Firm Size Distortions and the Productivity Distribution: Evidence from France

Luis Garicano* Claire Lelarge† John Van Reenen‡

Online Appendix

Contents

*Department of Management, Department of Economics and Centre for Economic Performance, London School of Economics, and CEPR (luis.garicano@gmail.com).

†Centre for Economic Performance (foreign affiliate, claire.lelarge@ensae.fr)

‡Centre for Economic Performance, London School of Economics, Houghton Street, London, WC2A 2AE, United Kingdom, CEPR and NBER (j.vanreenen@lse.ac.uk).

A Proof of Propositions 1 and 2

Recall that firm optimization yields a piecewise continuous and differentiable firm size function $n^*(\alpha)$ defined by equations (6a), (6b), (6c) and (6d). The equilibrium requires finding w, $\alpha_{\min}$ and α_r such that occupational choice (equation (5)), supply equals demand (equation (7)) and managerial choice between paying or not taxes (equation (4)) holds.[1] We can rewrite these equations as:[2]

$$
\begin{aligned}
\alpha_{\min} f(n^*(\alpha_{\min})) - wn^*(\alpha_{\min}) - w &= 0 && (19)\\
\int_{\underline{\alpha}}^{\alpha_{\min}} \phi(\alpha)\mathrm{d}\alpha - \int_{\alpha_{\min}}^{\alpha_c(w,N)} n^*(\alpha)\phi(\alpha)\mathrm{d}\alpha - \int_{\alpha_c(w,N)}^{\alpha_r} N\phi(\alpha)\mathrm{d}\alpha - \int_{\alpha_r}^{\alpha_{\max}} n^*(\alpha)\phi(\alpha)\mathrm{d}\alpha &= 0 && (20)\\
\alpha_r f(n^*(\alpha_r)) - w\tau n^*(\alpha_r) - F - (\alpha_r f(N) - wN) &= 0 && (21)
\end{aligned}
$$

where $\alpha_{\max}$ is finite or infinite, and where n^* has been replaced by its (constant) value between α_c and α_r. In vector form, these three equations write (ignoring N to reduce notation): $g(\alpha_{\min}, w, \alpha_r; \tau, F) = 0$.

Let (τ_0, F_0) be a specific value for the variable and fixed costs of the regulation and $\mathbf{x} = (\alpha_{\min}, w, \alpha_r)$. By the implicit function theorem there exists a function $h : A \to R$ on a neighborhood A of (τ_0, F_0) such that $(\mathbf{x}) = h(\tau, F)$ and $g(h(\tau, F), \tau, F) = 0$ for every $(\tau, F) \in A$, as long as $D_{\mathbf{X}} g(\mathbf{x}_0, \tau_0, F_0)$ is non singular (which is true as we show in the next sub-section).

Note that all functions are differentiable in the parameters in these equilibrium conditions except that n^* is not differentiable at α_r. Nevertheless all three equilibrium conditions are differentiable in α_r: the first is not a function of α_r; the second is differentiable in α_r despite the discontinuity of n^*, and the third is differentiable in α_r.

A.1 Comparative Statics with Respect to τ

To obtain the comparative statics described we compute $dh/d\tau = -\left(D_{\mathbf{X}} g(\mathbf{x}, \tau)\right)^{-1} D_\tau g(\mathbf{x}, \tau)$. First, computing derivatives and using the envelope theorem, $D_{\mathbf{X}} g(\mathbf{x}, \tau)$ simplifies to:

$$
D_{\mathbf{X}} g(\mathbf{x}, \tau) = \begin{pmatrix} f(n_{\min}) & -n_{\min} - 1 & 0 \\ \phi(\alpha_{\min})(1 + n_{\min}) & -I_w & \phi(\alpha_r)\left(n_r - N\right) \\ 0 & -\left(\tau n_r - N\right) & f(n_r) - f(N) \end{pmatrix} \tag{22}
$$

where:

$$
\begin{aligned}
I_w &= \int_{\alpha_{\min}}^{\alpha_c} \frac{\partial n^*(\alpha)}{\partial w} \phi(\alpha)\mathrm{d}\alpha + \int_{\alpha_r}^{\alpha_{\max}} \frac{\partial n^*(\alpha)}{\partial w} \phi(\alpha)\mathrm{d}\alpha \\
&= \int_{\alpha_{\min}}^{\alpha_c} \left(\frac{1}{\alpha f''(n^*)}\right) \phi(\alpha)\mathrm{d}\alpha + \int_{\alpha_r}^{\alpha_{\max}} \left(\frac{\tau}{\alpha f''(n^*)}\right) \phi(\alpha)\mathrm{d}\alpha < 0
\end{aligned}
$$

The second equality follows from equation (6d), $n^*(\alpha, w, \overline{\tau}) = f'^{-1}\left(\frac{\overline{\tau} w^*}{\alpha}\right)$.

[1] For compactness, α_c is here directly replaced with $\alpha_c(w, N) = \left(n^*\right)^{-1}(N) = \frac{w}{f'(N)}$.

[2] It is useful to also write equations (19) and (21) in terms of firm size (using $\alpha_{\min} = \frac{w}{f'(n_{\min})}$ and $\alpha_r = \frac{\tau w}{f'(n_r)}$):

$$
\begin{aligned}
f(n_{\min}) &= (1 + n_{\min}).f'(n_{\min}) \\
f(n_r) - f(N) &= f'(n_r).\left[n_r - \frac{N}{\tau} + \frac{F}{\tau.w}\right]
\end{aligned}
$$

Second, we have:

$$D_\tau g(\mathbf{x},\tau) = \begin{pmatrix} 0 \\ -I_\tau \\ -wn_r \end{pmatrix}, \tag{23}$$

where I_τ can be expressed as:

$$\begin{aligned} I_\tau &= \int_{\alpha_r}^{\alpha_{\max}} \frac{\partial n^*(\alpha)}{\partial \tau}\phi(\alpha)\mathrm{d}\alpha \\ &= \int_{\alpha_r}^{\alpha_{\max}} \left(\frac{w}{\alpha f''(n^*)}\right)\phi(\alpha)\mathrm{d}\alpha < 0 \end{aligned}$$

Then from $-\left(D_X g(\mathbf{x},\tau)\right)^{-1} D_\tau g(\mathbf{x}_0,\tau_0)$ and simplifying we have:

$$\frac{d\alpha_{\min}}{d\tau} = \frac{(n_{\min}+1)\left[\underbrace{I_\tau}_{<0}.(f(n_r)-f(N)) - w.n_r.(n_r-N).\phi(\alpha_r)\right]}{D} < 0 \tag{24}$$

$$\frac{dw}{d\tau} = \frac{f(n_{\min})\left[I_\tau.(f(n_r)-f(N)) - w.n_r.(n_r-N).\phi(\alpha_r)\right]}{D} < 0 \tag{25}$$

$$\frac{d\alpha_r}{d\tau} = \frac{f(n_{\min})n_r\left[\tau.I_\tau - w.I_w\right] - NI_\tau f(n_{\min}) + wn_r(1+n_{\min})^2\phi(\alpha_{\min})}{D} > 0 \tag{26}$$

In these equations, the determinant D of matrix $D_{\mathbf{X}} g(\mathbf{x},\tau)$ is:[3]

$$D = f(n_{\min}).\left[\underbrace{-I_w}_{>0}.(f(n_r)-f(N)) + (\tau n_r - N).\phi(\alpha_r).(n_r-N)\right] + \phi(\alpha_{\min}).(f(n_r)-f(N)).(1+n_{\min})^2 > 0$$

Equations 24 and 25 are straighforward and hold true a fortiori starting from $\tau = 1$, the situation without taxes. Points (i) and (ii) of Proposition 1 follow immediately.

Equation 26 also holds true since $I_\tau < 0$ and :

$$\begin{aligned} \tau.I_\tau - w.I_w &= \int_{\alpha_r}^{\alpha_{\max}} \left(\frac{\tau.w}{\alpha f''(n^*)}\right)\phi(\alpha)\mathrm{d}\alpha - \int_{\alpha_{\min}}^{\alpha_c}\left(\frac{w}{\alpha f''(n^*)}\right)\phi(\alpha)\mathrm{d}\alpha - \int_{\alpha_r}^{\alpha_{\max}}\left(\frac{\tau.w}{\alpha f''(n^*)}\right)\phi(\alpha)\mathrm{d}\alpha \\ &= -\int_{\alpha_{\min}}^{\alpha_c}\left(\frac{w}{\alpha f''(n^*)}\right)\phi(\alpha)\mathrm{d}\alpha > 0 \end{aligned}$$

This results proves point (iii) since the (unconditionnal) bulge generated in the distribution at N is $\int_{\alpha_c}^{\alpha_r}\phi(\alpha)\mathrm{d}\alpha$. Differentiating with respect to τ, we get:

$$\begin{aligned} \frac{d\int_{\alpha_c}^{\alpha_r}\phi(\alpha)\mathrm{d}\alpha}{d\tau} &= \phi(\alpha_r)\frac{d\alpha_r}{d\tau} - \phi(\alpha_c)\frac{d\alpha_c}{d\tau} \\ &= \phi(\alpha_r)\frac{d\alpha_r}{d\tau} - \phi(\alpha_c)\frac{1}{f'(N)}\frac{dw}{d\tau} > 0 \end{aligned}$$

Last, point (iv) is less straightforward. The partial equilibrium effect of taxes on size is immediate as:

$$\frac{\partial n}{\partial \tau} = \frac{w}{\alpha f''(n^*)} < 0.$$

[3] $D > 0$ shows that $D_X g(\mathbf{x},\tau)$ is non-singular.

However, the general equilibrium effect is the possible reversal induced by the reduction in wages calculated above:

$$\left.\frac{dn}{d\tau}\right|_{\alpha>\alpha_r} = \frac{\partial n}{\partial \tau} + \frac{\partial n}{\partial w}\frac{dw}{d\tau} = \frac{1}{\alpha f''(n^*)}\left(w + \tau\frac{dw}{d\tau}\right) \tag{27}$$

and thus $\left.\frac{dn}{d\tau}\right|_{\alpha>\alpha_r} < 0$ iff $w + \tau\frac{dw}{d\tau} > 0$. We get equivalently:

$$\begin{aligned} w + \tau\frac{f(n_{\min})\left[I_\tau.(f(n_r) - f(N)) - w.n_r.(n_r - N).\phi(\alpha_r)\right]}{D} &> 0 \\ f(n_{\min}).\left[\tau.I_\tau - w.I_w\right].\left[f(n_r) - f(N)\right] + w.\left[\phi(\alpha_{\min}).\left[f(n_r) - f(N)\right].(1 + n_{\min})^2 - \phi(\alpha_r).f(n_{\min}).N.(n_r - N)\right] &> 0 \end{aligned}$$

This holds true because the first term is positive since $\tau.I_\tau - w.I_w > 0$ as previously shown. The second term is also positive since we can use the conditions at $\alpha_{\min}$ and α_r to get:

$$\begin{aligned} \phi(\alpha_{\min}).\left[f(n_r) - f(N)\right].(1 + n_{\min})^2 &- \phi(\alpha_r).f(n_{\min}).N.(n_r - N) \geq 0 \\ &\iff \frac{\phi(\alpha_{\min})}{f'(\alpha_{\min})}.(1 + n_{\min}) \geq \frac{\phi(\alpha_r)}{f'(\alpha_u)}.\frac{n_r - N}{n_r - \frac{N}{\tau} + \frac{F}{\tau.w}} \\ &\iff \alpha_{\min}.\phi(\alpha_{\min}).\underbrace{(1 + n_{\min})}_{\geq 1} \geq \alpha_r.\phi(\alpha_r).\underbrace{\frac{n_r - N}{\tau.n_r - N + \frac{F}{w}}}_{\leq 1} \end{aligned}$$

This holds true because $\alpha.\phi(\alpha)$ is assumed to be decreasing in α ("scarcity of talent").

A.2 Comparative Statics with Respect to F

Derivations are analogous with respect to F. We have:

$$D_F g(\mathbf{x}, F) = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \tag{28}$$

such that, using the same notations as before:

$$\frac{d\alpha_{\min}}{dF} = -\frac{1}{D}.(n_{\min} + 1).\phi(\alpha_r).(n_r - N) < 0 \tag{29}$$

$$\frac{dw}{dF} = -\frac{1}{D}.f(n_{\min}).\phi(\alpha_r).(n_r - N) < 0 \tag{30}$$

$$\frac{d\alpha_r}{dF} = \frac{1}{D}.\left[(n_{\min} + 1)^2\phi(\alpha_{\min}) - I_w f(n_{\min})\right] > 0 \tag{31}$$

Last, in contrast to the comparative statics with respect to τ, we get:

$$\left.\frac{dn}{dF}\right|_{\alpha>\alpha_r} = \frac{\partial n}{\partial F} + \frac{\partial n}{\partial w}\frac{dw}{dF} = 0 + \frac{\tau}{\alpha f''(n^*)}.\frac{dw}{dF} > 0$$

B Details of the Baseline Estimated Model

This Appendix describes the details of the model laid out in sections I.C and II.

B.1 Firms' Objective Function and Labor Demand

In our baseline, single input model, the objective function of firms is:

$$\pi(\alpha) = \max_n \begin{cases} \alpha n^\theta - wn & \text{if } n \leq N \\ \alpha n^\theta - w\tau n - F & \text{if } n > N \end{cases} \tag{32}$$

Optimal labor demand is easily derived from the first order conditions:

$$n^*(\alpha) = \begin{cases} \left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}}.\alpha^{\frac{1}{1-\theta}} & \text{if } \alpha \in [\alpha_{\min};\alpha_c] \\ N & \text{if } \alpha \in]\alpha_c;\alpha_r] \\ \left(\frac{\theta}{\tau.w}\right)^{\frac{1}{1-\theta}}.\alpha^{\frac{1}{1-\theta}} & \text{if } \alpha \in]\alpha_r;1] \end{cases} \tag{33}$$

The indifference condition at $\alpha_{\min}$ is:

$$\alpha_{\min}.n_{\min}^\theta - w.n_{\min} = w \tag{34}$$

Using the relation between ability and size, we obtain: $n_{\min} = \frac{\theta}{1-\theta}$.

B.2 Density of the Firm Size Distribution

Using the previous variable change, we derive the density of the theoretical firm size distribution as:

$$\chi^*(n) = \begin{cases} \underbrace{\frac{c_\alpha}{p}(1-\theta)\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}}}_{=C} n^{-\beta} & \text{if } n_{\min} \leq n < N = n^*(\alpha_c) \\ \frac{1}{p}\int_{\alpha_c}^{\alpha_r}\phi(\alpha)\mathrm{d}\alpha = \delta & \text{if } n = N \\ 0 & \text{if } N < n < n_r \\ \underbrace{\frac{c_\alpha}{p}(1-\theta)\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}}}_{=C} \underbrace{\tau^{-\frac{\beta-1}{1-\theta}}}_{=T} n^{-\beta} & \text{if } n^*(\alpha_r) = n_r \leq n \end{cases}$$

where p is the proportion of entrepreneurs across all agents in the economy.
The quantities δ and C can be further specified. First, δ can be expressed in terms of firm size rather than ability, using the following variable change: $n(\alpha) = \left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}}.\alpha^{\frac{1}{1-\theta}}$ (see equation 33). We get:

$$\begin{aligned} \delta &= \int_{\alpha_c}^{\alpha_r}\phi(\alpha)\mathrm{d}\alpha \\ &= \int_{\left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}}.\alpha_c^{\frac{1}{1-\theta}}}^{\left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}}.\alpha_r^{\frac{1}{1-\theta}}} \frac{c_\alpha}{p}.(1-\theta).\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}}.s^{-\beta}\mathrm{d}s = \int_N^{\tau^{\frac{1}{1-\theta}}.n_r} \frac{c_\alpha}{p}.(1-\theta).\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}}.s^{-\beta}\mathrm{d}s \\ &= \frac{c_\alpha}{p}.\frac{1-\theta}{\beta-1}.\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}}.\left(N^{1-\beta} - \tau^{-\frac{\beta-1}{1-\theta}}.n_r^{1-\beta}\right) = \frac{C}{\beta-1}.\left(N^{1-\beta} - T.n_r^{1-\beta}\right) \end{aligned}$$

Moreover, add-up constraints on the probability density function χ^* provides an alternative expression for δ:

$$\begin{aligned} \delta &= 1 - \frac{c_\alpha}{p}.\frac{1-\theta}{\beta-1}.\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}}.\left((n_{\min})^{1-\beta} - N^{1-\beta} + \tau^{-\frac{\beta-1}{1-\theta}}.n_r^{1-\beta}\right) \\ &= 1 - \frac{C}{\beta-1}\left(\left(\frac{\theta}{1-\theta}\right)^{1-\beta} - N^{1-\beta} + T.n_r^{1-\beta}\right) \end{aligned}$$

Taken together, the two previous expressions imply: $\frac{C}{\beta-1} = \left(\frac{1-\theta}{\theta}\right)^{1-\beta}$ and $\delta = \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . \left(N^{1-\beta} - T.n_r^{1-\beta}\right)$

We obtain a simplified expression for the theoretical firm size distribution, which corresponds to equation (8) in the main text:

$$\chi^*(n) = \begin{cases} \left(\frac{1-\theta}{\theta}\right)^{1-\beta} (\beta-1)n^{-\beta} & \text{if } n_{\min} \le n < N \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} \left(N^{1-\beta} - Tn_r^{1-\beta}\right) & \text{if } n = N \\ 0 & \text{if } N < n < n_r \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} (\beta-1)Tn^{-\beta} & \text{if } n_r \le n \end{cases} \tag{35}$$

B.3 Empirical Model and Proof of Lemma 1

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

$$n(\alpha, \varepsilon) = n^*(\alpha).e^{\varepsilon} \tag{36}$$

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

$$\begin{aligned} \mathbb{P}(x < n|\varepsilon) &= \begin{cases} 0 & \text{if } n.e^{-\epsilon} \le n_{\min} \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} .(\beta-1). \int_{n_{\min}}^{n.e^{-\epsilon}} x^{-\beta} \mathrm{d}x & \text{if } n_{\min} \le n.e^{-\epsilon} < N \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . \underbrace{(\beta-1). \int_{n_{\min}}^{N} x^{-\beta} \mathrm{d}x}_{=\left(\frac{\theta}{1-\theta}\right)^{1-\beta} - N^{1-\beta}} + \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . \left(N^{1-\beta} - T.n_r^{1-\beta}\right) & \text{if } N \le n.e^{-\epsilon} \le n_r \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . \left(\left(\frac{\theta}{1-\theta}\right)^{1-\beta} - T.n_r^{1-\beta}\right) + \left(\frac{1-\theta}{\theta}\right)^{1-\beta} .(\beta-1).T. \int_{n_r}^{n.e^{-\epsilon}} x^{-\beta} \mathrm{d}x & \text{if } n_r \le n.e^{-\epsilon} \end{cases} \\ &= \begin{cases} 0 & \text{if } \ln(n) - \ln(n_{\min}) \le \varepsilon \\ 1 - \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . \left(n.e^{-\varepsilon}\right)^{1-\beta} & \text{if } \ln(n) - \ln(N) < \varepsilon \le \ln(n) - \ln(n_{\min}) \\ 1 - \left(\frac{1-\theta}{\theta}\right)^{1-\beta} .T.n_r^{1-\beta} & \text{if } \ln(n) - \ln(n_r) \le \varepsilon \le \ln(n) - \ln(N) \\ 1 - \left(\frac{1-\theta}{\theta}\right)^{1-\beta} .T. \left(n.e^{-\varepsilon}\right)^{1-\beta} & \text{if } \varepsilon \le \ln(n) - \ln(n_r) \end{cases} \end{aligned}$$

Assuming that ε is a Gaussian noise with mean 0 and variance σ, and denoting by φ the Gaussian pdf and by Φ the Gaussian cdf, we can compute the unconditional probability as:

$$\begin{aligned} \forall n > 0, \quad \mathbb{P}(x < n) &= \int_{\mathbb{R}} \mathbb{P}(x < n|\varepsilon) \frac{1}{\sigma}.\varphi\left(\frac{\varepsilon}{\sigma}\right) \mathrm{d}\varepsilon \\ &= \int_{\ln(n)-\ln(N)}^{\ln(n)-\ln(n_{\min})} \left[1 - \frac{C}{\beta-1}.n^{1-\beta}.e^{\varepsilon.(\beta-1)}\right] \frac{1}{\sigma}.\varphi\left(\frac{\varepsilon}{\sigma}\right) \mathrm{d}\varepsilon + \int_{\ln(n)-\ln(n_r)}^{\ln(n)-\ln(N)} \left[1 - \frac{C}{\beta-1}.T.n_r^{1-\beta}\right] \frac{1}{\sigma}.\varphi\left(\frac{\varepsilon}{\sigma}\right) \mathrm{d}\varepsilon \\ &\quad + \int_{-\infty}^{\ln(n)-\ln(n_r)} \left[1 - \frac{C}{\beta-1}.T.n^{1-\beta}.e^{\varepsilon.(\beta-1)}\right] \frac{1}{\sigma}.\varphi\left(\frac{\varepsilon}{\sigma}\right) \mathrm{d}\varepsilon \\ &= \underbrace{\Phi\left(\frac{\ln(n) - \ln(n_{\min})}{\sigma}\right)}_{=A(n)} - \underbrace{\frac{C}{\beta-1}.T.n_r^{1-\beta}. \left[\Phi\left(\frac{\ln(n) - \ln(N)}{\sigma}\right) - \Phi\left(\frac{\ln(n) - \ln(n_r)}{\sigma}\right)\right]}_{=B(n)} \\ &\quad \underbrace{- \frac{C}{\beta-1}.n^{1-\beta}.e^{\frac{\sigma^2}{2}.(\beta-1)^2}. \left[\Phi\left(\frac{\ln(n) - \ln(n_{\min})}{\sigma} - \sigma.(\beta-1)\right) - \Phi\left(\frac{\ln(n) - \ln(N)}{\sigma} - \sigma.(\beta-1)\right)\right]}_{=C(n)} \\ &\quad \underbrace{- \frac{C}{\beta-1}.T.n^{1-\beta}.e^{\frac{\sigma^2}{2}.(\beta-1)^2}.\Phi\left(\frac{\ln(n) - \ln(n_r)}{\sigma} - \sigma.(\beta-1)\right)}_{=D(n)} \end{aligned} \tag{37}$$

One can easily check that this function is strictly increasing (straightforward from the way we constructed it), with limits 0 in 0 and 1 in $+\infty$:

$$\begin{aligned}
A(n) &\underset{n\to+\infty}{\longrightarrow} 1 & A(n) &\underset{n\to 0}{\longrightarrow} 0 \\
B(n) &\underset{n\to+\infty}{\longrightarrow} Cst\times(1-1)=0 & B(n) &\underset{n\to 0}{\longrightarrow} Cst\times(0-0)=0 \\
C(n) &\underset{n\to+\infty}{\longrightarrow} 0\times(1-1)=0 & C(n) &\underset{x\to 0}{\longrightarrow} +\infty\times(0-0)=0\ (*) \\
D(n) &\underset{n\to+\infty}{\longrightarrow} 0\times 1=0 & D(n) &\underset{n\to 0}{\longrightarrow} +\infty\times 0=0\ (*)
\end{aligned}$$

To solve the two problematic cases, marked with $(*)$, let us consider $\mathcal{F}(n)$ defined for $\mathcal{F}\in\mathbb{R}$ as:

$$\begin{aligned}
\mathcal{F}(n) &= n^{1-\beta}.\Phi\left(\frac{\ln(n)}{\sigma}+\mathcal{F}\right)=\frac{\Phi\left(\frac{\ln(n)}{\sigma}+\mathcal{F}\right)}{n^{\beta-1}} \\
\text{(L'Hôpital's rule)}\quad &\underset{n\to 0}{\sim} \frac{\frac{1}{n.\sigma}\varphi\left(\frac{\ln(n)}{\sigma}+\mathcal{F}\right)}{(\beta-1).n^{\beta-2}} \\
&\underset{n\to 0}{\sim} \frac{1}{\sigma.\sqrt{2\pi}.(\beta-1)}.e^{-\frac{F^2}{2}}.\underbrace{n^{1-\beta}.e^{-\frac{1}{2}.\frac{(\ln n)^2}{\sigma^2}}.e^{-\frac{1}{2}.\mathcal{F}.\frac{\ln(n)}{\sigma}}}_{=e^{\ln(n).\left(1-\beta-\frac{1}{2}.\frac{\mathcal{F}}{\sigma}-\frac{\ln(n)}{2.\sigma^2}\right)}} \\
&\underset{n\to 0}{\longrightarrow} 0
\end{aligned}$$

Last, the density corresponding to the previous CDF is given by:

$$\begin{aligned}
\chi(n) =\ & \frac{1}{\sigma.n}.\varphi\left(\frac{\ln(n)-\ln(n_{\min})}{\sigma}\right)-\frac{1}{\sigma.n}.\frac{C}{\beta-1}.T.n_r^{1-\beta}\left[\varphi\left(\frac{\ln(n)-\ln(N)}{\sigma}\right)-\varphi\left(\frac{\ln(n)-\ln(n_r)}{\sigma}\right)\right] \\
&-\frac{C}{\beta-1}.n^{-\beta}.e^{\frac{\sigma^2}{2}.(\beta-1)^2}.(1-\beta).\left(\Phi\left(\frac{\ln(n)-\ln(n_{\min})}{\sigma}-\sigma.(\beta-1)\right)-\Phi\left(\frac{\ln(n)-\ln(N)}{\sigma}-\sigma.(\beta-1)\right)\right) \\
&-\frac{C}{\beta-1}.n^{-\beta}.e^{\frac{\sigma^2}{2}.(\beta-1)^2}.\frac{1}{\sigma}.\left(\varphi\left(\frac{\ln(n)-\ln(n_{\min})}{\sigma}-\sigma.(\beta-1)\right)-\varphi\left(\frac{\ln(n)-\ln(N)}{\sigma}-\sigma.(\beta-1)\right)\right) \\
&-\frac{C}{\beta-1}.n^{-\beta}.T.e^{\frac{\sigma^2}{2}.(\beta-1)^2}.\left[(1-\beta).\Phi\left(\frac{\ln(n)-\ln(n_r)}{\sigma}-\sigma.(\beta-1)\right)+\frac{1}{\sigma}.\varphi\left(\frac{\ln(n)-\ln(n_r)}{\sigma}-\sigma.(\beta-1)\right)\right] \qquad (38)
\end{aligned}$$

Using the fact that $e^{\frac{\sigma^2}{2}.(\beta-1)^2}.n^{1-\beta}.\varphi\left(\frac{\ln(n)-\ln(X)}{\sigma}-\sigma.(\beta-1)\right)=X^{1-\beta}.\varphi\left(\frac{\ln(n)-\ln(X)}{\sigma}\right)$ and simplifying, we get:

$$\begin{aligned}
\chi(n) =\ & \frac{1}{\sigma.n}.\frac{C}{\beta-1}.\left(N^{1-\beta}-T.n_u^{1-\beta}\right)\varphi\left(\frac{\ln(n)-\ln(N)}{\sigma}\right) \\
&-\frac{C}{\beta-1}.n^{-\beta}.e^{\frac{\sigma^2}{2}.(\beta-1)^2}.(1-\beta).\left(\Phi\left(\frac{\ln(n)-\ln(n_{\min})}{\sigma}-\sigma.(\beta-1)\right)-\Phi\left(\frac{\ln(n)-\ln(N)}{\sigma}-\sigma.(\beta-1)\right)\right) \\
&-\frac{C}{\beta-1}.n^{-\beta}.T.e^{\frac{\sigma^2}{2}.(\beta-1)^2}.(1-\beta).\Phi\left(\frac{\ln(n)-\ln(n_u)}{\sigma}-\sigma.(\beta-1)\right) \qquad (39)
\end{aligned}$$

We use standard ML techniques to estimate the parameters in equation (39). Note that we obtain an estimate of C from this procedure from which we can, in principle recover an estimate of the coefficient θ. 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 θ in this way, suggesting it was not well identified: this is in particular due to the fact that we only estimate the conditional size distribution for firms having 10 (or 5) to 1,000 employees (while we expect θ to be identified from the curvature of the distribution "on the left", for the smallest firms). In the baseline specifications, we use various calibrated values of θ, with 0.8 being our preferred value. Note that the value of the ln-likelihood in column (4) of Table 1 shows that very large values of θ are rejected by the data, because the obtained ln-likelihood drops. See online Appendix section C.3 for alternative strategies for the estimation of θ.

B.4 Estimation of the Fixed Cost Associated with Regulations

The fixed cost is identified and estimated from the indifference equation at α_r:

$$\alpha_r.N^\theta - w.N = \alpha_r.n_r^\theta - w.\tau.n_r - F$$

Since from equation (33) we can express α_r in terms of n_r, we get:

$$\frac{F}{w} = N - \frac{\tau}{\theta}.n_r.\left[\left(\frac{N}{n_r}\right)^\theta - 1 + \theta\right] \tag{40}$$

B.5 Derivation of the Additional Structural Parameters Required for Counterfactual Analysis

B.5.1 Case with Fully Flexible Wages

Given our estimates of β, we have to assume that there exists an upper bound $n_{\max}$ (which is typically set to 10,000, or alternative values in online Appendix Figure A10 to test robustness) in terms of firm size. Therefore, we slightly re-scale the firm size distribution in the following way:

$$\chi(n) = \begin{cases} \frac{\beta-1}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}}.n^{-\beta} & \text{if } n_{\min} \le n < N \\ \frac{N^{1-\beta} - T.n_r^{1-\beta}}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}} & \text{if } n = N \\ 0 & \text{if } N < n < n_r \\ T.\frac{\beta-1}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}}.n^{-\beta} & \text{if } n_r \le n \end{cases} \tag{41}$$

Furthermore, we normalize the ability level $\alpha_{\max}$ corresponding to $n_{\max}$ to 1 without loss of generality, such that $n_{\max} = \left(\frac{\theta}{\tau.w}\right)^{\frac{1}{1-\theta}}$. Applying the variable change formula, we get:

$$\frac{\beta - 1}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}} = \frac{c_\alpha}{p}.(1-\theta).\frac{n_{\max}^{\beta-1}}{T}$$

In this expression, p is the proportion of entrepreneurs in the economy and is still unknown. It can however be estimated using the labor market equilibrium equation (version with fully flexible wages) as:

$$\begin{aligned} p &= \frac{\text{number of firms}}{\text{number of firms} + \text{number of workers in firms}} \\ &= \frac{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}}{\left(n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}\right) + \frac{\beta-1}{\beta-2}.\left(n_{\min}^{2-\beta} - N^{2-\beta}\right) + N.\left(N^{1-\beta} - T.n_r^{1-\beta}\right) + \frac{\beta-1}{\beta-2}.T.\left(n_r^{2-\beta} - n_{\max}^{2-\beta}\right)} \end{aligned}$$

Therefore c_α can be computed as:

$$c_\alpha = \frac{T.n_{\max}^{1-\beta}}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}} \frac{p.(\beta-1)}{1-\theta} \tag{42}$$

and we have: $\beta_\alpha = \frac{\beta-\theta}{1-\theta}$.

B.5.2 Case with Rigid Wages

Let u be the unemployment rate on the labour market. The indifference equation at $\alpha_{\min}$ is altered as:

$$w.(1-u) = Y_{\min} - w.n_{\min} \tag{43}$$

Rearranging terms, we get a new expression for $n_{\min}$:

$$n_{\min} = \frac{\theta}{1-\theta}.(1-u) \tag{44}$$

Second, the labor market equation is:

$$\frac{\beta-1}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}}.\left[\frac{n_{\min}^{2-\beta} - N^{2-\beta}}{\beta-2} + \frac{N}{\beta-1}.\left(N^{1-\beta} - T.n_r^{1-\beta}\right) + T.\frac{n_r^{2-\beta} - n_{\max}^{2-\beta}}{\beta-2}\right]$$

$$= \frac{(1-p).(1-u)}{p} = \frac{1-\hat{p}}{\hat{p}} \tag{45}$$

where $p = \frac{\sharp\text{firms}}{\sharp\text{agents}}$ and $\hat{p} = \frac{\sharp\text{firms}}{\sharp\text{firms}+\sharp\text{workers in firms}} = p.\frac{\sharp\text{agents}}{\sharp\text{agents}-\sharp\text{unemployed workers}}$. This will affect how c_α is estimated.

B.5.3 Estimating the Degree of Wage Rigidity

As discussed in the text in order to calculate the welfare implications of the regulations we need to calibrate the parameter a, which indicates the degree to which real wages are flexible downwards in response to the regulation. The simplest case is complete flexibility ($a = 1$), but this is unrealistic in the French context of high minimum wages and powerful trade unions.

We therefore consider three ways of calibrating a.[4] Our preferred method is to use the average structural unemployment rate between 2002 and 2007 for the US and France as estimated by OECD (2015).[5] These were 8.308% and 5.686% leading to a difference of 2.622 percentage points. We use the US as a benchmark as it is the least regulated labor market in the industrial world so the unemployment rate could be considered as due to inescapable search and matching frictions. We then look for the value of a consistent with the model and its estimated parameters. This turns out to be an estimate of 0.70, i.e. the French real wage adjusts downwards by 70% of what it would do in the flexible wage economy (the US).

As a second alternative we exploit the institutional features of the French minimum wage. Unlike the US this is explicitly updated and indexed to past changes in the consumer price index. Aeberhardt et al (2012) estimate empirically the effect of the French minimum wage on the entire wage distribution. They look at the annual uprating of the minimum wage and what impact it has on different quantiles of the wage distribution. Using unconditional quantile regressions they estimate that the minimum wage is passed through 100% for those at or below the 10th percentile of the wage distribution; had a 40% pass through rate at for those between the 10th to 70th percentiles and had no affect for those in the top three deciles. This implies a value of $a = 0.67$, very similar to the previous method, which is reassuring. Hence we take $a = 0.70$ as our baseline case.

Finally, we consider using the degree of wage inflexibility assumed by the official macro-economic model of the French Finance Ministry (e.g. de Loubens and Thornary, 2010). This model estimates a partial adjustment of real wages of 62% after two years in response to an exogenous price shock (e.g. the oil price) which corresponds to an even greater inflexibility $a = 0.62$[6]

In Table 3 in the main Text, we show all four estimates, focusing on the baseline $a = 0.70$ case.

[4]In Garicano et al (2013) we also consider the opposite case of complete real wage rigidity which generates extremely large welfare losses).

[5]This are closest years to our sample period we could get from the OECD with a consistent time series.

[6]Note however that this (very standard) macro-model predicts full adjustment in the *very* long run (i.e. $a = 1$), so we have a preference for the other two methods. Nevertheless it serves roughly as an "upper bound" of the welfare loss that is less extreme than assuming complete wage inflexibility.

B.6 Equilibrium Without Regulation

In the unregulated economy, the profit maximization program is simply:

$$\pi(\alpha) = \max_n \alpha.n^\theta - w.n \tag{46}$$

$$n^*(\alpha) = \left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}}.\alpha^{\frac{1}{1-\theta}} \tag{47}$$

The indifference condition at $\alpha_{\min}$ is:

$$\alpha_{\min}.\left(n_{\min}\right)^\theta - w.n_{\min} = w$$

Using the relation between ability and size, we obtain as before: $n_{\min} = \frac{\theta}{1-\theta}$.
The density of the firm size distribution is given by:

$$\chi(n) = \frac{\beta-1}{\left(n_{\min}\right)^{1-\beta} - \left(n_{\max}\right)^{1-\beta}}.n^{-\beta} \tag{48}$$

$$\frac{\beta-1}{\left(n_{\min}\right)^{1-\beta} - \left(n_{\max}\right)^{1-\beta}} = \frac{c_\alpha}{p}.(1-\theta).\left(\frac{\theta}{w}\right)^{\frac{\beta-1}{1-\theta}} \tag{49}$$

where maximum firm size is measured at $\alpha_{\max} = 1$ as:

$$n_{\max} = \left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}}$$

This implies:

$$p.\frac{\beta-1}{\left(n_{\min}\right)^{1-\beta} - \left(n_{\max}\right)^{1-\beta}} = c_\alpha.(1-\theta).\left(n_{\max}\right)^{\beta-1}$$

The labor market equation writes simply:

$$p.\frac{\beta-1}{\left(n_{\min}\right)^{1-\beta} - \left(n_{\max}\right)^{1-\beta}}.\int_{n_{\min}}^{n_{\max}} \left[n^{1-\beta} + n^{-\beta}\right] \mathrm{d}n = 1 \tag{50}$$

This implies that $n_{\max}$ can be expressed implicitly as a function of $n_{\min}$:

$$\frac{\left(n_{\min}\right)^{1-\beta} - \left(n_{\max}\right)^{1-\beta}}{\beta-1} + \frac{\left(n_{\min}\right)^{2-\beta} - \left(n_{\max}\right)^{2-\beta}}{\beta-2} = \frac{\left(n_{\max}\right)^{1-\beta}}{c_\alpha.(1-\theta)} \tag{51}$$

All these equations enable to solve the model in both the "regulated" and "unregulated" cases, with flexible, or rigid (indexed by a) wages. In the latter case however, we rely on the additional relation introduced in definition 2 (relating wages in the rigid economy to wages in the fully flexible economies, both regulated and unregulated), which enables to retrieve the additional parameter (the unemployment rate, u).

C Extensions and Robustness Checks

C.1 Positive Spillovers Created by Regulation?

One of the empirical puzzles we found was that the fixed cost of the regulation was estimated to be negative. Such a finding can be rationalized by the existence of 'informational spillovers' within the firm from the regulation. The idea is that the reporting requirements of the regulation force a firm to keep much better track of workers' remuneration and conditions of employment. Collecting systematic information in this way may help firms reduce fixed costs by being able to better spot where there are problems - for example, with retention and recruitment, as well as potentially identify inefficiencies (e.g. through matching records on hours worked with performance of work units). Firms may develop these abilities in-house or they could purchase specialist software to help them do this, such as Enterprise Resource Planning.[7] These more systematic methods of monitoring the firm's employees and using this information to improve productivity are skin to the structured managerial practices of the type described by Bloom and Van Reenen (2007). For smaller firms, the overall costs of adopting these changes would still outweigh the benefits absent regulation - otherwise firms would already have adopted them. But being forced to collect and report information by the regulation could spill-over to other parts of a firm's operations potentially generating some small net fixed benefits.

To see this formally, assume as above that complying with the regulation involves a fixed cost $F \geq 0$ and a variable cost $\tau \geq 0$. Suppose that, by providing information to management, complying with information produces a positive internal spillover to the firm, describe by a concave function in size $G(n), G'(n) > 0$, $G(0) = 0$, $G''(n) < 0$. Assume that $F + w(1+\tau)n \geq G(n)$ for all n, so that only firms forced to comply enjoy the spillover. This ensures firms do not impose on themselves the regulatory costs voluntarily.

With these new functional forms a firm's profit becomes:

$$\pi(\alpha) = \max_n \alpha n^\theta - w\overline{\tau}n - F + 1_{\{n>N\}}G(n), \qquad \text{with} \begin{cases} \overline{\tau} = 1, F = 0 \text{ if } n \leq N \\ \overline{\tau} = \tau \geq 1; F \geq 0 \text{ if } n > N \end{cases} \tag{52}$$

To see that this type of spillovers can generate a negative intercept, approximate the concave function $G(n)$ with the function:

$$G(n) = \begin{cases} an \text{ for } n \leq N' \\ aN' \text{ for } n > N' \end{cases}$$

with $N' < N$. This functional form yields the result exactly: $\pi(\alpha) = \max_n \alpha n^\theta - w\overline{\tau}n - F + 1_{\{n>N\}}aN'$: for a sufficiently large, it will yield a negative estimate for the intercept of the fixed cost $F - aN'$ (where $1_{\{n>N\}}$ is an indicator variable that the firm is over the regulatory threshold, so $n > N$).

In general, this will be approximately true for a $G(n)$ concave and large enough.

This idea is related to an older literature which has discussed how regulations and other external cost shocks can have unexpected benefits in terms of fixed efficiency savings (e.g. the "induced innovations" of Hicks, 1932, or bad shocks that lead managers to reduce X-inefficiencies as discussed by Leibenstein, 1966). If firms are not optimizing these regulatory shocks could actually create positive welfare effects. However, in the context of our optimizing model, the net welfare impact of the regulations are still negative despite some of the offsetting fixed cost savings via such information spillovers.

C.2 Sensitivity of Estimates to the Gaussian Specification for Measurement Errors

How sensitive are our estimates to the Gaussian specification of the measurement error component? To investigate this issue, let us assume that the error term is distributed according to an unspecified pdf f, with associated cdf F. We only impose that f is sufficiently thin tailed, such that the expectation of e^ε is finite:

$$\int_{\mathbb{R}} e^{\varepsilon(\beta-1)} f(\varepsilon) \mathrm{d}\varepsilon = M < \infty \tag{53}$$

and that

$$1 - F(\ln(n)) = o\left(n^{1-\beta}\right) \text{ when } n \to +\infty \tag{54}$$

[7]Bloom, Garicano, Sadun and Van Reenen, (2014) discuss more details of the economics and organizational consequences of ERP introduction. Companies such as SAP and Oracle license these software systems to help companies keep better track of their human resources, procurement, customers, etc.

The unconditional cdf of n can be rewritten as[8]:

$$\begin{aligned}\forall n > 0, \quad \mathbb{P}(x < n) \quad &= \quad F\left(\ln(n) - \ln\frac{\theta}{1-\theta}\right) - \frac{C}{\beta-1} T n_r^{1-\beta}\left[F(\ln(n) - \ln(N)) - F(\ln(n) - \ln(n_r))\right] \\ &\frac{C}{\beta-1} n^{1-\beta}\left[\int_{\ln(n)-\ln(N)}^{\ln(n)-\ln\frac{\theta}{1-\theta}} e^{\varepsilon(\beta-1)} f(\varepsilon)\mathrm{d}\varepsilon + T\int_{-\infty}^{\ln(n)-\ln(n_r)} e^{\varepsilon(\beta-1)} f(\varepsilon)\mathrm{d}\varepsilon\right]\end{aligned}$$

Under the assumptions above, we can show that:

$$1 - \mathbb{P}(x < n) \underset{n\to+\infty}{\sim} \frac{C}{\beta-1} T M n^{1-\beta}$$

Therefore, for large values of n, the behavior of the optimization criterion is only affected by the distribution of ε *via* the term M. Conditional maximum likelihood estimation of the upper part of the distribution (unaffected by the "bulge" and the "valley") provides robust estimators β thanks to the slope. We can then obtain a robust estimate of the composite term $T.M$ comparing the mass of firms in the lower and upper part of the distribution, respectively. Assuming that f is centered on 0 and symmetric, one can show[9] that $M \geq 1$. In the Gaussian case, $\int_{\mathbb{R}} e^{\varepsilon(\beta-1)}\varphi(\varepsilon)\mathrm{d}\varepsilon = e^{\frac{\sigma^2}{2}(\beta-1)^2}$ which we estimate to be 1.0047 in our case. This means that the Gaussian assumption only induces a very small re-scaling of the T parameter via the M term. It is therefore very conservative with respect to τ, since the smallest value for τ would be obtained with no re-scaling at all. This lower bound for τ is estimated to be $\left(0.912 \times e^{\frac{\sigma^2}{2}(\beta-1)^2}\right)^{-\frac{1-\theta}{\beta-1}} \approx 1.022$, very close to the result presented in the main part of the text (1.023). Without any re-scaling at all, we would obtain an upper bound for T ($\widehat{T} = 0.951 > 0.948$) and a lower bound on τ ($\widehat{\tau} = 1.012 < 1.013$) which are in fact very close to the results presented in the main part of the text.

C.3 Using Information from the Productivity Distribution

As discussed in the main text, we consider three alternative routes of estimating θ. Our baseline is simply calibration, but we document here the two alternative strategies reported in Table 1: (i) estimates from the production function and (ii) using the TFP-size relationship. When we use these econometric methods to estimate θ, we take into account the variance around the estimation of θ in calculating the correct variance-covariance matrix by block-bootstrapping.

C.3.1 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, 2011).[10]

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 1995-2007 (using the unbalanced panel) by each of the four-digit manufacturing industries in our dataset. We also did the same for the retail sector

[8] The derivations below can alternatively be established in terms of the pdf.

[9] This is because in the case distribution of a distribution that is uniformly distributed over $[-x; x]$ (where $x > 0$), we have:

$$\frac{1}{2.x}\int_{-x}^{x} e^{\varepsilon(\beta-1)}\mathrm{d}\varepsilon = \frac{e^{x(\beta-1)} - e^{-x(\beta-1)}}{2x(\beta-1)} > 1$$

Therefore, if f is centered on 0 and symmetric, then it can be approximated by below by a bound that is arbitrarily close to 1.

[10] Results are all available on request.

and the business services sector. The production functions take the form (in each industry):

$$\ln y_{it} = \beta_n \ln n_{it} + \beta_k \ln k_{it} + \beta_m \ln m_{it} + \omega_{it} + \tau_t + \eta_{it} \tag{55}$$

where y = output, n = labour, k= capital, m = materials ω is the unobserved productivity shock, τ_t is a set of time dummies and η 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.

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 (55) by using the observed factor shares in revenues. We do this assuming constant returns to scale, so $\beta_n = \frac{wn}{py}$; $\beta_m = \frac{p^m m}{py}$and $\beta_k = 1 - \frac{wn}{py} - \frac{p^m m}{py}$ where p^m is 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 managerial ability, α. 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.[11] So it is unclear whether this would make too much of a practical difference to our results.

C.3.2 Using the Relationship Between Firm Size and TFP to Estimate θ

First, recall from sub-section 2.2 the relationship between firm size and TFP:

$$n^*(\alpha) = \begin{cases} 0 & \text{if } \alpha < \alpha_{\min} \\ \left(\frac{\theta}{w}\right)^{1/(1-\theta)} \alpha^{1/(1-\theta)} & \text{if } \alpha_{\min} \leq \alpha \leq \alpha_c \\ N & \text{if } \alpha_c \leq \alpha < \alpha_r \\ \left(\frac{\theta}{w}\right)^{1/(1-\theta)} \tau^{-1/(1-\theta)} \alpha^{1/(1-\theta)} & \text{if } \alpha_r \leq \alpha < \infty \end{cases}$$

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

$$n^*(\alpha) = \begin{cases} 0 & \text{if } \alpha < \alpha_{\min} \\ \left(\frac{\theta}{w}\right)^{1/(1-\theta)} \alpha^{1/(1-\theta)} e^{\varepsilon} & \text{if } \alpha_{\min} \leq \alpha \leq \alpha_c \\ N e^{\varepsilon} & \text{if } \alpha_c \leq \alpha < \alpha_r \\ \left(\frac{\theta}{w}\right)^{1/(1-\theta)} \tau^{-1/(1-\theta)} \alpha^{1/(1-\theta)} e^{\varepsilon} & \text{if } \alpha_r \leq \alpha < \infty \end{cases}$$

Or

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

[11] An alternative approach would be to follow de Loecker (2011) 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.

Combining these together:

$$\ln n = \ln n_1 I_{\{\alpha_{\min} \leq \alpha \leq \alpha_c\}} + \ln n_2 I_{\{\alpha_c \leq \alpha < \alpha_r\}} + \ln n_3 I_{\{\alpha_r \leq \alpha\}} \tag{56}$$

where I is an indicator function for a particular "regime". We use the previous firm-level estimates of TFP (α) to estimate equation (56) and provide an alternative estimates of parameter θ. Results are reported in Table 1, column (5).

C.4 Using Alternative Datasets and Concepts of Employment

As discussed in section III.B in the main Text, we also estimate our baseline model using alternative datasets and concepts of employment. In labor laws, the concept of employment in most regulations is a full-time equivalent concept (LAMY, 2010), with many special cases related to different types of workers (sick leaves, apprenticeship, etc.) and based on type of contract (full time *vs.* part time). This central information is unfortunately not available precisely in any dataset.

- The employment measure in the FICUS relates to a headcount of the number of workers in the firm averaged over the four quarters of the fiscal year in France. A headcount of employees is taken on the last day of the fiscal year (usually December 31st) and on the last day at the end of each of the previous three quarters. Employment is then the simple arithmetic average over these four days: this "smoothing" of headcounts renders this concept closer to full time equivalent jobs.

- We also experiment with the measure of full time equivalent that is available in the DADS files (Panel B of Table A9). The main problem here is that full time equivalents are computed from "non-annexes" jobs only, i.e. those lasting more than 30 days with more than 1.5 hours worked per day, or those that were associated with (total) gross earnings higher than three times the monthly minimum wage. As reported in the methodological documentation of the DADS, in 2002, the "non-annexes" jobs represented 76.6% of all jobs, which means that almost a quarter of the most short-end, low paid jobs were removed. This is unfortunate as these jobs are in fact covered by labor laws. Note also that the type of contract (full time vs. part time) is not observed in the DADS data: it is estimated from the 75^{th} percentile of the distribution of hours paid/worked per job (in firms having 20 to 1,000 workers). Unfortunately, this strategy only provides approximations of full time equivalents because the implementation of the mandatory 35 hour work week in 2000 was complex and differentiated across firms, which renders the relation between hours and full time equivalents heterogeneous across firms.

- We also estimate our model with a simple headcount on December 31th, that is available in the DADS (Panel C of online Appendix Table A9).

The results of using alternative datasets and definitions of employment are presented in online Appendix Table A5 and online Appendix Figure A9. The alternative datasets differ mainly in terms of the mass point and discontinuity at the threshold at 49. The the spike as well as the discontinuity are more pronounced in FICUS than in DADS. In the DADS files, they are more pronounced in terms of headcounts rather than full time equivalents (which seems unsurprising from the way they are computed). Nevertheless a bulge and shift of the power law in clearly present in all panels in online Appendix Figure A9.

Online Appendix Table A5 shows that these differences only affect the estimated fixed cost component of the regulation, while the variable cost part is remarkably stable across all datasets and concepts. This is not surprising, because the main component of the estimated tax generates a shift of the distribution rather than just mass points or discontinuities. Finally, online Appendix Figure A9 shows that the fit of our model is broadly similar across all three concepts of employment.

C.5 Robustness of the Results to the Definition of the Size of the Largest Firm in the Economy

Although the empirical maximum firm size is 86,587 in the data, we choose for the welfare calculation to focus on costs to firms up to 10,000 since there are on average only 5 firms per year having a size greater than 10,000 (out of an average of 170,000 firms with positive employment in manufacturing industries). In online Appendix Figure A10 we show that our quantitative estimates of welfare and employment are not much changed when we vary our assumption about the upper bound of firm size (using alternative values of 1,000, 5,000 and 10,000 employees) with the exception of job losses at large firms, which are halved when we only consider firms up to 1,000 employees.

D Allowing for Capital-Labor Substitution: Underlying CES Model in Section V.D

This Appendix presents the details of the CES model underlying section V.D and its estimation. Figures 13 and 14 in the main Text correspond to the estimation results for different calibrated values of the elasticity of substitution, η.

D.1 CES Specification of Firms' Objective Function

With a CES production function, the objective function of firms is:

$$\pi(\alpha) = \max_{n,k} \begin{cases} \alpha f(n,k) - wn - rk & \text{if } n \leq N \\ \alpha f(n,k) - w\tau n - rk - F & \text{if } n > N \end{cases} \tag{57}$$

with $f(n,k) = (\lambda.n^{\rho} + (1-\lambda).k^{\rho})^{\frac{\theta}{\rho}}$ and where w is workers' wage, n is the number of workers, r is capital (marginal) cost, and k is the number of units of capital. Moreover, $\theta, \lambda \in (0,1)$ and $\rho \in \mathbb{R}$.
Optimal input demand in each regime (but not at the threshold) is then determined by the first order conditions:

$$[n] : \alpha\theta\left(\lambda.n^{\rho} + (1-\lambda).k^{\rho}\right)^{\frac{\theta}{\rho}-1} \lambda n^{\rho-1} - w\overline{\tau} = 0 \;\text{ with } \begin{cases} \overline{\tau} = 1 & \text{if } n < N \\ \overline{\tau} = \tau & \text{if } n > N \end{cases} \tag{58}$$

$$[k] : \alpha\theta\left(\lambda.n^{\rho} + (1-\lambda).k^{\rho}\right)^{\frac{\theta}{\rho}-1} (1-\lambda).k^{\rho-1} - r = 0 \tag{59}$$

Solving for this system of equations, we get the optimal demand for inputs in each regime (again, these hold everywhere, except at the threshold, N):

$$k^* = \alpha^{\frac{1}{1-\theta}}.\left[\frac{\theta(1-\lambda)}{r}\right]^{\frac{1}{1-\theta}}.\left(\lambda.\left(\frac{\gamma}{\overline{\tau}}\right)^{\eta-1} + (1-\lambda)\right)^{\frac{\theta-\rho}{\rho.(1-\theta)}} \tag{60}$$

$$n^* = \alpha^{\frac{1}{1-\theta}}.\left[\frac{\theta\lambda}{w.\overline{\tau}}\right]^{\frac{1}{1-\theta}}.\left(\lambda + (1-\lambda).\left(\frac{\overline{\tau}}{\gamma}\right)^{\eta-1}\right)^{\frac{\theta-\rho}{\rho.(1-\theta)}} \tag{61}$$

where $\gamma = \frac{r}{w}.\frac{\lambda}{1-\lambda}$ and $\eta = \frac{1}{1-\rho} \in \mathbb{R}^{+}$ is the elasticity of substitution.

D.2 Normalization and Calibrated Values

First, the labor-capital ratio is obtained by taking the ratio of equations (60) and (61):

$$\frac{n^*}{k^*}=\left[\frac{r.\lambda}{w\overline{\tau}.(1-\lambda)}\right]^{\eta}=\left(\frac{\gamma}{\overline{\tau}}\right)^{\eta} \tag{62}$$

This ratio is normalized to 1 for small firms (having $\overline{\tau}=1$) without loss of generality: this simply defines the measurement unit of capital (to the quantity of capital allocated per worker in small, untaxed firms). This calibration is attractive since capital is only reported at historical costs in firms' balance sheets, such that capital intensity is typically difficult to measure empirically. Note that this implies that γ is always 1 for $\eta>0$[12], such that λ will simply be estimated from the ratio of prices: $\lambda=\frac{w}{w+r}$.

Second, we calibrate the labor share in small firms to 0.6 (NIPA). This restriction together with the previous choice of measurement unit for capital fully determines λ and implies that wages are three times as high as capital price:

$$\frac{w.n^*}{Y}=\theta.\frac{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}}{\frac{\lambda}{1-\lambda}\gamma^{\eta-1}+1}=\theta.\frac{w}{w+r}=\theta.\lambda \tag{63}$$

where the second equality holds for $\eta>0$ and follows from the definition of γ and from equation 62. θ is calibrated to 0.8, such that $\lambda=0.75$ and $w=3.r$.
Notice that this also determines the value of $n_{\min}$ in case of fully flexible prices (wages). Indeed, the indifference equation at $\alpha_{\min}$ writes:

$$\begin{aligned} w &= Y_{\min}-w.n_{\min}-r.k_{\min} \\ &= Y_{\min}.\left(1-\frac{\gamma^{\eta-1}.\theta.\lambda}{\lambda\gamma^{\rho\eta}+(1-\lambda)}-\frac{\theta.(1-\lambda)}{\lambda\gamma^{\rho\eta}+(1-\lambda)}\right) \end{aligned}$$

Rearranging terms, we get:

$$n_{\min}=\frac{\theta}{1-\theta}.\frac{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}}{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}+1}=\frac{1}{1-\theta}.\frac{w.n^*}{Y}. \tag{64}$$

D.3 Estimated Density for the Firm Size Distribution

The density of the firm size distribution is given by:

$$\chi(n)=\begin{cases}\frac{\beta-1}{n_{\min}^{1-\beta}}.n^{-\beta} & \text{if } n_{\min}\leq n<N \\ \frac{N^{1-\beta}-T.n_r^{1-\beta}}{n_{\min}^{1-\beta}} & \text{if } n=N \\ 0 & \text{if } N<n<n_r \\ T.\frac{\beta-1}{n_{\min}^{1-\beta}}.n^{-\beta} & \text{if } n_r\leq n\end{cases} \tag{65}$$

Note that this estimation problem is strictly similar to the baseline case, except that $n_{\min}$ is calibrated to a different value, and that the quantity T has a different interpretation:

[12] This normalization makes clear that the CES specification formally reduces to the baseline single factor model (w',τ') for small and large firms, with:

$$\begin{aligned} w' &= w+r \\ \tau' &= \frac{\tau.w+\tau^{\eta}.r}{w+r} \end{aligned}$$

At the threshold N however, for ability ranging from α_c to α_r, the two problems are different, since except in the Leontief case (see online Appendix section D.7 for more details in this specific case), output is not linear in α.

$$T^{\frac{1}{\beta-1}} = \tau^{-\frac{1}{1-\theta}}. \left[\frac{\lambda + (1-\lambda).\tau^{\eta-1}.\gamma^{-(\eta-1)}}{\lambda + (1-\lambda).\gamma^{-(\eta-1)}}\right]^{\frac{\theta-\rho}{\rho.(1-\theta)}} \tag{66}$$

Equation (66) is solved numerically to retrieve the estimate of τ.
In practice however, and as in the baseline model, we assume that there exists an upper bound $n_{\max}$ (which is typically set to 10,000) in terms of firm size. Therefore, we slightly re-scale the firm size distribution in the following way:

$$\chi(n) = \begin{cases} \frac{\beta-1}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}}.n^{-\beta} & \text{if } n_{\min} \leq n < N \\ \frac{N^{1-\beta} - T.n_r^{1-\beta}}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}} & \text{if } n = N \\ 0 & \text{if } N < n < n_r \\ T.\frac{\beta-1}{n_{\min}^{1-\beta} - T.n_{\max}^{1-\beta}}.n^{-\beta} & \text{if } n_r \leq n \end{cases} \tag{67}$$

D.4 Estimation of the Fixed Cost Associated with Regulations

In this section, we show how to compute the fixed cost ($\frac{F}{w}$) associated with the regulation as a function of the estimated parameters. As previously, the computation is mainly based on the indifference relation at α_r:

$$\alpha_r.f(N, K) - w.N - r.K = \alpha_r.f(n_r, k_r) - w.\tau.n_r - r.k_r - F \tag{68}$$

where K is optimal capital demand for a firm of productivity α_r choosing to remain at an employment level of N, while k_r is optimal capital demand for a firm of productivity α_r choosing to grow and pay taxes (remember profits are the same in both cases).
k_r is easily computed as $n_r.\frac{\tau^\eta}{\gamma^\eta}$. Since γ^η is calibrated to unity, this is simply:

$$k_r = n_r.\tau^\eta \tag{69}$$

Notice also that α_r can be expressed as a function of n_r:

$$n_r = \alpha_r^{\frac{1}{1-\theta}}. \left[\frac{\theta\lambda}{w.\tau}\right]^{\frac{1}{1-\theta}}. \left(\lambda + (1-\lambda).\left(\frac{\tau}{\gamma}\right)^{\eta-1}\right)^{\frac{\theta-\rho}{\rho.(1-\theta)}}$$

$$\iff \alpha_r = n_r^{1-\theta}.\frac{w.\tau}{\theta.\lambda}. \left(\lambda + (1-\lambda).\left(\frac{\tau}{\gamma}\right)^{\eta-1}\right)^{\frac{\rho-\theta}{\rho}} \tag{70}$$

K can be computed from the first order condition of the optimization program for "constrained" firms:

$$\max_K \left\{\alpha_r. \left(\lambda.N^\rho + (1-\lambda).K^\rho\right)^{\frac{\theta}{\rho}} - w.N - r.K\right\} \tag{71}$$

The first order condition is:

$$\alpha_r.\theta.(1-\lambda).K^{\rho-1} \left(\lambda.N^\rho + (1-\lambda).K^\rho\right)^{\frac{\theta}{\rho}-1} = r \tag{72}$$

Using the previous expression for α_r and remembering that $r = \frac{1-\lambda}{\lambda}.\gamma.w$, we get:

$$n_r^{1-\theta}.K^{\rho-1}. \left(\frac{1 + \frac{1-\lambda}{\lambda}.\tau^{\eta-1}.\left(\frac{1}{\gamma}\right)^{\eta-1}}{N^\rho + \frac{1-\lambda}{\lambda}.K^\rho}\right)^{\frac{\rho-\theta}{\rho}} = \frac{\gamma}{\tau} \tag{73}$$

Equation (73) can be solved numerically to get an estimate of K.
Lastly, fixed costs can be computed as:

$$\begin{aligned}\frac{F}{w} = N \quad &+ \quad \frac{1-\lambda}{\lambda}.\gamma.K \\ &- \quad n_r.\tau.\left(1+\tau^{\eta-1}.\frac{1-\lambda}{\lambda}.\left(\frac{1}{\gamma}\right)^{\eta-1}\right).\left[\frac{\theta-1}{\theta}+\frac{1}{\theta}.\frac{\left(N^\rho+\frac{1-\lambda}{\lambda}.K^\rho\right)^{\frac{\theta}{\rho}}}{n_r^\theta.\left(1+\tau^{\eta-1}.\frac{1-\lambda}{\lambda}.\left(\frac{1}{\gamma}\right)^{\eta-1}\right)^{\frac{\theta}{\rho}}}\right]\end{aligned} \tag{74}$$

D.5 Estimation of the Additional Structural Parameters Required for Counterfactual Analysis

D.5.1 Case with Fully Flexible Prices

To compute alternative equilibria (under the fully flexible price assumption), we are still missing the scaling parameter of the ability distribution, c_α, and the amount of capital per agent in the economy, k_{TOT}. Given our calibration choices[13], this quantity will entirely capture the new margin of adjustment when we will allow for different values of η.

As in the baseline model, we normalize $\alpha_{\max}$ to 1 without loss of generality, such that, using the relation between ability and size and applying the variable change formula, we get [14]:

$$\frac{\beta-1}{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}} = \frac{c_\alpha}{p}.(1-\theta).\frac{n_{\max}^{\beta-1}}{T}$$

In this expression, p is the proportion of entrepreneurs in the economy and is still unknown. It can however be estimated using the labor market equilibrium equation (version with fully flexible wages) as:

$$\begin{aligned} p \quad &= \quad \frac{\text{number of firms}}{\text{number of firms} + \text{number of workers in firms}} \\ &= \quad \frac{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}}{\left(n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}\right)+\frac{\beta-1}{\beta-2}.\left(n_{\min}^{2-\beta}-N^{2-\beta}\right)+N.\left(N^{1-\beta}-T.n_r^{1-\beta}\right)+\frac{\beta-1}{\beta-2}.T.\left(n_r^{2-\beta}-n_{\max}^{2-\beta}\right)}\end{aligned}$$

Therefore c_α can be computed as:

$$c_\alpha = \frac{T.n_{\max}^{1-\beta}}{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}}\frac{p.(\beta-1)}{1-\theta} \tag{75}$$

and we have: $\beta_\alpha = \frac{\beta-\theta}{1-\theta}$, as in the baseline specification.

Last, we can compute the total stock of capital in the economy (normalized by the total number of agents),k_{TOT}, via the stock of capital utilized by firms in each size class.
The two easy cases are the following:

[13] Especially the normalization $n/k = 1$ for small firms. An alternative but strictly equivalent approach would have been to calibrate the total amount of capital and allow the parameters γ, λ (or r/w) to adjust. This would however make the comparison with the Leontief case less straightforward (see section D.7).

[14] This also determines prices since $r = w/3$ and:

$$w = \theta.\frac{T^{\frac{1-\theta}{\beta-1}}}{n_{\max}^{1-\theta}}.\lambda^{\frac{\theta}{\rho}}.\left(1+\frac{1-\lambda}{\lambda}.\left(\frac{1}{\gamma}\right)^{\eta-1}\right)^{\frac{\theta-\rho}{\rho}}$$

- Small firms in $[n_{\min}; N[$

$$p.\frac{\beta-1}{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}}.\int_{n_{\min}}^{N}\left(\underbrace{\gamma^{-\eta}}_{=1}\right)n^{1-\beta}\mathrm{d}n = p.\frac{\beta-1}{\beta-2}.\frac{n_{\min}^{2-\beta}-N^{2-\beta}}{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}} \tag{76}$$

- Large firms in $[n_r; n_{\max}]$

$$p.\tau^{\eta}.T.\frac{\beta-1}{\beta-2}.\frac{n_r^{2-\beta}-n_{\max}^{2-\beta}}{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}}$$

The difficulty lies in the computation of the stock of capital $k(\alpha)$ used by firms at N. In this range, there is no closed form for $k(\alpha)$, but the FOC of the profit maximization program delivers a closed form expression for the inverse of this function, $\alpha(k)$:

$$\alpha(k) = \frac{r}{1-\lambda}.\frac{1}{\theta}.k^{1-\rho}.\left[\lambda.N^{\rho}+(1-\lambda).k^{\rho}\right]^{1-\frac{\theta}{\rho}}$$

We use a variable change to get an expression of the integral to compute:

$$\int_{\alpha_c}^{\alpha_r} k(\alpha).c_{\alpha}.\alpha^{-\beta_\alpha}\mathrm{d}\alpha = c_{\alpha}.\left(\frac{r}{1-\lambda}\frac{1}{\theta}\right)^{1-\beta_\alpha}\int_{N}^{K} k \quad\times\quad \left[\lambda.N^{\rho}+(1-\lambda).k^{\rho}\right]^{(\beta_\alpha-1).\frac{\theta}{\rho}-\beta_\alpha} \left[(1-\rho).\lambda.\left(\frac{N}{k}\right)^{\rho}+(1-\lambda).(1-\theta)\right]\mathrm{d}k \tag{77}$$

This integral is then evaluated numerically.

D.5.2 Case with "Rigid" Wages but Where the Price of Capital is Fully Flexible

Let u be the unemployment rate on the labour market. The indifference equation at $\alpha_{\min}$ is altered as:

$$w.(1-u) = Y_{\min} - w.n_{\min} - r.k_{\min}$$

Rearranging terms, we get a new expression for $n_{\min}$, replacing former equation (64):

$$n_{\min} = \frac{\theta}{1-\theta}.(1-u).\frac{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}}{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}+1} = \frac{1}{1-\theta}.(1-u).\frac{w.n^*}{Y}$$

Second, the labor market equation is:

$$\frac{\beta-1}{n_{\min}^{1-\beta}-T.n_{\max}^{1-\beta}}.\left[\frac{n_{\min}^{2-\beta}-N^{2-\beta}}{\beta-2}+\frac{N}{\beta-1}.\left(N^{1-\beta}-T.n_r^{1-\beta}\right)+T.\frac{n_r^{2-\beta}-n_{\max}^{2-\beta}}{\beta-2}\right]$$
$$=\frac{(1-p).(1-u)}{p}=\frac{1-\hat{p}}{\hat{p}} \tag{78}$$

where $p=\frac{\sharp\text{firms}}{\sharp\text{agents}}$ and $\hat{p}=\frac{\sharp\text{firms}}{\sharp\text{firms}+\sharp\text{workers in firms}}=p.\frac{\sharp\text{agents}}{\sharp\text{agents}-\sharp\text{unemployed workers}}$. This will affect how c_α is estimated. The derivation of k_{TOT} is unaltered: in the previous equation (77), only the estimation of c_α is affected by the new labor market equation.

D.6 Equilibrium without Regulation

In the unregulated economy, the profit maximization program is simply:

$$\pi(\alpha) = \max_{n,k}\alpha.\left(\lambda.n^{\rho}+(1-\lambda).k^{\rho}\right)^{\frac{\theta}{\rho}} - w.n - r.k \tag{79}$$

First order conditions are:

$$
\begin{aligned}
[n] \quad &: \quad \alpha\theta\left[\lambda.n^{\rho}+(1-\lambda).k^{\rho}\right]^{\frac{\theta}{\rho}-1}.\lambda.n^{\rho-1}=w && (80)\\
[k] \quad &: \quad \alpha\theta\left[\lambda.n^{\rho}+(1-\lambda).k^{\rho}\right]^{\frac{\theta}{\rho}-1}.(1-\lambda).k^{\rho-1}=r && (81)
\end{aligned}
$$

We obtain optimal input demand as:

$$
\begin{aligned}
\frac{n}{k} &= \left[\frac{r.\lambda}{w.(1-\lambda)}\right]^{\eta}=\gamma^{\eta} && (82)\\
k &= \alpha^{\frac{1}{1-\theta}}.\left[\frac{\theta(1-\lambda)}{r}\right]^{\frac{1}{1-\theta}}.\left(\lambda.\gamma^{\eta-1}+(1-\lambda)\right)^{\frac{\theta-\rho}{\rho.(1-\theta)}} && (83)\\
n &= \alpha^{\frac{1}{1-\theta}}.\left[\frac{\theta\lambda}{w}\right]^{\frac{1}{1-\theta}}.\left(\lambda+(1-\lambda).\left(\frac{1}{\gamma}\right)^{\eta-1}\right)^{\frac{\theta-\rho}{\rho.(1-\theta)}} && (84)
\end{aligned}
$$

The density of the firm size distribution is given by:

$$
\begin{aligned}
\chi(n) &= \frac{\beta-1}{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}}.n^{-\beta} && (85)\\
\frac{\beta-1}{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}} &= \frac{c_{\alpha}}{p}.(1-\theta).\left(\frac{\theta.\lambda}{w}\right)^{\frac{\beta-1}{1-\theta}}.\lambda^{\frac{(\theta-\rho)(\beta-1)}{\rho(1-\theta)}}.\left[1+\frac{1-\lambda}{\lambda}.\left(\frac{1}{\gamma}\right)^{\eta-1}\right]^{\frac{(\theta-\rho)(\beta-1)}{\rho(1-\theta)}} && (86)
\end{aligned}
$$

The minimum firm size is obtained as previously as:

$$
n_{\min}=\frac{\theta}{1-\theta}.\frac{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}}{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}+1}=\frac{\theta}{1-\theta}.\left[1-\frac{1}{\frac{\lambda}{1-\lambda}.\gamma^{\eta-1}+1}\right] \qquad (87)
$$

Last, maximum firm size is measured at $\alpha_{\max}=1$ as:

$$
n_{\max}=\left(\frac{\theta.\lambda}{w}\right)^{\frac{1}{1-\theta}}.\lambda^{\frac{\theta-\rho}{\rho(1-\theta)}}.\left[1+\frac{1-\lambda}{\lambda}.\left(\frac{1}{\gamma}\right)^{\eta-1}\right]^{\frac{\theta-\rho}{\rho(1-\theta)}} \qquad (88)
$$

which implies:

$$
p.\frac{\beta-1}{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}}=c_{\alpha}.(1-\theta).(n_{\max})^{\beta-1} \qquad (89)
$$

The labor market equation writes simply:

$$
p.\frac{\beta-1}{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}}.\int_{n_{\min}}^{n_{\max}}\left[n^{1-\beta}+n^{-\beta}\right]\mathrm{d}n=1 \qquad (90)
$$

This implies that $n_{\min}$ can be expressed implicitly as a function of $n_{\max}$:

$$
\frac{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}}{\beta-1}+\frac{(n_{\min})^{2-\beta}-(n_{\max})^{2-\beta}}{\beta-2}=\frac{(n_{\max})^{1-\beta}}{c_{\alpha}.(1-\theta)} \qquad (91)
$$

The capital market equation is:

$$
p.\frac{\beta-1}{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}}.\gamma^{-\eta}\int_{n_{\min}}^{n_{\max}}n^{1-\beta}\mathrm{d}n=k_{TOT} \qquad (92)
$$

Using the previous relations, this can be rewritten as:

$$\frac{(n_{\max})^{1-\beta}}{c_\alpha.(1-\theta)}.\left[1-k_{TOT}.\left(\frac{\lambda}{1-\lambda}\right)^{\frac{\eta}{1-\eta}}.\left[\frac{(1-\theta).n_{\min}}{\theta-(1-\theta).n_{\min}}\right]^{\frac{\eta}{\eta-1}}\right]=\frac{(n_{\min})^{1-\beta}-(n_{\max})^{1-\beta}}{\beta-1} \tag{93}$$

All these equations enable to solve the model in both the "regulated" and "unregulated" cases, with flexible or rigid wages. In the latter case however, as in online Appendix sections B.5 and B.6, we also use the additional relation introduced in the definition 2 of the main Text (relating wages in the rigid economy to wages in the fully flexible economies, both regulated and unregulated), which enables to retrieve the additional parameter (the unemployment rate, u).

D.7 Relation with the Baseline, Single Input Model

The baseline, single input model features $\rho \to -\infty$, $\eta \to 0$. However, it does not map directly into the Leontief limiting case of this section. To see this more clearly, let us rewrite the model with two inputs combined into a Leontief production function. As previously, the measurement unit of capital is chosen such that, in small firms (below the threshold), $n^*/k^* = 1$. In the Leontief case however, if this relation holds for small firms, then it also holds for all other firms, since the optimal capital to labor ratio is fully constrained by technology. Therefore, the profit function can be written as:

$$\pi(\alpha)=\max_n\begin{cases}\alpha.n^\theta-\underbrace{(w+r)}_{=\tilde{w}}.n & \text{if } n\le N\\ \alpha.n^\theta-\underbrace{(w+r)}_{=\tilde{w}}.\underbrace{\left(\frac{\tau.w+r}{w+r}\right)}_{=\tilde{\tau}}.n-F & \text{if } n>N\end{cases} \tag{94}$$

In the case of flexible prices, the analogy with the baseline case in the main text is straightforward. There are two important differences however, which prevents to interpret the Leontief case as the limit to the baseline, single factor model:

- The indifference condition defining $\alpha_{\min}$ is altered, since equation 5 in the main Text is altered in the following way:

$$\alpha_{\min}.\left(n^*(\alpha_{\min})\right)^\theta-(w+r).n^*(\alpha_{\min})=w \tag{95}$$

  which implies: $n_{\min}.(1-\theta)=\theta.\frac{w}{w+r}$. In the Leontief case, we still have that in small firms $\frac{w.n^*}{Y}=\theta.\frac{w}{w+r}$, such that the calibration in section D.2 drives both the theoretical minimum firm size, which differs from the baseline case (in main text), and the split of costs between w and r.

- Most importantly, equilibrium conditions on the capital market in both the regulated and unregulated economies imply that the number of firms is the same (and therefore, the total number of workers, which equates the aggregate supply of capital in the two cases). In other words, in the Leontief case with two factors of production, there is no adjustment via new entrants.

These two features explain why the CES extension does not map perfectly with the baseline, single factor model, though the difference in the obtained results is small.

The case where we assume that wages w are rigid but capital prices r are fully flexible also departs from the baseline case with rigid wages presented in the section I.E of the main Text.
As previously, the indifference equation at $\alpha_{\min}$ is altered, as well as $n_{\min}$:

$$\alpha_{\min}.\left(n^*(\alpha_{\min})\right)^\theta-(w+r).n^*(\alpha_{\min})=(1-u).w \tag{96}$$

which implies: $n_{\min}.(1-\theta) = (1-u).\theta.\frac{w}{w+r}$.
The labor market equations in both the regulated and unregulated economies are identical to the case with only one input described in the main text but equilibrium on the capital market however generates an important difference between the Leontief case with two inputs and the single input case. In the regulated economy with rigid wages, equilibrium on the capital market requires:

$$k_{TOT} = (1-u).\int_{\underline{\alpha}}^{\alpha_{\min}} c_\alpha.\alpha^{-\beta_\alpha}\mathrm{d}\alpha \tag{97}$$

In the unregulated economy, equilibrium on the capital market requires that:

$$k_{TOT} = \int_{\underline{\alpha}}^{\alpha_{\min}^0} c_\alpha.\alpha^{-\beta_\alpha}\mathrm{d}\alpha \tag{98}$$

This implies that $\alpha_{\min}^0 < \alpha_{\min}$, i.e. there is a *deficit* of entry in the equilibrium with regulations. Using the indifference conditions at $\alpha_{\min}$ and $\alpha_{\min}^0$ respectively, one can show that this deficit is driven by the rise of the price of capital, $r > r_0$.
This difference with the baseline model is however not a robust feature of the Leontief framework with two inputs: for example, allowing the aggregate supply of capital to be increasing in capital price would reduce entry in the undistorted equilibrium. Alternatively, moving away from the Leontief case and allowing η to increase rises r_0, which reduces entry in the undistorted economy (and raises unemployment very quickly in the regulated economy).

E Details of the Simulated Dynamic Model in Section V.F

Our baseline approach, presented in the sections I to IV of the main Text, is static. Although such a parsimonious approach can be interpreted as the outcome of long term adjustments, it might miss important empirical patterns that arise in a dynamic environment. In section V.F, we therefore consider a generalization of our model to allow dynamics which is more realistic but obviously more complex. All technical details are enclosed in this appendix.[15]

We make several important extensions to the basic static framework. First, we allow managerial ability (TFP) to change over time. Managerial ability is idiosyncratic when firms startup due to heterogeneous talent, but after a firm is founded it is subject each period to a shock which will usually cause it to want to change size. Second, we allow for adjustment costs in labor (and capital).[16] Thus, a firm will not always adjust immediately to its long-run frictionless optimal size.

Given this additional complexity, we can only obtain limited analytical results so we rely on numerical methods to simulate the steady state distributions of firm size and growth (e.g. Judd, 1998). In this appendix, we focus on the partial equilibrium effects of introducing adjustment costs (on the behavior of firms). We do not attempt to include general equilibrium effects as we have done in the static model as doing so would introduce substantial additional complexity to the problem without, we feel, delivering any additional analytical insights over and above the static model. Consequently, factor prices are treated as exogenous, common to all firms and time invariant.

E.1 Production, Costs and Regulation

Firm output y is determined by productivity α, capital k, labor n and materials m with common factor prices r, w and p_m respectively. Firm-specific values are indicated by an i sub-script: $y_{it} = \alpha_{it}k_{it}^{\beta_k}n_{it}^{\beta_n}m_{it}^{\beta_m}$. Span of control problems mean that ($\beta_k + \beta_n + \beta_m = \theta < 1$). Profits are defined as revenues less costs (variable and a fixed overhead cost, O). Materials adjust at no cost, but there are adjustment costs for capital ($C_k(k_{it}, k_{it-1})$) and for labor $C_n(n_{it}, n_{it-1})$

[15] The code for running the dynamic simulations is in Matlab available on request.
[16] In this extension, we include both capital and materials as additional factors of production to labor.

that take the following quadratic form:

$$C_n(n_{it}, n_{it-1}) = \gamma_n n_{it-1}(\frac{n_{it} - n_{it-1}}{n_{it-1}} - \delta_n)^2 \text{ and } C_k(k_{it}, k_{it-1}) = \gamma_k k_{it-1}(\frac{k_{it} - k_{it-1}}{k_{it-1}} - \delta_k)^2$$

Labor and capital are defined by the accumulation equations $n_{it} = (1-\delta_n)n_{it-1}+H_{it}$ and $k_{it} = (1-\delta_k)k_{it-1}+I_{it}$where H_{it} is hiring, $I_{it}$is gross investment, δ_n is the separation rate and δ_k is the depreciation rate of capital. Normalizing output prices to unity, flow profits are:

$$\begin{aligned}\pi_{it} &= \alpha_{it} k_{it}^{\beta_k} n_{it}^{\beta_n} m_{it}^{\beta_m} - \gamma_n n_{it-1}(\frac{n_{it} - n_{it-1}}{n_{it-1}} - \delta_n)^2 - \gamma_k k_{it-1}(\frac{k_{it} - k_{it-1}}{k_{it-1}} - \delta_k)^2 - r k_{it} \\ &\quad - w\overline{\tau} n_{it} - p_m m_{it} - F - O\end{aligned}$$

As in the baseline static model we assume that in the regulated economy there is a variable cost (τ) and fixed cost (F) that has to be paid when firms cross the discrete employment threshold at 50 employees.

E.2 Timing, Shocks and Equilibrium

Agents choose to become entrepreneurs and enter until the expected value of entry is equal to the sunk entry costs (S). Upon entry firms draw their management ability from a known distribution $\phi(\alpha)$. There is an endogenous profit threshold such that some firms will immediately exit after observing their ability as they cannot cover their fixed costs (as in e.g. Melitz, 2003). In contrast to the static model we allow managerial ability/TFP to evolve over time according to an AR(1) process: $\ln \alpha_{it} = \rho_A \ln \alpha_{it-1} + u_{it}$. We assume u_{it} is i.i.d. normal with mean zero and variance σ_u. Firms decide their choices of hiring and investment (the policy correspondences) based on this shock, adjustment costs and the economic environment they face.[17]

Given the firm's three state variables - productivity α, capital k, and labor n - and the discount factor ϕ, we can write a value function (dropping i-subscripts for brevity) as:

$$\begin{aligned}V(\alpha_t, k_t, n_t) &= \max[V^c(\alpha_t, k_t, n_t), 0] \\ V^c(\alpha_t, k_t, n_t) &= \max_{k_{t+1}, n_{t+1}} y_t - r k_t - w\overline{\tau} n_t - p_m m_t - C_n(n_{t+1}, n_t) - C_k(k_{t+1}, k_t) - F - O \\ &\quad + \phi E_t V(\alpha_{t+1}, k_{t+1}, n_{t+1})\end{aligned}$$

where the first maximum reflects the decision to continue in operation or exit (where exit occurs when $V^c < 0$), and the second (V^c for "continuers") is the optimization of capital and labor conditional on operation. Firms begin with no factor inputs. Hence, the free entry condition is $S = \int V(\alpha, n_0, k)dG(\alpha)$. We solve for the steady-state equilibrium ensuring that the expected cost of entry equals the expected value of entry given optimal factor demand decisions. This equilibrium is characterized by a distribution of firms in terms of their state values α, k, n plus their optimized choice of materials m.

We choose a set of calibration values drawn from our data and the general literature to simulate the model, which are given in online Appendix Table A6. The results seem robust to reasonable changes in the exact values of these calibration values. For example, the labor adjustment cost parameter has generally be found to be lower than the capital adjustment cost parameter (e.g. Bloom, 2009), but it is not settled on exactly how much lower.

The initial distribution of TFP (α) is a power law. Firms have zero values of all factor inputs when they enter the economy. We discretize the state space for L, K and α into exponential bins for purposes of the numerical simulation. This means for TFP that there is a transition matrix between discrete points of the grid with fixed

[17] In addition to the endogenous exit decision, we also allow for an exogenous death rate that is modelled as an iid shock that causes some (small) margin of firms go out of business. This is not a necessary feature of the model as the TFP process is AR(1) rather than random walk, hence even a firm with a very high TFP draw will tend to revert to the average over the long run.

transition probabilities. We use results in Tauchen (1986) to approximate the AR(1) TFP process by this discrete transition matrix. TFP is the only source of stochastic variation shifting the environment. Firms adjust their factor inputs in response to changes in this state variable taking into account expectations, adjustment costs and their current state. Labor and capital are the other two state variables. For TFP we use a 30 by 30 grid. For employment we use a much finer grid to see the effect of regulation (100 points in the baseline). We use a coarser grid for capital (10 points), but not much hinges upon this (except the run-times for our MatLab code). We then generate the value function for entrants and incumbents. Using the contraction mapping (e.g. Stokey and Lucas, 1986) we iterate until we obtain a fixed point of the value function. The policy correspondences for hiring workers and capital investment are formed from the optimal choices given these value functions. Materials is simply the static period by period level based on the first order condition, as we are assuming no adjustment costs for this variable factor.

We simulate data for 20,000 firms over 100 periods (years). This time span was sufficient for the data to settle down to a steady state. We drop the first 75 "training years" and use the last 25 years to form the steady state in online Appendix Figures A11 to A14. In the various experiments, we observe how these change when we alter key model parameters such as the level of adjustment costs and the regulatory tax parameters. When making changes we simply re-calibrate the parameter (e.g. magnitude of the adjustment cost) and re-simulate the economy again as previously.

E.3 Results

Online Appendix Figure A11 shows the firm size distribution (in employment terms) in the steady state for our baseline calibration. Just as we saw in the static model there is a discontinuity at 49 employees with a large spike breaking the firm size distribution. To the right of the threshold there is then a valley before the power law continues again. Interestingly, there is a positive mass of firms in this valley, as there is in the real world data, even though there is no measurement error as we had in the baseline static model. This is essentially because of the adjustment cost structure as we will discuss below.

Online Appendix Figure A12 shows the impact of changing the level of the regulatory tax. In Panel A we simulate an economy where the tax is 8 times larger than in the baseline case (so 16% instead of 2%). As expected, the higher regulatory tax increases the spike at 49 and the length of the valley to the right of the threshold. There are now some employment cells with no firms to the right of the 50 employee cut-off. By contrast, Panel B reduces the regulatory tax to 25% of the baseline (i.e. to 0.5% instead of 2%). Here, the spike and the valley are almost muted as we would expect.

The intuition from the static model of changing the firm size distribution as a function of the regulatory tax are therefore robust to our dynamic extension. But the more interesting results are in online Appendix Figure A13 which changes the level of the labor adjustment cost parameter. In Panel A we double the adjustment cost. In this experiment we still observe a small spike at 49, but now the valley to the right of the regulatory threshold is much more populated with firms and the sharp discontinuity is blurred. By contrast, in Panel B we reduce the labor adjustment costs to about 13% of the baseline case (i.e. 0.010 instead of 0.075) and we see that there is a much smaller mass of firms in the valley and a much larger spike.

The intuition for this is that all else equal, firms try to avoid the valley - in the static model for example, no firm would optimally choose this point between N and n_r. But with in a dynamic equilibrium with labor adjustment costs, firms may find it optimal to transit through the valley. To see this assume that current managerial ability/productivity would put an incumbent firm to the left of the regulatory threshold and consider what would happen if such a firm received a large positive productivity shock. The firm would like to move to a larger size[18] and earn more profits by hiring more workers. However, since there are convex adjustment costs, spending a period in the valley that is sub-optimal from a static point of view may be optimal from a dynamic point of view. Large jumps in employment are extremely costly because of adjustment costs so landing for a period or two in the valley may be less costly than

[18] Although the firm expects to eventually converge back down to average productivity, the positive productivity shock is persistent (recall $\rho_A = 0.9$ in the baseline calibration).

paying these large adjustment costs to make a large immediate expansion. Hence when we increase adjustment costs, these considerations become more important and the mass of firms in the valley rises.

We can see this even more clearly when we consider the dynamic adjustment properties of the model and compare them with the actual data. Online Appendix Figure A14 shows the proportion of firms who increase employment in the next year as a function of firm size in the current year. Panel A shows the raw data. We see that there is a big fall in the probability of increasing employment for firms to the left of the regulatory threshold. Panel B shows the same graph for the simulated data in our baseline calibration (of online Appendix Figure A11). As with the raw data there is much less incentive to grow when a firm is just below the threshold. Panel C shows the same graph in the calibration that uses higher adjustment costs (online Appendix Figure A13, Panel A). We see the same qualitative picture, but with a lower level of employment changes and less volatility as we would expect from the higher adjustment costs. The sharpness of the fall in growth rates to the left of the threshold becomes more muted.[19]

To conclude, when there are labor adjustment costs, the previous exercises show that steady state simulations reproduce the intuitions underlying the static model (and key empirical features of the data). There is a bulge of firms with employment just to the left of the regulatory threshold, a valley to the right of the threshold and then a continuation of the power law of the firm size distribution. Just as in the static model, the size of the bulge and the width of the valley are increasing in the magnitude of the regulatory "tax". An additional insight from the dynamic model is that there is a positive mass of firms in the valley, even when all firms are fully optimizing. In the static model we rationalized the positive mass in the valley in terms of measurement error, whereas in reality we believe a mixture of adjustment costs and measurement error helps explain the valley. Although a useful exercise, ultimately the dynamic model does not delivery fundamentally new insights compared to the static model, so we believe the considerably simpler approach of our basic model has much to recommend it. Nevertheless, estimating the regulatory costs through Simulated Method of Moments on the dynamic model, would be an interesting extension that we leave for future work.

F More Details of Some Size-Related Regulations in France

The size-related regulations evolve over time and are defined in four groups of laws. The Code du Travail (labor laws), Code du Commerce (commercial law), Code de la Sécurité Social (social security) and in the Code Général des Impôts (fiscal law). The main bite of the 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 10 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), or, more administratively and exhaustively, LAMY (2010).

F.1 Main Labor Regulations

The unified and official way of counting employees has been defined since 2004[20] in the Code du Travail[21], articles L.1111-2 and 3. Exceptions to the 2004 definition are noted in parentheses in our detailed descriptions of all the regulations below. Employment is taken over a reference period which from 2004 depends on the precise date when

[19] An alternative way to dynamically model the regulatory tax would be to consider it as a purely sunk cost (or sunk cost plus variable cost) as in Gourio and Roys (2014). This has similar predictions on the firm size distribution and simplifies the dynamics compared to the model presented here (as there are no convex adjustment costs). However, the raw data on employment dynamics do not seem compatible with this approach. Panel A of online Appendix Figure A1 shows how employment adjusts around the threshold at 49 for all firms and then the sub-set of firms who had (in the past five years) already had 50 or more employees (Panel B). If the regulatory cost was sunk the latter group - who presumably have already paid the sunk cost - should not bunch much around 49 employees. However, whether we look at firms who make no change in employment in Panel A, we observe a big spike at 49 employees for both groups. This does not seem consistent with a simple sunk cost story.

[20] Before that date, the concept of firm size was different across labor regulations.

[21] The text is available at:
`http://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006072050&dateTexte=20120822`

employment is estimated and covers the 365 previous days. There are precise rules over how to fractionally count part-year workers, part-time workers, trainees, workers on sick leave, etc. For example, say a firm employs 10 full-time workers every day but in the middle of the year all 10 workers quit and are immediately replaced by a different 10 workers. Although in the year as a whole 20 workers have been employed by the firm the standard regulations would mean the firm was counted as 10 employee firm. In this case this would be identical to the concept used in our main data (FICUS, see the discussion in section C.4 above).
Last, almost all of these regulations strictly apply to the firm level, which is where we have the FICUS data. Some case law has built up, however, which means that a few of them are also applied to the group level. We report below regulations that were stable over time, but some unreported regulations were altered, introduced or removed during our estimation period. The precise tracking of these events is very difficult in practice.

From 200 employees:

- Obligation to appoint nurses (Code du Travail, article R.4623-51)
- Provision of a place to meet for union representatives (Code du Travail, article R.2142-8)

From 50 employees:

- Monthly reporting of the detail of all labor contracts to the administration (Code du Travail, article D.1221-28)
- Obligation to establish a staff committee ("comité d'entreprise") with business meeting at least every two months and with minimum budget = 0.3% of total payroll (Code du Travail, article L.2322-1-28, threshold exceeded for 12 months during the last three years)
- Obligation to establish a committee on health, safety and working conditions (CHSC) (Code du Travail, article L.4611-1, threshold exceeded for 12 months during the last three years)
- Appointing a shop steward if demanded by workers (Code du Travail, article L.2143-3, threshold exceeded for 12 consecutive months during the last three years)
- Obligation to establish a profit sharing scheme (Code du Travail, article L.3322-2, threshold exceeded for six months during the accounting year within one year after the year end to reach an agreement)
- Obligation to do a formal "Professional assessment" for each worker older than 45 (Code du Travail, article L.6321-1)
- Higher duties in case of an accident occurring in the workplace (Code de la sécurité sociale and Code du Travail, article L.1226-10)
- 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 (Code du Travail, articles L.1235-10 to L.1235-12; threshold based on total employment at the date of the redundancy)

From 25 employees:

- Duty to supply a refectory if requested by at least 25 employees (Code du Travail, article L.4228-22)
- Electoral colleges for electing representatives. Increased number of delegates from 25 employees (Code du Travail, article L.2314-9, L.2324-11)

From 20 employees:

- Formal house rules (Code du Travail, articles L.1311-2)
- Contribution to the National Fund for Housing Assistance;

- Increase in the contribution rate for continuing vocational training of 1.05% to 1.60% (Code du Travail, articles L.6331-2 and L.6331-9)
- Compensatory rest of 50% for mandatory overtime beyond 41 hours per week

From 11 employees:

- Obligation to conduct the election of staff representatives(threshold exceeded for 12 consecutive months over the last three years) (Code du Travail, articles L.2312-1)

From 10 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 R.2531-7 and 8 of the General Code local authorities, Code général des collectivités territoriales);
- Increase the contribution rate for continuing vocational training of 0.55% to 1.05% (threshold exceeded on average 12 months).

Note that, in additions to these regulations, some of the payroll taxes are related to the number of employees in the firm.

F.2 Accounting rules

The additional requirements depending on the number of employees of entreprises, but also limits on turnover and total assets are as follows (commercial laws, Code du Commerce, articles L.223-35 and fiscal regulations, Code général des Impôts, article 208-III-3):

From 50 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 sales are more than 3.1 million euros, applicable rules of the current year).

From 10 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).

References

Ackerberg, Daniel, Lanier Benkard, Stephen Berry and Ariel Pakes (2007) in Edward Leamer and James Heckman (eds) "Econometric tools for analyzing market outcomes" *Handbook of Econometrics Vol 6A*, 63, 4172-4276, Amsterdam: Elsevier.

Armbruster, Heidi, Steffan Kinkel, Gunter Lay and Someka Maloca (2005) "Techno-organizational innovation in European manufacturing industries" *European Manufacturing Survey Bulletin*, 1, 1-16.

Bloom, Nicholas (2009) “The Impact of Uncertainty Shocks” *Econometrica,* 77(3), 623–685.

Bloom, Nicholas, Raffaella Sadun and John Van Reenen (2014) “Management as a Technology”, LSE mimeo.

Bond, Stephen and John Van Reenen (2007) “Micro-econometric models of investment and employment” Chapter 65 in Heckman, J. and Leamer. E. (eds) *Handbook of Econometrics Volume 6A* (2007) 4417-4498.

Cooper, Russell and John Haltiwanger (2006) “On the nature of Capital Adjustment Costs” *Review of Economic Studies*, 73, 611-633.

Foster, Lucia, John Haltiwanger and Chad Syverson (2008) “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review,* 98(1): 394-425.

Hicks, John (1932) *The Theory of Wages.* London: Macmillan.

Leibenstein, Harvey (1966) “Allocative Efficiency vs. X-Efficiency” *American Economic Review,* 56 (3): 392–415.

Stokey, Nancy and Robert Lucas (1986) *Recursive Methods in Economic Dynamics*, Cambridge: Harvard University Press.

Stokey, Nancy (2009) *The Economics of Inaction: Stochastic Control Models with Fixed Costs*, Princeton: Princeton University Press.

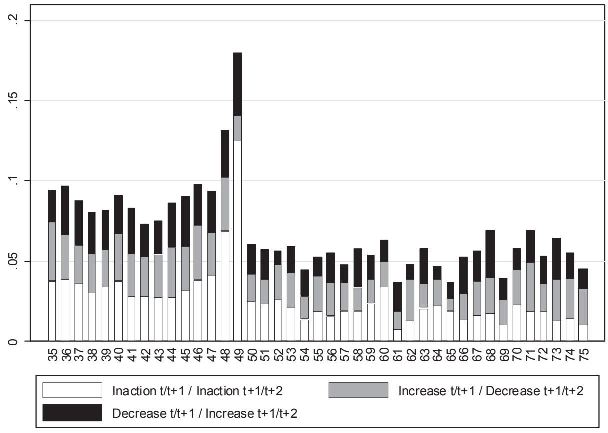
Syverson, Chad (2011) “What determines productivity?” *Journal of Economic Literature*, 49(2) 326-65.

Tauchen, George (1986) “Finite State Markov Chain Approximation to univariate and vector autoregressions” *Economics Letters*, 20: 177-181.

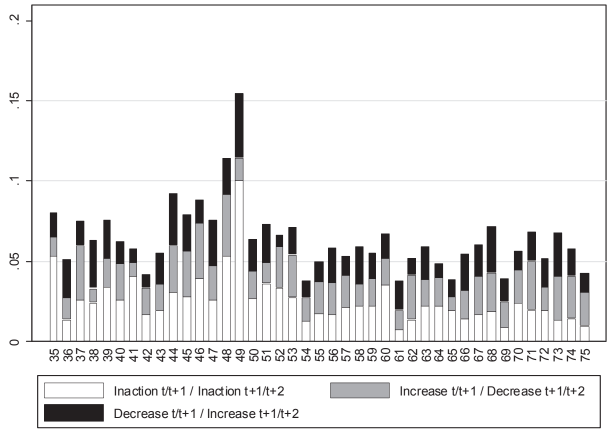
APPENDIX FIGURES AND TABLES

Figure A1: Inaction Rate over a Two Year Period (Between t and t+2)

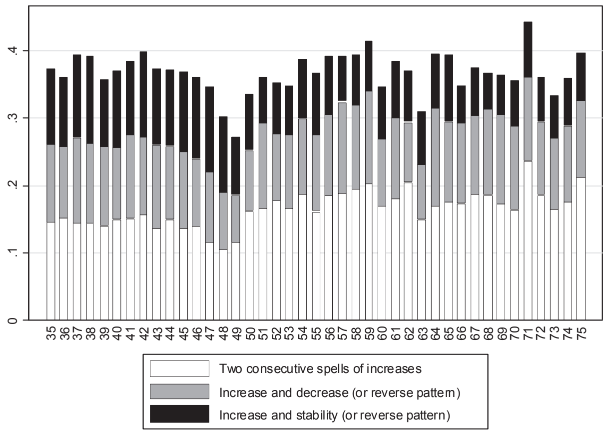
Panel A: All Firms



Panel B: Conditional on Having Crossed the 49 Threshold in the Past 5 Years (Sunk vs. Fixed Costs)



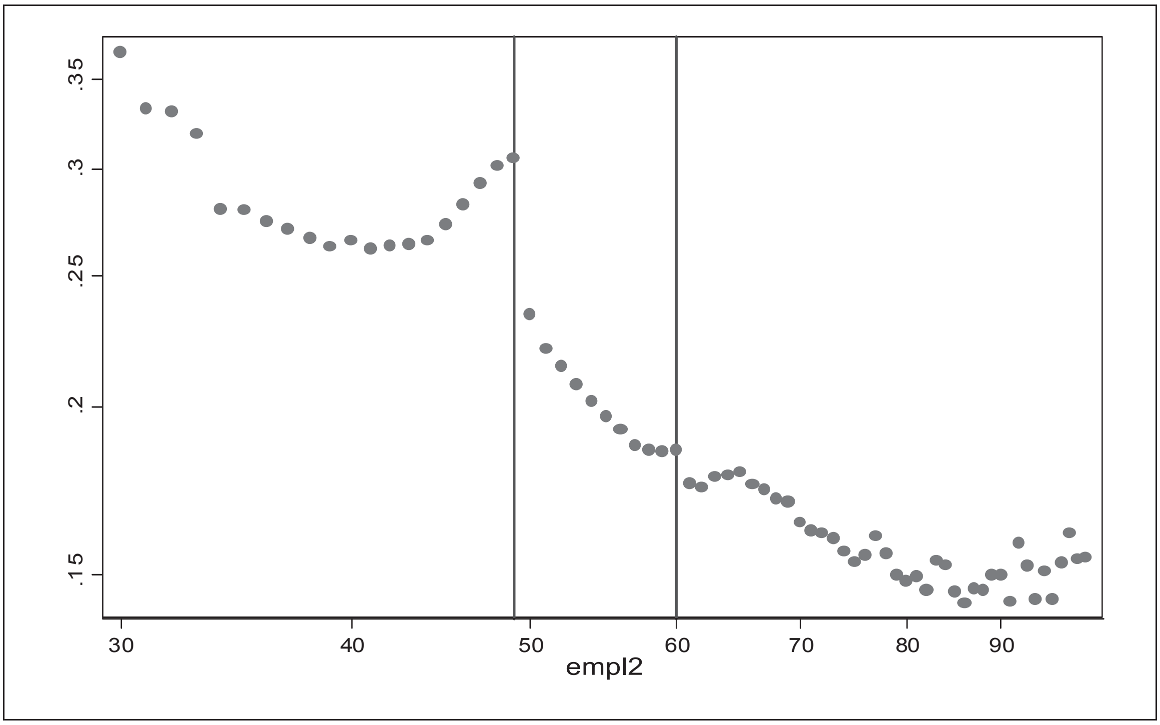
Panel C: Patterns of Firm Growth over a Two Year Period (Between *t* and *t*+2)



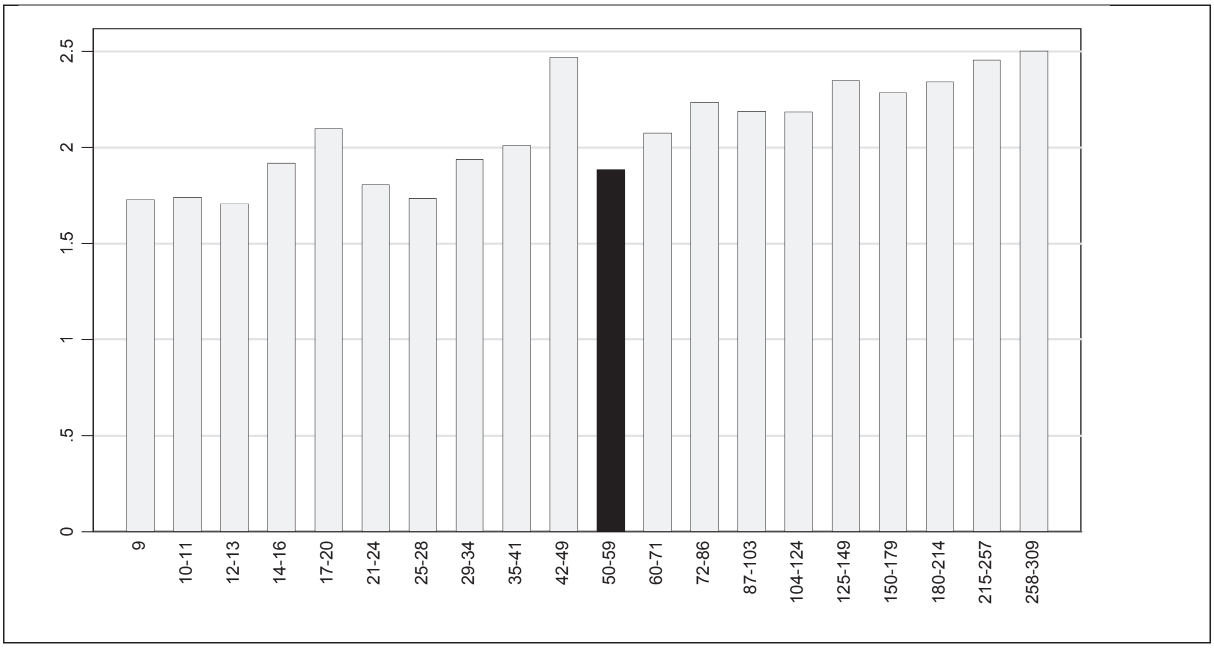
Notes: This Figure shows employment transitions over a two year period, broken down by size class. In Panels A and B the "inaction rate" is the share of firms having strictly the same size at dates *t* and *t+ 2*. We break this down into firms who have stayed at exactly the same size for two years (white bar) vs. firms who got larger than shrank back to their original size (grey bar) vs. firms who shrank and then grew back to their original size (black bar). Panel A is for the full sample and Panel B condition on the subset of firms who had 50 or more employees at least once in the previous 5 years. Panel C describes the patterns of firms growth over a two year period for firms at different employment sizes that are larger at time t+2 as compared with time t. The grey bar are for firms who increased then decreased (or decrease then increased). The white bar is the proportion of firms who grew in both periods, the grey bar is for firms who increased and then did not change (or who did not change and then increased).

Figure A2: Average Spell Length in an Employment Size Bin

Panel A: Raw Firm Size Bins



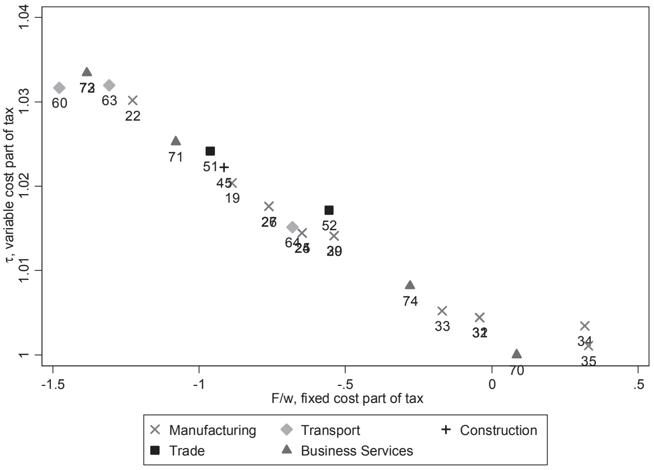
Panel B: Exponential Firm Size Bins



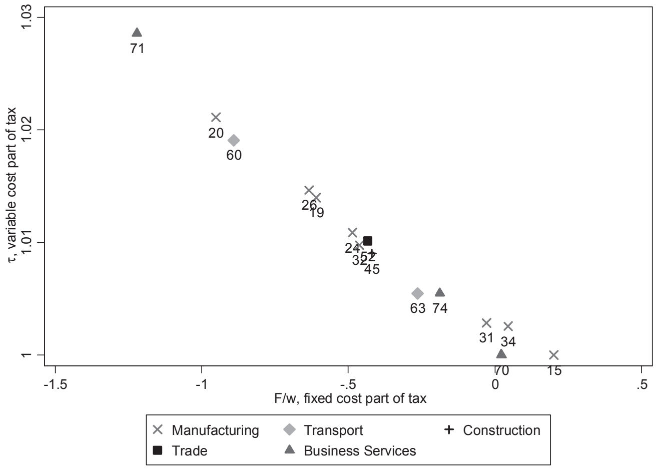
Notes: Manufacturing firms observed between 1995 and 2007 (censored and uncensored spells - results are robust to the removing of censored spells). To compute durations in each state, we only observe firm size at an annual frequency, so we assume that a firm which we observe at size 30 at date t, and at size 35 at date t+1 has transited through sizes 31, 32, 33, 34 during year t/t+1, and we adjust spells in each state accordingly. Panel A gives durations by employment sizes and Panel B presents the same but in exponential bins.

Figure A3: Heterogeneity of Results by Two Digit Sectors

Panel A: θ Calibrated to 0.8



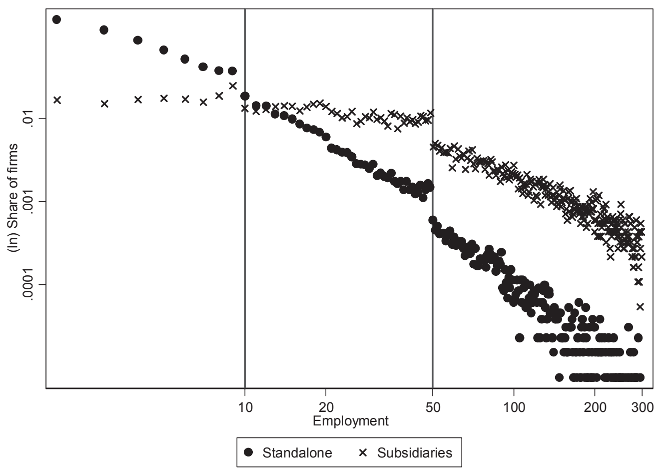
Panel B: Industry Specific θ's



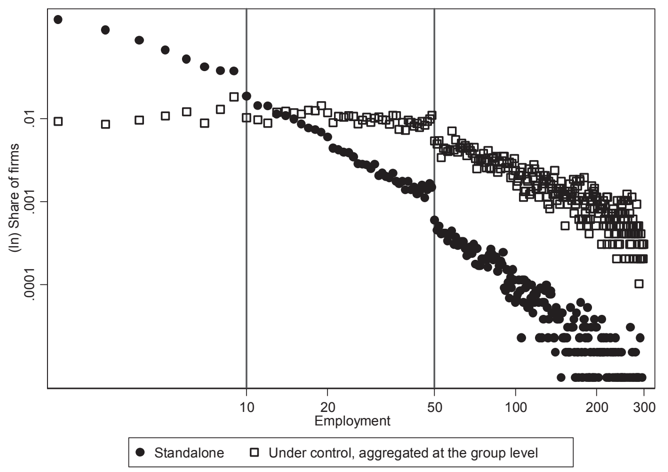
Notes: These are the results from industry-specific estimation on the same lines as column (1) of Table 1 in panel A and column (6) of Table 1 in panel B. Industry codes correspond to the first two digits of the statistical classification of economic activities in the European Community (NACE).

Figure A4: Corporate Restructuring in Response to the Regulation? Independent Firms vs. Corporate Groups in 2000

Panel A: Standalone Firms vs. Affiliates of Larger Groups



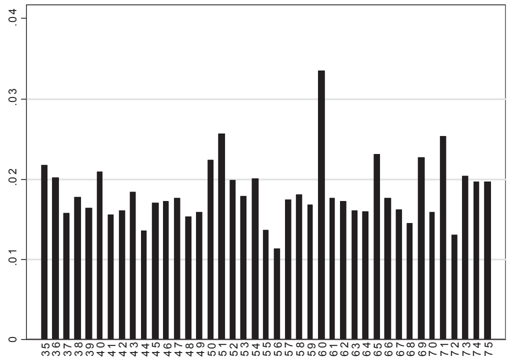
Panel B: Standalone Firms vs. Groups (i.e. All Affiliates Aggregated at the Group Level)



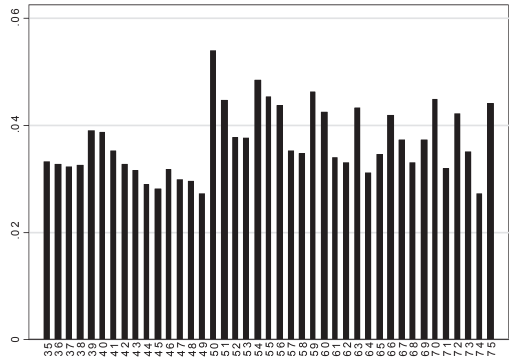
Sources: FICUS, 2000 and LIFI, 2000.
Notes: "Standalone" are independent firms that are not subsidiaries or affiliates of larger groups (blue dots). In Panel A we compare these to affiliates of larger groups with size measured at the affiliate level. A broken power law is visible in both distributions. In Panel B we repeat the standalone distribution but now compare this to affiliates aggregated to the group level (in France, we do not count overseas employees). Although there is a break in the power law for both type of firms it is stronger for the standalone firms as we would expect, i.e. the subsidiaries are not driving the results.

Figure A5: Descriptive Statistics on Entry and Exit

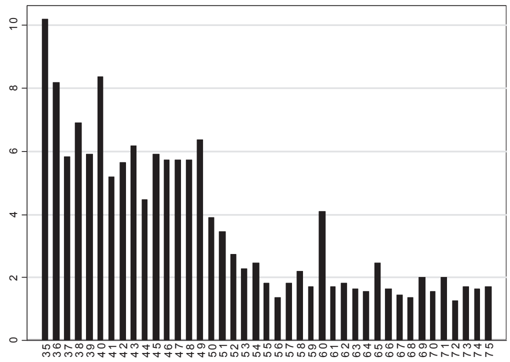
Panel A: Share of Entry Between *t*-1 / t in Each Size Bin



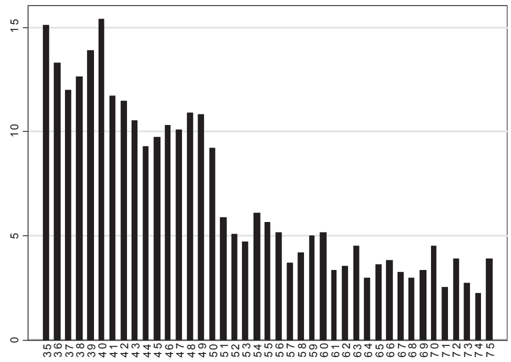
Panel B: Share of Exit t / t+1 in Each Size Bin



Panel C: Size Distribution at Entry (Average Number of Entrants in Each Size Bin)

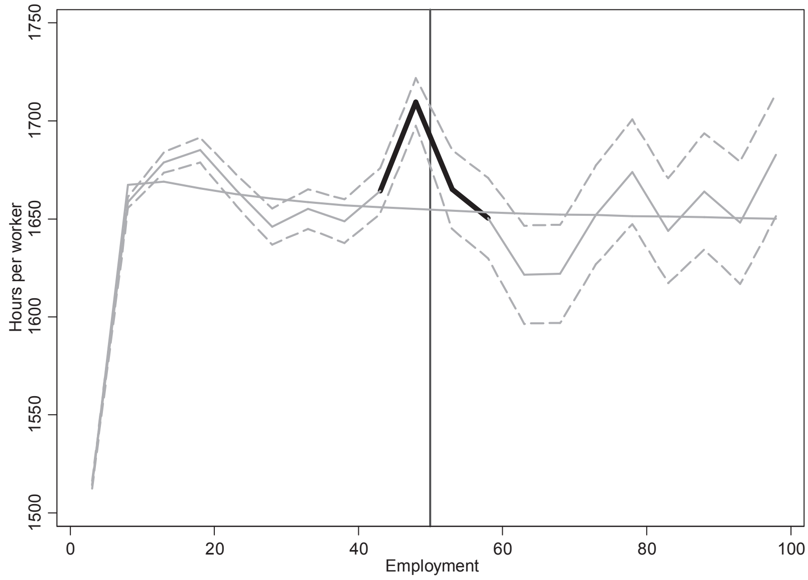


Panel D: Size Distribution at Exit (Average Number of Exiters in Each Size Bin)



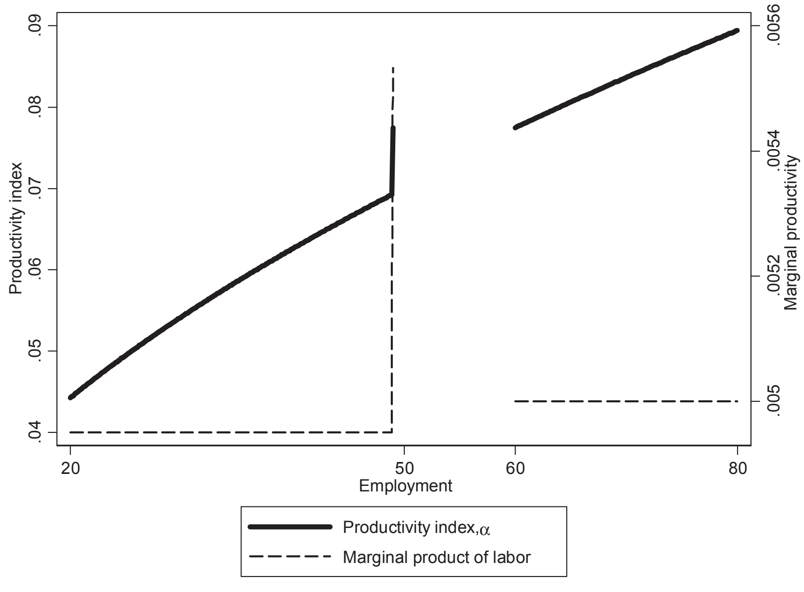
Notes: Manufacturing firms, 1995-2007, FICUS. Size is measured at time t in all panels.

Figure A6: Annual Hours per Worker



Notes: annual average hours per worker - combined FICUS and DADS data for 2002.

Figure A7: MRPL and Marginal Revenue Productivity of Labor in our Model

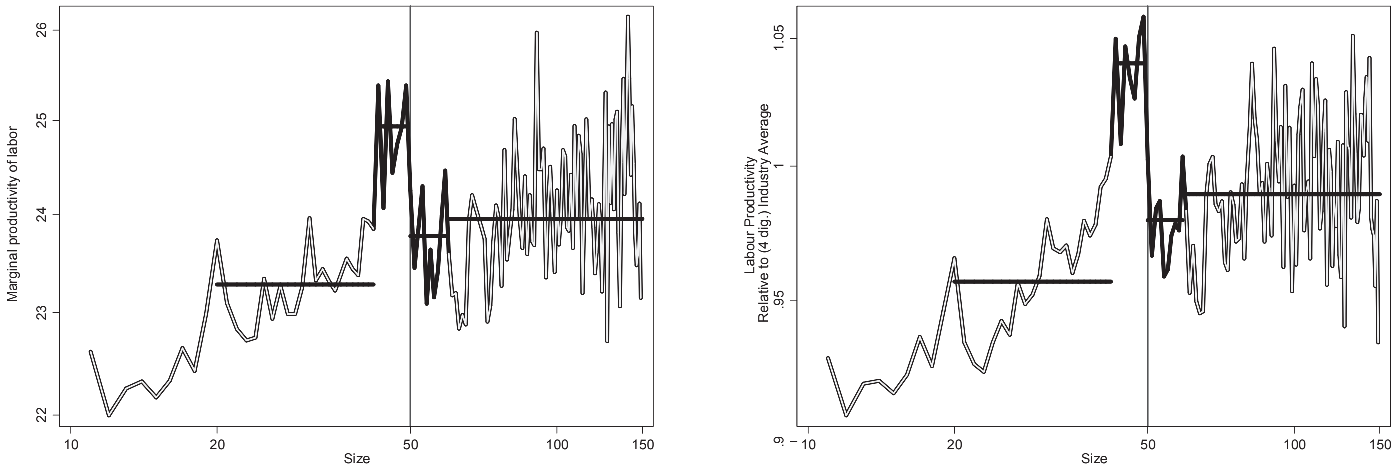


Notes: This figure shows the relation (correspondence) between productivity (α, left axis) or marginal product of labor ($\alpha.\theta.(n^*(\alpha))^{\theta-1}$, right axis) and size. The relations are defined as $\alpha.\theta.(n^*(\alpha))^{\theta-1} = w$ for $n < N$ and $\alpha.\theta.(n^*(\alpha))^{\theta-1} = \tau.w$ for $n > n_r$

Figure A8: Marginal Revenue Productivity and Firm Size (Value Added per Worker Relative to Industry Average)

Panel A: Value added per worker relative to industry average

Panel B: Empirical counterpart of theory $\boldsymbol{\alpha.\theta.(n^*(\alpha))^{\theta-1}}$

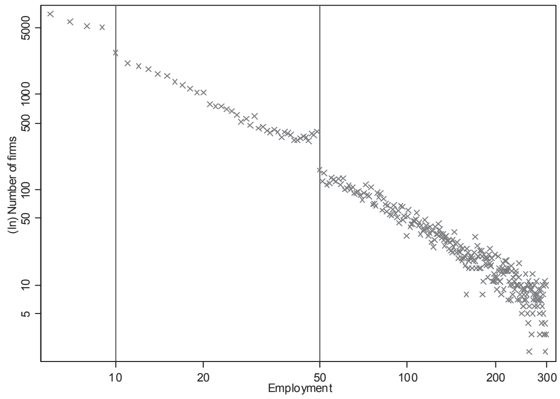


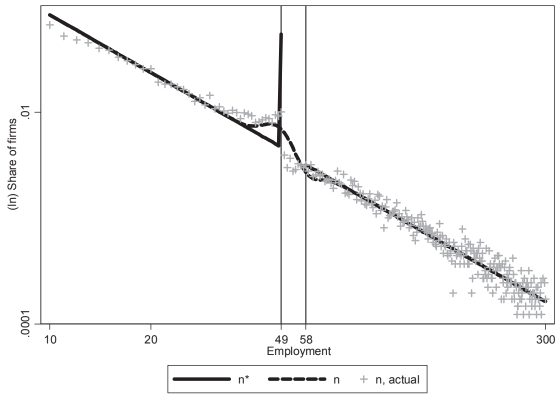
Notes: Manufacturing firms in year 2000. In panel A the marginal revenue productivity of labor is measured by value added per worker (relative to the four digit industry average) by firm size. The "Hsieh-Klenow" estimate of the implicit tax distortion is the difference between average productivity for firms between 20 and 42 employees (€23,280) and average productivity for firms between 60 and 150 employees (€23,960). This implies a log difference of 0.02866 or 2.866%. In Panel B, the marginal revenue productivity of labor is measured by $\alpha.\theta.(n^*(\alpha))^{\theta-1}$ by firm size. The "Hsieh-Klenow" estimate of the implicit tax distortion is the difference between average productivity for firms between 20 and 42 employees (0.957) and average productivity for firms between 60 and 150 employees (0.989). This implies a log difference of 0.03326 or 3.326%.

Figure A9: Model fit Using Alternative Datasets and Definitions of Employment

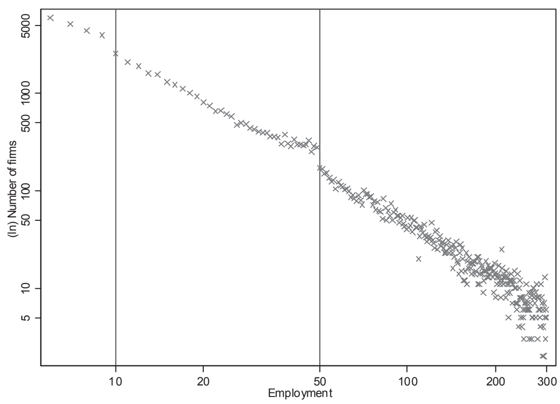
Log-log Plot *Model Fit*

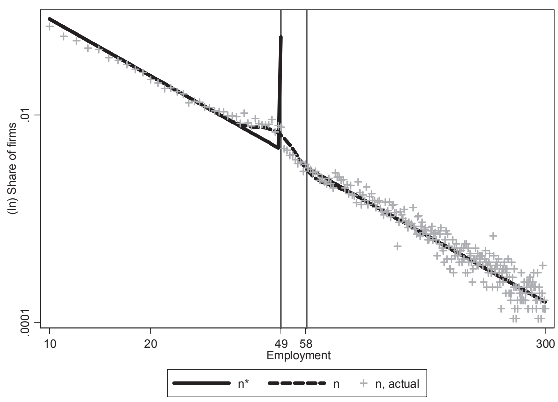
Panel A: Fiscal Source (Corporate Tax Collection to Fiscal Administration, FICUS 2002)
Arithmetic Average of Quarterly Head Counts



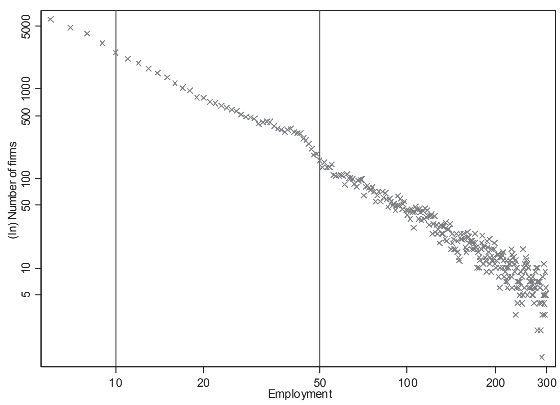


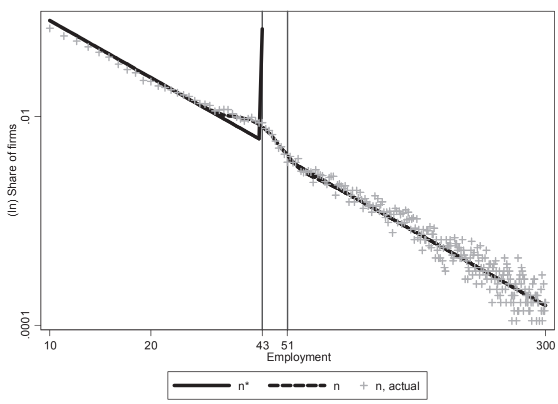
Panel B: Payroll Tax Reporting to Social Administration (DADS 2002)
“Declared” Workers on Dec. 31st: Cross-Sectional (and Fractional) Count





Panel C: Payroll Tax Reporting to Social Administration (DADS 2002)
"Full-Time Equivalent" (FTE), Computed by the French Statistical Institute

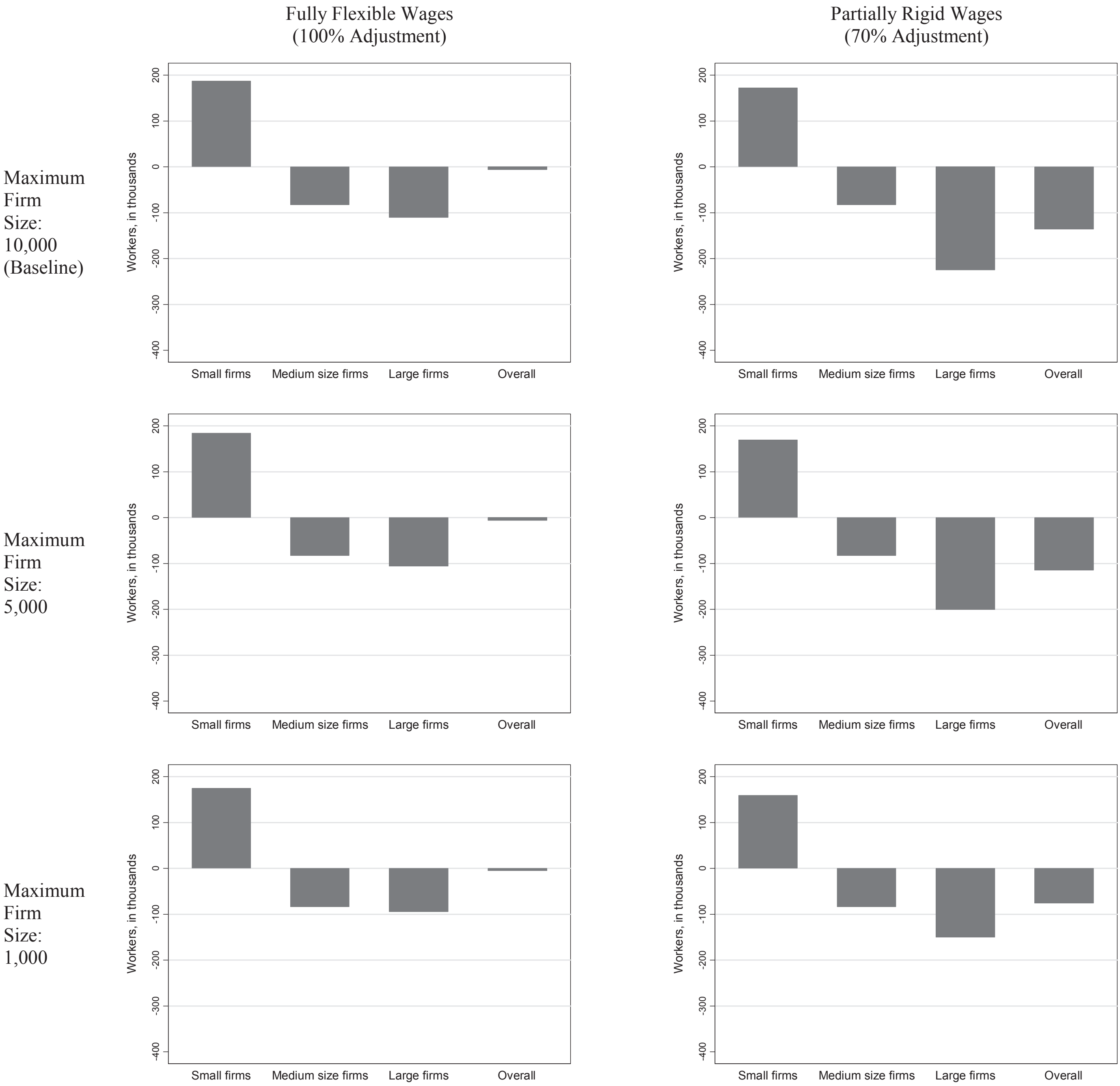




Notes: Data sources are indicated in the headers of the table. Full time equivalents (panel C) are only available from 2002 onwards in the files, such that all distributions have been computed in 2002. All firms in this figure had an industry affiliation in manufacturing industries at creation date. Full time equivalents in panel C are calculated by the statistical institute using as a benchmark, the 75th percentile of the series of hours per workers in each 2 digit industries and size classes, and are bounded above by 1, which produces the small shift to the left.

Figure A10: Welfare and Employment Reallocation Analysis under Alternative Assumptions for the Upper Bound of Firm Size

Panel A: Employment, Gross Changes

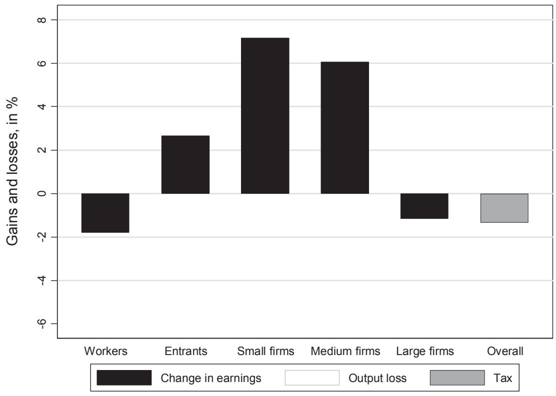


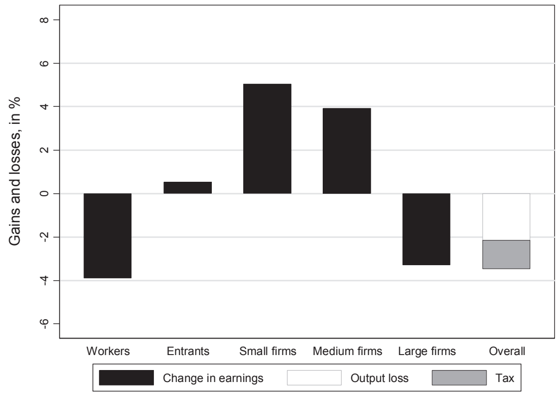
Notes: This is based on the baseline of Table 1 column (1), under various assumptions for maximum firm size. All definitions are similar to those in Figure 10.

Panel B: (Expected) Percentage Changes in Welfare (Income from Wages or Profits)

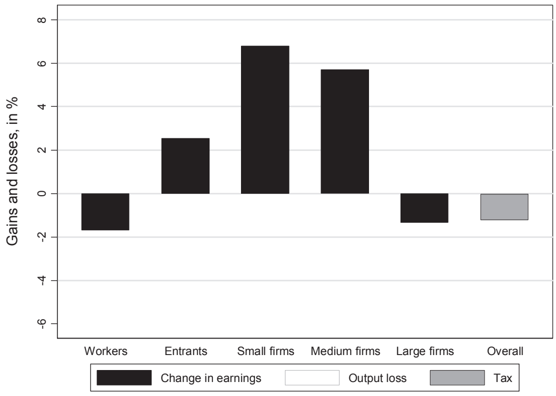
Fully Flexible Wages (100% Adjustment) | Partially Rigid Wages (70% Adjustment)

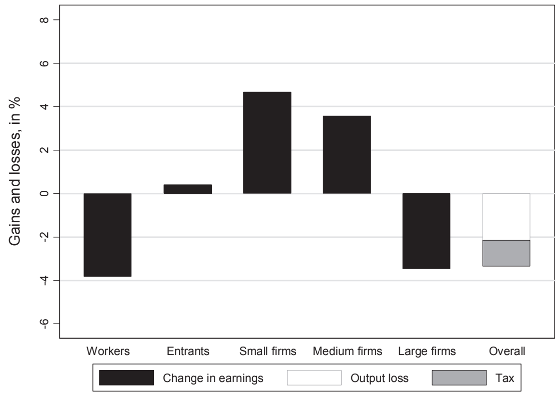
Maximum Firm Size: 10,000 (Baseline)



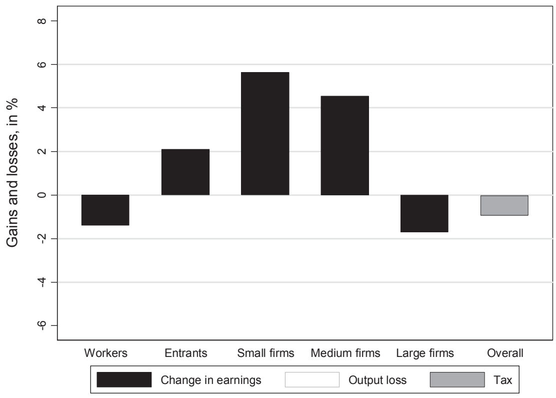


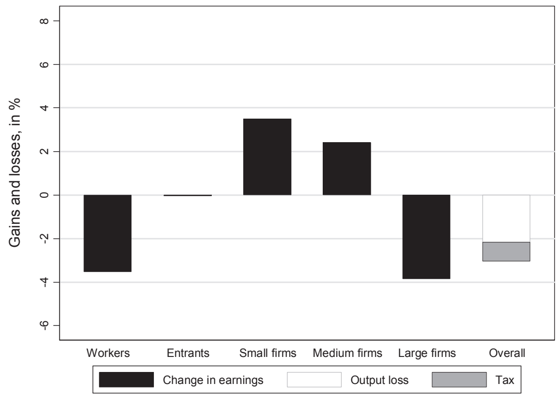
Maximum Firm Size: 5,000





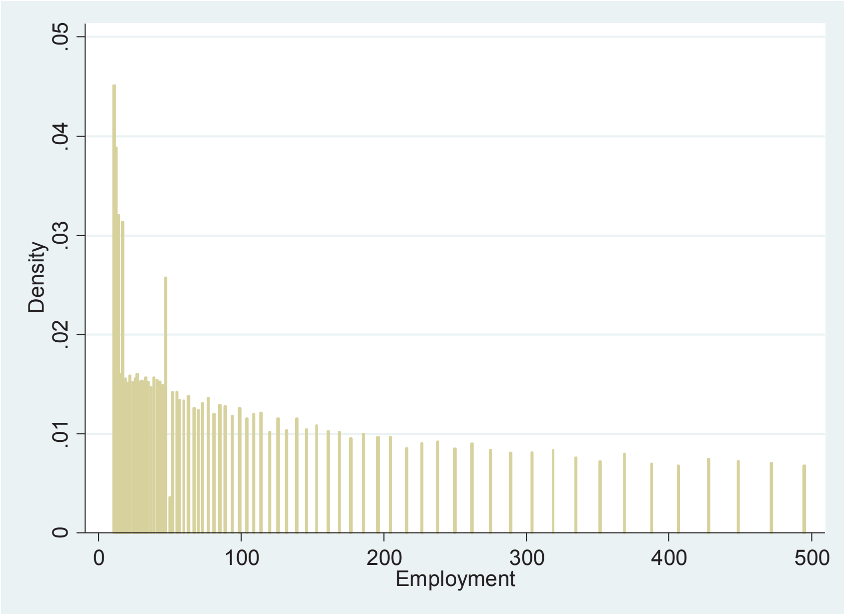
Maximum Firm Size: 1,000





Notes: This is based on the baseline of Table 1 column (1), under various assumptions for maximum firm size. All definitions are the same as those in Figure 10.

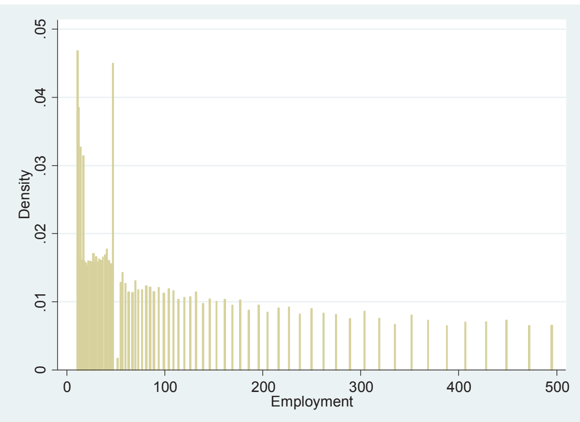
Figure A11: Firm Size Distribution from the Simulated Dynamic Model: Baseline Specification



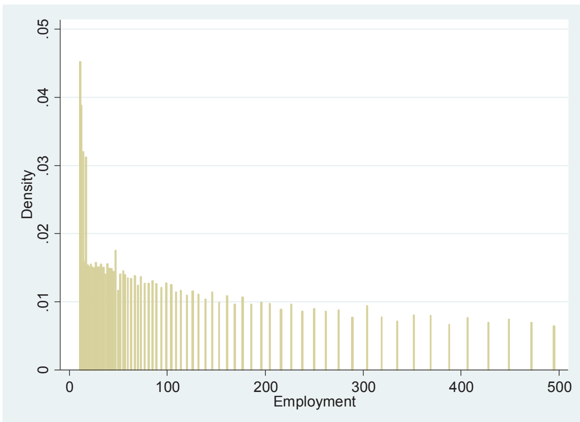
Notes: This is the histogram of the firm size distribution, i.e. the proportion of firms across different size classes. We run the simulation model for 100 years over 20,000 firms and keep only the last 25 years (but which time the distributions appeared to be very stable). Regulatory variable tax = 2%; regulatory fixed cost, F/w = 0.2; adjustment cost parameter for labor=0.075. Appendix E gives the full model details and other calibration values.

Figure A12: Changing the Magnitude of the Regulatory Tax

Panel A: Increase Regulatory Tax by 8 Times as Much as Baseline Case



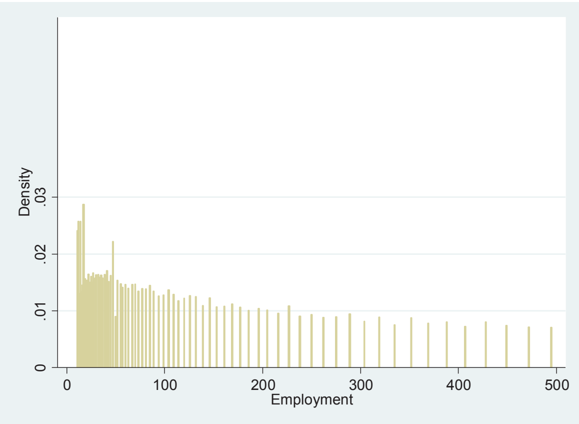
Panel B: Lowering the Regulatory Tax to 25% of the Level in Baseline Case



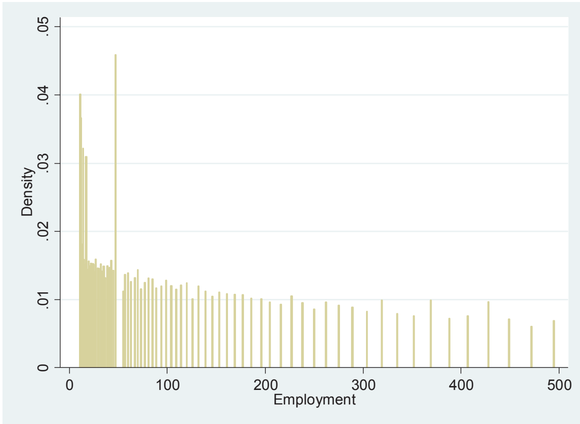
Notes: This is the histogram of the firm size distribution, i.e. the proportion of firms across different size classes. The simulation model the same as baseline in Figure 15 except in Panel A the regulatory variable tax = 16% (instead of 2% as in the baseline case) and in Panel B the regulatory variable tax is lowered to 0.5%.

Figure A13: Changing the Magnitude of the Adjustment Costs

Panel A: Increasing Labor Adjustment Costs Smooths over the "Valley"



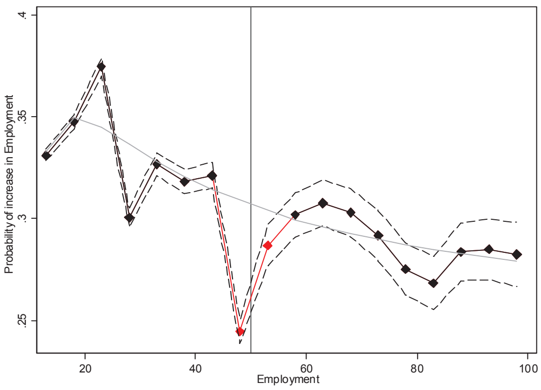
Panel B: Reducing Adjustment Costs Makes Hump Higher and Valley Wider



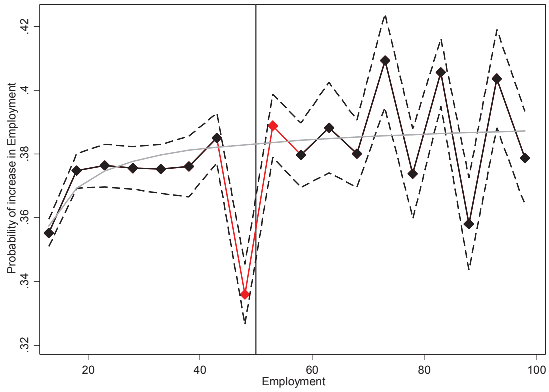
Notes: This is the histogram of the firm size distribution, i.e. the proportion of firms across different size classes. The model is the same as baseline in Figure 15 except in Panel A except the labor adjustment cost parameter is doubled to 0.150 (instead of 0.075 in baseline) and in Panel B labor adjustment costs are lowered to 0.010.

Figure A14: Percentage of Firms Increasing Employment in Actual and Simulated Data

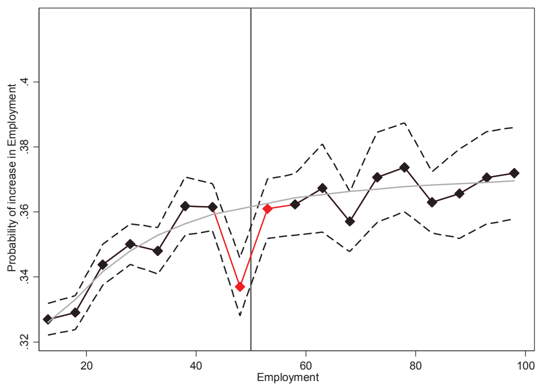
Panel A: Actual Data



Panel B: Simulated Data, Baseline



Panel C: Simulated Data, High Adjustment Costs



Notes: These graphs examine the proportion of firms whose employment in the next period as a function of employment in the current period. Panel A does this for the actual data and Panels B and C do the same for the simulated data. Panel B is the baseline calibration (as in Figure 15) and Panel C uses a higher adjustment cost (as in Figure 17, Panel A). In the empirical data we do this for all firms and for all available years. 95% confidence intervals shown.

APPENDIX TABLES

Table A1: Parameter Estimates in Different Years

Sample	(1) 1995-2007	(2) 1995	(3) 2000	(4) 2005
β, power law	1.815	1.766	1.800	1.834
	(0.055)	(0.053)	(0.054)	(0.061)
n_r, upper employment threshold	58.014	58.483	59.271	60.297
	(1.381)	(4.052)	(2.051)	(1.310)
σ, variance of measurement error	0.102	0.114	0.121	0.143
	(0.023)	(0.069)	(0.033)	(0.025)
τ-1, implicit tax, variable cost	0.017	0.022	0.023	0.017
	(0.006)	(0.013)	(0.008)	(0.008)
F/w, implicit tax, fixed cost	-0.711	-0.930	-0.941	-0.592
	(0.269)	(0.484)	(0.338)	(0.379)
Mean (Median) # of employees	55.0 (24)	57.6 (25)	55.8 (24)	53.9 (23)
Observations	517,985	39,638	41,067	37,954
Firms	74,183	39,638	41,067	37,954
Ln Likelihood	-2,314,827.4	-180,281.0	-184,128.7	-168,337.0

Notes: Parameters estimated by ML with standard errors below in parentheses (clustered at the four digit level). Estimation is on unbalanced panel 1995-2007 of population of French manufacturing firms with 10 to 1,000 employees. θ, the scale parameter, is calibrated to 0.8 in all columns. In column (4), 2005 is simply chosen for symmetry (2007 would deliver very similar results).

Table A2: Variation in Estimates Across Different Sectors

Industry	(1) Manufacturing	(2) Transport	(3) Construction	(4) Trade	(5) Business services
β, power law	1.800	1.857	2.345	2.129	1.972
	(0.054)	(0.098)	(0.122)	(0.085)	(0.079)
n_r, upper employment threshold	59.271	62.139	56.916	55.863	53.370
	(2.051)	(4.134)	(1.869)	(3.057)	(1.333)
σ, measurement error	0.121	0.150	0.083	0.072	0.045
	(0.033)	(0.053)	(0.022)	(0.035)	(0.015)
τ-1, implicit tax, variable cost	0.023	0.035	0.020	0.026	0.008
	(0.008)	(0.007)	(0.005)	(0.010)	(0.002)
F/w, implicit tax, fixed cost	-0.941	-1.397	-0.878	-1.206	-0.358
	(0.338)	(0.332)	(0.225)	(0.458)	(0.102)
Mean (Median) # of employees	55.8	48.5	30.0	36.1	47.1
	(24)	(23)	(17)	(19)	(20)
Observations	41,067	10,907	23,506	41,071	30,125

Notes: Parameters estimated by ML with standard errors below in parentheses (clustered at the four digit level). Estimation is on unbalanced the population of French firms with 10 to 1,000 employees in year 2000. These estimates of the implicit tax are based on a calibrated value for θ of 0.8.

Table A3: Parameter Estimates Across Industries in the Specification with 2 Breaks (at 9 and 49 Workers)

Industry:	(1) Manufacturing industries	(2) Transport	(3) Construction	(4) Trade	(5) Business services
β, power law	1.828	1.835	2.355	2.291	2.080
	(0.095)	(0.064)	(0.100)	(0.132)	(0.079)
σ, measurement error	0.079	0.076	0.078	0.079	0.078
	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)
τ_1-1, implicit tax, variable cost,	0.001	0.005	~0	~0	0.003
threshold at 9	(0.001)	(0.001)			(0.001)
F_1/w, implicit tax, fixed cost	0.013	-0.023	0.017	0.024	-0.010
threshold at 9	(0.011)	(0.017)	(0.002)	(0.003)	(0.014)
n_{r1}	10.361	10.295	10.288	10.534	10.364
threshold at 9	(0.073)	(0.187)	(0.088)	(0.102)	(0.144)
τ_2-1, implicit tax, variable cost	0.017	0.029	0.019	0.016	0.002
threshold at 49	(0.008)	(0.009)	(0.005)	(0.009)	(0.004)
F_2/w, implicit tax, fixed cost	-0.732	-1.315	-0.845	-0.639	0.005
threshold at 49	(0.401)	(0.458)	(0.257)	(0.478)	(0.189)
n_{r2}	56.241	55.880	56.517	57.029	56.288
threshold at 49	(0.456)	(0.225)	(0.370)	(0.901)	(0.668)

Notes: Parameters estimated by ML with standard errors below in parentheses (clustered at the four digit level). Estimation is on unbalanced the population of French firms with 5 to 1,000 employees in year 2000. θ, the scale parameter, is calibrated to 0.8 in all columns. In columns (3) and (4), the variable part of tax at first threshold is set to 1 to reach convergence, indicating that the implicit tax is very close to zero.

Table A4: Welfare and Distributional Analysis in the Two Threshold Specification

Panel A: Fully Flexible Wages

(Regulated Economies - Same Unregulated Economy) Variable		(1) 2 thresholds	(2) Threshold at 10 only	(3) Threshold at 50 only
1.	Unemployment rate	0%	0%	0%
	Implied wage flexibility	**100%**	**100%**	**100%**
2.	Percentage of firms avoiding the first regulation, δ_1	5.748%	5.748%	-
3.	Percentage of firms avoiding the second regulation, δ_2	2.106%	-	2.112%
4.	Percentage of firms paying the first regulation	45.277%	45.273%	
5.	Percentage of firms paying the second regulation (on top of the first)	10.296%	-	10.322%
6.	Change in labor costs (wage reduction) for small firms (below 9)	-1.341%	-0.067%	-1.273%
7.	Change in labor costs (wage reduction but first tax) for medium firms (between 10 and 49)	-1.279%	-0.005%	-1.273%
8.	Change in labor costs (wage reduction but first and second taxes) for large firms (above 50)	0.404%	-0.005%	0.410%
9.	Excess entry by small firms (percent increase in number of firms)	5.602%	0.279%	5.280%
10.	Change in size of small firms (below 9)	6.705%	0.336%	6.367%
11.	Change in size of medium firms (between 10 and 49)	6.394%	0.0250%	6.367%
12.	Change in size of large firms (above 50)	-2.022%	0.0250%	-2.048%
13.	Annual welfare loss (as a percentage of GDP):			
	a. First implicit tax (firms having more than 10 workers)	0.055%	0.055%	-
	b. Second implicit tax (firms having more than 50 workers)	0.932%	-	0.932%
	b. Output loss	0.013%	4.648e-04%	0.012%
	c. Total (Implicit Taxes + Output loss)	1.000%	0.056%	0.944%
14.	Winners and losers:			
	a. Change in expected wage for those who remain in labor force	-1.341%	-0.067%	-1.273%
	b. Average gain by entering entrepreneurs of small firms	1.999%	0.101%	1.899%
	c. Average profit gain by small unconstrained firms	5.364%	0.269%	
	d. Average profit gain by firms constrained at 9	5.103%	0.008%	5.094%
	e. Average profit gain by medium firms (10 to 48)	4.863%		
	f. Average profit gain by firms constrained at 49	4.415%	-0.027%	4.491%
	g. Change in profit for large firms (above 50)	-0.903%		-0.912%

Panel B: Partially Rigid Wages

(Regulated Economies - Same Unregulated Economy) Variable		(1) 2 thresholds	(2) Threshold at 10 only	(3) Threshold at 50 only
1.	Unemployment rate	1.989%	0.101%	1.889%
	Implied wage flexibility	**70%**	**70%**	**70%**
2.	Percentage of firms avoiding the first regulation, δ_1	5.653%	5.743%	-
3.	Percentage of firms avoiding the second regulation, δ_2	2.072%	-	2.079%
4.	Percentage of firms paying the first regulation	44.529%	45.234%	
5.	Percentage of firms paying the second regulation (on top of the first)	10.126%	-	10.160%
6.	Change in labor costs (wage reduction) for small firms (below 9)	-0.938%	-0.047%	-0.891%
7.	Change in labor costs (wage reduction but first tax) for medium firms (between 10 and 49)	-0.876%	0.015%	-0.891%
8.	Change in labor costs (wage reduction but first and second taxes) for large firms (above 50)	0.807%	0.015%	0.792%
9.	Excess entry by small firms (percent increase in number of firms)	5.556%	0.279%	5.275%
10.	Change in size of small firms (below 9)	4.696%	0.235%	4.454%
11.	Change in size of medium firms (between 10 and 49)	4.379%	-0.076%	4.454%
12.	Change in size of large firms (above 50)	-4.037%	-0.076%	-3.961%
13.	Annual welfare loss (as a percentage of GDP):			
	a. First implicit tax (firms having more than 10 workers)	0.055%	0.055%	-
	b. Second implicit tax (firms having more than 50 workers)	0.931%	-	0.915%
	b. Output loss	1.620%	0.081%	1.539%
	c. Total (Implicit Taxes + Output loss)	2.606%	0.137%	2.469%
14.	Winners and losers:			
	a. Change in expected wage for those who remain in labor force	-2.947%	-0.148%	2.798%
	b. Average gain by entering entrepreneurs of small firms	0.390%	0.020%	0.372%
	c. Average profit gain by small unconstrained firms	3.752%	0.188%	
	d. Average profit gain by firms constrained at 9	3.491%	-0.073%	3.563%
	e. Average profit gain by medium firms (10 to 48)	3.251%		
	f. Average profit gain by firms constrained at 49	2.803%	-0.313%	2.961%
	g. Change in profit for large firms (above 50)	-2.515%		-2.442%

Notes: The two panels are based on the estimates of Table A3 column (1), under the additional assumption that maximum firm size is 10,000. In column (1), we use our estimates of cost parameters for the two thresholds. In column (2), we cancel the second threshold while in column (3), we cancel the first threshold. Other notations follow the same logic as those in Table 3 (the baseline case with one threshold). In particular, column 1 in panel (B) replicates the calibration exercise of column 2 in Table 3.

Table A5: Estimates Obtained Using Alternative Datasets and Definitions of Employment

Dataset	(1) FICUS, 2002	(2) DADS workers "declared" on Dec. 31st, 2002	(3) DADS, "Full-time equivalent" As estimated by Insee, 2002
θ, scale parameter	0.8	0.8	0.8
N, threshold	49	49	49
β, power law	1.808	1.820	1.822
	(0.056)	(0.065)	(0.053)
n_r, upper employment threshold	57.950	58.361	63.053
	(2.267)	(1.011)	(1.417)
σ, variance of measurement error	0.105	0.130	0.239
	(0.035)	(0.019)	(0.035)
τ-1, implicit tax, variable cost	0.020	0.021	0.022
	(0.008)	(0.008)	(0.011)
F/w, implicit tax, fixed cost	-0.806	-0.860	-0.732
	(0.311)	(0.365)	(0.492)
Mean (Median) # of employees	55.4 (24)	54.4 (24)	54.4 (24)
Observations (firms)	40,637	36,576	36,141
Ln Likelihood	-181,940	-162,970	-160,702.

Notes: Parameters estimated by ML with standard errors below in parentheses (clustered at the industry, four digit level). Estimation is on population of French manufacturing firms with 10 to 1,000 employees, in the year 2002 (rather than 2000 in baseline because the proxy for full time equivalents in the DADS files is only available from 2002 onwards).

Table A6: Values of Parameters used for Simulation in Dynamic Extension

Parameter	Mnemonic	Value	Rationale
TFP AR(1) process	ρ_A	0.9	Cooper and Haltiwanger (2006)
Capital share of gross output	β_k	1/6	FICUS
Labor share of gross output	β_n	1/3	FICUS
Material share of gross output	β_m	1/2	FICUS
Returns to scale	θ	0.8	Basu and Fernald (1997) and main text
Adjustment cost parameter for capital	γ_k	0.150	Bloom et al. (2014)
Adjustment cost parameter for labor	γ_n	0.075	Half of capital adjustment cost
Discount Factor	ϕ	0.91	Based on real interest rates
Capital depreciation rate	δ_k	10%	Bond and Van Reenen (2007)
Sunk cost of entry	S	90%	Bloom et al (2014)

Notes: Baseline values of parameters used in the numerical simulations across all model runs. Fixed costs of production normalized to 100. We use τ = 1.02 and F/w = -0.94 as estimated in the paper.