"Firm Size Distortions and the Productivity Distribution: Evidence from France"
A Discussion and some Extensions

by Luis Garicano, Claire Lelarge and John Van Reenen

(First version: 28/05/2022, this version: 11/08/2023)

I. Introduction

We have had many interactions with various colleagues about our analysis of the impact of firm size contingent regulations in France (henceforth, "GLVR").[1] These discussions have often provided a better understanding of the work. This notes tries to compile some of these discussions and our responses, as some issues are often raised, particularly with respect to the data, and may be of interest for future research on related topics or methodologies.

In Section II we discuss some issues around the data, in Section III the econometric methodology and draw some conclusions in Section IV.

II. Misreporting of employment by firms

II.1 Differences between DADS and FICUS

One aspect that is raised repeatedly is that the firm size distributions which emerge from alternative French datasets seem to be different. Fortunately, this problem will soon disappear, since the concept of employment has been clarified and unified in the recent PACTE law[2], while a new administrative information system has been constructed as the unique source of information capturing the new legal concept of employment (DSN).

[1]Garicano, Luis, Claire Lelarge, and John Van Reenen. "Firm size distortions and the productivity distribution: Evidence from France." *American Economic Review* 106, no. 11 (2016): 3439-3479.

[2] Since January 1, 2020, the PACTE law has grouped or raised the social thresholds and reduced the legal obligations they represent for employers. These are now mainly grouped around three thresholds (11, 50, 250 employees) and various obligations have been reduced (such as the establishment of internal regulations, contributions to the National Housing Fund and the provision of and the provision of catering facilities). The calculation of headcount, which determines whether or not the social thresholds are exceeded, is now harmonized (the "social security" headcount): For a given year, it corresponds to the average number of employees for each month of the previous calendar year. The law also provides that a workforce threshold is considered to have been crossed only when it has been reached for five consecutive years. On the other hand, if a workforce threshold is crossed downwards during a year, the employer is immediately exempt from the obligations in question. With the implementation of the Nominative Social Declaration (DSN), the distribution of "social security" headcount is available since 2018.
https://www.strategie.gouv.fr/sites/strategie.gouv.fr/files/atoms/files/fs-2021-rapport-comite_suivi_et_evaluation_loi_pacte-septembre_0.pdf
https://www.legifrance.gouv.fr/loda/id/JORFTEXT000038496102/

This is fortunate, because it should put an end to (in our view) the rather sterile discussions about why the two main sources of data for employment analysis is France - DADS and FICUS - are different.

In short, our takeaways are the following:

- The concepts of employment in the DADS and in FICUS are different. In FICUS, it is recorded as an average of end-of quarter headcount.[3] In contrast, DADS is the headcount at the end of the calender year,[4] which can clearly be different. The job-level version of DADS also contains a comprehensive list of all contracts, with start date, end date and hours. In principle, it is possible to try and reconstruct the FICUS job concept in DADS, but in practice, hours are missing for around 10-12% of the records for the period of interest. This makes it hard to compare the two data sources precisely.
- The national statistical institute in France (INSEE) imputes hours according to the firms' industry affiliation, size bin and worker occupation. This delivers an approximation of full-time equivalent (FTE) employment at the job level.[5]

In our 2016 paper, we took the pragmatic approach of estimating our model on both datasets, obtaining very similar results (see Figure 6 and Online Appendix C4). Many researchers have also implemented this after our paper and also obtained similar results on alternative datasets such as DIANE (e.g. Poutvaara et al., 2015).

[3] « Effectif moyen du personnel: Il s'agit de l'ensemble des personnes titulaires d'un contrat de travail, rémunérées directement par l'entreprise. L'effectif moyen est égal à la moyenne arithmétique des effectifs à la fin de chacun des trimestres de l'exercice comptable. Pour le calcul du nombre d'apprentis il est fait abstraction de ceux liés à l'entreprise par un contrat d'apprentissage établi dans les conditions prévues à l'article L.177 du livre 1er du Code du travail. Les handicapés à retenir sont ceux reconnus comme tels par la Commission départementale technique d'orientation et de reclassement professionnel (COTOREP) ».

[4] « Nombre de salariés inscrits dans l'établissement au 31 décembre de l'année; effectif déclaré par l''entreprise ».

[5] Askenazy, Breda and Pecheu (ABP, 2022) aim to reconstruct from the job level information and imputed hours an alternative FTE proxy of the variable made available by Insee. They replicate our analyses in GLVR an unpublished working paper and obtain estimates that are very similar to ours, which is reassuring as it suggests that our quantitative results do not depend on the precise measurement of employment. Whether their proxy is better than what we used is debatable, since over the period, different sized-based regulations were based on different concepts of employment. To be precise, ABP Section 2 (p.13) shows that direct measures of (new) Works Councils and profit-sharing schemes are significantly higher after the 50 employee threshold in FICUS. This seems to demonstrate that the FICUS employment threshold is indeed salient for setting up these institutions. By contrast, ABP's constructed DADS measure of employment does not seem predictive. This suggests that the FICUS measure is more relevant, *at least for this regulation,* than the new variable they propose. Furthermore, their methodology remains limited by the fact that it is based on hours, that are imputed in about 10% of the records. Finally, ABP do not clearly explain how they reconstruct whether labor contracts are at 39 hours per week or 35 hours per week. This information is not in DADS. Despite the legal norm at 35 hours per week, firms have local agreements with workweeks at 39 hours. They say, "We call FTE size the number of hours worked in a firm during the 12 last months divided by the number of hours corresponding to a full-time job." But this number varies at the firm level and is not available in the DADS.

Figure I: Replication of GLVR Figure 6

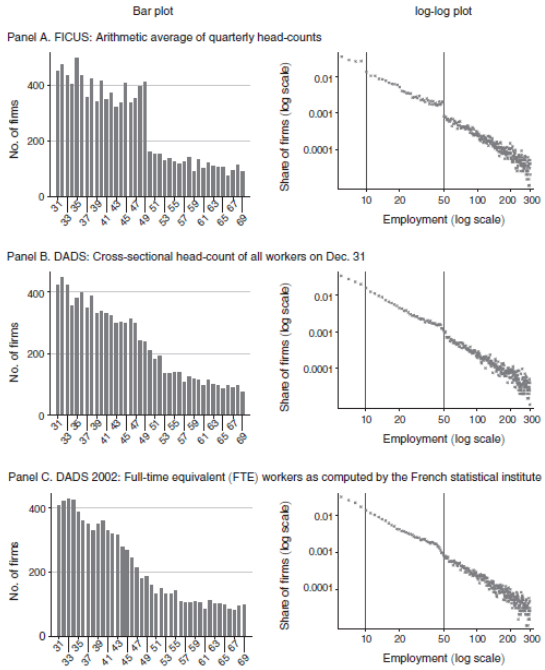


FIGURE 6. THE EFFECT ON THE MEASURED FIRM SIZE DISTRIBUTION USING ALTERNATIVE DATASETS (FICUS AND DADS) AND DEFINITIONS OF EMPLOYMENT

Notes: Data sources are from FICUS (corporate tax collection reported to fiscal administration) and DADS (payroll tax collection reported to social administration); see text and online Appendix C4 for more details on datasets. Panels A and B relate to the year 2000, and panel C relates to the year 2002 (as FTE calculation is unavailable in earlier years). All firms in manufacturing. FICUS employment is head-count at end of quarter averaged over four quarters (panel A). Panel B is head-count on December 31. FTE calculated in panel C by the French Statistical Agency INSEE over the available sample of workers (76 percent of all workers). This excludes jobs lasting less than 30 days or less than 1.5 hours worked per day, or those that were associated with (total) gross earnings lower than three times the monthly minimum wage (using as a benchmark, the seventy-fifth percentile of the series of hours per workers in each two-digit industries and size classes, and are bounded above by 1).

Figure II: Replication of GLVR Table A5

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).

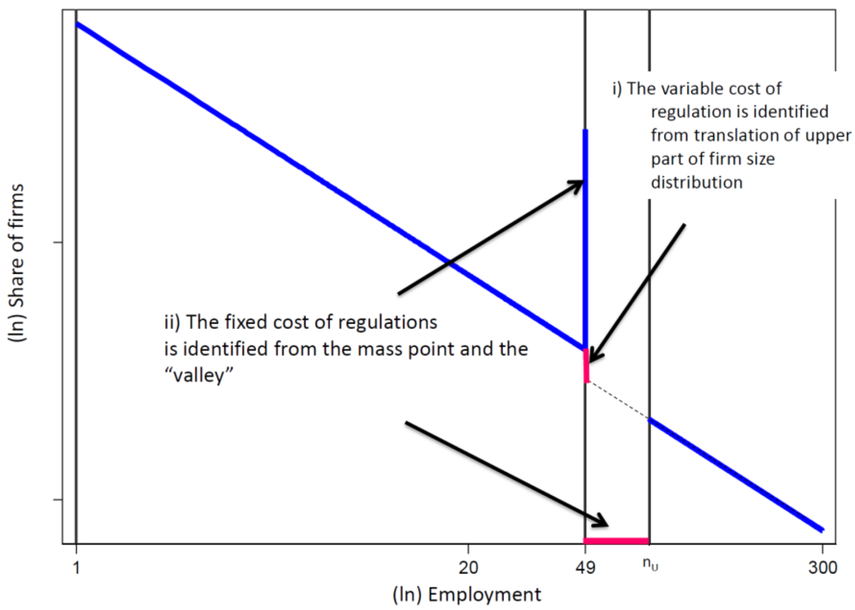
II.2 Our baseline methodology addresses measurement error

A strength of our method (compared to the bunching literature in Public Finance/Labor) is that we do not rely purely on identification at a discontinuity, so that our methodology is robust to many forms of measurement error.

As discussed in the paper, the main welfare losses we identify are *not* from the bunching before 50 employees, but rather from the downward shift in the firm size distribution for larger firms in the region of the 50 cut-off. Figure III aims to clarify this point.

The estimates of the two regulatory tax parameters can be recovered from three features of the distribution. First, the downward shift of the power law after 49 employees. Second, the bulge of firms just before the regulatory threshold at 50 employees, and third, the width of the valley in the size distribution between 49 employees and where the power law recovers at n_u. An implicit variable tax, τ, without any fixed cost results in a parallel shift of the power law. By contrast if there were only a fixed cost F of the regulation we should only see a bulge, and a valley, but no shift down in the power law after n_u. Hence, the existence of a downward shift in the firm size distribution after the regulatory threshold is powerful evidence of a variable cost component of the regulation over and above any fixed cost. In GLVR, we find that this is the main contributing factor to the welfare costs of the regulation.

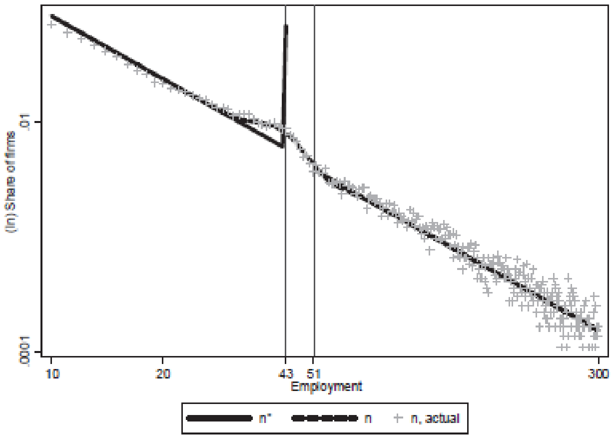
Figure III: Identifying the Impact of the regulation



Notes: *This is the power law of firm size in log-log space, using the same calibration values in GLVR Figure 4.*

In addition, we estimated a version of the model where we allowed the data to determine where the effective bunching point was. We replicate this analysis and report full details in Section III below. This method does not rely on knowing that the regulation binds specifically at the measured employment value of 50 but allows for measurement error / misreporting. This is illustrated in GLVR Figure A9 Panel C (replicated below in Figure IV). Empirically, we found that the bunching was 49 in FICUS (see Figure II), but came earlier in DADS, at 43 employees.

Figure IV: Replication of GLVR, Figure A9 Panel C



Notes: *DADS data from 2002 using FTE calculated from INSEE.*

II.3 Extension: A Model of Misreporting Firm Size in FICUS

Finally, an attractive feature of our approach is that our results and interpretation are robust to underreporting as long as non-compliance is costly. To show this, we augment our model with a new "regime" in which firms can cheat by under-reporting their number of workers, but with a risk of being detected and of being fined for that behavior.

To simplify, we assume that the probability p of detection is homogenous across firms and that the fine in case of being caught lying is increasing in how much greater the firm is over the 50 threshold (or that these two assumptions are reasonable approximations). We also assume that the largest firms will certainly get caught if they lie (if you have over 5,000 workers, you cannot realistically pretend you have 50). Specifically, with probability $(1-p)$ the firm gets away with misreporting, while with probability p it pays a variable fine τ_{nc} (so the more you lie about size, the more you pay) plus a fixed fine f_{nc}.[6] This implies that the profit function generalizes Equation (1) in GLVR to be:

[6] At the cost of some more notation, it is relatively easy to extend the model so that the probability of getting caught is also increasing in size. For example, assume that the probability of being detected is increasing in a firm's size and approximate this with a first order (linear) polynomial expansion in the neighborhood of N: $p_i \approx u.n_i + v$. The expected profit in case of non-compliance is thus:

$$\begin{aligned} \pi^{nc}(\alpha) &\approx \alpha.n^{\theta} - w.n - \underbrace{(u.n + v)}_{p}.F \\ &\approx \alpha.n^{\theta} - w.\underbrace{\left(1 + \frac{u}{w}\right)}_{\tau_{nc}}.n - \underbrace{v.F}_{F_{nc}} \end{aligned}$$

$$\pi(\alpha) = \max_n \begin{cases} \alpha . n^\theta - w.n & \text{if} \quad n \leq N \\ \alpha . n^\theta \underbrace{- w.n - p.(\tau_{nc} - 1).w.n}_{-w.\tau_{nc}.n} - \underbrace{p.f_{nc}}_{F_{nc}} & \text{if} \quad n > N \text{ and non} - \text{compliance} \\ \alpha . n^\theta - w.\tau_c.n - F_c & \text{if} \quad n > N \text{ and compliance} \end{cases}$$

where n corresponds here to the *true* employment of the considered firm (that we do not observe in the data). There are four possible parameter configurations as shown in Table I below, only one of which is interesting (highlighted in yellow).

Table I: Parameters defining regimes in the Costly Compliance Model

$\tau_{nc} < \tau_c$	$F_{nc} \leq F_c$	No firm would comply: Impossible
$\tau_{nc} < \tau_c$	$F_{nc} > F_c$	The large firms would not comply, but the intermediate would: Impossible.
$\tau_{nc} \geq \tau_c$	$F_{nc} > F_c$	All firms would comply
$\tau_{nc} \geq \tau_c$	$F_{nc} \leq F_c$	This is the interesting case

Assuming that the last case applies, this implies that the mapping between productivities and true (but unobserved) optimal employment becomes:

$$n(\alpha) = \begin{cases} \left(\frac{\theta}{w}\right)^{\frac{1}{1-\theta}} . \alpha^{\frac{1}{1-\theta}} & \text{for} \quad \alpha_{\min} \leq \alpha < \alpha_b \\ N & \text{for} \quad \alpha_b \leq \alpha < \alpha_{nc} \\ \left(\frac{\theta}{\tau_{nc}.w}\right)^{\frac{1}{1-\theta}} . \alpha^{\frac{1}{1-\theta}} & \text{for} \quad \alpha_{nc} \leq \alpha < \alpha_c \\ \left(\frac{\theta}{\tau_c.w}\right)^{\frac{1}{1-\theta}} . \alpha^{\frac{1}{1-\theta}} & \text{for} \quad \alpha_c \leq \alpha \end{cases}$$

The *true* firm size distribution becomes:

$$\pi(\alpha) = \max_n \begin{cases} \alpha . n^\theta - w.n & \text{if} \quad n \leq N \\ \alpha . n^\theta - w.\tau_{nc}.n - F_{nc} & \text{if} \quad n > N \text{ and non} - \text{compliance} \\ \alpha . n^\theta - w.\tau_c.n - F_c & \text{if} \quad n > N \text{ and compliance} \end{cases}$$

$$\chi^*(n) = \begin{cases} \left(\frac{1-\theta}{\theta}\right)^{1-\beta}.(\beta-1).n^{-\beta} & \text{if} \quad \frac{\theta}{1-\theta} \le n < N \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta}.\left(N^{1-\beta} - T_{nc}.n^{1-\beta}\right) & \text{if} \quad n = N \\ 0 & \text{if} \quad N < n < n_{nc} \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta}.(\beta-1).T_{nc}.n^{-\beta} & \text{if} \quad n_{nc} \le n < \left(\frac{\tau_c}{\tau_{nc}}\right)^{\frac{1}{1-\theta}}.n_c \\ 0 & \text{if} \quad \left(\frac{\tau_c}{\tau_{nc}}\right)^{\frac{1}{1-\theta}}.n_c \le n < n_c \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta}.(\beta-1).T_c.n^{-\beta} & \text{if} \quad n_c \le n \end{cases}$$

with $T_{nc} = \tau_{nc}^{-\frac{\beta-1}{1-\theta}}$ and similarly $T_c = \tau_c^{-\frac{\beta-1}{1-\theta}}$.

The declared firm size distribution under FICUS is therefore:

$$\chi(n) = \begin{cases} \left(\frac{1-\theta}{\theta}\right)^{1-\beta}.(\beta-1).n^{-\beta} & \text{if} \quad \frac{\theta}{1-\theta} \le n < N \\ \underbrace{\left(\frac{1-\theta}{\theta}\right)^{1-\beta}.\left(N^{1-\beta} - T_{nc}.n_{nc}{}^{1-\beta}\right)}_{\text{do not cheat}} + \underbrace{\left(\frac{1-\theta}{\theta}\right)^{1-\beta}.\left(T_{nc}.n_{nc}{}^{1-\beta} - T_c.n_c{}^{1-\beta}\right)}_{\text{do cheat}} & \text{if} \quad n = N \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta}.(\beta-1).T_c.n^{-\beta} & \text{if} \quad n_c \le n \end{cases}$$

This formula shows several useful things. First, our estimation procedure in GLVR is *unaffected* in this setting, only the *interpretation* of the mass point changes. Second, the estimate and interpretation of the variable cost of regulation, τ_c, is *unaffected*. Since this variable cost is the main source of the aggregate welfare loss, the insights of GLVR are not changed by this extension to allowing non-compliance. Third, only the mapping between the mass point at N and our estimate of the fixed cost is changed. Since only a fraction of the mass point should in principle be considered as a regulatory cost, this implies that our current estimates are an upper bound on the true fixed cost of interest. But this upper bound is very low (-0.94) and contributes only a trivial amount to the total welfare cost of the regulation. The reason is that removing the mass point would only affect around 100 firms (the excess mass at 49), for less than 10 workers (n_r = 59 in GLVR Table 1), so just 1,000 workers compared to around 3.6 million in the manufacturing sector (in 2000). By contrast, reducing the variable cost affects all large firms (total employment around 2.5 million, about 70% in aggregate). See

GLVR Table 3 and the discussion around it that gives a rigorous analysis of these welfare losses.

III. Identification of the costs of regulation from the "broken power law"

A frequent point of misunderstanding of our methodology is that some researchers relate it to a regression discontinuity design. Although we are using nonlinearities for identification, the comparison is misleading.

The point of the paper is to show that the assignment variable (firm size) is *not* continuous around the cutoff point: firms *do* change their employment levels (the running variable) in order to avoid the regulation. Hence, the McCrary Density Test fails.

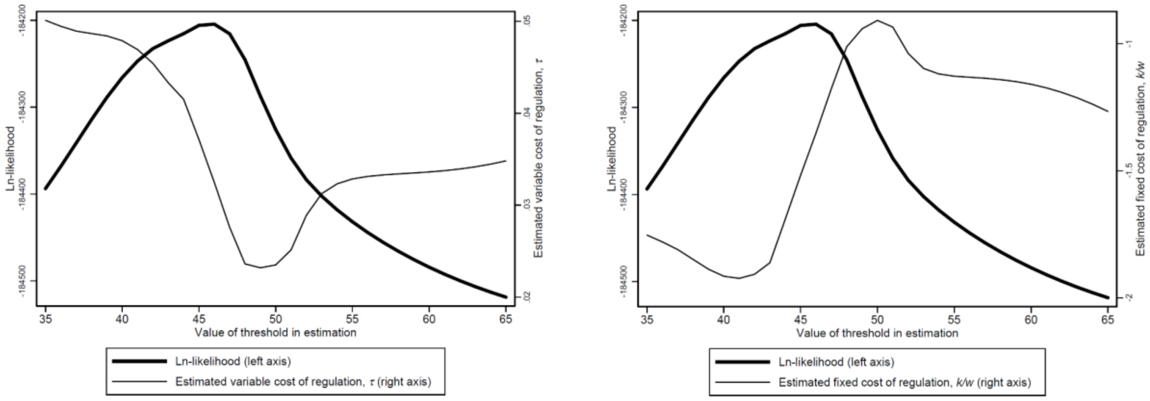
III.1 "Placebo thresholds": Maximum likelihood approach to rule out the wrong candidates

Nonetheless, there are some benefits to the idea of trying to detect the precise location of the cutoff point by using "placebo thresholds". The correct way of implementing this strategy is to interpret it in the maximum likelihood estimation framework.
More precisely, the model must be estimated for different candidate values of this cutoff, and the value which maximizes the log-likelihood is the estimate to be retained. The same holds for the other parameters (β, τ).

Note that the model can numerically be estimated for "wrong" thresholds, thus also delivering associated estimates of (β, τ). These should be discarded if the threshold does not maximize the log-likelihood.

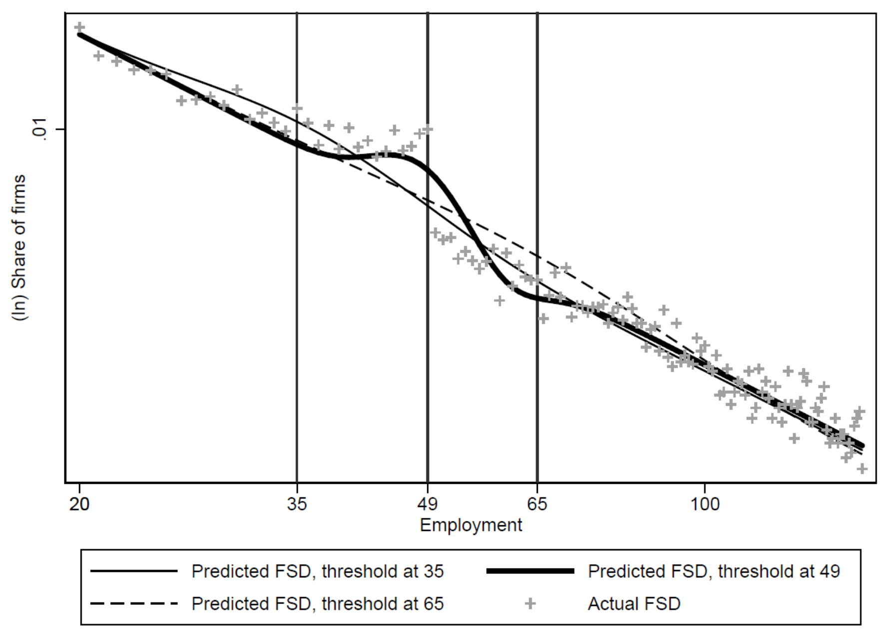
Figures V and VI illustrate the approach. The maximum of the likelihood is attained between 48 and 49, consistent with Figure A9 in GLVR's Online Appendix (Figure V). When using the wrong thresholds, the log-likelihood is low, the model is rejected by likelihood ratio tests, and the fit of the model is poor (Figure VI).

Figure V: Identifying the Impact of the regulation, using alternative values of the threshold between 35 and 65



Notes: *This shows the results of using FICUS data from the year 2000 and implements the model we estimate in Table 1 column (1) of GLVR, for regulatory thresholds varying between 35 and 65. The maximum of the likelihood is attained between 48 and 49, consistent with Figure A9 in GLVR's Online Appendix. The likelihood ratio test statistic comparing the model with the value of the regulatory threshold maximizing the likelihood and the model where the threshold is constrained at 35 attains the value of 378.438. Under the null hypothesis that the constrained model is valid, this statistic is distributed as a* $\chi^2_{(1)}$*, but the p-value is below 10E-10, which rejects the null hypothesis. Similarly, the likelihood ratio test statistic comparing the model with the value of the regulatory threshold maximizing the likelihood and the model where the threshold is constrained at 65 is 628.345 (p-value < 10E-10), which also rejects this constrained model.*

Figure VI: Implied Firm size distribution from different models: Wrong thresholds deliver a poor fit



Notes: *These are the predicted firm size distributions from models with different assumptions over the thresholds. GLVR used the actual threshold at 49 employees (bold lines) and we compare with (wrong) alternatives of 35 and 65. The actual data are in crosses.*

III.2 More flexible frameworks.

Identification in our 2016 paper relies heavily on the comparison between actual firm size data and a counterfactual power law distribution with the implicit tax parameters estimated from the observed distortions. In our paper, the only distortions that we consider are: (i) the mass point at the threshold; (ii) the "valley" to the right of the threshold, and (iii) the downwards homogenous shift (translation) of the upper part of the firm size distribution.

It would be easy to generalize the paper in this respect if different distortions are suspected. For example, it is worth noting that a richer (polynomial) specification of the regulation costs would naturally generate "curvature" in the power law. Generalizing GLVR Equation (1) in the paper, we obtain:

$$\pi(\alpha) = \begin{cases} \alpha . n^{\theta} - w . n & \text{if} \quad n \leq N \\ \alpha . n^{\theta} - w . \left[\sum_{k=0}^{N} \tau_k . n^k\right] & \text{if} \quad n > N \end{cases}$$

which implies that GLVR Equation (8) becomes:

$$\chi^*(n) = \begin{cases} \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . (\beta - 1) . n^{-\beta} & \text{if} \quad n < N \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} . \left[N^{1-\beta} - \left(\sum k . \tau_k . n_r^{k-\theta}\right)^{\frac{1-\beta}{1-\theta}}\right] & \text{if} \quad n = N \\ 0 & \text{if} \quad N < n < n_r \\ \left(\frac{1-\theta}{\theta}\right)^{1-\beta} \frac{\beta-1}{1-\theta} \left(\sum k . (k-\theta) . \tau_k . n^{k-\theta-1}\right) \left(\sum k . \tau_k . n^{k-\theta}\right)^{\frac{\theta-\beta}{1-\theta}} & \text{if} \quad n_r \leq n \end{cases}$$

Our intuition is that a "quadratic France" would reveal further distortions of this type.

IV. Conclusions

We are pleased that GLVR stimulated much discussion amongst academics, the media and policy makers. This note tries to clarify some of the issues that have been raised regarding the data, methodology and empirical results. Certainly, there is room for further work investigating the phenomenon – as well as many more aspects of regulation and economic performance.

References

Askenazy, Philippe, Thomas Breda, and Vladimir Pecheu (2022) “Under-Reporting of Firm Size Around Size-Dependent Regulation Thresholds: Evidence from France.” AMSE Working Paper 2211.

Garicano, Luis, Claire Lelarge, and John Van Reenen. (2016) "Firm size distortions and the productivity distribution: Evidence from France." *American Economic Review* 106, no. 11 (2016): 3439-3479.

Poutvaara, Panu, Marcus Drometer, Romuald Méango, Till Nikolka, Daniel Leithold, Katrin Oesingmann, Sabine Rumscheidt, Daniela Wech (2015) “Effet sur l’emploi des seuils sociaux en France et en Allemagne” IFO Center for International Institutional Comparisons and Migration Research. https://www.senat.fr/fileadmin/Fichiers/Images/delegation/entreprise/Effet-emploi-seuils_sociaux_Ifo_Francais-V2.pdf