, and a small, but significant proportional payroll tax of 0.04%. Shocks to total factor productivity are estimated to be 5% per year. The mean growth of productivity is small, consistent with the small mean employment growth in our data, and the estimated exit rate is around 3% per year, to fit the Pareto distribution. The full model requires a measurement error of around 2% or, on average, one worker around the threshold. In spite of its parsimony, the model is able to reproduce reasonably well all the targeted moments, and, in particular, the discontinuities in both the unconditional and conditional distributions. A graphical illustration is provided in the bottom panel of Figure 7.

Turning to the restricted versions of the model that have only the sunk cost or the tax, we first note that the parameters that summarize firm dynamics (μ, σ, λ) are reassuringly stable. The model with only the sunk cost delivers a slightly higher estimate of

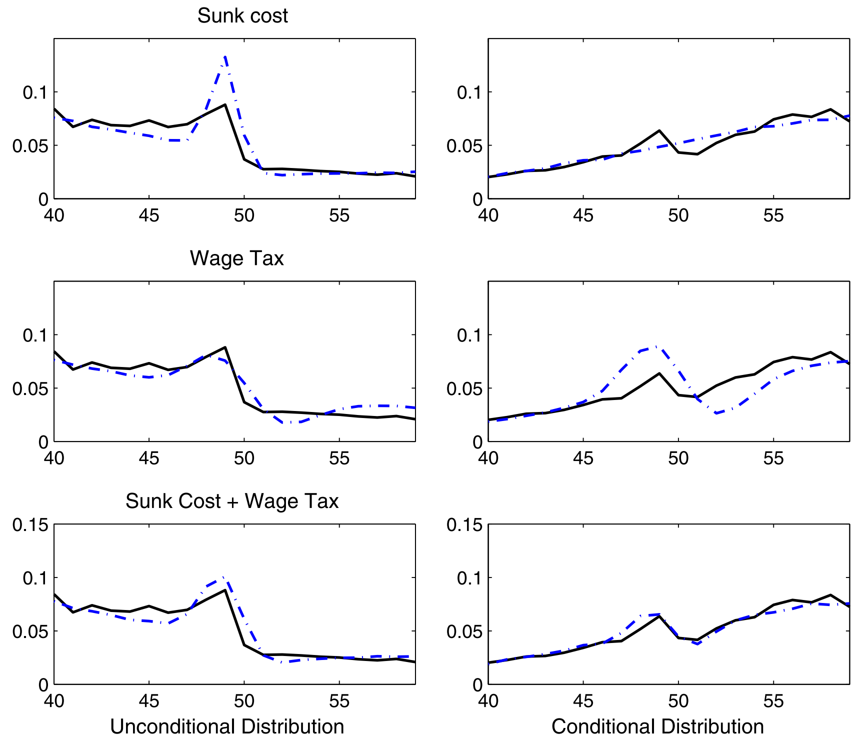

FIGURE 7. Distribution of firm employment (between 40 and 59 employees) in the data and in different variants of the model. The full line is the data; dashed lines represent different models. The left column is the unconditional distribution; the right column is the conditional distribution. The top panel is the model with sunk cost only; the middle panel is the model with payroll tax only; the bottom panel is the full model with both sunk cost and payroll tax. The distributions are normalized by the total number of firms between 40 and 59 employees.

the cost, and the fit is only mildly worse. The main difference, clearly visible in Figure 7, is that this model has more difficulty fitting the conditional distribution. The intuition is that in the absence of any per-period cost and without any measurement error, the conditional distribution should not have any spike at 49. In principle, measurement error could help, because some firms that are classified as having been above 55 employees in the past were never actually above 55 employees and, consequently, some remain bunched at 49. However, this mechanism does play an important role because of the small estimated amount of measurement error. Our small proportional tax helps reconcile the model, and the data as can been seen by comparing the conditional distributions in the fourth and fifth columns of Table 5 (or the top and bottom panels of Figure 7).

The model with the tax only fits much worse: the minimized criterion is twice larger and the model cannot fit well the discontinuity in the two distributions, as seen in Figure 7. The tax is estimated to be larger, around 0.15%. However, this value is lower than the taxes that are actually set in the law, which presents an apparent puzzle. One possible interpretation is that some of these regulations are indeed not as costly as they appear and represent benefits that are valued by the workers. Measurement error is larger,

Table 6. Firm size distribution in the model and the data.

	Fraction of Firms			Employment Share		
	>200	>500	>1000	>200	>500	>1000
Data	0.4711	0.1615	0.0674	0.8326	0.6191	0.4722
Model	0.4097	0.1259	0.0516	0.8193	0.6296	0.5158

Notes: This table reports the number of firms (resp. the employment in firms) with more than 200, 500, or 1000 employees, divided by the number of firms (resp. employment) with more than 100 employees, in the data and in the full model.

because this is the only way the model can generate a nonempty (conditional or unconditional) distribution to the right of 50. However, measurement error leads to counterfactual implications, such as a high number of firms with 50 employees, and the shape of the distribution does not match the data (Figure 7).

We finally examine the ability of the model to account for the large firms' size distribution. Table 6 reports the number of firms (and the total employment in firms) above 200, 500, and 1000 employees, normalized by the number of firm with more than 100 employees (resp. the total employment in firms with more than 100 employees). Although these moments are not directly targeted in the estimation, our model does a reasonable job. The model overestimates somewhat the share of large firms—a limitation that can be traced back to the failure of Zipf's law for very large firms.

5. Policy experiments

In the previous section, we estimated the regulatory cost as perceived by firms. In this section, we use our estimates to infer the aggregate effect of the regulation on output, employment, and average productivity. From the point of view of a social planner, the regulation misallocates labor across firms and, hence, reduces productivity. Moreover, the regulation affects the incentives of firms to enter. To demonstrate this, we embed our partial equilibrium estimation into a general equilibrium framework and use it to simulate the response of the economy if the regulation were to be removed. We first discuss the general equilibrium framework for our experiments; then we present and discuss the results.

5.1 *General equilibrium framework*

We incorporate endogenous entry and labor supply to the model by embedding our firm dynamics in a general equilibrium framework as in Hopenhayn and Rogerson (1993). Since this model is well known, we describe it only briefly here. First, there is a representative agent with utility function

$$\int_0^\infty e^{-\rho t}\left(\log(C_t) - B\frac{N_t^{1+\phi}}{1+\phi}\right) dt.$$

Here B reflects the preference for leisure and ϕ is the inverse of the Frisch labor supply elasticity. This agent supplies work to the market at wage w_t and buys or sells assets at interest rate r_t. In equilibrium, the only assets are the firms. We consider a steady-state stationary equilibrium: there is no aggregate variation, since a law of large numbers applies, and macroeconomic aggregates are constant. As a result, the interest rate is constant, $r_t = r = \rho$.

For a given wage, we can solve the value function $V(z, s; w)$, the employment policy function $n(z, s; w)$ and the threshold for paying the sunk cost, $z^*(w)$, and the stationary distribution $f(z, s; w)$ as in Section 3. We have added the wage as an explicit argument to these functions to emphasize the dependence. Firms are assumed to enter at a cost k. Since all firms enter with a productivity z_0, the free entry condition reads

$$k = V(z_0, 0; w). \tag{13}$$

Denote the flow of firms that enter per unit of time as E and denote as $Mf(z, s; w)$ the stationary distribution of firms, where M is the mass of firms and f is the probability density function. With exogenous exit at rate λ, the flow of entrants per unit of time E must equal λM in a stationary equilibrium.

Total output is then given by

$$Y = M \sum_{s=0}^{1} \int_{-\infty}^{\infty} e^z n(z, s; w)^\alpha f(z, s; w)\, dz \tag{14}$$

and total labor is

$$N = M \sum_{s=0}^{1} \int_{-\infty}^{\infty} n(z, s; w) f(z, s; w)\, dz. \tag{15}$$

Labor supply satisfies the first order condition

$$BCN^\phi = w \tag{16}$$

and the goods market equilibrium reads

$$C + Ek + \tau w \int_{-\infty}^{\infty} n(s, 1; w) f(s, 1; w)\, ds + \lambda F \int_{-\infty}^{\infty} f(s, 1; w)\, ds = Y. \tag{17}$$

A stationary equilibrium is given by $\{Y, C, E, M, w, N\}$ such that $E = \lambda M$ and the equations (13)–(17) are satisfied, and the value function, policy function, and cross-sectional distribution $V(z, s)$, z^*, $f(z, s)$ are obtained as in Section 3.

In this model, the free entry condition pins down the equilibrium wage. Given this wage, the number of firms adjusts the scale of the economy so that labor demand equals labor supply; that is, there is a perfectly elastic supply of firms.

We close by mentioning three issues. First, we need to take a stand on whether the regulation cost is a real resource cost (that must be deducted from the resource constraint, as we assumed in equation (17)) or is a transfer (which is rebated lump sum to households and, hence, disappears from equation (17)). In reality it is likely that both components are present. Hence we present the results for the two possible assumptions. Second, we only calculate steady states and abstract from transitional dynamics.

We believe this is appropriate to examine the long-run effects of the regulation on employment and productivity, which are the key questions of interest, but of course this makes the welfare comparison inaccurate. Third, our calculations have little to say on the desirability of the regulations themselves since we do not model the benefits of the regulation. The goal for us it to understand if there are significant costs to the regulation.

5.2 *Results*

We first discuss the calibration of the macroeconomic parameters; then we present our results.

5.2.1 *Calibration* We set the entry cost k and the initial productivity z_0 so that (i) the wage is normalized to 1, as assumed in our estimation, and (ii) the average firm size matches the French data (7.5 employees per firm). We also need to parametrize labor supply preferences. We set an elasticity of labor $\phi = 1$ (see Chetty (2012) for a discussion) and set B such that total employment is 0.25. These are standard values in the macroeconomics literature.

5.2.2 *Results* Table 7 presents the results of our experiments. This table reports the full model result in the bottom row as well as some partial results that are helpful in understanding the mechanism. Specifically, the third row considers the case where labor supply is perfectly inelastic, and the first and second rows assume that entry is inelastic. In this case, we perform the experiment as follows: starting from the equilibrium with the regulation and with elastic entry, we remove the regulation and calculate the equilibrium, discarding the free entry condition and simply assuming that the number of firms M remains constant. The first and second rows differ in the assumed labor supply elasticity.

The first row can be interpreted as the pure productivity gain from reallocation; that is, how much of an increase in output can we obtain, holding total employment constant, simply by reallocating labor across firms. This is the solution to the allocation problem,

$$Y(N) = \max_{\{n(z)\}_{z=-\infty}^{\infty}} \int_{-\infty}^{\infty} e^z n(z)^\alpha f(z)\, dz$$

$$\text{s.t.} \quad \int_{-\infty}^{\infty} n(z) f(z)\, dz \leq N,$$

Table 7. Policy experiments: steady-state effect of removing the regulation.

Experiment	Y	N	w	M	C
Inelastic labor, inelastic entry	0.2720	0	0.2936	0	0.2878
Elastic labor, inelastic entry	0.2738	0.0028	0.2927	0	0.2898
Inelastic labor, elastic entry	−0.0191	0	0.0025	−0.8514	0.0295
Elastic labor, elastic entry	−0.0326	−0.0135	0.0025	−0.8647	0.0160

Notes: This table reports the steady-state percentage change in output, employment, the real wage, the number of firms, and consumption if the sunk cost and the tax are both eliminated.

where $n(z)$ is the employment of firms with productivity z. The gain in total output, holding total labor constant, is 0.27%, which is significant.

One way to understand this experiment is to decompose it into two steps. In the first step, the regulation is removed, but the wage is kept constant; in the second step, the wage adjusts (upward) so as to bring employment back to its initial level. In the first step, output and employment increase by 0.84% and 0.87% respectively, as many medium-sized firms grow by going over the threshold and, hence, demand more labor. Average labor productivity falls sightly, as many firms that were previously constrained in their employment are now able to increase it. In the second step, the wage rises and reduces the labor demand of both very large firms and very small firms, which end up shrinking. We note that this result goes some way toward addressing the observation that France has relatively less medium-sized firms than comparable countries (see Bartelsman, Haltiwanger, and Scarpetta (2009) or Bartelsman, Scarpetta, and Schivardi (2005), among others).

The second row reports the results if one allows labor supply to adjust; this has essentially no effect on the results. Labor rises slightly due to lower taxes, but the higher effective productivity of the economy has no direct effect on labor supply, since these preferences are compatible with balanced growth (hence a pure increase in productivity has perfectly offsetting income and substitution effects).

The third and fourth rows allow for elastic entry. Perhaps surprisingly, but consistent with Fattal Jaef (2012), this yields quite different results. Allowing the number of firms to adjust reduces dramatically—even overturns—the steady-state output gains. Since firms close to the threshold can grow, the economy needs fewer firms, which economizes on entry costs. Overall, output actually falls slightly in the new steady state, but the reduced entry costs imply that consumption rises. If labor supply is elastic, the wealth gains from removing the threshold further reduce labor supply and output. However, this effect is fairly small.

Table 8 presents some additional policy experiments to better understand the results. (These experiment assume elastic labor and elastic entry, as in the fourth row of Table 7.) First, our results are somewhat smaller if the regulation cost is a transfer rather than a real resource cost. The main difference is a smaller wealth effect, leading to smaller changes in consumption and employment. Second, one might ask, "How

Table 8. Variants on the policy experiments.

Experiment	Y	N	w	M	C
Benchmark	−0.0326	−0.0135	0.0025	−0.8647	0.0160
Regulations are transfers	−0.0396	−0.0056	0.0025	−0.8569	0.0081
Entry cost is in labor units	0.0238	−0.0221	0.0016	−0.8105	0.0238
Apply the regulation to all firms	−2.5039	−0.0335	−2.4502	−7.8654	−2.4565
Apply the regulation above 75 employees	−0.0063	−0.0020	0.00134	−0.1711	0.0033
Remove only the sunk cost	−0.0269	−0.0170	0.0019	−0.8105	0.0188

Notes: This table reports the steady-state percentage change in output, employment, the real wage, the number of firms, and consumption if the sunk cost and the tax are both eliminated, for some variants of the model.

much of the efficiency gain can be achieved by extending the threshold to 75 employees rather than 50?" The answer is, not much: consumption (output net of entry and regulation costs) rises by only about 0.0033%, about one-fifth of what is obtained by removing the regulation altogether.[12] Third, the motivation for the phase-in of the regulation at 50 employees is that it is too costly to impose the compliance cost on small firms. We can evaluate this argument by considering the counterfactual: What would happen if all firms were subject to the regulation? With free entry, this would have dramatic effects on the number of firms. For instance, the effect of imposing the sunk cost on everyone would be to reduce output by 2.50%, with the number of firms declining by a whooping 7.86%. It is safe to say, then, that applying the regulation to all firms would be quite costly, which suggests that the phase-in is perhaps not such a bad policy. We also consider a variant of the model where entry costs are not paid in terms of goods but in terms of labor. In this case, the reduction in the number of firms is smaller, leading to a higher total output. Finally, as shown in the last row, and not surprisingly given our estimates, the vast majority of the gains from removing the regulation comes from removing the sunk cost, rather than the payroll tax. We found that all the results discussed in this table are robust.

One criticism of these experiments is that the free entry assumption is too extreme. In this spirit, Figure 8 presents the results where we vary the elasticity of supply of firms. To do so, we extend this model by relaxing the assumption that entry is perfectly elastic at cost k. To generate an upward-sloping supply of entrants to the economy and we sup-

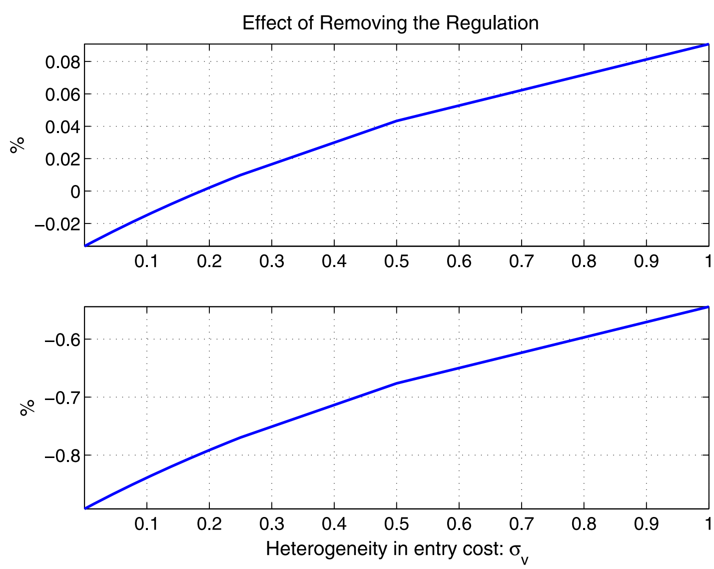

FIGURE 8. Comparative statics of policy experiments: steady-state percentage change in total output (top panel) and in the number of firms (bottom panel) when the regulation is removed, as a function of the standard deviation of entry costs.

[12] Of course, if the threshold is pushed sufficiently high, the gains converge to those obtained by fully eliminating the thresholds; but this convergence is slow.

pose that in each period there is a pool N of potential entrants, which differ in their entry cost. The entry cost is distributed according to the cumulative distribution function H. In a given period, only potential entrants with an entry cost below $V(z_0, 0; w)$ will enter. Denote by k^* the threshold value for k: $V(z_0, 0; w) = k^*$. Then the flow of entrants E is $NH(k^*)$.

We parametrize the cumulative distribution function (c.d.f.) H as a log-normal distribution with standard deviation σ_v. This parameter captures the heterogeneity of entry costs and, hence, the (inverse) elasticity of supply of entrants. For each value of σ_v, we recalibrate the model and run the policy experiments. Figure 8 shows that as we reduce σ_v, the results approach the fourth row, where entry is perfectly elastic: there is a large decline in the number of firms M and a smaller increase, or even a decrease, in output Y. As we increase heterogeneity in entry costs σ_v and, hence, reduce the elasticity of firms, we see a smaller reaction in the number of firms and a larger increase in output. In the limiting case, we go back to the second row of the table. It is, however, difficult to pin down a realistic value for σ_v from cross-section data alone.

6. Robustness

This section first discusses how our results are affected by some parameter choices, in particular, the return to scales α, then provides estimates of the cost of the regulation by sector, and finally compares our results to those of the contemporaneous, closely related study of Garicano, Lelarge, and Van Reenen (2013).

6.1 *Sensitivity analysis*

We do not estimate the curvature of the production function α, which is difficult to identify in our data. There is substantial disagreement on the value of this parameter, which captures not only the labor share, but also decreasing returns to scale and the elasticity of demand if firms' outputs are imperfect substitutes. Finally this parameter must be adjusted to account for the absence of capital in our model. If capital is flexible, the elasticity of demand is infinite, and there is constant return to scale, then α should equal 1. Relaxing any of these assumptions leads to a lower α.[13] Rather than defending very strongly a particular value of α, in this section, we report (in Table 9) the results from estimating our benchmark model for different values of α.

With higher α, the effect of productivity shocks on employment is amplified. As a result, the model with higher α requires smaller shocks to match the observed volatility of employment growth. The effect on the estimated cost of the regulation is ambiguous: on the one hand, the benefits to growing are larger, which means the estimation should likely require a larger cost to fit the observed inaction; on the other hand, the shocks

[13] Specifically, if the elasticity of demand is $1/\varepsilon$ and the production function is $y = zk^\alpha n^\nu$, and if capital is flexible, the reduced-form profit function is $\pi(z, n) = z^{(1-\varepsilon)/(1-\alpha(1-\varepsilon))} n^{\nu(1-\varepsilon)/(1-\alpha(1-\varepsilon))}$, while if capital is fixed, $\pi(z, n) = z^{1-\varepsilon} n^{\nu(1-\varepsilon)}$ (in both cases, up to a multiplicative constant). If $\varepsilon = 0$ (perfectly elastic demand) and $\alpha + \nu = 1$ (constant returns to scale), and capital is flexible, then the reduced-form profit function is linear in n. However, if $\varepsilon > 0$ or $\alpha + \nu < 1$ or capital is fixed, there are decreasing returns to n.

TABLE 9. Different values of α: parameter estimates.

	λ	μ	σ	σ_{mrn}	F	τ
$\alpha = 0.66$	0.0301	0.0030	0.0520	0.0211	1.0281	0.0004
	(0.0027)	(0.0004)	(0.0028)	(0.0023)	(0.0758)	(0.0001)
$\alpha = 0.5$	0.0300	0.0044	0.0766	0.0210	1.4634	0.0004
	(0.0026)	(0.0006)	(0.0038)	(0.0025)	(0.0147)	(0.0001)
$\alpha = 0.85$	0.0296	0.0013	0.0230	0.0177	0.4516	0.0001
	(0.0024)	(0.0002)	(0.0012)	(0.0021)	(0.0546)	(0.00002)

Notes: This table reports the estimated parameters when we preset the parameter α at either 0.66 (benchmark), 0.5, or 0.85.

TABLE 10. Sensitivity analysis: policy experiments.

	Y	N	w	M	C
Benchmark	−0.0326	−0.0135	0.0025	−0.8647	0.0160
Lower $\mu = 0\%$	−0.0392	−0.0236	0.0013	−0.6709	0.0249
Higher $\mu = 0.5\%$	−0.0126	0.0027	0.0037	−0.5763	0.0010
Lower $\sigma = 6\%$	−0.0353	−0.0183	0.0006	−0.6917	0.0189
Higher $\sigma = 8\%$	−0.0068	0.0043	0.0066	−0.4672	0.0023
Lower $\lambda = 4.5\%$	−0.0168	0.0002	0.0025	−0.6800	0.0024
Higher $\lambda = 5.5\%$	−0.0432	−0.0238	0.0025	−0.8768	0.0263

Notes: This table reports the steady-state effect of removing the regulation on output, employment, the number of firms, the real wage, and consumption for different parameter values (in percentage change). The sunk cost and wages are kept constant (not reestimated) as we change the parameters.

are now smaller, which implies a smaller cost is enough to reconcile the model and the data. It turns out that this second effect dominates, so that the estimated sunk cost is significantly smaller with larger α. The other parameters are largely unchanged.

These estimates in turn affect the policy experiments. Given the lower estimated sunk cost, the model with higher α leads to smaller gains from removing the regulation. For instance, the first row is 0.13% with the higher α, instead of 0.27%, while the lower α implies a gain of 0.40%. These results suggest that future research on estimating the relevant value of α would be useful for our exercise.

As a sensitivity analysis, we also provide in Table 10 the effect of removing the regulation for different parameter values. In particular, we show that a lower mean growth leads to larger effects, and so does a lower standard deviation or a higher exit rate. The intuition is that all these parameter changes tend to make firms smaller on average, so that there are more firms close to the threshold for which the regulation bites more. One important implication is that if we underestimate σ because our model abstracts from labor adjustment costs, we will understate the effect of the regulation.

6.2 *Sectoral results*

In this section, we study how our results vary by sector. We consider three broad sectors—manufacturing, construction, retail—and we further consider the smaller hospitality industry (lodging and restaurants). In principle, this sectoral heterogeneity is

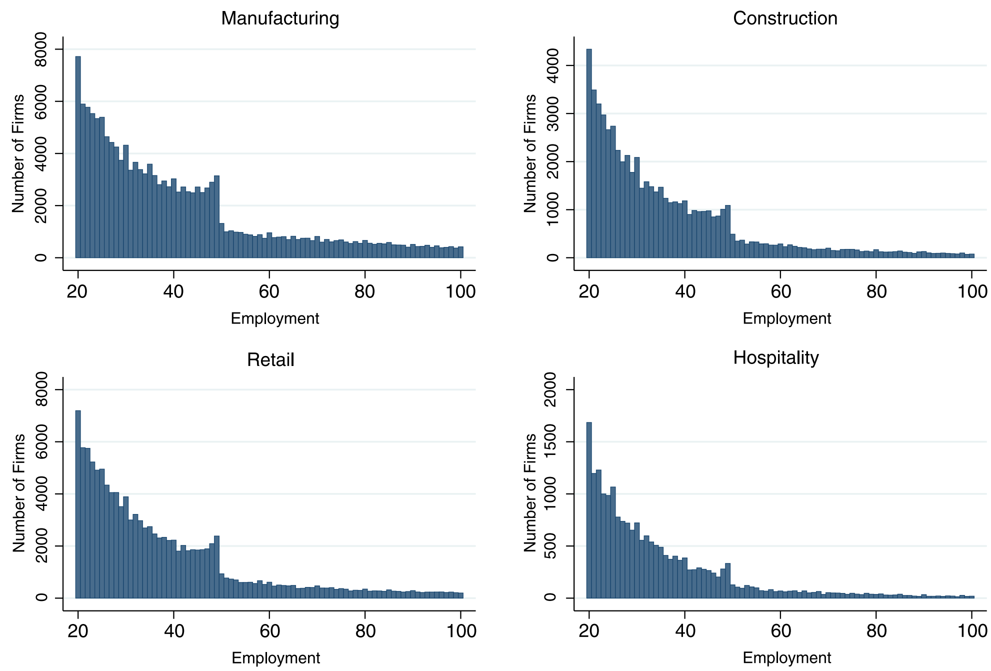

FIGURE 9. Distribution of firm employment, by sector, between 20 and 100 employees.

TABLE 11. Sectoral results: parameter estimates.

	λ	μ	σ	σ_{mrn}	F	τ
All	0.0301	0.0030	0.0520	0.0211	1.0281	0.0004
	(0.0027)	(0.0004)	(0.0028)	(0.0023)	(0.0758)	(0.0001)
Manufacturing	0.0211	0.0020	0.0416	0.0218	0.9021	0.0004
	(0.0017)	(0.0004)	(0.0009)	(0.0018)	(0.1315)	(0.0001)
Construction	0.0230	0.0011	0.0291	0.0148	0.6059	0.0001
	(0.0040)	(0.0007)	(0.0020)	(0.0017)	(0.1454)	(0.00002)
Retail	0.0276	0.0025	0.0480	0.0270	0.9475	0.0008
	(0.0026)	(0.0006)	(0.0039)	(0.0016)	(0.1095)	(0.0002)
Hospitality	0.0226	0.0015	0.0370	0.0483	0.8959	0.0002
	(0.0036)	(0.0010)	(0.0025)	(0.0004)	(0.1863)	(0.0001)

Notes: This table reports the parameter estimates of the full model, for the entire economy and for each sector, with the associated standard error (in parentheses).

useful in validating the mechanism since there could be technological differences across sectors. However, as shown in Figure 9, the discontinuity in the firm size distribution is observed in all sectors and, perhaps surprisingly, it is similar across the different sectors.

We estimate our model in each sector separately: Table 11 reports the parameter estimates for each sector and Table 12 reports the fit of the model for each sector. For these estimations, we set the parameter α in each sector according to the labor share observed in each sector. Unsurprisingly, the labor share is estimated to be larger in con-

Table 12. Sectoral results: model fit.

	Manufacturing		Construction		Retail		Hospitality	
	Data	Model	Data	Model	Data	Model	Data	Model
Mean $\Delta \log n$	0.0058	0.0058	0.0049	0.0046	0.0061	0.0062	0.0031	0.0030
Std. dev. $\Delta \log n$	0.1382	0.1382	0.1344	0.1344	0.1465	0.1465	0.1580	0.1580
Power law coefficient	2.2103	2.2140	2.3159	2.3604	2.2596	2.3398	2.0721	2.2421
Density of firms in each bin, unconditional								
40–46	0.0719	0.0651	0.0747	0.0671	0.0724	0.0666	0.0761	0.0667
47–49	0.0790	0.0810	0.0758	0.0782	0.0800	0.0860	0.0723	0.0669
50–52	0.0304	0.0337	0.0308	0.0332	0.0307	0.0343	0.0292	0.0450
53–59	0.0241	0.0286	0.0225	0.0280	0.0230	0.0247	0.0233	0.0282
Density of firms in each bin, conditional								
40–46	0.0275	0.0266	0.0314	0.0249	0.0248	0.0262	0.0344	0.0295
47–49	0.0459	0.0607	0.0515	0.0505	0.0521	0.0693	0.0703	0.0489
50–52	0.0452	0.0411	0.0504	0.0497	0.0472	0.0402	0.0495	0.0544
53–59	0.0763	0.0726	0.0678	0.0750	0.0755	0.0697	0.0571	0.0691
Residuals		96.48		43.77		91.70		73.14

Notes: This table reports the data and model moments for each sector, and the minimized criterion.

struction and hospitality.[14] For reasons explained in the previous section, a larger labor share implies, ceteris paribus, a smaller volatility of shocks and a lower cost of misallocation. The main differences in the moments between sectors are that (a) the volatility of employment is larger in retail, and especially in hospitality; (b) while the unconditional distributions are strikingly similar across sectors, the conditional distributions have almost no spike in the bin $[47, 49]$ in manufacturing and construction, while there is a clear spike in retail and especially in hospitality.

The estimation reconciles these model features with the data by postulating a larger volatility process in manufacturing and in retail. The sunk cost is smaller in construction than in the other three sectors. Hospitality is subject to a large measurement error (i.e., shocks may be less persistent). Given the importance of the conditional spike in retail, the tax is also estimated to be larger (0.8%).

6.3 *Comparison with Garicano, Lelarge, and Van Reenen (2013)*

This section compares our results with those of the contemporaneous, related study of Garicano, Lelarge, and Van Reenen (2013). Both papers focus on the same striking fact, use similar data,[15] and analyze it using a simple structural model. However, we obtain different estimates of the regulation cost and the output gains from removing distortions. What is the source of these differences?

We see three key differences between our papers. First, our model is dynamic and distinguishes between sunk costs and per-period costs, while their model is static and,

[14]The values are 0.69 for manufacturing, 0.78 for construction, 0.66 for retail, and 0.74 for hospitality.

[15]The period studied and some details of data construction differ, however.

hence, assumes that the regulation costs are paid each period if a firm operates with more than 50 employees. This difference, while conceptually interesting,[16] is not critical in explaining why our study finds much smaller effects: even when we consider the model with only a payroll tax, our estimated tax is significantly smaller. This leads to the second difference, which lies in our estimation strategy. Our view is that the key feature of the data is the sharp discontinuity around 50, which we believe we ought to match precisely. Hence, we target moments based on the distribution around the threshold. Their strategy is to evaluate the model using the entire size distribution. This is an attractive approach that is also more ambitious. In the absence of tax, their model implies a Pareto distribution over the entire range of firm size. One concern is that the estimated tax may end up accounting for deviations of the observed size distribution from the Pareto distribution. Indeed, it is well known that the Pareto distribution does not fit well for small firms and for large firms, and this is observed in many countries, including countries that do not have an explicit size-dependent regulation. We believe this explains in part why they estimated a large tax. Moreover, because they attempted to fit the entire firm size distribution, their model fits more poorly around the threshold of 50 employees, even with significant "measurement error." On the other hand, our approach is also sensitive to model misspecification, in particular regarding the dynamic moments. For instance, we abstract from other frictions such as adjustment costs or financial constraints which may interact with our estimation.

Finally, regarding the policy experiments, their paper finds a similar effect on output (0.02%) of the regulation when wages are flexible. The larger output loss reported (0.8%) is driven by the assumption that the proceeds of the tax are wasted. However, rather than this standard flexible wage experiment, Garicano, Lelarge, and Van Reenen (2013) preferred to emphasize an exercise that assumes fixed wages and generates a much larger effect: output goes down by around 4%. The mechanism is the following: after-tax wages are assumed to be completely rigid. As a result, a higher payroll tax increases the pre-tax wage (paid by the employer) one-for-one. A very simple calculation approximates their results. The production function implies an elasticity of labor demand equal to 5. Since the tax is estimated at 1.3%, and applies to 70% of the population (those who work in a firm with more than 50 employees), an elasticity of 5 implies an employment effect of 4.5% and an output effect of 3.6%. Hence, these large effects are unrelated to reallocation, but are due to the distortionary effect of payroll taxes when after-tax wages are rigid.

We are somewhat skeptical of this calculation. First, a standard finding in the public finance literature regarding the incidence of payroll taxes is that these taxes are passed through to workers, suggesting that after-tax wages eventually adjust. The reduction in after-tax wages would decrease the size of the employment effect.[17] Second, their argument applies to all payroll taxes—and of course, the 1.3% that is estimated is only a small

[16]The sunk cost model has additional implications beyond the conditional distribution that we have emphasized in our estimation. Because growing above 50 employees is now an investment, it is sensitive to expectations about the future and to uncertainty.

[17]The high legal minimum wage in France (which affects directly around 15% of the workforce) might partially prevent this adjustment.

part of labor taxes in France (which, combining standard payroll taxes, the personal income tax, and value added tax (VAT) add up to well over 40% at the margin). Their rigid wage model implies a huge sensitivity of employment to taxes. By comparison, Prescott (2002) argued that the difference in hours worked between the United States and France could be accounted for solely through distortionary taxes in a neoclassical model, assuming that government spending was perfect substitute with private consumption. But even his model implies a much lower sensitivity of employment to taxes: an increase in labor taxes of 1 percentage point reduces employment by around 0.8%, or about six times less than in Garicano, Lelarge, and Van Reenen (2013).[18] Overall, it is unclear that the assumptions of complete wage rigidity and a highly elastic labor demand are the most appropriate ones for this question.

7. Conclusion

Our paper studies a particular regulation that clearly distorts the firm size distribution, leading to an obvious misallocation of labor—a channel that has been emphasized in the recent literature. Our paper, hence, provides a "case study" that is complementary to broader macroeconomic approaches (Hopenhayn and Rogerson (1993), Restuccia and Rogerson (2008), and Buera, Kaboski, and Shin (2011)).

We show that the regulation can be modeled as a combination of a sunk cost and a payroll tax. Our model fits the size distribution discontinuity around the threshold well and it also fits the smaller discontinuity in the conditional distribution. We obtain plausible estimates of the costs of the regulation. Removing the regulation leads to an increase of output close to 0.3%, holding employment and the number of firms fixed. These effects are an order of magnitude smaller if firm entry is elastic.

There are several interesting extensions. First, incorporating labor adjustment costs or search frictions would be useful to take into account the fact that it is difficult, and costly, for a firm to control its labor force perfectly. Second, introducing in the model other margins of adjustment, such as hours worked or capital, would generate some factor substitution close to the threshold: if it is costly to increase employment, firms may react by using other inputs. Third, the regulation may have additional costs to the extent that it makes it costly for firms to experiment. Finally, given the limitations of our data for small firms, we have abstracted from the existence of other thresholds (at 10 and 20 employees), but incorporating them would be useful to quantify the total effect of these regulations on the firm life cycle.

Appendix

Appendix A.1 presents distribution of employment around the threshold. Appendix A.2 presents the proofs of some model results and formulas.

[18]This value is calculated starting from zero taxes, as in Garicano, Lelarge, and Van Reenen (2013) or as in our paper. The elasticity is twice larger if we calculate it at the current level of taxes.

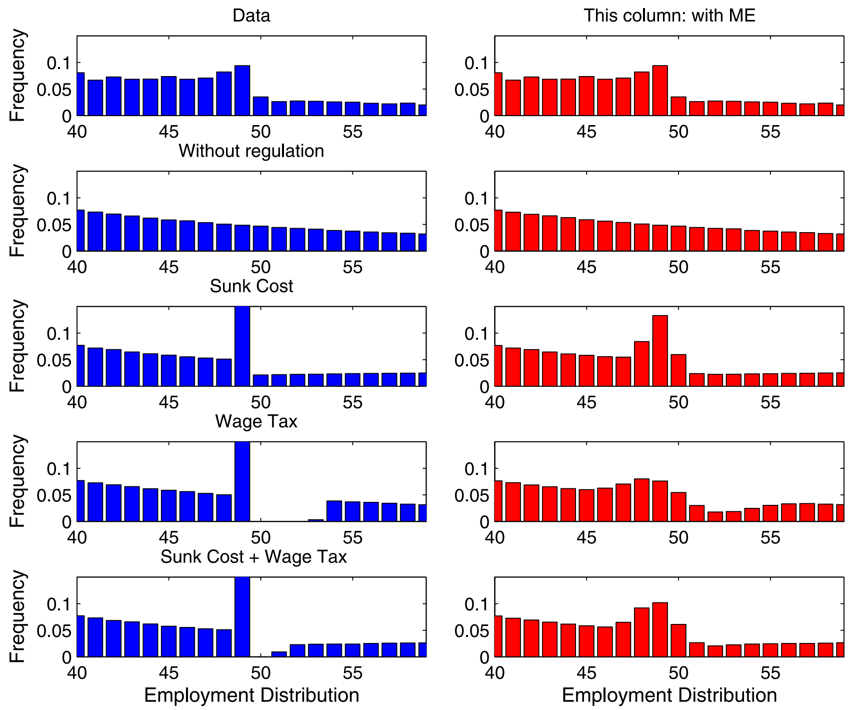

FIGURE 10. Distribution of firm employment (between 40 and 59 employees). The distribution is normalized by the total number of firms that have between 40 and 59 employees.

A.1 *Distribution of employment around the threshold*

Figure 10 presents the firm size distribution around the threshold at our estimated parameter values for the data, the model without any regulation, the model with only a sunk cost, the model with only a wage tax, and the full model. For completeness, the left panel shows the distribution without measurement error while the right panel shows the model with measurement error. Figure 11 provides the same experiments for the unconditional distribution. Both pictures use our parameter estimates. Note how the wage tax model generates a significant "hole" in the unconditional distribution without measurement error, and a very large "valley" after 52 employees even with measurement error. This is in sharp contrast to the data.

Also notable is the fact that the very small tax that we estimate is enough to generate a significant difference in the conditional distribution between the model with only sunk cost—which generates no spike at 49—and our full model—where there is no one at 50 in the conditional distribution without measurement error, and a significant spike at 49. All these features are then smoothed by measurement error.

A.2 *Proofs*

A.2.1 *Proof of Proposition* First, note that the function $V(\cdot, 1)$ is twice continuously differentiable (see Stokey (2008, Chapter 5.6) for a proof). Using the previously com-

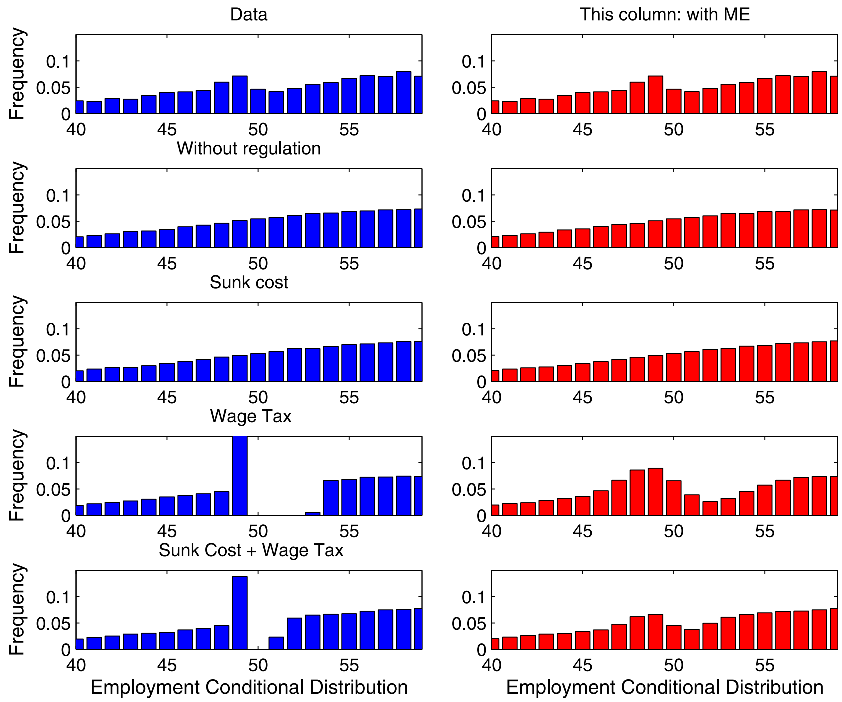

FIGURE 11. Distribution of firm employment (between 40 and 59 employees), conditional on having had more than 55 employees in the past. The distribution is normalized by the total number of firms that have between 40 and 59 employees.

puted $\pi(z, 1)$ gives

$$
\begin{aligned}
V(z^*, 1) &= \frac{1}{J}\left[\int_{z^*}^{\infty} e^{R_2(z^*-z)}\pi(z,1)\,dz + \int_{-\infty}^{z^*} e^{R_1(z^*-z)}\pi(z,1)\,dz\right] \\
&= \frac{1}{J}\left[\int_{z^*}^{\infty} e^{R_2(z^*-z)}\pi^a(z)\,dz + \int_{\overline{z}}^{z^*} e^{R_1(z^*-z)}\pi^a(z)\,dz\right. \\
&\quad \left. + \int_{\underline{z}}^{\overline{z}} e^{R_1(z^*-z)}\pi^b(z)\,dz + \int_{-\infty}^{\underline{z}} e^{R_1(z^*-z)}\pi^b(z)\,dz\right].
\end{aligned}
$$

Define, for all $x \le z^*$,

$$
\begin{aligned}
H(x, z^*) &\equiv E_x\left[\int_0^{T(z^*)} e^{-(r+\lambda)t}\pi(z_t, 0)\,dt + e^{-(r+\lambda)T(z^*)}\big(V(z^*, 1) - F\big)\right] \\
&= \frac{1}{J}\left[\int_x^{z^*} e^{R_2(x-z)}\pi(z,0)\,dz + \int_{-\infty}^{x} e^{R_1(x-z)}\pi(z,0)\,dz\right. \\
&\quad \left. - e^{R_2(x-z^*)}\int_{-\infty}^{z^*} e^{R_1(z^*-z)}\pi(z,0)\,dz\right] \\
&\quad + e^{R_2(x-z^*)}\big(V(z^*, 1) - F\big).
\end{aligned}
$$

Then $V(x,0)=\sup_{z^*\geq x} H(x,z^*)$. Note that $H(x,z^*)$ is twice continuously differentiable. The first order condition (FOC) for a maximum at $z^*\geq \overline{z}$ is

$$\begin{aligned}
0 &\leq H_{z^*}\bigl(x,z^*\bigr)\\
&= \frac{1}{J}\biggl[e^{R_2(x-z^*)}\pi\bigl(z^*,0\bigr)+R_2e^{R_2(x-z^*)}\int_{-\infty}^{z^*} e^{R_1(z^*-z)}\pi(z,0)\,dz\biggr]\\
&\quad+\frac{1}{J}\biggl[-e^{R_2(x-z^*)}\pi\bigl(z^*,0\bigr)-R_1e^{R_2(x-z^*)}\int_{-\infty}^{z^*} e^{R_1(z^*-z)}\pi(z,0)\,dz\biggr]\\
&\quad-R_2e^{R_2(x-z^*)}\bigl(V\bigl(z^*,1\bigr)-F\bigr)+e^{R_2(x-z^*)}V_{z^*}\bigl(z^*,1\bigr)\\
&= e^{R_2(x-z^*)}\biggl[\frac{R_2-R_1}{J}\int_{-\infty}^{z^*} e^{R_1(z^*-z)}\pi(z,0)\,dz-R_2\bigl(V\bigl(z^*,1\bigr)-F\bigr)+V_{z^*}\bigl(z^*,1\bigr)\biggr],
\end{aligned}$$

with equality if $z^*>\overline{z}$. Hence,

$$\begin{aligned}
&V\bigl(z^*,1\bigr)\\
&\quad=\frac{1}{J}\biggl[\int_{z^*}^{\infty} e^{R_2(z^*-z)}\pi^a(z)\,dz+\int_{\overline{z}}^{z^*} e^{R_1(z^*-z)}\pi^a(z)\,dz\\
&\qquad+\int_{\underline{z}}^{\overline{z}} e^{R_1(z^*-z)}\pi^b(z)\,dz+\int_{-\infty}^{\underline{z}} e^{R_1(z^*-z)}\pi^b(z)\,dz\biggr]\\
&\qquad\times V_z\bigl(z^*,1\bigr)\\
&\quad=\frac{R_2}{J}\int_{z^*}^{\infty} e^{R_2(z^*-z)}\pi(z,1)\,dz+\frac{R_1}{J}\int_{-\infty}^{z^*} e^{R_1(z^*-z)}\pi(z,1)\,dz\\
&\quad=\frac{R_2}{J}\int_{z^*}^{\infty} e^{R_2(z^*-z)}\pi^a(z)\,dz\\
&\qquad+\frac{R_1}{J}\biggl[\int_{\overline{z}}^{z^*} e^{R_1(z^*-z)}\pi^a(z)\,dz+\int_{\underline{z}}^{\overline{z}} e^{R_1(z^*-z)}\pi^b(z)\,dz\\
&\qquad+\int_{-\infty}^{\underline{z}} e^{R_1(z^*-z)}\pi^b(z)\,dz\biggr].
\end{aligned}$$

Plugging in the FOC gives

$$(R_1-R_2)\int_{\overline{z}}^{z^*} e^{R_1(z^*-z)}\bigl[\pi^a(z)-\pi^b(z)\bigr]\,dz+R_2JF=0,$$

which simplifies to

$$R_1\int_{\overline{z}}^{z^*} e^{R_1(z^*-z)}\bigl[\pi^a(z)-\pi^b(z)\bigr]\,dz+(r+\lambda)F=0.$$

It is easy to see that there exists a unique value of z^* that satisfies the preceding equality. Moreover, one can compute these integrals easily given our formulas for $\pi^a(z)$ and $\pi^b(z)$.

A.2.2 *Alternative derivation of optimal policy using dynamic programming* We start by writing the Hamilton–Jacobi–Bellman equation satisfied by V:

$$(r+\lambda)V(z,1)=\pi(z,1)+\mu V_z(z,1)+\frac{\sigma^2}{2}V_{zz}(z,1) \tag{18}$$

for any z and

$$(r+\lambda)V(z,0)=\pi(z,0)+\mu V_z(z,0)+\frac{\sigma^2}{2}V_{zz}(z,0) \tag{19}$$

for $z<z^*$. Note that $\pi(z,0)$ and $\pi(z,1)$ are only C^1 (continuously differentiable): the second derivative is discontinuous at $z=\underline{z}$ for $\pi(\cdot,0)$ and $\pi(\cdot,1)$, and at $z=\overline{z}$ for $\pi(\cdot,1)$.

The boundary conditions are given by value matching

$$V\bigl(z^*,1\bigr)=V\bigl(z^*,0\bigr)-F \tag{20}$$

and by the smooth pasting condition

$$V_z\bigl(z^*,1\bigr)=V_z\bigl(z^*,0\bigr). \tag{21}$$

The general solution of the associated homogeneous ordinary differential equation (ODE) (i.e., without the term $\pi(\cdot,\cdot)$) is $A_1e^{R_1z}+A_2e^{R_2z}$, where R_1 and R_2 are the roots of the quadratic

$$\frac{\sigma^2}{2}X^2+\mu X-(r+\lambda)=0, \tag{22}$$

that is, $R_2=\frac{-\mu+\sqrt{\mu^2+2(r+\lambda)\sigma^2}}{\sigma^2}>0$ and $R_1=\frac{-\mu-\sqrt{\mu^2+2(r+\lambda)\sigma^2}}{\sigma^2}<0$.

The specific forms of $\pi(z,0)$ and $\pi(z,1)$ make it possible to find particular solutions. Starting with the first equation, we guess that

$$\begin{aligned}\widetilde{V}(z,0)&=b_0e^{z/(1-\alpha)}\quad \text{for } z<\underline{z}\\ &=b_1e^z+b_2\quad \text{for } z>\underline{z}\end{aligned}$$

is a solution of (19), for constants b_0, b_1, and b_2 to be determined.

The $\widetilde{V}$ satisfies the ODE for $z<\underline{z}$, provided that b_0 solves

$$(r+\lambda)b_0=\left(\frac{\alpha}{w}\right)^{\alpha/(1-\alpha)}(1-\alpha)+\mu\frac{b_0}{1-\alpha}+\frac{\sigma^2}{2}\frac{b_0}{(1-\alpha)^2}$$

or

$$b_0=\frac{\left(\frac{\alpha}{w}\right)^{\alpha/(1-\alpha)}(1-\alpha)}{r+\lambda-\frac{\mu}{1-\alpha}-\frac{\sigma^2}{2(1-\alpha)^2}}.$$

For $z > \underline{z}$, we require that

$$(r+\lambda)\big(b_1 e^z + b_2\big) = e^z \underline{n}^\alpha - w\underline{n} + \mu b_1 e^z + \frac{\sigma^2}{2} b_1 e^z,$$

that is,

$$b_2 = -\frac{w\underline{n}}{r+\lambda},$$
$$b_1 = \frac{\underline{n}^\alpha}{r+\lambda-\mu-\dfrac{\sigma^2}{2}}.$$

Combining the results, the general solution of the ODE for $V(z,0)$ is

$$\begin{aligned}
V(z,0) &= \widetilde{V}(z,0) + A_1 e^{R_1 z} + A_2 e^{R_2 z} \\
&= \frac{\left(\dfrac{\alpha}{w}\right)^{\alpha/(1-\alpha)}(1-\alpha)}{r+\lambda-\dfrac{\mu}{1-\alpha}-\dfrac{\sigma^2}{2(1-\alpha)^2}} e^{z/(1-\alpha)} + A_1 e^{R_1 z} + A_2 e^{R_2 z} \quad \text{for } z < \underline{z} \\
&= \frac{\underline{n}^\alpha}{r+\lambda-\mu-\dfrac{\sigma^2}{2}} e^z - \frac{w\underline{n}}{r+\lambda} + A_1 e^{R_1 z} + A_2 e^{R_2 z} \quad \text{for } z \geq \underline{z}.
\end{aligned}$$

Turning to the ODE for $V(z,1)$, we again look for a solution, which we guess as

$$\begin{aligned}
\widetilde{V}(z,1) &= e^{z/(1-\alpha)} b_3 \quad \text{for } z < \underline{z} \\
&= e^z b_4 + b_5 \quad \text{for } \overline{z} > z > \underline{z} \\
&= e^{z/(1-\alpha)} b_6 + b_7 \quad \text{for } z > \overline{z}.
\end{aligned}$$

The scalars b_3, b_4, b_5, b_6, and b_7 must satisfy

$$b_3 = \frac{\left(\dfrac{\alpha}{w}\right)^{\alpha/(1-\alpha)}(1-\alpha)}{r+\lambda-\dfrac{\mu}{1-\alpha}-\dfrac{\sigma^2}{2(1-\alpha)^2}} = b_0,$$

$$b_4 = \frac{\underline{n}^\alpha}{r+\lambda-\mu-\dfrac{\sigma^2}{2}} = b_1,$$

$$b_5 = -\frac{w\underline{n}}{r+\lambda} = b_2,$$

$$b_6 = \frac{\left(\dfrac{\alpha}{w(1+\tau)}\right)^{\alpha/(1-\alpha)}(1-\alpha)}{r+\lambda-\dfrac{\mu}{1-\alpha}-\dfrac{\sigma^2}{2(1-\alpha)^2}} = \frac{b_0}{(1+\tau)^{\alpha/(1-\alpha)}},$$

$$b_7 = -\frac{c_f}{r+\lambda},$$

and the general solution is

$$V(z,1)=\widetilde{V}(z,1)+A_3e^{R_2z}+A_4e^{R_1z}.$$

Finally we need to determine A_1, A_2, A_3, A_4, and z^*. A standard argument implies that $A_3=0$ (the investment option values goes to 0 if $z\to\infty$). Moreover, $A_4=0$ since as $z\to-\infty$, the firm value remains finite. Last, $A_1=0$ for the same reason. The two scalars A_2 and z^* are thus determined by the following system of two equations in two unknowns:

$$\begin{aligned}\widetilde{V}\big(z^*,1\big)&=\widetilde{V}\big(z^*,0\big)+A_2e^{R_2z^*}-F,\\ \widetilde{V}_z\big(z^*,1\big)&=\widetilde{V}_z\big(z^*,0\big)+A_2R_2e^{R_2z^*}.\end{aligned}$$

Given the formulas for $\widetilde{V}$ and that $z^*>\overline{z}>\underline{z}$, this can be rewritten as

$$\begin{aligned}e^{z^*/(1-\alpha)}b_6+b_7&=\frac{\underline{n}^\alpha}{r+\lambda-\mu-\dfrac{\sigma^2}{2}}e^{z^*}-\frac{w\underline{n}}{r+\lambda}+A_2e^{R_2z^*}-F,\\ e^{z^*/(1-\alpha)}\frac{b_6}{1-\alpha}&=\frac{\underline{n}^\alpha}{r+\lambda-\mu-\dfrac{\sigma^2}{2}}e^{z^*}+A_2R_2e^{R_2z^*}.\end{aligned}$$

This characterizes entirely the solution. It is easy to verify that this yields the same results as those obtained in the main text using the theoretical results of Stokey (2008).

A.2.3 *Derivation of the stationary cross-sectional distribution* To solve for f, first note that the general solution of the ODE (11) is

$$f(z,0)=D_0e^{\beta_1z}+D_1e^{\beta_2z},$$

where $\beta_1<0<\beta_2$ are the two real roots of the characteristic equation

$$\lambda=-\mu X+\frac{\sigma^2}{2}X^2.$$

The ODE must be solved separately on each interval. Given that f is a density, the exponential terms that do not go to 0 as $|z|\to\infty$, must disappear. This yields the simpler form

$$\begin{aligned}f(z,0)&=C_1e^{\beta_2z}\quad\text{for } z<z_0\\ &=C_2e^{\beta_1z}+C_3e^{\beta_2z}\quad\text{for } z^*>z>z_0\end{aligned}$$

and

$$\begin{aligned}f(z,1)&=C_4e^{\beta_2z}\quad\text{for } z<z^*\\ &=C_5e^{\beta_1z}\quad\text{for } z>z^*.\end{aligned}$$

The boundary conditions can then be expressed as a system of five linear equations in five unknowns. First, f is a p.d.f., that is, its integral is 1:

$$\frac{C_1}{\beta_2}e^{\beta_2 z_0}+\frac{C_2}{\beta_1}\left(e^{\beta_1 z^*}-e^{\beta_1 z_0}\right)+\frac{C_3}{\beta_2}\left(e^{\beta_2 z^*}-e^{\beta_2 z_0}\right)+\frac{C_4}{\beta_2}e^{\beta_2 z^*}-\frac{C_5}{\beta_1}e^{\beta_1 z^*}=1.$$

Second, $f(\cdot, 0)$ is continuous at z_0:

$$C_1 e^{\beta_2 z_0}=C_2 e^{\beta_1 z_0}+C_3 e^{\beta_2 z_0}.$$

Third, $f(\cdot, 0)$ is continuous at z^*:

$$C_2 e^{\beta_1 z^*}+C_3 e^{\beta_2 z^*}=0.$$

Fourth, $f(\cdot, 1)$ is continuous at z^*:

$$C_5 e^{\beta_1 z^*}=C_4 e^{\beta_2 z^*}.$$

And finally the boundary condition at z^*:

$$-\frac{\sigma^2}{2}\left(C_2\beta_1 e^{\beta_1 z^*}+C_3\beta_2 e^{\beta_2 z^*}\right)=\lambda\left(\frac{C_4}{\beta_2}e^{\beta_2 z^*}-\frac{C_5}{\beta_1}e^{\beta_1 z^*}\right).$$

Solving this system of equations yields the analytical formula for f shown in the main text.

References

Altonji, J. and L. Segal (1996), "Small-sample bias in GMM estimation of covariances structures." *Journal of Business & Economic Statistics*, 14 (3), 353–366. [394]

Axtell, R. (2001), "Zipf distribution of US firm sizes." *Science*, 293 (5536), 1818–1820. [383]

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2009), "Cross-country differences in productivity: The role of allocation and selection." Working Paper 15490, National Bureau of Economic Research. [380, 401]

Bartelsman, E., S. Scarpetta, and F. Schivardi (2005), "Comparative analysis of firm demographics and survival: Evidence from micro-level sources in OECD countries." *Industrial and Corporate Change*, 14 (3), 365–391. [401]

Blundell, R., L. Pistaferri, and I. Preston (2008), "Consumption inequality and partial insurance." *American Economic Review*, 98 (5), 1887–1921. [393]

Buera, F., J. Kaboski, and Y. Shin (2011), "Finance and development: A tale of two sectors." *American Economic Review*, 101 (5), 1964–2002. [379, 408]

Cahuc, P. and F. Kramarz (2004), "De la précarité à la mobilité: vers une sécurité sociale professionnelle." La documentation française. [380]

Ceci-Renaud, N. and P. Chevalier (2011), "L'impact des seuils de 10, 20 et 50 salariés sur la taille des entreprises françaises." *Economie et Statistique*, 437, 29–45. [380]

Chetty, R. (2012), "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply." *Econometrica*, 80 (3), 969–1018. [400]

Cooper, R., J. Haltiwanger, and J. Willis (2007), "Search frictions: Matching aggregate and establishment observations." *Journal of Monetary Economics*, 54, 56–78. [394]

Di Giovanni, J., A. Levchenko, and R. Ranciere (2011), "Power laws in firm size and openness to trade: Measurement and implications." *Journal of International Economics*, 85 (1), 42–52. [383]

Dixit, A. K. and R. S. Pindyck (1994), *Investment Under Uncertainty*. Princeton University Press, Princeton. [379, 391]

Fattal Jaef, R. N. (2012), "Entry, exit and misallocation frictions." Report, International Monetary Fund. [401]

García-Santana, M. and J. Pijoan-Mas (2011), "Small scale reservation laws and the misallocation of talent." Working Paper 8242, Centre for Economic Policy Research. [380]

Garicano, L., C. Lelarge, and J. Van Reenen (2013), "Firm size distortions and the productivity distribution: Evidence from France." Technical Report 7241, IZA. [380, 403, 406, 407, 408]

Gourieroux, C., A. Monfort, and E. Renault (1993), "Indirect inference." *Journal of Applied Econometrics*, 8 (S1), S85–S118. [393]

Guner, N., G. Ventura, and Y. Xu (2008), "Macroeconomic implications of size-dependent policies." *Review of Economic Dynamics*, 11 (4), 721–744. [380]

Hall, P. and J. Horowitz (1996), "Bootstrap critical values for tests based on generalized-method-of-moments estimators." *Econometrica*, 64 (4), 891–916. [394]

Hopenhayn, H. A. and R. Rogerson (1993), "Job turnover and policy evaluation: A general equilibrium analysis." *Journal of Political Economy*, 101 (5), 915–938. [379, 398, 408]

Hsieh, C.-T. and P. Klenow (2009), "Misallocation and manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124 (4), 1403–1448. [379]

Lucas, R. E. (1978), "On the size distribution of business firms." *Bell Journal of Economics*, 9 (2), 508. [386]

Prescott, E. C. (2002), "Prosperity and depression." *American Economic Review*, 92 (2), 1–15. [408]

Restuccia, D. and R. Rogerson (2008), "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic Dynamics*, 11 (4), 707–720. [379, 380, 408]

Stokey, N. L. (2008), *The Economics of Inaction: Stochastic Control Models With Fixed Costs*. Princeton University Press, Princeton. [379, 389, 390, 409, 414]

Submitted December, 2012. Final version accepted October, 2013.