Firm Size Distortions and the Productivity Distribution: Evidence from France*

Luis Garicano† Claire Lelarge‡ John Van Reenen§

April 3, 2012

Abstract

Preliminary. A major empirical challenge in economics is to identify how regulations (such as firing costs) affect economic efficiency. Almost all countries have regulations that increase costs when firms cross different size thresholds. We show how these size-contingent regulations can be used to identify the equilibrium and welfare effects of regulation through combining a new model with the firm-level distributions of size and productivity. Our framework adapts the Lucas (1978) model to a world with size-contingent regulations and applies this to France where there are sharp increases in firing costs when firms employ 50 or more workers. Using administrative data on the population of firms 2002 through 2007, we show how this regulation has major effects on the distribution of firm size (a “broken power law”) and productivity. We then econometrically recover the key parameters of the model in order to estimate the costs of regulation which appear to be non-trivial.

Keywords: Firm size, productivity, labor regulation, power law

JEL Classification: L11, L51, J8, L25

---

*Preliminary and Incomplete. We would like to thank Samuel Bentolila, Nick Bloom, Jesus Fernandez-Villaverde, Esteban Rossi-Hansberg and participants of seminars in Chicago, Columbia GSB, CREST, Insead, Nantes, NBER, Oxford and Stanford GSB. We particularly our formal discussants Panle Jia and Ariel Pakes. Finance for this project has been kindly provided by the Economic and Social Research Council for the Centre for Economic Performance, LSE International Growth Centre and the European Commission.

†London School of Economics, Centre for Economic Performance and CEPR
‡CREST, Insee
§London School of Economics, Centre for Economic Performance, CEPR and NBER
1 Introduction

A recent literature has documented empirically how distortions that raise the cost of labor or capital affect aggregate productivity though misallocations of resources from more productive to less productive firms. As Restuccia and Rogerson (2008) argued, more efficient firms may have “too little” output or employment allocated to them due to various distortions in their economies. Hsieh and Klenow (2009) have argued that these misallocations go a long way towards explaining the gap in aggregate productivity between the US, China, and India. In this paper, we focus on understanding the impact and the size of one specific distortion on the French firm size distribution: regulations sharply increasing labor costs when firms reach 50 workers.

The idea that misallocations of resources may be partly behind the productivity gap is attractive in understanding the differences between the US and Europe. According to the European Commission (1996) the average production unit in the EU employed 23% less workers than in the US. As Figure 1 shows, there appear to be far fewer French firms which are able to grow to the same scale as the productive US firms. Figure 1 shows two interesting patterns. First, there is a large bulge in the number of firms with employment just below 50 workers in France, but no such bulge in the distribution of American firms. Second, there is a much larger share of very large firms in the US - the US has many more firms with over 2,500 employees than does France. This paper focuses on the first of those patterns, although we plan to examine the absence of very large French firms in later work.

Labor legislation in France substantially increases firing costs when firms employ 50 or more workers. Specifically, firms with 50 or more employees must formulate a “social plan,” which is designed to facilitate reemployment, through training, etc. As a result, the costs of employing workers also rise (see Bertola and Bentolila, 1990) at that threshold. Figure 2 shows that indeed the legislation binds, so that there is a clear threshold effect at precisely 50 firms. This uses 2007 data, but the discontinuity in the size distribution at 50 workers is present in all years of our data (see below).

What are the distortions in the size distribution, in the productivity distribution, and on aggregate productivity that result from those distortions? Our approach relies on revealed preference and on positive sorting effects. Some firms that would have been larger without the regulation choose to remain below the

---

1 See also Parente and Prescott (2000), Bloom and Van Reenen (2010) and Petrin and Sivadasan (2010).

2 In development economics many scholars have pointed to the “missing middle”, i.e. a preponderance of very small firms in poorer countries compared to richer countries (see Banerjee and Duflo, 2005, or Jones, 2011). For example, in the late 1980s in India large firms were banned from producing about 800 product groups (Little et al, 1987). Many explanations have been put forward for this such as financial development, taxes, human capital, lack of competition in product markets, and social capital. One possibility, related to our approach, is size related labor regulations. Besley and Burgess (2004), for example, suggest that labor regulation is one of the reasons why the formal manufacturing sector is much smaller in some Indian states compared to others.

3 Bartelsman et al (2009) examine misallocation using micro-data across many OECD countries and make a similar point. In particular, they find that the “Olley Pakes” (1996) covariance between size and productivity is much smaller in France (0.24 in their Table 1) and other European countries compared to the US (0.51 in their Table 1).
legal threshold to avoid these additional costs. In this paper we aim to identify these firms, calculate their counterfactual size, and use this observation to infer the cost of this legislation.

There has been extensive discussion of the importance of Employment Protection Legislation (EPL) for unemployment and more recently productivity (e.g. Layard and Nickell, 1999; OECD, 2009). The OECD, World Bank and other agencies have developed various indices of the importance of these regulations based on examination of laws and (sometimes) surveys of managers. It is very hard, however, to see how these can be quantified as “adding up” the legal and regulatory provisions has a large arbitrary component. A contribution of our paper is to offer a methodology for quantifying the “tax equivalent” of a regulation, albeit in the context of a specific model.

There are different views on the underlying sources of heterogeneity in firm productivity. We follow Lucas (1978) in taking the stand that managerial talent is the primitive, and that the economy-wide observed resource distribution is, as Manne (1965) felicitously put it, “a solution to the problem: allocate productive factors over managers of different ability so as to maximize output.” Managers make discrete decisions or solve problems (Garicano, 2000). Making better decisions, or solving problems that others cannot solve, raises everyone’s marginal product. This means that, in equilibrium, better managers must be allocated more resources. In fact, absent decreasing returns to managerial talent, the best manager must be allocated all resources. Given limits to managerial time or attention, the better managers are allocated more workers and more capital to manage. This results in a “scale-of-operations” effect whereby differences in talent are amplified by the resources allocated. Lucas (1978) first explored these effects in an equilibrium setting.

Consequently better managers, that is those that for whatever reason are able to generate more productivity, should be allocated (or equivalently, should choose) larger firm sizes. When managers are confronted with legislation that introduces a cost of acquiring a size that is beyond a certain threshold, they may choose to stay below the threshold and stay at an inefficiently small size. By studying the productivity of these marginal managers, we are able to estimate the cost of the legislation, the distortions in them, and thus the welfare cost of the legislation for the entire firm size distribution.

We start by setting up a simple model of the allocation of a single factor, labor to firms in a world where

---

4In a model of this kind, the source of decreasing returns are on the production size, and are linked to limits to managerial time. For our purposes here, as Hsieh and Klenow (2009) show, this source of decreasing returns is equivalent to having the decreasing returns come (as is more common in recent literature following Melitz, 2003) from the utility side.

5Such a scale of operations effects is at the heart of Rosen's (1982) theory of hierarchies, where efficiency units of labor controlled (and not just number of bodies) matter, and also in Garicano and Rossi-Hansberg (2006) where there is limited quantity-quality substitutability so that matching between workers and managers takes place. Empirically, this technology has been used to explain a wide-range of phenomena, most recently to calibrate the impact of scale of operations effects on CEO wages (Gabaix and Landier, 2008).

6Many empirical papers have shown that deregulation (e.g. Olley and Pakes, 1996), higher competition (e.g. Syverson, 2004) and trade liberalization (e.g. Pavcnik, 2002) have tended to improve reallocation by increasing the correlation between firm size and productivity.
there are decreasing returns to managerial talent. We use it to study the effect of a step change in labor costs after a particular size and show that there are four main effects:

1. Equilibrium wages fall as a result of the reduction in the demand for workers (i.e. some of the tax incidence falls on workers)

2. Firm size increases for all firms below the threshold as a result of the general equilibrium effect on wages

3. Firm size reduces to precisely the regulatory threshold for a set of firms that are not productive enough to justify incurring the regulatory costs

4. Firm size reduces proportionally for all firms that are productive enough to incur the additional cost of regulation.

We use the model to guide our estimation of the impact of these costs. The theory tells us there is a deviation from the ‘correct’ firm size distribution as a result of the regulation. That is, we expect to see a departure from the usual power law firm size distribution as firms bunch up below the threshold (50 workers). Given factors such as measurement error, the observed empirical departure from the power law is not just at 49 workers but at a slightly smaller firm size. Similarly, there is not precisely zero mass to the left of the cut-off (at 50), but rather a “valley” were there are significantly fewer firms than we would expect from an unbroken power law. Then, at some point the distribution becomes again a power law, with a lower intercept. The jump in the power law identifies the size of the distortion.

Our results are consistent across specifications, and the cost of these regulations is quite precisely estimated. We find that these regulations operate mainly as a variable cost, and are equivalent to a 1% increase in wages across the distribution. There are thus large misallocations of resources that follow from the implicit tax. Of course, it does not follow that the welfare loss is of this magnitude: society may value the increase in job security of workers working in medium and large size firms as sufficiently important to compensate such large welfare losses. Determining this, however, requires knowing the cost in terms of lost output, which is what our approach delivers.

The most closely related paper to ours is Braguinsky, Branstetter and Regateiro (2011) who seek to explain the shift to the left in the Portuguese firm size distribution in the context of the Lucas model with labour regulations. Their calibrations also show substantial effects of the regulations on lower aggregate productivity.

---

7 See Axtell (2001), Sutton (1997) and Gabaix, (2009). There is a large literature on the size and productivity distribution of firms in macro, trade, finance and IO. Appropriately, the first major study in this area was by Gibrat (1931) who studied French industrial firms, the main focus of the empirical part of our paper.

8 On the quantitative theory side Guner et al (2006, 2008) also consider a Lucas model with size-contingent regulation. They calibrate this to uncover sizeable welfare losses. Unlike our paper and Barguinsky et al (2011), however, there is no econometric application.
ity. A difference between our paper and theirs is that we exploit the sharp discontinuity at employment size
50 evident in the law and our data to identify the structural parameters of our model and the implicit cost
of regulation, whereas there does not appear to be such a sharp break in their data on Portuguese firms\textsuperscript{9}.

The structure of the paper is as follows. Section 2 describes our theory and some extensions. Section 3
describes the institutional setting and data. Section 4 describes the empirical strategy we use to map the
theory into the data. Section 5 contains the main results, which come in two parts. First we show that the
main empirical predictions of the model in terms of the size and productivity distribution are consistent with
the data. Secondly, we estimate the parameters of the structural model and use this to show that the costs of
the regulation are non-trivial. We present various extensions and robustness tests in Section 6 before drawing
some conclusions in the final section.

2 Theory

We aim to estimate the distortions in the productivity distribution and the reallocation effect that result
from an implicit tax on firm size that starts at a particular threshold. Our strategy relies on analyzing the
choices of those firms that prefer to stay at a lower size in order to avoid the tax. Having done that, we will
be able to estimate the general equilibrium effects of the tax through the changes in firm size.

We study regulatory effects on the firm size distribution and on the productivity distribution in the
simplest possible version of Lucas’ model. There is only one input in production, labor, and a single sector.
The primitive of the model is the distribution \( \phi(c) \) of ‘managerial ability’ \( \alpha \) (which we will measure as Total
Factor Productivity, \( TFP \)), with cdf \( \Phi : R^+ \to [0, 1] \). Ability is defined and measured by how much an agent
can raise a team’s output: a manager who has ability \( \alpha \) and is allocated \( n \) workers produces \( y = \alpha f(n) \). Larger
teams produce more, \( f' > 0 \), but given e.g. limited managerial time, there are decreasing returns to the firm
scale that a manager can manage, \( f'' < 0 \).

The key difference between our setting and the original Lucas model is that, in our application, there is
a tax on firm size, which imposes a wedge between the wage the worker receives and the cost to the firm.\textsuperscript{10}
Since termination costs are generally denominated in years of salary, we assume this cost is a proportional
increase in wage costs, taking the form of a labor tax. Moreover, this tax does not grow in a smooth way,
but instead it begins hitting firms after they reach a given size \( N \).

\textsuperscript{9}Braguinsky et al (2011) attribute this to the sheer multitude of size-related regulations in Portugal that makes it hard to
identify any sharp cut-off in the size distribution.

\textsuperscript{10}In our application the ‘tax’ involves an extra marginal cost and also a fixed cost component. However, previous studies of
this problem, such as particularly Kramarz and Michaud (2003) show that the fixed cost component are second order relative
to the marginal cost component.
2.1 Individual Optimization

Let $\pi(\alpha)$ be the profits obtained by a manager with skill $\alpha$ when he manages a firm at the optimal size. These profits are then given by:

$$\pi(\alpha) = \max_n \alpha f(n) - w\tau n \left\{ \begin{array}{ll} \tau = 1, k = 0 & \text{if } n < N \\ \tau = \tau, k = k & \text{if } n \geq N \end{array} \right.$$  \hspace{1cm} (1)

where $w$ is the worker’s wage, $n$ is the number of workers, $k$ is the fixed cost that must be incurred over threshold $N$, and $\tau$ is the tax, which only applies for firm over a minimum threshold of $N$ (50 workers in our application).

Firm size at each side of the threshold is then determined by first order condition:

$$\alpha f'(n^*) - \tau w = 0, \quad \text{with } \tau = \tau \text{ if } n \geq N$$  \hspace{1cm} (2)

so that $n^* = f'^{-1}(\frac{\tau w}{\alpha})$. Note that $\partial n/\partial \alpha > 0$, $\partial n/\partial \tau < 0$ and $\partial n/\partial w < 0$.

The size constraint is reached at size $N$ and managerial ability $\alpha_c$ (sub-script “c” for “constrained”) is given by:

$$\alpha_c = \frac{w}{f'(N)}$$  \hspace{1cm} (3)

Firms can legally avoid being hit by the regulation by choosing to remain small. The cost of this avoidance is increasing in the talent ($\alpha$) of the individual, and thus at a given ability level, given a choice between staying at $n = N$ and avoiding the tax, managers choose to pay the tax. The ability of the “marginal manager” that is unconstrained ($\alpha_u$) is defined by the indifference condition between remaining small and jumping to be a larger firm and paying the regulatory tax as:

$$\alpha_u f(N) - wN = \alpha_u f(n^*(\alpha_u)) - w\tau n^*(\alpha_u) - k$$  \hspace{1cm} (4)

where $n^*(\alpha_u)$ is the optimal firm size for an agent of skill $\alpha_u$. We call this threshold $\alpha_u$, where $u$ denotes the boundary of the unconstrained firms, and the firm size $n^*(\alpha_u)$ is denoted $n_u$.

2.2 Equilibrium

The most skilled individuals choose to be manager-entrepreneurs, since they benefit from their higher ability in two ways. First, for a given firm size $n$, they earn more. Second, the most skilled individuals hire a larger team, $n(\alpha)$. We denote the ability threshold between managers and workers as $\alpha_{\text{min}}$. 


A competitive equilibrium is defined as follows:

**Definition 1** Given a distribution of managerial talent $\phi(\alpha)$, a per worker labor tax $\tau$ that binds all firms of size $n \geq N$, and a production function $\alpha f(n)$, a competitive equilibrium consists of:

(i) a wage level $w$ paid to all workers

(ii) an allocation $n(\alpha)$ that assigns a firm of size $n$ to a particular manager of skill $\alpha$

(iii) a triple of cutoffs $\{\alpha_{\text{min}}, \alpha_c, \alpha_u\}$, such that $W = [0, \alpha_{\text{min}}]$ is the set of workers, $M_1 = [\alpha_{\text{min}}, \alpha_c]$ is the set of unconstrained, untaxed managers, $M_2 = [\alpha_c, \alpha_u]$ is the set of constrained, $n^* = N$, but untaxed managers, and $M_2 = [\alpha_u, \infty]$ is the set of taxed managers such that:

1. No agent wishes to change occupation from manager to worker or to change from unconstrained to constrained.

2. The choice of $n(\alpha)$ for each manager $\alpha$ is optimal given their skills, taxes $\tau$ and wages $w$;

3. Supply of labor equals demand for labor

Start with condition (1): an agent prefers to be a worker if $w > \alpha f(n) - wn$, or a manager if $w < \alpha f(n) - wn$, and thus we have:

$$\alpha_{\text{min}} f(n) - wn = w$$

Equilibrium condition (2), from the first order condition (2) implies that firm sizes are given by:

- $n(\alpha) = 0$ if $\alpha < \alpha_{\text{min}}$
- $n(\alpha) = f^{-1}\left(\frac{w}{\alpha}\right)$ if $\alpha_{\text{min}} < \alpha < \alpha_c$
- $n(\alpha) = N - 1$ if $\alpha_c < \alpha < \alpha_u$
- $n(\alpha) = f^{-1}\left(\frac{\tau w}{\alpha}\right)$ if $\alpha_u < \alpha < \infty$

Thus we have four categories of agents as the following figure shows:

**Equilibrium partition of individuals into workers and firm types by managerial ability, $\alpha$**

Finally, from condition (3), equilibrium requires that markets clear- that is the supply and demand of workers must be equalized. The supply of workers is $\Phi(\alpha_{\text{min}})$, and the demand of workers by all available...
managers, $\int_{\alpha_{\text{min}}}^{\infty} n(\alpha) d\Phi(\alpha)$, where $n(\alpha)$ is the continuous and piecewise differentiable function given as above, thus:

$$\Phi(\alpha_{\text{min}}) = \int_{\alpha_{\text{min}}}^{\infty} n(\alpha) d\Phi(\alpha)$$ (10)

Solving the model involves finding four parameters, the cutoff levels $\alpha_{\text{min}}, \alpha_c, \alpha_u$, and the equilibrium wage $w$. For this we use the four equations (3), (4), (5) and (10).

The equilibrium is unique; the following proposition characterizes the comparative statics in the equilibrium:

**Proposition 1** The introduction of a tax/variable cost $\tau$ of hiring workers starting at firm size $N$ has the following effects:

1. Reduces equilibrium wages as a result of the reduction in the demand for workers
2. Increases firm size for all firms below the threshold, $[\alpha_{\text{min}}, \alpha_c]$, as a result of the general equilibrium effect that reduces wages
3. Reduces firm size to the threshold $N$ for all firms that are constrained, that is those in $[\alpha_c, \alpha_u]$
4. Reduces firm size for all firms that are taxed $[\alpha_c, \infty]$

**Example.** Consider a power law, $\phi(\alpha) = \frac{0.6}{\alpha^{0.6}}$ and returns to scale parameter of $\theta = 0.9$. Figure 4 shows the firm size distribution for a firm size cut-off at 50 employees, and an employment tax of 1%. As in the distribution in the data, there is a spike at 49 employees that breaks the power law. Figure 5 reports the productivity distribution $\alpha$ as a function of firm size $n$. It shows that we should expect a spike in the productivity distribution at the point in which the regulation starts to bind. Essentially the maximum bar of this graph is the most productive firm that is affected by the regulation. We can trace the firm size simply by moving horizontally to the right in the graph.

### 2.3 Empirical Implications

The econometric work that follows aims to use the theory as a guide to estimate the welfare losses that result from this regulation. As is well known, the firm size distribution generally follows a power law (see e.g. Axtell, 2001). Lucas (1978) shows that Gibrat’s law implies that the returns to scale function must be $f(n) = n^\theta$, and that for it to be consistent with a power law, the managerial ability or productivity distribution must also be power, $\phi(\alpha) = c_\alpha \alpha^{-\beta_\alpha}$ with the constants $c_\alpha > 0$ and $\beta_\alpha > 0$. This is not a bad approximation: the
distribution of TFP is somewhere between log normal and power, but the fit for a power distribution is good for a large fraction of the data.

In this case, from the first order conditions, firm sizes are given, for a given wage, by:

\[ n^* = \left( \frac{\alpha \theta}{w^\tau} \right)^{1/(1-\theta)} \]  \hspace{1cm} (11)

A firm below the tax threshold \( N \) chooses a firm size \( n^* = f^{-1}(\frac{w}{\alpha}) \); while a firm above the threshold chooses \( n^*_t = f^{-1}(\frac{\tau w}{\alpha}) \). In the empirical section we rely on this relationship between TFP and firm size \( n \) to estimate the counterfactual welfare if the labor regulation \( \tau \) were to be removed:

\[ n^*(\alpha) = \begin{cases} \left( \frac{\alpha \theta}{w^\tau} \right)^{1/(1-\theta)} = c_1 \alpha^{1/(1-\theta)} = n_1(\alpha) & \text{if } \alpha_{\min} < \alpha < \alpha_c \\ N - 1 = n_1(\alpha_c) & \text{if } \alpha_c \leq \alpha < \alpha_u \\ \left( \frac{\alpha \theta}{w^\tau} \right)^{1/(1-\theta)} = c_1 \tau^{-1/(1-\theta)} \alpha^{1/(1-\theta)} = n_1(\alpha) \tau^{-1/(1-\theta)} < n_1(\alpha) & \text{if } \alpha_u \leq \alpha \end{cases} \]

with \( c_1 = \left( \frac{\alpha}{w^\tau} \right)^{1/\theta} \) and \( c_2 = c_1 \tau^{-\frac{1}{1-\theta}} \). In \([\alpha_{\min}, \alpha_c]\) we find firms that are not directly affected by the distortion. The only impact of the regulation comes through the general equilibrium impact because of lower wages \( w \). This induces some low-ability individuals to become small firms rather than remain as workers (i.e. the regulatory distortion creates too many entrepreneurial small firms). At \([\alpha_c, \alpha_u]\) we find the constrained firms: firms that, given the choice between paying the labor cost \( \tau w^* \) and choosing their optimal size and paying \( w^* \) but staying at size \( n < 50 \), prefer to stay below 50. Once productivity exceeds a higher threshold \( \alpha_u \) firms are sufficiently productive that they pay the tax in order to produce at a higher level.

Given our assumption above that the distribution of \( \phi(\alpha) \) follows a power law, \( \phi(\alpha) = c_\alpha \alpha^{-\beta_\alpha} \) the distribution of firm sizes \( \chi(n) \) is also power (apart from the threshold), since by the change of variable formula, \( \chi(n) = \phi(\alpha(n)) \ast n^{-\theta} w^{\frac{1}{1-\theta}} \) (omitting the threshold). The power law on \( n^* \) is then given by:

\[ \chi^*(n) = \begin{cases} c_\alpha (1 - \theta) \cdot c_1^{\beta - 1} \cdot n^{-\beta} & \text{if } n < N - 1 = n_1(\alpha_c) \\ \int_{\alpha_u}^{\alpha_c} \phi(\alpha) d\alpha = \delta & \text{if } n = N - 1 = n_1(\alpha_c) \\ 0 & \text{if } N - 1 < n < n_u = n_2(\alpha_u) \\ c_\alpha (1 - \theta) \cdot c_2^{\beta - 1} \cdot n^{-\beta} & \text{if } n_2(\alpha_u) = n_u \leq n \end{cases} \]

where \( \beta = \beta_\alpha (1 - \theta) + \theta \) and \( \delta \) is the mass of firms whose size is distorted- these are firms that choose to stay below the firm size threshold, rather than getting to a large size and paying the additional labor cost, \( \tau \).

The adding up constraints on \( \delta \) can be written more conveniently in the size \( n \) space rather than the ability space. After some straightforward manipulation, relegated to Appendix A, we can rewrite the pdf of \( n^* \) as:
\( \chi^*(n) = \begin{cases} 
(\frac{1}{\theta})^{1-\beta}.(\beta - 1).n^{-\beta} & \text{if } \frac{\theta}{\theta_U} < n < 49 = n_1(\alpha_c) \\
(\frac{1}{\theta})^{1-\beta}.(49^{1-\beta} - T.n_1^{1-\beta}) & \text{if } n = 49 = n_1(\alpha_c) \\
0 & \text{if } 49 < n < n_u = n_2(\alpha_u) \\
(\frac{1}{\theta})^{1-\beta}.(\beta - 1).T.n^{-\beta} & \text{if } n_2(\alpha_u) = n_u \leq n 
\end{cases} \) (14)

where \( T = \tau - \frac{\beta - 1}{1 - \theta} \). The upper employment threshold, \( n_u \), is unknown and must be estimated, and so are \( \beta \), the power law term, and \( T \). Note that the scale parameter in the power law \( \beta \) is unaffected by the law; instead, in the log-log space, the labor regulations generate a parallel shift in the firm size distribution measured by \( T \). Thus the key empirical implication is that the tax can be recovered from the jump \( T \) in the power law.

In Section 4, we propose an empirical model in which we introduce an error term in the model so that we can take it to the data. Such empirical model must account for two departures in Figure 2 from the predictions in the theory:

1. The departure from the power law does not start at \( N \), but slightly earlier: there is a bump in the distribution starting at around 46 workers.

2. The region immediately to the right of \( N \) does not have zero density, but rather there are some firms with positive employment levels just to the right of the regulatory cut-off, \( N \).

The model we propose to account for these departures features a measurement error. The justification for such mis-measurement is straightforward: the measurement of firm size that we have is not exactly the same one as the one used to determine whether a firm is subject to the regulation or not. From the perspective of the regulation, the relevant concept of employment is the number of workers at the precise date where the collective dismissal is announced. Our measure of firm size is the mandatory item that is reported in the firm’s fiscal accounts - the arithmetic mean of the workforce at the end of the quarter of the fiscal year.\(^{11}\)

3 Empirical Strategy

How costly is the employment protection legislation? We uncover this cost through revealed preference. Essentially, our approach is to identify the “constrained firms”, those which legally avoid the regulation by remaining too small, and identifying them. Once we have done this, we can calculate what they would have produced in the counterfactual world and thus we have an estimate of the cost of the regulation. In this section we explain how we apply our theoretical framework to the data we just reviewed.

\(^{11}\)Fiscal definition, article 208-III-3 du Code General des Impots.
3.1 Empirical Model

Recall that our starting point is the pdf of $n^*$, which is, according to the theory, given by equation (14).

Employment is measured with error so we assume that rather than observing $n^*(\alpha)$ we observe:

$$n(\alpha, \varepsilon) = n^*(\alpha)e^\varepsilon$$

Where the measurement error $\varepsilon$ is unobservable. In the data we observe the distribution of $n$, and thus obtaining the likelihood function requires that we obtain the density function of $n$. The law of $n|\varepsilon$, has support on $[e^\varepsilon; +\infty]$. The conditional cumulative distribution function is given by (see Appendix A):

$$P(x < n|\varepsilon) = \begin{cases} 0 & \text{if } \ln(n) - \ln\left(\frac{\theta}{1-\theta}\right) < \varepsilon \\ 1 - \left(\frac{1-\theta}{\theta}\right)^{1-\beta} \cdot (n.e^{-\varepsilon})^{1-\beta} & \text{if } \ln(n) - \ln(49) < \varepsilon \leq \ln(n) - \ln\left(\frac{\theta}{1-\theta}\right) \\ 1 - \left(\frac{1-\theta}{\theta}\right)^{1-\beta} \cdot T \cdot n_{u1}^{1-\beta} & \text{if } \ln(n) - \ln(n_{u1}) < \varepsilon \leq \ln(n) - \ln(49) \\ 1 - \left(\frac{1-\theta}{\theta}\right)^{1-\beta} \cdot T \cdot (n.e^{-\varepsilon})^{1-\beta} & \text{if } \varepsilon \leq \ln(n) - \ln(n_{u1}) \end{cases}$$

Integrating over $\varepsilon$ we can compute the unconditional CDF simply as:

$$\forall n > 0, \quad P(x < n) = \int_{\mathbb{R}} P(x < n|\varepsilon) \frac{1}{\sigma} \varphi\left(\frac{\varepsilon}{\sigma}\right) d\varepsilon.$$

In Appendix A we show that no further constraints on the parameters are required for this object to be a CDF:

**Lemma 1** Let $\varepsilon$ be normally distributed with mean 0 and variance $\sigma$ so that the measurement error is log normal. Then the function $P(x < n)$ is a cumulative distribution function, that is strictly increasing in $n$, with $\lim_{n \to 0} P = 0$ and $\lim_{n \to \infty} P = 1$ for all feasible values of all parameters, $\sigma, T, \beta, n_u$.

Thus taking the derivative of $P$ formulated in this way we can obtain the density of the observed $n$. Given such a density, it is straightforward to estimate the parameters of the model by maximum likelihood. Specifically, the maximum likelihood estimation yields estimates of the parameters: $\hat{\sigma}, \hat{T}, \hat{\beta}, \hat{n}_u$.

Figure 6 shows the difference between the pure model where employment was measured without error and the true model where there is measurement error. The solid (blue) line shows the firm size distribution under the pure model of Section 2 (same as Figure 4) whereas the hatched line shows the firm size distribution under the regulation.
when we allow for measurement error. The smoothness of the hump around 50 will depend on the degree of measurement error - Figure 6 shows that if we increase the measurement error to $\sigma = 0.5$ instead of $\sigma = 0.15$ it is almost impossible to visually identify the effects of the regulation.

### 3.2 Identification and Inference

ML estimation over the size distribution allows us to obtain most of the parameters of interest. Intuitively, the slope line in Figure 6 (which is the same before and after the cut-off) identifies $\beta$, the the power law parameter. The composite parameter $T = \tau^{-\frac{\beta-1}{1-\theta}}$ which is a function of our key object of interest the implicit tax, $\tau$, is identified from three related features of the data. First, the downward shift of the power law slope around 49 employees, the “intercept” as it where. Second, the hump of firms just before the regulatory threshold at 49 employees and third the “hole” in the size distribution between 49 employees and where the power law recovers at $n_u$. The larger is the implicit tax, the greater will be the size of the intercept, the hump of firms at the regulatory threshold and the hole in the firm size distribution. The fixed costs $k/w$, are identified from the indifference equation (4) and the measurement error from the size of the random deviations of size from the broken power law.

Given these results and the definition of $T = \tau^{-\frac{\beta-1}{1-\theta}}$, we still need an estimate for the returns to scale parameter $\theta$. There are several ways of obtain this important parameter. In principle, it can be obtained from the size distribution itself, and we will discuss this in the extensions. This method relies on rather strong assumptions over the identity of the smallest firms and the indifference condition between being a worker and a firm. As we will discuss below, the data is not rich enough to estimate $\theta$ from the size distribution alone, so we consider several alternatives in order to examine the empirical robustness of our estimates of $\tau$. Our first approach is to calibrate $\theta$ from existing estimates. Since this is well recognized to be an important parameter in the macro reallocation literature there is a sizeable literature here. Basu and Fernald (1997) show a large number of estimates based on US data and suggest a value of 0.8 is reasonable. Atkeson and Kehoe (2005) using a version of the Lucas model with organizational capital use a $\theta = 0.85$. Most calibrations seem to take a value of around 0.8 (e.g. Guner et al, 2006, use a $\theta = 0.802$ for Japan. We also consider more extreme values of $\theta = 0.5$ and $\theta = 0.9$.

A second approach is to use information from the production function. Since we have rich data on firms in our sample we can estimate production functions and from the sum of the coefficients on the factor inputs estimate returns to scale. Appendix C details how we do this using a variety of methods adapting Levinsohn and Petrin (2003), Olley and Pakes (1996) in addition to the more standard Solow residual approach. A final and related method is to use the relationship between size and TFP to back out an estimate of the returns
to scale.

For each of these estimates of $\theta$, we have an estimate of the implicit tax of regulation as:

$$\hat{\tau} = \hat{T} \cdot \frac{1-\hat{\theta}}{\beta-1} \quad (15)$$

We obtain standard errors for the estimate of the tax using the delta method. To be precise, let $\hat{\Theta} = (\hat{\beta}, \hat{T}, \hat{\theta})'$ be the vector of asymptotically Gaussian estimates obtained in the previous steps. The asymptotic variance-covariance matrix of this vector can be written as:

$$V(\hat{\Theta}) = 
\begin{pmatrix}
V_{\hat{\beta}} & \text{Cov}_{\hat{\beta}, \hat{T}} & 0 \\
\cdot & V_{\hat{T}} & 0 \\
\cdot & \cdot & V_{\hat{\theta}}
\end{pmatrix}$$

Such that an estimator of the asymptotic variance of $\hat{\tau}$ is:

$$V(\hat{\tau}) = (\frac{\partial \tau}{\partial \beta}, \frac{\partial \tau}{\partial T}, \frac{\partial \tau}{\partial \theta}).V(\hat{\Theta}).(\frac{\partial \tau}{\partial \beta}, \frac{\partial \tau}{\partial T}, \frac{\partial \tau}{\partial \theta})'$$

### 3.3 Welfare calculations

Total output of a given team, managed by manager of skill $\alpha$ and with size $n^*_\tau(\alpha)$ given by replacing the equilibrium wages $w$ in $n^*(\alpha)$ as given by (12) is:

$$y(\alpha, N, \tau, k) = \alpha f(n^*_\tau(\alpha)) - k - w* n^*_\tau(\alpha) + w n^*_\tau(\alpha) \begin{cases} 
\tau = 1, k = 0 & \text{if } n^*_\tau(\alpha) < N \\
\tau = \tau & \text{if } n^*(\alpha, \tau) \geq N
\end{cases}$$

$$= \alpha f(n^*_\tau(\alpha)) - k - (\tau - 1)w n^*_\tau(\alpha) \begin{cases} 
\tau = 1, k = 0 & \text{if } n^*_\tau(\alpha) < N \\
\tau = \tau & \text{if } n^*(\alpha, \tau) \geq N
\end{cases}$$

and thus total output is given by:

$$Y(N, k, \tau) = \int_{\alpha_{\min}}^{\alpha_{\max}} \alpha f(n^*_\tau(\alpha))d\Phi(\alpha) + 
\int_{\alpha_{\min}}^{\alpha_{\max}} \alpha f(N-1)d\Phi(\alpha) + 
\int_{\alpha_{\max}}^{\alpha_{\infty}} \alpha f(n^*_\tau(\alpha)) - k - w* n^*_\tau(\alpha)(\tau - 1)$$

And thus the welfare losses are given as follows:
\[ \Delta Y = \int_{\alpha_{fb \min}}^{\alpha_{fb \min}} \alpha [f(n^*_f(\alpha))] d\Phi(\alpha) + \int_{\alpha_{fb \min}}^{\alpha_{fc}} \alpha [f(n^*_f(\alpha)) - f(n)] d\Phi(\alpha) \]

\[ + \int_{\alpha_{fc}}^{\alpha_{u}} \alpha [f(N - 1) - f(n)] d\Phi(\alpha) + \]

\[ + \int_{\alpha_{u}}^{\infty} \alpha \{[f(n^*_f(\alpha)) - f(n)] - k - w \ast n^*_f(\alpha)(\tau - 1)\} d\Phi(\alpha) \]  

(16)

where \( n \) is the first best team size and \( \alpha_{fb \min} \) is the first best cutoff between workers and managers.

The welfare losses are then the result of adding up three effects:

1) There is an output gain given by the gain in size of small firms resulting from the general equilibrium effect on wages (wages go down) of the firing costs.

2) There is a first “local” output loss, that is the result of the firms that would have had optimal size but instead are constrained at \( N - 1 \) (49 workers).

3) Finally there is the global loss in output resulting from the smaller firm size of all firms in the economy which now have a smaller firm size than otherwise as a result to the tax on labor.

Once we have our estimates of all the model parameters, we turn back to this calculation to obtain the output losses by obtaining the counterfactual firm size distribution in the world without tax and the total production loss in this economy due to the tax distortion.

4 Institutional Setting and Data

4.1 Institutions: The French Labor market and Employment Costs

France is renowned for having a highly regulated labor market (see Abowd and Kramarz, 2003; Kramarz and Michaud, 2010). What is less well known is that most of these laws only bind on a firm when it reaches a particular employment size threshold. By far the most important size threshold is when a firm hits fifty employees - at this point of number of labour market regulations bind regarding the firm’s ability to adjust its labor. Although there are some regulations that bind when a firm (or less often, a plant) reaches a lower threshold such as 10, 20 or 25 employees, 50 is generally agreed by labour lawyers and business people to be the critical threshold when costs rise significantly (see Appendix C)\textsuperscript{13}.

Perhaps the most important of these is a set of regulations introduced under a major piece of legislation in 1989. This required firms with 50 or more employees to formulate a “social plan” before laying off 10 or more workers (a “collective termination”). This social plan must place a limit on the total number of

\textsuperscript{13} http://www.travail-emploi-sante.gouv.fr/informations-pratiques,89/lames-pratiques,91/licenciement,121/le-plan-de-sauvegarde-de-l-emploi,1107.html

14
terminations, lay out plans to facilitate reemployment of terminated workers and will typically insist on an extensive retraining program. Union representatives or personnel delegates and the departmental director of the Ministry of Labor must also be informed of the plan. Two public meetings of the works council ("comité d’entreprise") must be organized with an interval between the meetings of 2–4 weeks depending upon the number of terminations proposed. The works council may require the firm to hire a consulting accountant (at the company’s expense) to help the council with its analysis. During this period, the Ministry of Labor must be continuously informed of the proceedings, the plan, and the names of the proposed terminated workers. In addition to these firing costs in the 1989 law, there are some other pieces of regulation that bite at size 50 (see Appendix C).

How important are such provisions for firms? It is hard to know directly, as the opportunity cost of managerial time involved in preparing for such eventualities may be very great. Our framework is designed to recover the costs of such regulations. We treat such firing costs as an increase in the cost of labour. Firms face future shocks which will require them to adjust labor. Firms facing such a firing cost will effectively face a much higher cost in the eventuality that they face a negative shock. This affects the decision to hire and is (in expected value terms) very much like a labor tax. Since our analysis is fundamentally cross sectional we will model the firing cost as a labour tax.

There are other laws affecting French firms, so in one sense we are estimating a lower bound to the cost of regulation. But we are alert to the problem that some of the data is also affected by other laws which may also have a size-related threshold. Discussions with the labor ministry, lawyers, unions and business people confirm that the threshold of 50 is the most important one in France, so it makes sense to begin our analysis here.

4.2 Data

Our main dataset is administrative data covering the universe of French firms between 2002 and 2007. These hold about 2.2m observations per year, but we restrict our estimation sample to the ca 200,000 of them are active in manufacturing industries (NACE2 class 15 to 35; 227 four-digit industries). These are the (mandatory) fiscal returns of all French firms ("FICUS") and are the appropriate level for analysis as it is on this administrative unit ("entreprise") that the main laws pertain to. In addition to accurate information on employment (average number of workers in last quarter of the fiscal year), FICUS contains balance sheet information on labor, capital, investment, wage bills, materials, four digit industry affiliation, zipcode, etc. that are important in estimating productivity. More details of the dataset are given in the Appendix.

We take several approaches to estimating productivity. Our baseline results use the Levinsohn and Petrin
version of the Olley and Pakes (1996) method of using a control function approach to deal with unobserved productivity shocks and selection when estimating production functions. Because we have a panel of firms we can implement this and estimate the production function coefficients. The details of these regressions are reported in Appendix C. There are several issues with this approach (see Ackerberg et al, 2007) to estimating production functions so we also estimate TFP using a variety of other methods (see Appendix C for details).

5 Results

5.1 Qualitative analysis of the data

Before moving to the econometrics we first examine some qualitative features of the data to see whether they are consistent with our model. Many commentators have expressed skepticism about the quantitative importance of employment regulations as it is sometimes hard to observe any clear change in the size distribution around important legal thresholds, so we first focus on this issue. Figure 7 presents the empirical distribution of firm size around the cut-off of 50 employees for two datasets. The dataset we use (FICUS), the fiscal files of the French tax administration, is the population dataset of the universe of French firms that forms the basis of our econometric work. Panel 7.1 in Figure 7 is the same as Figure 2 except now on 2002 data instead of 2007 data. The qualitative features are clearly stable across years: just as in Figure 2 there is a sharp discontinuity in size precisely at 50 employees which is strong non-parametric evidence for the importance of the regulation. There are just over 400 firms with exactly 49 employees and then only about 130 with 50 employees. Importantly, the distribution which declines from 31 employees flattens after about 44 employees, just before the stacking up at 49 employees then dropping off a sharp cliff when size hits 50. The top right hand side of Figure 7 shows this in log-log space clearly indicating the evidence of a “broken power law”.

The next panel of Figure 7 compares FICUS with another dataset, DADS (Déclaration Annuelle de Données Sociales), that is also frequently typically used by labor economists. DADS is a worker-level dataset containing information on occupation (see Figure 7.2), wages and demographics. In Panel 7.2 we aggregate employment up to the appropriate level for each FICUS firm. This enables us to investigate different measures of employment such as employment dated on 31st December. The discrete jump at 50 shows up here almost as clearly as the FICUS data. The bottom panels of Figure 7 uses Full-Time Equivalents which shows less of

14 For example, Schivardi and Torrini (2008) and Boeri and Jimeno (2005) on Italian data, Braguinsky et al (2011) on Portuguese data or Abidoye et al (2010) on Sri Lankan data. The authors find that there is slower growth just under the threshold consistent with the regulation slowing growth, but they find relatively little effect on the cross-sectional distribution. This may be because of the multitude of regulations, variable enforcement or measurement error in the employment data (see sub-section 2.3).
a jump than the straight count of employees in the previous panels (the main labor laws relate to the number of workers rather than full-Time Equivalents, so this is expected). Figure 7 illustrates the importance of good data - one of the reasons that other studies have not identified such a clear discontinuity around the regulatory threshold is that they may have been using data with even greater measurement error than our own and not using the full population (as we do).

Figure 8 shows the firm size distribution over a larger range between 1 and 1,000 employees. Overall, firm size seems to approximate a power law in the employment size distribution prior to the bulge around 50. After 50, there is a sharp fall in the number of firms and the line more flat than expected before resuming what looks like another power law. Broadly, outside a “distorted” region around 50 employees, one could describe this pattern a “broken power law” with the break at 50. The finding of the power-law for firm size in France is similar to that for many other countries and has been noted by other authors (e.g. Giovanni et al, 2010; Giovanni and Levchenko, 2010), but the finding of the break in the law precisely around the main labor market regulation has not (yet) been documented in the academic literature, except in Ceci-Renaud and Chevalier (2011). As is well known the power law fits rather less well for the very small firms. Additionally, there does appear to be some break in the power law at firm size 10 and possibly as smaller one at firm size 20. This corresponds to the size thresholds from other pieces of labor and accounting regulations (see Appendix D). In order to avoid conflating these issues we focus our analysis on firms with 20 or more employees in the rest of the paper. We can generalize the methods used here to other breaks in the Power Law which we will exploit in future versions.

The distribution of TFP is presented in Figure 9. This shows that the mean level of TFP is higher in each size class of firm which is what we would expect from the model. Our basic model, following Lucas, has the implication that more talented managers leverage their ability over a greater number of workers.

We cut the same data in a slightly different way in Figure 10 plotting the mean TFP levels by firm size. Panel A does this for firms between 5 and 100 employees whereas Panel B extends the threshold out to firms with up to 1000 employees. In all panels productivity appears to rise monotonically with size, although there is more heteroskedacity for the larger firms as we would expect because there are fewer firms in each bin. The relationship between TFP and size is broadly log-linear. What is particularly interesting for our purposes, however, is the “bulge” in productivity just before the 50 employee threshold. We mark these points in red. This looks consistent with our model where some of the more productive firms who would have been just over 50 employees in the counterfactual world, choose to be below 50 employees to avoid the cost of the regulation. Firms just below the cut-off are a mixture of firms who would have had a similar employment level without

---

15See Howell (2002) for examples of how to estimate these types of distributions. More generally see Bauke (2007) for ways of consistently estimating power laws.
the implicit tax and those firms whose size is distorted by the size-related regulation.

We exploit the relationship between size and TFP to identify the \( \theta \), returns to scale parameter in the empirical estimates.

### 5.2 Econometric Implementation of the model

#### 5.2.1 Main results

The key parameters are estimated from the size distribution of firms using the ML procedure described above. We begin in Table 1 with a set of baseline results using calibrated values of \( \theta \) for the entire sample of French manufacturing firms 2002-2007\(^{16}\). Given the large sample sizes the results are very precise.

We begin by using a calibrated value of \( \theta = 0.8 \) from Basu and Fernald (1997) and Guner et al (2006, 2008) in column (1). The slope of the power law is about 1.8 and highly significant. The upper employment threshold is estimated to be about an employment level of 58 and we obtain a standard deviation of the measurement error of just over 0.10, which suggests significant, but not major amounts of mismeasured employment.

Turning the estimates of the tax equivalent costs of the regulation we obtain. The estimate of \( T \) which is determined in part by the implicit tax of 0.048 which implies that the regulation increases variable costs by 1.3 percentage points (\( \tau = 1.013 \)). This is a moderately large and important effect. By contrast the fixed cost component of the regulation is insignificant at the 5% level, incorrectly signed and small in magnitude (for a 100 worker firm the implies estimates are less than one half of 1% of the wage bill). Figure 11 shows the data and the fit of the model using the estimated parameters. Although not perfect, we seem to do a reasonable job at mimicking the size distribution around the threshold when allowing for measurement error.

If we constrain the model to set the fixed costs to zero, the results are stable, but the implied value of the variable tax falls to 1.004 with a standard error of 0.001. This does suggest that the unconstrained model may have some desirable features.

Column (2) of Table 1 considers an alternative calibration of \( \theta = 0.85 \) from Atkeson and Kehoe (2005). The results are very stable, although as expected the estimate of the marginal tax falls from 1.3% to 1%. This is because the importance of the distortions of the tax depend on returns to scale. When returns to scale are close to unity the most efficient firms have a very large share of output, so it only takes a small distortionary tax to have a large effect on the size distribution. As decreasing returns set in it takes a much larger estimate of \( \tau \) to rationalize any given distorted distribution (the “intercept”, “hump” and “hole” in

\(^{16}\)We use a sample of firms with between 10 and 1,000 employees correcting our estimates for censoring at the lower and upper known thresholds. We do this because there are other regulations that bite at 10 employees. The results are robust to reasonable extensions around these thresholds.
firm size). Column (3) considers $\theta = 0.5$ and column (4) a $\theta = 0.9$. The first case is formally equivalent to a minimum firm size of unity which is empirically consistent with the data and implies a large tax of 3.3 percentage points. The second case reduces the implicit tax to 0.4 percentage points.

Table 2 uses estimates of the returns to scale parameter directly estimated from a production function (see Table A1 and Appendix C). The first column reproduces our baseline results from column (1) of Table 1. Column (2) contains the results using the estimations from the production function giving a value of $\theta = 0.855$ that is highly significant. Other parameter estimates remain stable and we obtain an estimate of $\tau = 1.010$ only slightly lower than the baseline case of $\tau = 1.013$. Column (3) uses the TFP estimates from the production function to estimate the TFP-size relationship as in Figure 10. We recover an estimate of $\theta$ from the slope of this relationship which is $\theta = 0.799$ and re-estimate all the parameters. This generates a generally stable values with an estimate of $\tau = 1.013$.

5.3 Industry Heterogeneity

Holding the parameters constant across industries is an attempt to focus on the macro-economic consequences of the regulation. But there is nothing in our approach we requires we do this. Consequently we have investigated various ways of allowing the coefficients to vary across industries.

We begin with simply splitting the industries into high tech and low tech following OECD definitions (these are based on R&D intensity). The estimates of parameters are given in columns (4) and (5) of Table 2 and the analogous production function estimates are in the last two columns of Table A1. There does appear to be significant heterogeneity with the estimated implicit tax insignificantly different from zero in the high tech sectors and bearing more heavily in low-tech sectors (1.3%).

Next, in Table 3, we examine the other main sectors outside manufacturing. We examine heterogeneity in greater detail in Table 3 using calibrated values of $\theta = 0.80$ and $\theta = 0.85$. The first two columns repeat the baseline estimates using these values. We estimate the models for the other three large sectors of the French economy outside manufacturing in the rest of the table - Transport (columns (3) and (4)), Construction (columns (5) and (6)), wholesale and distribution Trade (columns (7) and (8)) and Business Services (columns (9) and (10)). The implicit tax seems to be more important in both Transport (2.5%) and Construction (2.0%) than it was in manufacturing (1.3%). In business services, by contrast, the regulation seems to be estimated to be insignificantly different from zero.

Finally, we estimated the production functions separately by three digit sectors and used the full ML technique with estimated production functions as in column (2) of Table 2. This allows the scale ($\theta$) and all other parameters to be freely estimated. Some of the industries have insufficient number of firms to do
this estimation but we are still able to do this for a large number. The results are in Figure A1 which again
demonstrates a substantial degree of heterogeneity with some sectors with estimates of the implied tax from
near zero to over 50%.

6 Extensions and Robustness

In this section we consider several extensions to our framework and robustness tests of the results.

6.1 Estimates of GDP and welfare loss

We have not yet completed a full welfare evaluation, but an indication of the importance of the regulation
can be gauged from a back of the envelope calculation based on how many firms around the threshold are
distorted. We estimate that $\delta = 0.05$ indicating that about 0.05% of French firms are distorted. This is a
small number but we estimate that these firms lose about 35% of their output so this contributes to a lowering
of GDP of 0.5%. This is likely to be an underestimate of the full regulatory cost because we are not inter alia
taking into account (1) the additional cost of the tax for firms above the upper threshold ($n_u$) and (2) the
distortion arising from the artificially low wage because of the incidence of the regulation on workers. There
is an offsetting positive effect on output because of the greater number and size of firms below the constraint
due to a lower equilibrium wage ($n_c$). Future work will expand these calculations. include the full welfare
calculations of equation (16).

6.2 Changing the organizational structure of corporations

An obvious way in which a business group could respond to the regulation is by splitting itself into smaller
subsidiaries. Thus a firm which wished to grow to 50 employees could split itself into two 25 employee firms
controlled by the group CEO. There are costs to such a strategy - the firm will have to file separate fiscal and
legal accounts, demonstrate that the affiliates are operating autonomously and suffer from greater problems
of loss of control. The authorities are well aware of such strategies of large firms pretending to be small in
order to avoid regulation and there are hefty fines and prison terms for executives seeking to do this.

Nevertheless, one way to check for this issue is to split the sample into those firms that are stand alone
businesses and those that are part of larger groups. If groups could simply split themselves into smaller
subsidiaries when they crossed the threshold of 50 employees we should expect to see no discontinuity in
group size around the threshold. Figure 12 splits the size distribution into subsidiaries which are standalone
and those which are part of larger groups. For the latter we aggregate employment to the group level. We
can see a clear discontinuity around 50 employees for the group size (as well as the standalone firms). This suggests that corporate restructuring does not fully undo the regulation\textsuperscript{17}.

6.3 Other margins of adjustment to the regulation

The simplest version of the model focuses on the decision over firm size based on employment. However, there are many other possible margins of adjustments that firms could take to avoid the regulation. This can be allowed for in the model by re-writing output as $y = \alpha[h(n, x)]^\theta$ instead of $y = \alpha n^\theta$ where $x$ are the other factors of production such as hours or human capital. If there was perfect substitutability between labor and these other factors then the firm could avoid the size-distortion we have discussed. More realistically when there is imperfect substitution the firm can mitigate some of the costs of the regulation through substitution. Of course, having to sub-optimally adjust into other factors of production is going to generate some welfare loss by itself.

The first way that the firm could adjust is by making its workforce work harder rather than expand the number of employees. We find clear evidence that firms respond in this way in Figure 13 as the number of annual hours increases just before the threshold of 50 employees. This is a combination of firms making workers do more overtime hours and substituting towards full time workers and away from part-timers. A second method would be by substituting across workers of different occupational types, such as employing workers who are of higher quality with more human capital. Figure 14 shows that change in the share of the main three skill groups in French firms across firm size (managers - the most skilled group, manual workers - the least skilled group and clerical workers - the middle skill group). Panel A shows the share of managers (excluding the CEO). This share seems to rise with firm size, but there is a clear change in the pattern around the threshold with firms choosing to increase their proportion of managers just after the regulatory threshold. Panel B shows almost a mirror image for manual workers - firms seem to reduce their reliance on less skilled workers around the threshold. The middle group of workers in Panel C is relatively unaffected (the smaller residual groups look broadly like Panel C). This indicates a pattern whereby instead of expanding the quantity of workers a as it nears the threshold, firms will increase the quality of employees by substituting away from low skilled manuals to more skilled managers. This enables them to increase output without necessarily increasing employment and paying the extra regulatory cost.

We also looked at other possible margins of adjustment. We see an increase in capital intensity around the threshold and a greater use of outsourced workers\textsuperscript{18}, which is also consistent with firms trying to avoid

\textsuperscript{17} A more extreme reaction of the firm would be to engage in franchising. This has some further costs as the CEO no longer has claims over the residual profits of the franchisee and loses much control. In any case, franchising is rare in manufacturing.

\textsuperscript{18} These outsourced workers will enter as an intermediate input and are therefore not included in total employment.
the regulatory costs.

Since we observe all these margins we are able to take account of them in our estimation of the production function. They should therefore not in principle bias our estimates of TFP. If the firm was able to perfectly substitute into these other types of activity the regulation would have little welfare effect as the firm would be able adjust around the regulation by smoothly adjusting into other factors.

In summary, firms do appear to be adjusting to the regulation around the threshold, in particular by attempting to increase hours and skills rather than raw labor when they get close to 50 employees. This is reassuring as it suggests that firm size is not just being misreported to avoid the regulation - firms are genuinely changing their activities in a theoretically expected direction.

6.4 Insurance Benefits of EPL?

Lazear (1990) argues that firms can respond to employment protection regulation in a “Coasian” manner by contracting around the regulation. Since workers enjoy greater insurance from negative shocks through the increased firing costs they should be prepared to accept lower wages as a compensating differential. With sufficiently flexible and enforceable contracts, the cost to the firm of being above 50 may not matter. Contracts may be very incomplete of course, causing downward nominal wage rigidity especially in France where there are strong unions and high minimum wages. Low trust between employers and employees and credit constraints for workers could also limit such contractual flexibility.

We examine this empirically by looking at wages around the threshold (see Figure 15). If there are such insurance benefits we would expect wages to be lower after the threshold of 50 employees. As expected the wage is upward sloping, but there does not appear to be a significant fall in wages after the regulatory threshold. Therefore we do not think that firms are able to completely “contract around” the regulation in this manner to offset the costs of the implicit labor tax.

6.5 Growth Analysis near the threshold

In a dynamic sense, the knowledge that going over 50 employees will generate a large increase in costs should affect the growth behavior of firms near the threshold. Since firms are hit by idiosyncratic shocks (e.g. less quits than usual), they may inadvertently get caught in the high regulation regime. Thus there is a strong disincentive to grow for firms which are approaching the threshold size. We can examine this issue by looking at growth dynamics at different size thresholds. Indeed we find that the growth process appears to be distorted near the threshold with firms just below the threshold significantly less likely to grow and firms just above the threshold significantly more likely to shrink.
More subtly, the difficulty of growing beyond a threshold may affect firms dynamic decisions. For example, if firms learn about their TFP through growing then the tax on size will reduce the learning process and cause a dynamic further. Also, if firms are concerned for stochastic reasons they may cross the threshold they may cut their size back even further below 49 in order to be well away from the threshold.

7 Conclusions

How costly is labor market regulation? This is a long-debated subject in policy circles and economics. We have tried to shed light on this issue by introducing a structural methodology that combines a simple theoretical general equilibrium approach based on the well known Lucas (1978) model of the size and productivity distribution of firms. We introduce size-specific regulations into this model, exploiting the fact that in most countries EPL only bites when firms cross specific size thresholds. We show how such a model generates predictions about the changes in the size and productivity distribution and moreover, can be used to generate an estimate of the implicit tax of the regulation. Intuitively, firms will optimally choose to remain small to avoid the regulation, so the size distribution becomes distorted with “too many” firms just below the size threshold and “too few” firms just above it. Furthermore, the distribution of productivity is also distorted: some of those firms just below the cut-off are “too productive” as they have been prevented from growing to their optimal size by the regulation. We show how the regulation creates welfare losses by (i) allocating too little employment to more productive firms who choose to be just below the regulatory threshold, (ii) allocating too little employment to more productive firms because they bear the implicit labor tax (whereas small firms do not) and (iii) through reducing equilibrium wages (due to some tax incidence falling on workers) this encourages too many individuals to become small entrepreneurs rather than working as employees for more productive entrepreneurs.

We implement this model on the universe of firms in the French private economy. France has onerous labor laws which bite when a firm has 50 employees, so is ideally suited to our framework. We find that the qualitative predictions of the model fit very well: (i) there is a sharp fall off in the firm size distribution precisely at 50 employees resembling a “broken power law” and (ii) there is a bulge in productivity just to the left of the size threshold. Having good employment measures over the population of firms helps a lot.

We then estimate the key parameters of the theoretical model from the firm size distribution. Our approach delivers quite a stable and robust cost of the employment regulation which seems to place an additional cost on labor in the range of 1-3% of the wage, but with considerable variation across industries.

This is just the start of our research program. We need to do a lot more testing of the results and extensions to the greater institutional complexity of the labor market. We believe that our approach is a simple, powerful
and potentially fruitful way to tackle the vexed problem of the impact of regulation on modern economies. Size-contingent regulations are ubiquitous and our methodology can be used for other regulations, other parts of the size distribution, other industries\textsuperscript{19} and other countries.

\textsuperscript{19}For example, the retail sector has a large number of size-contingent regulations with “big boxes” being actively discouraged in many countries and US cities (e.g. Bertrand and Kramarz, 2002, or Bally and Solow, 2001).
References


A Omitted Proofs

A.1 Adding up constraint on $\delta$

How do we derived equation (14) from equation (13)? The firm size distribution is given by the broken power law in equation (13):

$$
\chi^*(n) = \begin{cases} 
    c_\alpha c_1 (1 - \theta) c_1^{\beta-1} n^{-\beta} & \text{if } n(\alpha_{\text{min}}) < n < 49 = n_1(\alpha_c) \\
    \int_{\alpha_c}^{\alpha_0} \xi(\alpha) d\alpha = \delta & \text{if } n = 49 = n_1(\alpha_c) \\
    0 & \text{if } 49 < n < n_u = n_2(\alpha_u) \\
    c_\alpha (1 - \theta) c_2^{\beta-1} n^{-\beta} & \text{if } n(\alpha_u) = n_u \leq n
\end{cases}
$$

We have two additional restrictions:

- The constraints on $\delta$ can be re-written more conveniently in terms of regimes 1 and 2 as (the two are equivalent up to a variable change):

$$
\delta = \int_{n_1(\alpha_c)}^{n_1(\alpha_u)} c_\alpha (1 - \theta) c_1^{\beta-1} n^{-\beta} dn = \int_{n_1(\alpha_c)}^{n_1(\alpha_u)} c_\alpha (1 - \theta) c_1^{\beta-1} n^{-\beta} dn \\
= \int_{n_2(\alpha_c)}^{n_2(\alpha_u)} c_\alpha (1 - \theta) c_2^{\beta-1} n^{-\beta} dn = \int_{n_2(\alpha_c)}^{n_2(\alpha_u)} c_\alpha (1 - \theta) c_2^{\beta-1} n^{-\beta} dn \\
= \frac{c_\alpha (1 - \theta)}{\beta - 1} c_1^{\beta-1} \left( 49^{1-\beta} - \int_{n_u}^{49} \frac{\theta - 1}{\beta - 1} n^{1-\beta} dn \right)
$$

- This is a pdf, so this adds up to 1 (with support on $[\theta/(1 - \theta); +\infty]$):

$$
\delta = 1 - c_\alpha (1 - \theta) c_1^{\beta-1} \left( \frac{49^{1-\beta}}{1 - \theta} - c_\alpha (1 - \theta) c_2^{\beta-1} \frac{n_u^{1-\beta}}{1 - \theta} \right) \\
= 1 - c_\alpha \frac{1 - \theta}{\beta - 1} c_1^{\beta-1} \left( \left( \frac{\theta}{1 - \theta} \right)^{1-\beta} - 49^{1-\beta} + \frac{1}{\beta - 1} \frac{n_u^{1-\beta}}{1 - \theta} \right)
$$

Taken together, these relations imply:

$$
\delta = C \left( 49^{1-\beta} - T n_u^{1-\beta} \right) = 1 - C \left[ \left( \frac{\theta}{1 - \theta} \right)^{1-\beta} - \left( 49^{1-\beta} - T n_u^{1-\beta} \right) \right]
$$

Therefore:

$$
C = \left( \frac{1 - \theta}{\theta} \right)^{1-\beta} > 1 \quad (17)
$$

$$
\chi^*(n) = \begin{cases} 
    \left( \frac{1 - \theta}{\theta} \right)^{1-\beta} (\beta - 1) n^{-\beta} & \text{if } \frac{\theta}{1 - \theta} < n < 49 = n_1(\alpha_c) \\
    \left( \frac{1 - \theta}{\theta} \right)^{1-\beta} (49^{1-\beta} - T n_u^{1-\beta}) & \text{if } n = 49 = n_1(\alpha_c) \\
    0 & \text{if } 49 < n < n_u = n_2(\alpha_u) \\
    \left( \frac{1 - \theta}{\theta} \right)^{1-\beta} (\beta - 1) T n^{-\beta} & \text{if } n(\alpha_u) = n_u \leq n
\end{cases}
$$

29
A.2 Proof of Lemma 1

When employment is measured with error, we can only observe the following quantity:

\[ n(\alpha, \varepsilon) = n^*(\alpha).e^\varepsilon \]

We can then write the conditional CDF of this variable denoted by \( x \) below:

\[
\Pr(x < n|\varepsilon) = \begin{cases} 
0 & \text{if } n.e^{\varepsilon} < 49 \\
\frac{(1-\beta)}{(1-\frac{\theta}{1-\beta})} \cdot (1 - n.e^{\varepsilon})^{1-\beta} & \text{if } \frac{\theta}{1-\beta} \leq n.e^{\varepsilon} < 49 \\
\left(\frac{(1-\beta)}{(1-\frac{\theta}{1-\beta})}\right)^{1-\beta} \cdot \left(1 - T.n_u^{1-\beta}\right) & \text{if } 49 \leq n.e^{\varepsilon} < n_u \\
\left(\frac{(1-\beta)}{(1-\frac{\theta}{1-\beta})}\right)^{1-\beta} \cdot \left(1 - T.n_u^{1-\beta}\right) + \left(\frac{(1-\beta)}{(1-\frac{\theta}{1-\beta})}\right)^{1-\beta} \cdot (\beta-1).T.\int_{n_u}^{n.e^{\varepsilon}} x^{-\beta} \, dx & \text{if } n_u \leq n.e^{\varepsilon} \\
\end{cases}
\]

This time we can compute the unconditional probability as:

\[
\forall n > 0, \quad \Pr(x < n) = \int\Pr(x < n|\varepsilon) \frac{1}{\sigma} \varphi \left(\frac{x}{\sigma}\right) \, dx
\]

\[
= \int_{\ln(n) - \ln(49)}^{\ln(n) - \ln(49)} \left[1 - C.n^{1-\beta}.e^{\varepsilon.(\beta-1)}\right] \frac{1}{\sigma} \varphi \left(\frac{x}{\sigma}\right) \, dx
\]

\[
+ \int_{-\infty}^{\ln(n) - \ln(n_u)} \left[1 - C.T.n_u^{1-\beta}.e^{\varepsilon.(\beta-1)}\right] \frac{1}{\sigma} \varphi \left(\frac{x}{\sigma}\right) \, dx
\]

\[
= \Phi \left(\frac{\ln(n) - \ln(49)}{\sigma}\right) - C.n^{1-\beta}.e^{\varepsilon.(\beta-1)^2}
\]

\[
\Phi \left(\frac{\ln(n) - \ln(49)}{\sigma}\right) - \Phi \left(\frac{\ln(n) - \ln(n_u)}{\sigma}\right)
\]

\[
- C.T.n_u^{1-\beta}.e^{\varepsilon.(\beta-1)^2}\left(1 - \Phi \left(\frac{\ln(n) - \ln(n_u)}{\sigma}\right) - \sigma.(\beta-1)\right)
\]

In fact there is no additional constraint in the parameters, because we can show that this function is strictly increasing (straightforward from the way we constructed it), with limits 0 in 0 and 1 in +\( \infty \):
We use standard ML techniques to estimate the parameters in equation (18). Instead, as discussed in the main text we generate estimates from three alternative routes (i) calibration, (ii) estimates from the production function and (iii) using the TFP-size relationship. We use this when estimating equation (18) together with equation (17). For example, a calibrated \( \theta = 0.5 \) implies that \( C = 1 \). When we use methods (ii) and (iii) and estimate \( \theta \) using the productivity distribution, we take into account the variance around the estimation of \( \theta \) in calculating the

\[
\begin{align*}
A(n) & \xrightarrow{n \to \infty} 1 & A(n) & \xrightarrow{n \to 0} 0 \\
B(n) & \xrightarrow{n \to \infty} Cst \times (1 - 1) = 0 & B(n) & \xrightarrow{n \to 0} Cst \times (0 - 0) = 0 \\
C(n) & \xrightarrow{n \to \infty} 0 \times (1 - 1) = 0 & C(n) & \xrightarrow{n \to \infty} +\infty \times (0 - 0) = 0 \quad (*) \\
D(n) & \xrightarrow{n \to \infty} 0 \times 1 = 0 & D(n) & \xrightarrow{n \to \infty} +\infty \times 0 = 0 \quad (*)
\end{align*}
\]

To solve the two problematic cases, marked with (*), let us consider \( F(n) \) defined as:

\[
F(n) = n^{1-\beta} \Phi \left( \frac{\ln(n)}{\sigma} + F \right) = \Phi \left( \frac{\ln(n)}{\sigma} + F \right) \quad \text{(L'Hôpital's rule)}
\]

\[
\begin{align*}
& \sim \frac{1}{n} \frac{1}{\pi \sqrt{2} \sqrt{(\beta - 1)n^{\beta - 2}}} \\
& \sim \frac{1}{\sigma \sqrt{2\pi} (\beta - 1)} e^{-\frac{1}{2}(\ln(n)-\ln(n_u))^2} \quad \text{for } n \to 0
\end{align*}
\]

Last, the density that corresponds to the CDF is given by:

\[
\begin{align*}
\chi(n) &= \frac{1}{\sigma \sqrt{\pi}} \varphi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} \right) - \frac{1}{\sigma \sqrt{\pi}} C.T.n_u^{1-\beta} \left[ \varphi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} \right) - \varphi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} \right) \right] \\
&\quad - C.n^{1-\beta} \frac{\pi^2}{2} (1-\beta)^2 \left( \Phi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} - \beta \right) - \Phi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} - (\beta - 1) \right) \right) \\
&\quad - C.n^{1-\beta} \frac{\pi^2}{2} (1-\beta)^2 \frac{1}{\sigma} \left[ \varphi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} - \beta \right) - \varphi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} - (\beta - 1) \right) \right] \\
&\quad - C.n^{1-\beta} T.C.n_u^{1-\beta} \frac{\pi^2}{2} (1-\beta)^2 \left[ (1-\beta) \Phi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} - \beta \right) + \frac{1}{\sigma^2} \varphi \left( \frac{\ln(n) - \ln(n_u)}{\sigma} - (\beta - 1) \right) \right] \quad (18)
\end{align*}
\]

We use standard ML techniques to estimate the parameters in equation (18).

Note that we obtain an estimate of \( C \) from this procedure from which we can, in principle recover an estimate of the coefficient \( \theta \) from equation (17). This is unlikely to be a powerful way of identifying the scale parameter, however. We found empirically that the likelihood was very flat when trying to estimate \( \theta \) in this way, suggesting it was not well identified. Instead, as discussed in the main text we generate estimates of \( \theta \) from three alternative routes (i) calibration, (ii) estimates from the production function and (iii) using the TFP-size relationship. We use this when estimating equation (18) together with equation (17). For example, a calibrated \( \theta = 0.5 \) implies that \( C = 1 \). When we use methods (ii) and (iii) and estimate \( \theta \) using the productivity distribution, we take into account the variance around the estimation of \( \theta \) in calculating the
correct variance-covariance matrix.
B  Least squares estimation of broken Power Law

We discuss here an alternative to our MLE approach. Taking as our starting point the power law for firm sizes, we can proceed as follows:

\[ \ln \chi(n) = \ln k - \beta \ln n + \delta(D_{n>n_u}) + \sum_{i=n_u}^{n_c} d_i \]  

(19)

where \( D_{n>n_u} \) is a dummy variable that turns on to 1 for firms above the threshold \( n_u \) and is zero otherwise, but we have added \( d_i \) dummies that pick up the average number of firms in the distorted size categories, i.e. between the upper \( (n_u) \) and lower \( (n_c) \) employment thresholds. Equation (19) is estimated subject to the constraint \( \sum_{i=n_u}^{n_c} d_i = 0 \).

Following Axtell (2001), we can estimate equation (19) through OLS\(^{20}\), conditional on the ‘structural breaks’ at \( n_c \) and \( n_u \). To find these structural break points, we follow Bai (1997) and Bai and Perron (1998) in their study of structural breaks in time series models. In our context, their result implies that for each partition \( \{1, \ldots, n_c\}, \{n_c, \ldots, n_u\}, \{n_u, \ldots\} \), one obtains the OLS estimators of \( \{k, \beta, \delta_1, \delta_2\} \) subject to constraint \( \sum_{i=n_u}^{n_c} d_i = 0 \).\(^{21}\) Letting the sum of squared errors generated by each of these partitions be \( SSE(n_u, n_c) \), our estimates of the ‘break points’, \( n_u \) and \( n_c \) are:

\[ (\hat{n}_u, \hat{n}_c) = \arg \min_{n_u, n_c} SSE(n_u, n_c) \]  

(20)

Bai and Perron (1998) show that, for a wide range of error specifications (including heteroskedastic like in our case) the break points are consistently estimated, and converge at rate \( \tilde{N} \), where \( \tilde{N} \) is the maximum firm size as long as \( n_u - n_c > \varepsilon \tilde{N} \), and \( n_c < n_u \), (the break points are asymptotically distinct) which is true in our framework since we know \( n_c < N < n_u \).

Armed with these parameter estimates we can the proceed to estimate \( \tau \) using the results above. One intuitive way of seeing the procedure is as follows. Fix the lower employment threshold (say 43) and estimate the power law (conservatively) only on the part of the employment distribution below this and on the upper part of the size distribution that is undistorted (say under 43 and over 100).\(^{22}\) This procedure generates a
mass of firms (entrepreneurs) displaced to the “bulge” in the distribution between $n_c$ and $N$ (i.e. 43 and 50) as shown in Figure 9. These firms are drawn from between $N$ and $n_u$, and since we know the counterfactual slope of the power paw over this region, we can reallocate these firms so as to minimize the deviation from this counterfactual power law. $n_u$ is estimated as the maximum employment bin which is attained in this procedure.

Rather than fixing $n_c$, the Bai and Perron (1998) procedure estimates this efficiently by minimizing a sum of squares criterion along with the other parameters in the model as in equation (20).

This procedure gives us all the parameters necessary to estimate the implicit cost of the regulation which we calculate is equivalent to a labor tax of around 26% ($\tau = 1.26$).
C Using information from the productivity distribution

C.1 Incorporating TFP into the estimation method

We can do much better if we have direct information on the TFP Distribution. Estimation is a challenge here (see next sub-section), but let us initially assume we have reliable on TFP. First, recall from equation (12) the relationship between firmsize and TFP:

\[
\begin{align*}
\nu^*(\alpha) &= \begin{cases} 
\left(\frac{\alpha \theta}{w}\right)^{1/(1-\theta)} & \text{if } \alpha < \alpha_c \\
N - 1 & \text{if } \alpha_c \leq \alpha < \alpha_u \\
\left(\frac{\alpha \theta}{w + \tau}\right)^{1/(1-\theta)} e^\varepsilon & \text{if } \alpha_u \leq \alpha
\end{cases}
\end{align*}
\]

The empirical model adds a stochastic error term to this to obtain:

\[
\begin{align*}
\nu^*(\alpha) &= \begin{cases} 
\left(\frac{\alpha \theta}{w}\right)^{1/(1-\theta)} e^\varepsilon & \text{if } \alpha < \alpha_c \\
(N - 1)e^\varepsilon & \text{if } \alpha_c \leq \alpha < \alpha_u \\
\left(\frac{\alpha \theta}{w + \tau}\right)^{1/(1-\theta)} e^\varepsilon & \text{if } \alpha_u \leq \alpha
\end{cases}
\end{align*}
\]

Or

\[
\begin{align*}
\ln n_1 &= \frac{1}{1-\theta} \ln \alpha + \frac{1}{1-\theta} \ln \left(\frac{\theta}{w}\right) + \varepsilon \\
\ln n_2 &= \ln(N - 1) + \varepsilon \\
\ln n_3 &= \frac{1}{1-\theta} \ln \alpha + \frac{1}{1-\theta} \ln \tau + \frac{1}{1-\theta} \ln \left(\frac{\theta}{w}\right) + \varepsilon
\end{align*}
\]

Combining these together:

\[
\ln n = \ln n_1 I_{\{\alpha < \alpha_c\}} + \ln n_2 I_{\{\alpha_c \leq \alpha < \alpha_u\}} + \ln n_3 I_{\{\alpha_u \leq \alpha\}}
\]

where \( I \) is an indicator function for a particular regime. If we have a measure of firm-specific \( \alpha \), TFP, then we can estimate equation (21). This is one way to obtain an estimate of \( \theta \) that is needed to calculate the implicit tax of regulation. Alternatively, we can estimate \( \theta \) directly as the returns to scale parameter directly from a production function. We show the results from both methods in Table 2.

C.2 Estimation of TFP

There is no one settled way of best estimating TFP on firm level data and there are many approaches suggested in the literature. Fortunately, at least at the micro-level, different methods tend to produce results where the correlation of TFP estimated by different methods is usually high (see Syverson, 2010).
In the baseline result we follow the method of Levinsohn and Petrin (2003) who propose extending the Olley and Pakes (1996) control function method to allow for endogeneity and selection. Olley and Pakes proposed inverting the investment rule to control for the unobserved productivity shock (observed to firm but unobserved to econometrician) that affects the firm’s decision over hiring (and whether to stay in business). Because of the problem of zero investment regimes (common especially among smaller firms that we use in our dataset) Levinsohn and Petrin (2003) recommended using materials as an alternative proxy variable that (almost) always takes an observed positive value.

We use this estimator to estimate firm-level production functions on French panel data 2002-2007 (using the unbalanced panel) by each of the four-digit manufacturing industries in our dataset. We also did the same for the retail sector and the business services sector. The production functions take the form (in each industry):

\[
\ln y_{it} = \beta_n \ln n_{it} + \beta_k \ln k_{it} + \beta_m \ln m_{it} + \omega_{it} + \tau_t + \eta_{it}
\]

where \( y \) = output, \( n \) = labour, \( k \) = capital, \( m \) = materials, \( \omega \) is the unobserved productivity shock, \( \tau_t \) is a set of time dummies and \( \eta \) is the idiosyncratic error of firm \( i \) in year \( t \). From estimating the parameters of the production function we can then recover our estimate of the persistent component of TFP. Note that TFP is always normalized within industry and year.

There are of course many problems with these estimation techniques. For example, Ackerberg et al (2006) focus on the problem of exact multicollinearity of the variable factors conditional on the quasi-fixed factors given the assumption that input prices are assumed to be common across firms. Ackerberg et al (2007) suggest various solutions to this issue.

We consider alternative ways to estimate TFP including the more standard Solow approach. Here we assume that we can estimate the factor coefficients in equation (22) by using the observed factor shares in revenues. We do this assuming constant returns to scale, so \( \beta_n = \frac{wn}{py} \), \( \beta_m = \frac{cm}{py} \) and \( \beta_k = 1 - \frac{wn}{py} - \frac{cm}{py} \) where \( c \) = the price of materials. We used the four digit industry factor shares averaged over our sample period for the baseline but also experimented with some firm-specific (time invariant) factor shares. As usual these alternative measures led to similar results.

A problem with both of these methods is that we do not observe firm-specific prices so the estimates of TFP as we only control for four digit industry prices. Consequently, the results we obtain could be regarded as only revenue-based TFPR instead of quantity-based TFPQ (see Hsieh and Klenow, 2009). TFPQ is closer to what we want to theoretically obtain as our estimate of \( \alpha \). In practice, there is a high correlation between these two measures as shown by Foster et al (2008) who have actual data on plant level input and output.
prices. So it is unclear whether this would make too much of a practical difference to our results.

An alternative approach would be to follow de Loecker (2010) and put more structure on the product market. For example, assuming that the product market is monopolistically competitive enables the econometrician in principle to estimate the elasticity of demand and correct for the mark-up implicit in TFPR to obtain TFPQ. We will pursue this in future work.
D More Details of some Size-Relation Labor Market Regulations in France

The main bite of labor (and some accounting) regulations comes when the firm reaches 50 employees. But there are also some other size-related thresholds at other levels. The main other ones comes at 10-11 employees. For this reason we generally trim the analysis below 12 employees to mitigate any bias induced in estimation from these other thresholds. For more details on French regulation see inter alia Abowd and Kramarz (2003) and Kramarz and Michaud (2010).

D.1 Labor Regulations

From fifty employees:

- Obligation to use a complex redundancy plan with oversight, approval and monitoring from Ministry of Labor in case of a collective redundancy for 9 or more employees (threshold based on total employment at time of redundancy). See text.

- Appointing a shop steward if demanded by workers (threshold exceeded for 12 consecutive months during the last three years);

- Obligation to establish a committee on health, safety and working conditions (HSC) and train its members (threshold exceeded for 12 months during the last three years)

- Obligation to establish a profit sharing (threshold exceeded for six months during the accounting year within one year after the year end to reach an agreement);

- Obligation to establish a staff committee with business meeting at least every two months (plant level: threshold exceeded for 12 months during the last three years )

From twenty-five employees:

- Duty to supply a refectory if requested by all employees;

- Electoral colleges for electing representatives. Increased number of delegates from 26 employees.

From twenty employees:

- Contribution to the National Fund for Housing Assistance;

- Increase the contribution rate for continuing vocational training of 1.05% to 1.60%
• Compensatory rest of 50% for mandatory overtime beyond 41 hours per week

From eleven employees:

• Allowance of at least six months salary if terminated without cause or serious;

• Obligation to conduct the election of staff representatives (threshold exceeded for 12 consecutive months over the last three years).

• From ten employees:

• Monthly payment of social security contributions, instead of a quarterly payment (according to the actual last day of previous quarter);

• Obligation for payment of transport subsidies (Article L. 2333-64 of the General Code local authorities);

• Increase the contribution rate for continuing vocational training of 0.55% to 1.05% (threshold exceeded on average 12 months).

D.2 Accounting rules

The additional requirements depending on the number of employees of enterprises, but also limits on turnover and total assets are as follows:

From fifty employees:

• loss of the possibility of a simplified presentation of Schedule 2 to the accounts (also if the balance sheet total exceeds 2 million or if the CA exceeds 4 million);

• requirement for LLCs, the CNS, limited partnerships and legal persons of private law to designate an auditor (also if the balance sheet total exceeds 1.55 million euros or if the CA is more than 3.1 million euros, applicable rules of the current year).

From ten employees:

• loss of the possibility of a simplified balance sheet and income statement (also if the CA exceeds 534 000 euro or if the balance sheet total exceeds 267 000 euro, applicable rule in case of exceeding the threshold for two consecutive years).
Table 1: Parameter estimates (calibrating returns to scale, $\theta$)

<table>
<thead>
<tr>
<th>Method</th>
<th>(1) Unconstrained ($\theta$ calibrated from Basu and Fernald, 1997)</th>
<th>(2) Unconstrained ($\theta$ calibrated from Atkeson and Kehoe, 2005)</th>
<th>(3) Unconstrained ($\theta$ calibrated at 0.5)</th>
<th>(4) Unconstrained ($\theta$ calibrated at 0.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$, scale parameter</td>
<td>0.8</td>
<td>0.85</td>
<td>0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>$\beta$, power law</td>
<td>1.822 (0.059)</td>
<td>1.822 (0.059)</td>
<td>1.822 (0.059)</td>
<td>1.829 (0.057)</td>
</tr>
<tr>
<td>$T = \frac{1-\beta}{\tau^{1-\theta}}$</td>
<td>0.948 (0.018)</td>
<td>0.948 (0.018)</td>
<td>0.948 (0.018)</td>
<td>0.965 (0.015)</td>
</tr>
<tr>
<td>$n_u$, upper employment threshold</td>
<td>57.898 (0.024)</td>
<td>57.898 (0.024)</td>
<td>57.898 (0.024)</td>
<td>52.562 (0.002)</td>
</tr>
<tr>
<td>$\sigma$, variance of measurement error</td>
<td>0.104 (0.025)</td>
<td>0.104 (0.025)</td>
<td>0.104 (0.025)</td>
<td>0.036 (0.008)</td>
</tr>
<tr>
<td>$\tau$, implicit tax, variable cost</td>
<td>1.013 (0.005)</td>
<td>1.013 (0.005)</td>
<td>1.033 (0.013)</td>
<td>1.004 (0.002)</td>
</tr>
<tr>
<td>$k/w$, implicit tax, fixed cost</td>
<td>-0.496 (0.257)</td>
<td>-0.372 (0.192)</td>
<td>-1.243 (0.653)</td>
<td>-0.201 (0.099)</td>
</tr>
<tr>
<td>Mean (Median) # of employees</td>
<td>55.7 (23)</td>
<td>55.7 (23)</td>
<td>55.7 (23)</td>
<td>55.7 (23)</td>
</tr>
<tr>
<td>Observations</td>
<td>238,701</td>
<td>238,701</td>
<td>238,701</td>
<td>238,701</td>
</tr>
<tr>
<td>Firms</td>
<td>57,008</td>
<td>57,008</td>
<td>57,008</td>
<td>57,008</td>
</tr>
<tr>
<td>Ln Likelihood</td>
<td>-1065936</td>
<td>-1065936</td>
<td>-1065936</td>
<td>-1066165</td>
</tr>
</tbody>
</table>

Notes: Parameters estimated by ML with standard errors below in parentheses (clustered at the four digit level). Estimation is on unbalanced panel 2002-2007 of population of French manufacturing firms with 10 to 1,000 employees. These estimates of the implicit tax are based on different estimates of $\theta$; the methods are indicated in the different columns.
Table 2: Parameter estimates (exploiting information from the Production Function to estimate returns to scale, $\theta$)

<table>
<thead>
<tr>
<th>Method</th>
<th>(1) Baseline (column (1) of Table 1)</th>
<th>(2) Using Production Function estimates</th>
<th>(3) TFP/Size relationship</th>
<th>(4) High-Tech Sectors (Using Production Function estimates)</th>
<th>(5) Low-Tech Sectors (Using Production Function estimates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta$, scale parameter</td>
<td>0.8</td>
<td>0.855</td>
<td>0.799</td>
<td>0.882</td>
<td>0.848</td>
</tr>
<tr>
<td>$\beta$, power law</td>
<td>1.822</td>
<td>1.822</td>
<td>1.822</td>
<td>1.625</td>
<td>1.864</td>
</tr>
<tr>
<td>$T = \tau^{1/\theta}$</td>
<td>0.948</td>
<td>0.948</td>
<td>0.948</td>
<td>0.997</td>
<td>0.929</td>
</tr>
<tr>
<td>$n_u$, upper employment threshold</td>
<td>57.898</td>
<td>57.899</td>
<td>57.898</td>
<td>50.000</td>
<td>58.328</td>
</tr>
<tr>
<td>$\sigma$, variance of measurement error</td>
<td>0.104</td>
<td>0.104</td>
<td>0.104</td>
<td>0.000</td>
<td>0.114</td>
</tr>
<tr>
<td>$\tau$, implicit tax, variable cost</td>
<td>1.013</td>
<td>1.010</td>
<td>1.013</td>
<td>1.001</td>
<td>1.013</td>
</tr>
<tr>
<td>$k/w$, implicit tax, fixed cost</td>
<td>-0.496</td>
<td>-0.359</td>
<td>-0.498</td>
<td>-0.026</td>
<td>-0.514</td>
</tr>
<tr>
<td>Mean (Median) # of employees</td>
<td>55.7 (23)</td>
<td>55.7 (23)</td>
<td>55.7 (23)</td>
<td>78.3 (29)</td>
<td>51.3 (23)</td>
</tr>
<tr>
<td>Observations</td>
<td>238,701</td>
<td>238,701</td>
<td>238,701</td>
<td>38,713</td>
<td>199,988</td>
</tr>
<tr>
<td>Firms</td>
<td>57,008</td>
<td>57,008</td>
<td>57,008</td>
<td>9,099</td>
<td>48,139</td>
</tr>
<tr>
<td>Ln Likelihood</td>
<td>-1065936</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: Parameters estimated by ML, with standard errors below in parentheses (clustered at the four digit level). Estimation is on unbalanced panel 2002-2007 of population of French manufacturing firms with 10 to 1,000 employees. These estimates of the implicit tax are based on different estimates of $\theta$; the methods are indicated in the different columns. Standard errors are calculated using bootstrap in columns (2) to (5). “Using TFP-Size relationship” calculates $\theta = 1 - \frac{\partial \ln n}{\partial \ln \alpha} \frac{\partial \ln \alpha}{\partial \ln n}$ where $\frac{\partial \ln n}{\partial \ln \alpha} \frac{\partial \ln \alpha}{\partial \ln n}$ is calculated from the coefficient of a regression of ln(TFP) on ln(employment) on firms with 10 to 45 workers. “Using the production function” calculates $\theta$ as the sum of the coefficients on the factor inputs (see Table A1). “High tech” sectors are based on R&D intensity as defined by the OECD.
Table 3: Variation in estimates across different sectors

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta )</td>
<td></td>
<td>0.8</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \beta ), power law</td>
<td>1.822</td>
<td>1.822</td>
<td>1.878</td>
<td>1.878</td>
<td>2.372</td>
<td>2.372</td>
<td>2.128</td>
<td>2.128</td>
<td>2.001</td>
<td>2.001</td>
</tr>
<tr>
<td>( T = \frac{1-\beta}{\tau^{1-\theta}} )</td>
<td>0.948</td>
<td>0.948</td>
<td>0.898</td>
<td>0.989</td>
<td>0.871</td>
<td>0.872</td>
<td>0.885</td>
<td>0.885</td>
<td>0.984</td>
<td>0.984</td>
</tr>
<tr>
<td>( n_u, ) upper employment threshold</td>
<td>57.898</td>
<td>57.898</td>
<td>55.312</td>
<td>55.318</td>
<td>57.874</td>
<td>57.873</td>
<td>57.151</td>
<td>57.151</td>
<td>58.254</td>
<td>58.256</td>
</tr>
<tr>
<td>( \sigma, ) measurement error</td>
<td>0.104</td>
<td>0.104</td>
<td>0.060</td>
<td>0.060</td>
<td>0.089</td>
<td>0.089</td>
<td>0.084</td>
<td>0.084</td>
<td>0.106</td>
<td>0.106</td>
</tr>
<tr>
<td>( \tau ), implicit variable tax</td>
<td>1.013</td>
<td>1.010</td>
<td>1.025</td>
<td>1.019</td>
<td>1.020</td>
<td>1.015</td>
<td>1.022</td>
<td>1.016</td>
<td>1.003</td>
<td>1.003</td>
</tr>
<tr>
<td>( \frac{k}{w}, ) implicit tax, fixed cost</td>
<td>-0.496</td>
<td>-0.372</td>
<td>-1.137</td>
<td>-0.850</td>
<td>-0.842</td>
<td>-0.631</td>
<td>-0.951</td>
<td>-0.711</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>Mean (Median) # of employees</td>
<td>55.7 (23)</td>
<td>55.7 (23)</td>
<td>48.2 (23)</td>
<td>48.2 (23)</td>
<td>29.3 (17)</td>
<td>29.3 (17)</td>
<td>36.8 (19)</td>
<td>36.8 (19)</td>
<td>45.3 (20)</td>
<td>45.3 (20)</td>
</tr>
<tr>
<td>Observations</td>
<td>238,701</td>
<td>238,701</td>
<td>70,479</td>
<td>70,479</td>
<td>159,440</td>
<td>159,440</td>
<td>255,812</td>
<td>255,812</td>
<td>205,835</td>
<td>205,835</td>
</tr>
<tr>
<td>Firms</td>
<td>57,008</td>
<td>57,008</td>
<td>14,487</td>
<td>14,487</td>
<td>41,768</td>
<td>41,768</td>
<td>66,848</td>
<td>66,848</td>
<td>61,906</td>
<td>61,906</td>
</tr>
</tbody>
</table>

Notes: Parameters estimated by ML with standard errors below in parentheses (clustered at the four digit level). Estimation is on unbalanced panel 2002-2007 of firms with 10 to 1,000 employees. These estimates of the implicit tax are based on different estimates of \( \theta \); the methods are indicated in the different columns.
**Figure 1: The Firm size distribution in the US and France**

Source: FICUS for France and Census for the US. Population databases of all firms.

Notes: This is the distribution of firms (not plants). Authors’ calculations
Figure 2: Number of Firms by employment size in France

Source: FICUS

Notes: This is the population of manufacturing firms in France with between 31 and 69 employees. This plots the number of firms in each exact size category (i.e. raw data, no binning). There is a clear drop when the employment regulation begins for firms with 50 or more employees.
Figure 3: Definitions of regimes in terms of managerial ability

Notes: This figure shows the definitions of different regimes in our model. Individuals with managerial ability below $\alpha_{\text{min}}$ choose to be workers rather than managers. Individuals with ability between $\alpha_{\text{min}}$ and $\alpha_c$ are “small firms” who (conditional on the equilibrium wage, which is lower under regulation) do not change their optimal size. Between $\alpha_c$ and $\alpha_u$ are individuals who are affected by the regulatory constraint and choose their firm size to be smaller than they otherwise would have been - we call these individuals/firms who are in a “distorted” regime. Individuals with ability above $\alpha_u$ are choosing to pay the tax rather than keep themselves small.
Figure 4: Theoretical Firm size distribution with regulatory constraint

Notes: This figure shows the theoretical firm size distribution with exponentially increasing bins. The tallest bar represents the point at which the size constraint bins. Parameters: $\beta_\alpha = 1.6$, $\tau = 1.01$, $n_u = 60$, $\theta = 0.9$, $\beta = 1.06$. 
Figure 5: Theoretical Relationship between TFP (managerial talent) and firm size

Notes: This figure shows the theoretical relationship between TFP and firm size. There is a mass of firms at employment size=50 where the regulatory constraint binds. Parameters: $\beta = 1.6$, $\tau = 1.01$, $n_u = 60$, $\theta = 0.9$, $\beta = 1.06$. 
Figure 6: The Theoretical Firm Size Distribution when employment is measured with error

Note: The solid (blue) line shows the theoretical firm size distribution (broken power law), $n^*$. The dashed line shows the new firm size distribution when we extend the model, to allow employment size to be measured with error with $\sigma = 0.15$. The solid dark line increases the measurement error to $\sigma = 0.5$.
Figure 7: The effect on the measured firm size distribution using Alternative Datasets and definitions of employment

**FICUS (2002): Fiscal source**
Arithmetic average of quarterly head counts.

**Bar plot**

**Log-log plot**

**DADS (2002): Payroll tax reporting to social administration**
"Declared" workers on Dec. 31st: cross-sectional "fractional" count, i.e. taking account of part-timers.

**Bar plot**

**Log-log plot**

**DADS (2002): Payroll tax reporting to social administration**
"Full-time equivalent", computed by the French statistical institute as \( \min \left\{ \frac{\text{hours}}{P\text{atten}(H_{\text{last}, \text{size}})}, 1 \right\} \)
Figure 8: Share of Firms by employment size
Notes: This figure plots the (kernel density smoothed) distribution of TFP (estimated by the Levinsohn Petrin method) across four size classes. TFP is relative to the four digit industry by year average.
Figure 10: TFP Distribution around the regulatory threshold of 50 employees

Panel A: Short Employment span

Panel B: Longer Employment span

Notes: This figure plots the mean level of TFP by firm employment size using an upper support of 100 (Panel A) or 500 (Panel B). A fourth order polynomial is displayed in both panels using only data from the "undistorted" points (shown in red).
Figure 11: Firm Size Distribution and Broken Power Law: Data and Fit of Model

Notes: This shows the difference between the fit of the model (dashed red line) which allows for measurement error with the actual data. We also include the “pure” theoretical predictions (in blue).
Figure 12: Corporate Restructuring in Response to the Regulation?

Notes: “Under Control” are firms who are subsidiaries of larger business groups (blue crosses). Black crosses reports employment aggregated to the business group level. This shows that the discontinuity at 50 is not simply due to subsidiaries of larger business groups as we can see the relationship even at the group level.
Figure 13: Adjustment in the hours margin around the threshold (annual hours per worker)

Notes: Annual average hours per worker - combined FICUS and DADs data. 95% confidence intervals shown.
Figure 14: Adjustment in Types of Labor

A. Share of Managers

B. Share of manual workers

C. Share of Clerical Workers

Notes: This looks at the share of the main three occupational groups using combined FICUS and DADs data. Managers and professionals are the most skilled group and manual workers are the least skilled group (clerical workers are in between). 95% confidence intervals shown.
Figure 15: Workers do not appear to be accepting significantly lower wages in return for “insurance” of employment protection.

Notes: Wages is the nominal wage (net of payroll tax) by employer size. 95% confidence intervals shown.
## Appendices

### Table A1: Production Function Estimation

<table>
<thead>
<tr>
<th></th>
<th>(1) Baseline</th>
<th>(2) High Tech Sectors</th>
<th>(3) Low Tech Sectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>0.739</td>
<td>0.756</td>
<td>0.735</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.011)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Capital</td>
<td>0.116</td>
<td>0.126</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.012)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Observations</td>
<td>219,938</td>
<td>35,233</td>
<td>184,705</td>
</tr>
<tr>
<td>Firms</td>
<td>53,127</td>
<td>8,410</td>
<td>44,931</td>
</tr>
</tbody>
</table>

**Notes:** Parameters estimated by Levinsohn-Petrin (2003) method. Estimation on unbalanced panel 2002-2007 of population of French manufacturing firms with 10 to 1,000 employees. “High tech” sectors are based on R&D intensity as defined by the OECD.
Figure A1: Heterogeneity of Results by three digit sector in manufacturing

Notes: These are the results from industry-specific estimation on the same lines as column (2) of Table 2