Inferring Capital-Labor Substitution from Firm-level Distortions

Flavien Moreau, UCLA*†

January 6, 2019

[Please click here to read the most recent version]

Abstract

I propose a novel identification strategy to estimate the elasticity of substitution between capital and labor at the firm level, taking advantage of a set of size-dependent policies. These policies create a break in the size distribution of French firms by employment. To the extent that firms around the threshold distort their capital-labor ratio, the elasticity of substitution can be recovered from firm-level distorted factor ratios and the amount of bunching. On the other hand, because firms have an incentive to under-report their size, part of the apparent distortions is only due to evasion. I derive closed-form formulas that link the elasticity to observed distortions at the firm level. I then aggregate these micro elasticities to analyze capital-labor substitution in the aggregate economy. I find that firm-level distortions are small, and therefore that the micro- elasticities are small and close to 0.1. I obtain an aggregate elasticity of 0.3, about half as small as the values usually assumed. Moreover, the discrepancies between self-reported employment and a measure of employment made by the administration provide a direct evidence that evasion accounts for a large fraction of the observed distortions.

JEL Codes: E25, E60, L11, L51

Keywords: Elasticity of substitution, Production function, Firm-level datasets

*Department of Economics, UCLA. Email: flavienm@ucla.edu

†I am indebted to Andy Atkeson, Adriana Lleras-Muney and François Geerolf for guidance and support throughout this project. I also thank John Asker, Jonathan Vogel, and David Baqaee for helpful comments and support at various stages of this project. I am thankful to Sciences Po Paris, where some of this work was conducted, to UCLA for funding, and to the Californian Center for Population Research for hospitality. I also want to thank employees at Ministry of Finance and at the French National Statistical Agency, who have helped me better understand the data. Finally, for providing access to admnistrative data, I am grateful to the Centre d'accès sécurisé aux données (CASD), which is supported by a public grant overseen by the French National Research Agency (ANR), as part of the « Investissements d'avenir » program (reference : ANR-10-EQPX-17). All errors are my own.

"Machinery and labour are in constant competition, and the former can frequently not be employed until labour rises. (...) The demand for labour will continue to increase with an increase of capital, but not in proportion to its increase; the ratio will necessarily be a diminishing ratio."

Ricardo, 1817, *On The Principles of Political Economy and Taxation*. (Ch. 31 On Machinery)

1 Introduction

Quantitative assessments of the phenomenon described by Ricardo's rely on a measure of the elasticity of substitution between capital and labor. This parameter plays a central role in macroeconomic analysis. This elasticity not only governs the distribution of the labor share, but also determines the impact of investment and innovation on long-run growth, or the incidence of corporate taxation. While the bulk of the evidence suggests that this aggregate elasticity ranges around $0.6 - 0.7$, a firm consensus remains elusive.[1]Whether this elasticity is greater or smaller than one, however, entails completely opposite economic implications.[2] For instance, in a widely cited article, Karabarbounis and Neiman (2014) estimate an aggregate elasticity of 1.25, with which they account for the reduction of the labor share with a decrease of capital prices. With an elasticity smaller than 1, a decrease of capital prices would have *increased* the labor share.

This paper proposes the first quasi-natural experimental design to estimate this key elasticity, using the distortions that several size-dependent regulatory thresholds create on the size distribution of French firms' and on their choices of inputs.[3] These distortions arise from firms' reactions to a combination of administrative and organizational regulations that become mandatory once a firm reaches a certain number of employees. I consider three main thresholds, standing respectively at 10, 20, and 50 employees. Standard profit maximization predicts that in the face of discontinuities in their cost functions, some firms will prefer to bunch before the threshold instead of growing past it. Unless their production functions are Leontief, that is, with fixed factors, it is optimal for these firms to substitute capital for labor in order to avoid incurring the regulation costs. As a consequence, the extent to which this substitution actually occurs empirically informs us about the elasticity of substitution at the firm-level.

The key intuition of this paper is to jointly exploit two moments of the data – the amount of bunching in firms' labor size and the distortion in capital intensity – in order to recover the elasticity of substitution and the cost of the policy. A strong response to the regulation arises if either the policy costs are large, or if substitution between capital and labor is very elastic. This intuition is

[1] See Antras (2004), Chirinko (2008), or León-Ledesma et al. (2010) for recent surveys of elasticity estimates.

[2] This important point was recognized by Hicks (1932)in the first analytical study of the elasticity of substitution between capital and labor.

[3] In contemporaneous and independent research, Benzarti and Harju (2018) follow an approach similar to mine, but using a regulatory threshold in Finland which is based on the amount of capital depreciation. Interestingly, they find micro elasticities that are exactly 0.

first illustrated in a simple model where I derive simple closed-form formulas. I use these formulas, combined with sufficient statistics, to produce a first set of estimates for the elasticity. I then extend this baseline model in two directions, in order to recognize that several frictions interfere with firm's input decisions and that, therefore, managers have an incentive to evade the regulations. I develop a more comprehensive, structural model that incorporates two extra margins: a labor friction that rationalizes the presence of firms in the dominated region after the policy thresholds, and the possibility for managers to misreport the true number of employees. This structural approach delivers similarly low firm-level elasticities.

These two additional margins are considered in order to address two common concerns with bunching designs, which can potentially bias my estimates. First, the economic effect of the policy can be *overestimated* if the observed response is mainly due to evasion or misreporting.[4] Second, the structural elasticity can be *underestimated* if the agents fail to notice or understand the policy, or if their decisions are dampened by frictions.[5] I take these two concerns seriously and build a framework designed to tackle both evasion and salience. To detect evasion, I confront two administrative datasets and use the gap between the alternative measures of employment at the firm level to detect evasion.[6] The first measure is self-reported in the appendix of the firms' tax returns. The firm size distribution calculated from these measures shows very visible bunching patterns before the policy thresholds, which rules out the concern that these policies fall unnoticed by the firm managers.[7] I then merge this first dataset with payroll data that is extensively checked by the social security administration. The bunching on this second dataset is much more muted, suggesting that misreporting might be an important issue. Instead of focusing on just one of these measures, my estimation procedure uses the information contained in the joint distribution. In particular, the distribution of the gap between these two measures is informative of the *perceived* cost of the policies on which the firm's manager based her hiring and investment decisions.

In the data, misreporting is concentrated among the firms located before the threshold. Evasion drives a large part of the bunching response. The bunching mass, once evasion is properly taken into account, is 20% to 25% smaller than the naive bunching mass based on the self-reported counts. The firm-level elasticities I obtain lie around 0.1 and are on the lower spectrum of the estimates in the literature. The elasticities are the highest in Business Services and Wholesale Trade, sectors who tend to be less capital intensive. I repeat this exercise at the various thresholds and in the main sectors of the economy. I find very little evidence of capital labor substitution, either when looking directly at distortions in capital labor ratios or by inspecting investment behavior

[4]See Kleven (2016): "in contexts in which evasion and avoidance responses are feasible, observed bunching can be large in elasticity terms" and also Saez (2010), Chetty et al. (2011), Kleven and Waseem (2013), Bastani & Selin (2014).

[5]See Saez (2010), Kleven and Waseem (2013)

[6]Garicano et al. (2016) used the second source as a robustness check but they do not examine the joint distribution of the labor measures.

[7]There is no shortage of evidence that firm managers feel strongly about these size-based regulations. For at least half a century (see Gattaz 1979), policy recommendations about these regulations emanating from French major trade groups or their representatives read like a litany of calls for drastic simplification or outright suppression. A law voted in Fall 2018 in the French National Assembly trimmed some of the regulations around the 20-employee threshold, a partial granting of these demands that fell short of a more thorough simplification.

in the threshold in a dynamic extension of my baseline framework.The conclusion of this exercise is that micro-level production functions are quite rigid and look more like Leontief production functions. As a consequence, flexibility at the aggregate level and in the longer-run must arise from reallocation between firms or from the entry of firms with radically different technologies. To assess the potential magnitude of these channels, I build up the aggregate production function from the micro production functions as in Baqaee and Farhi (2018) and Oberfield and Raval (2014). The extent of the reallocation across firms is empirically limited by the amount of heterogeneity within sectors. I find that given the amount of heterogeneity measured in the data, the aggregate elasticity of the economy is about 0.3. This is lower than the consensus of recent estimate, but similar to the values found in Lucas (1969). As a consequence, an active extensive margin, i.e. the entry and exit of firms, is important in order to obtain the elasticity of subsitution in the range commonly assumed.

There are several advantages to my approach. First, the quasi natural experiment setting is not vulnerable to the usual endogeneity issues and bias towards one that characterize the approaches based on aggregate time-series or cross-sectional variations. Second, the size of the French economy is large enough so that I can estimate the elasticities for 2-digit sectors, across which technologies and substitution can in principle be different. Another advange of my setting is to be able to estimate the elasticity at several points in the size distribution of firms, corresponding to the different regulatory thresholds. The 50-employee regulatory threshold is the most visible in the firm size distribution and the most notorious. I also consider the 10- and 20-employee thresholds. While there is hardly any reaction around the 20-employee threshold, firms do respond to the 10-employee threshold, and I use the larger number of firms in this range of employment to sharpen the power of my estimates. Third, my approach is flexible and can be adapted to other contexts of size-dependent policies since I derive closed-form formulas for the elasticity of substitution under a variety of wedge specifications. This is particularly useful in some contexts where the regulatory cost comes from a collection of policies, contrary to the usual implementation of the bunching estimator in the public finance literature where the precise cost of the policy is known.

This work builds on a number of contributions. While growth theory has conventionally relied on Cobb-Douglas production functions, which have a unitary elasticity (Solow, 1957), most empirical studies reject this knife-edge assumption (see Chirinko, 2008 for an extensive survey) and conclude that the range $0.4 - 0.6$ is a reasonable ballpark for this elasticity. These studies rely on either a combination of strong structural assumptions coupled with time-series variations or try to make use of plausibly exogenous cross-sectional variations in factor prices. Oberfield and Raval (2014) follow the second approach and obtain an aggregate elasticity of 0.6 that is obtained from aggregating up micro elasticities. I depart from their work in several dimensions. First, instead of estimating plant-level elasticities from cross-sectional and spatial variations in the factor prices, I use very different identification assumptions: a quasi-natural setting and bunching techniques. My setting is therefore not subject to the endogeneity issues in Oberfield and Raval (2014) which would bias the elasticity estimate towards one. This bias is driven by the correlation in the time and spatial variations of the wages, with the location and input decisions of firms (e.g. more labor-intensive firms locating in regions with lower wages). Second, and perhaps for this reason, I found elasticities that are half as small as theirs. My finding that elasticities are close to zero is congruent with recent work in the empirical literature on production networks, such as Atalay, 2017, Boehm et al., 2016, and Barrot and Sauvagnat, 2016, who all find structural elasticities of substitution in production that are significantly below one. In a recent work Doraszelski and Jaumandreu (2018)also structurally estimate small elasticities of substitution in an unbalanced panel of Spanish firms. This finding cast doubt on the ability of the decline in the price of equipment to serve as an explanation for the fall of the labor share, which would require an elasticity higher than one. More precisely, my results suggest that this mechanism is unlikely to occur at the *intensive* margin, that is, for firms already operating in the economy. In contrast, this mechanism could be important on the *extensive* margin of capital-labor substitution, i.e. through patterns of entry and exit, if newly created firms are able to more flexibly choose their factor ratios. Recent evidence, however, points to a decline entry and exit rates (Decker et al. 2017), which limits how much traction could be gained through this channel.

The use of bunching designs to estimate tax elasticities has blossomed in the public finance literature following the seminal work of Saez (2010) who studied kinks in taxpayers marginal income tax rates and found that the actual response was much smaller than the one that could be expected from the theory.[8] Notches, which create a discontinuity in the decision sets, elicit larger responses.[9] A recurrent concern, though, is whether the observed responses are due to evasion or not. As emphasized by Saez (2010) for the case of kinks, bunching designs are potentially vulnerable to misreporting and other manipulation, especially in absence of third-party reporting. [10] By jointly exploiting the two datasets, I establish that evasion is behind a large part of the bunching observed. By contrast, taking the self-reported data at face value leads to an overestimation of the true elasticity, as it not only overstates the amount of bunching but also inflates the capital intensity distortion by artificially shrinking the denominator. Closest to the spirit of my paper is Chen et al. (2017)'s investigation of a Chinese R&D subsidy program. They use a notch design in order to uncover the returns on investment and are also confronted with potential misreporting issues, that they control for indirectly using proxy variables.

Assessing the aggregate impact of size-dependent policies has been an active research topic in macroeconomics. Restuccia and Rogerson (2008) developped one of the first framework to study the economy-wide impact of the misallocation generated by these policies. Inferring the aggregate effects of these distortions from the firm size distribution is a popular diagnostic tool in the macro-development literature (see Hsieh and Olken (2014) for a survey, or Bachas and Soto (2018) for a recent application). These aggregate implications can be either inferred from wedges taken as

[8] see Kleven (2016)for an extensive survey.

[9] See Chetty (2012): "Audit studies show that self-employment income is frequently misreported on tax returns because of the lack of double reporting. (...) Unlike kinks, notches in budget sets, where a $1 change in earnings leads to a discontinuous jump in consumption, generate substantial behavioral responses. (...) *Notches are therefore a promising source of variation for identification of structural elasticities.* "

[10] See for example Carrillo et al. 2017 and reference therein for a study of the effect of third-party reporting on evasion.

primitives or directly traced back to specific policies. I follow the first approach and capture the bundle of policies with wedges. While these concerns have been progressively integrated in the applied and public finance literature (see Kleven 2016 and references therein) they have not been fully recognized in the large macroeconomic literature on misallocation, which cites the bunching of French firms around these regulatory thresholds as a canonical example of distortionary, size-dependent policies with potentially significant welfare-reducing effects (see Garicano et al. 2016 for a welfare analysis, and Hopenhayn (2014) for a survey of the misallocation literature). Because of its prominence in the macroeconomic literature, clarifying the role of evasion in the French is important.

Finally, the 50-employee threshold in France has attracted the attention of both policymakers and researchers. Recently, Garicano et al. (2016) and Gourio and Roys (2014) have studied the 50-employee threshold in France to estimate its welfare impact. While these two papers are concerned with the potential misallocation of factors induced by these policies, the focus of this paper is rather different: I am looking at these policies only to the extend that they provide a quasi-natural setting to estimate the elasticity of substitution. Moreover, Garicano et al. (2016) and Gourio and Roys (2014) reach divergent conclusions as to the significance of the welfare effects. In a dynamic model with sunk costs, Gourio and Roys (2014) find the aggregate effects to be negligible while Garicano et al. (2016) in their baseline specification estimate a welfare cost amounting to 3.4 of GDP. These large effects might seem at odds with most of the studies reviewed in Hsieh and Olken (2014), who express skepticism about the potency of size distortions on medium-size firms. I show that the strong bunching pattern used in their estimation could be in large part an artefact of misreporting.[11] Therefore, acknowledging the role of misreporting could help reconcile their findings with Hsieh and Olken 2014's conclusions. Unlike these studies I do not assume that the distribution of firms' productivities is Pareto, and relax assumptions on the production function in order to analyze capital labor substitution.

The rest of the paper is organized as follows. Section 2 lays down the general framework to aggregate micro elasticities into macro elasticities. Section 3 develops an heterogeneous firms model and derives closed-form formulas that link the elasticity of substitution and the policy cost with the bunching mass and the capital distortion patterns. Section 4 describes the institutional background and the data. Section 5 details the estimation and presents the main results. Section 6 confirms these findings by estimating a fully-fledge structural framework that incorporates friction and an additional margin for evasion, and Section 7 concludes.

2 The Aggregate Production Function and Firms' Micro-Elasticities

How does the substitution between capital and labor in the economy relates to elasticities of substitution at the firm level? To answer this question, I build explicit foundations for the aggregate

[11] This view is confirmed by ongoing work by Askenazy and Breda, who carefully investigate the institutional analysis of the misreporting phenomenon. The possibility that misreporting could be important was raised in Ceci-Renaud and Chevalier (2010).

production function in this economy using the rich framework developped by Baqaee and Farhi (2018) and a series of follow-up articles. It is important to recognize that the micro and macro elasticities can in theory be completely disconnected, a point was first raised by Houthakker (1955). In his framework, a continuum of firms using only Leontief technologies ($\sigma = 0$) aggregates into a Cobb-Douglas production function ($\sigma = 1$). This stark result was later extended by Levhari (1968). Any aggregate CES production functions can a emerge from a suitable distribution of Leontief units.

Several mechanisms contributing to making the aggregate elasticity of substitution potentially larger than the micro-elasticity: (i) substitution across production units, (ii) substitution between sectors, and (iii) entry of new units with different capital intensities. These possibilities have long been recognized.[12]. The first two channels are explicitly present in my framework, which has a nested CES structure.This framework delivers a simple expression of the aggregate elasticity, that only depends on the firm-level elasticities and sufficient statistics that capture the heterogeneity and shares of each sectors. [13] Final demand is a CES aggregator over the sectoral goods

$$Y = \left(\sum_n \omega_n Y_n^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}$$

and the output in sector n is a CES over the output of all the firms operating in that sector

$$Y_n = \left(\sum_i \omega_{ni} Y_{ni}^{\frac{\epsilon_n-1}{\epsilon_n}} \right)^{\frac{\epsilon_n}{\epsilon_n-1}}$$

where ϵ_n is the elasticity of demand, Y_{ni} the output of firm i in sector n and D_{ni} demand weights. where η is the elasticity of demand across sectors. Under these demand assumptions, the industry level elasticity is a weighted average of the micro elasticity and the demand elasticity. The more heterogeneity there is between units, the more potential for substitution across units. I suppose that firms in this economy have access to a CES production whose elasticity is fixed at the sector level but operate at potentially different factor ratios. Each firm maximizes profit taking all prices as given and market for all goods and factors clear. Therefore the welfare theorems apply and I can use Shephard's lemma to express the aggregate elasticity using the equilibrium the labor cost ratios

$$\alpha_{ni} \equiv \frac{rK_{ni}}{rK_{ni} + wL_{ni}}$$

[12] Hicks (1932) conjectured that in a multisector economy, the aggregate elasticity of substitution is greater, (a) the greater the intra-sectoral elasticity, (b) the greater the difference in factor intensity among sectors, (c) the greater the inter-commodity substitution by consumers, and (d) the greater the technological innovation that enhances intra-sectoral and inter-commodity substitution (see Miyagiwa and Papageorgiou, 2007). Arrow et al. (1961) suggest that "Given systematic intersectoral differences in the elasticity of substitution and in income elasticities of demand, the possibility arises that the process of economic development itself might shift the over-all elasticity of substitution".

[13] Reallocation between firms occurs as long as firms are heterogenous in a meaningful way. While reminiscent of the reallocation mechanism at play in Houthakker (1955), Levhari (1968), and Satō (1975), reallocation in my framework do not rely on distributional assumptions about firm's production *capacities*. See Oberfield and Raval (2014) for more detailed comments regarding this important distinction.

The exact industry elasticity is a linear combination of within and between plant capital-labor substitution:

$$\sigma_n^{ind} = (1 - \chi_n)\,\sigma_n + \chi_n \epsilon_n$$

where $\chi_n = \sum_i \frac{(\alpha_{ni} - \alpha_n)^2}{\alpha_n (1-\alpha_n)} \theta_{ni}$ is a measure of the dispersion of the factor cost ratios within sectors . Similarly the aggregate elasticity is a linear combination of the sectoral elasticities and a dispersion

$$\sigma_{agg} = (1 - \chi)\,\hat{\sigma} + \chi \eta \tag{1}$$

where $\hat{\sigma} = \sum_n \frac{\alpha_n (1-\alpha_n)\theta_n}{\sum_{n'} \alpha_{n'} (1-\alpha_{n'})\theta_n} \sigma_n$ and $\chi = \sum_n \frac{(\alpha_n - \alpha)^2}{\alpha\,(1-\alpha)} \theta_n$

The simple framework presented in the previou section aggregates sector in a symmetric fashion, regardless of the actual links between these sectors. Aggregate effects can be traced down to micro-elasticities, as they propagate through the input-output structure of the economy. In my baseline framework, all the firm-level productions factors have only two factors of productions. There is therefore no room for input-output linkages as they are no intermediary goods. In the appendix, I consider the full framework with intermediary goods. I do not observe the basket of intermediary goods materials at the firm-level but I observe spendings on materials. To understand the propagation effects I can the input-output composition and scale it by the materials at the firm-level.

3 Micro Elasticities of Substitution and Size-Dependent policies

This section develops a new method to relate moments in the data such as the mass of bunchers and the capital intensity distortion to the structural parameters of interest and in particular the elasticity of substitution. I start with a simple span of control model à la Lucas (1978) in order to derive transparently the theoretical effects of the policies in the firm's input decision. The basic intuition is that some firms will respond to the exogenous change in the implicit price of labor by staying below the threshold and overaccumulating capital. In this stark model, a mass of firms should bunch right before the threshold and there should be a valley with no firms after the threshold. This prediction is at odds with the observed firm distribution in which the bump is sprayed from 46 to 49 and there is a non-zero mass of firms with 50, 51, etc. employees. Because the discontinuity in the firm-size distribution is not as sharp as the one generated by this simple theoretical setting, Section 5 introduces frictions in the input decisions. These frictions reflect the uncertainty faced by the firm owner due to the turnover of employees – such as unability to hire or unexpected quits –, or can be thought of as "optimization errors" (Chetty, 2012).

Baseline Model. I start with the simplest model to study size-depend policies with only one factor – labor – and then build up the general framework to show how to decompose the bunching response along the different margins. I first show how the bunching mass relates to the structural parameters in a simple, static span of control model as in Lucas (1978) where labor is the only production factor. In this baseline model, size-dependent regulations creates what the bunching literature calls a *notch* in the profit function, that is, a change in the *average* cost of production.[14] Because the very same logic applies at each policy threshold, I write the model for a generic threshold fixed at N employees. In the empirical exercise I study the firms' behavior around each threshold, that is for $N \in \{10, 20, 50\}$.

In the simplest model, I consider firm i in sector n that uses only one input - labor - and revenue productivity, z_{ni}, enters multiplicatively. The before-tax revenues of the firms are

$$p_{in} Y_{ni} = z_{ni}^{1-\nu} \cdot L_{ni}^{\nu} \tag{2}$$

The rest of this section analyzes the decisions of this generic firm and therefore the subscripts can be dropped to ease the exposition. The productivity index z follows an exogenous distribution with cumulative distribution function $H(.)$. I do not place any parametric assumption on $H(.)$ and only assume that it is continuous. In practice, I estimate $H(.)$ non-parametrically by local polynomials, as explained in details in Section 4.[15]

A large amount of policies come into force at the thresholds. These policies are a combination of organizational, accounting, and human resources regulations. For instance, firms need to set up a profit-sharing scheme and subsidize a wokplace committee with a budget equal to at least 0.3% of the payroll. The workers committee are also awarded paid hours for training. Only a handful are expressed in monetary terms, it is thus hard to put a precise price tag on many of these measures individually. Most of these policies scale with the size of the firms. Appendix B contains a detailed list of all these policies. Therefore, I capture the *combined* burden created by this bundle of regulations thus be captured as a hassle cost $\Delta\tau$ that acts as an increase in the baseline tax rate and reduces the net profit of the manager. A similar approach can be found in Aghion et al. 2017 who study in details tax regime decisions of small French entrepreneurs.[16]

For a policy threshold standing at N employees, the profit function of the firm is defined piece-wise as follows:

$$\Pi(L; z) = \begin{cases} (1-\tau)\left[z^{1-\nu} L^{\nu} - wL\right] & \text{if } L < N \\ (1-\tau+\Delta\tau)\left[z^{1-\nu} L^{\nu} - wL\right] & \text{if } L \geq N \end{cases} \tag{3}$$

The profit function takes the usual form before the threshold and additional costs due to the regulations decrease the profits above the employee threshold. Even in such a simple setup, the consequences on the firms decisions are not completely straightforward. While the bunching literature distinguishes between two types of bunching designs, kinks and notches (Kleven 2016), our case does not fall neatly into either category: while the labor wedge on its own creates a kink,

[14] As opposed to "kinks", that are changes in the marginal cost of production.

[15] It has been customary to assume that firms' productivities follow a Pareto distribution. While this approximation is a parcimonious description for the full range of the firm size distribution, it is in fact rejected by the data (see Appendix, Figure **??** for details). Maintining the Pareto assumption would create significant misspecification bias in our estimate of the elasticity.

[16] In contrast, Garicano et al. (2016) choose to combine a labor wedge and a fixed cost and find that this fixed cost to be negative. Gourio and Roys (2014) choose a sunk cost in a dynamic model.

i.e. a change in the *marginal* cost function, the fixed cost creates a notch – a change in the average cost. From the first order condition, it is immediate that the labor decision is defined piecewise. The labor demand schedule is

$$L(z) = \begin{cases} z\left(\frac{\nu}{w}\right)^{\frac{1}{1-\nu}} & \text{if } z \leq \underline{z_0} \\ N & \text{if } z \in \left[\underline{z_0}, \bar{z_0}\right] \\ z\left(\frac{\nu}{w}\right)^{\frac{1}{1-\nu}} & \text{if } z \geq \bar{z_0} \end{cases} \tag{4}$$

where $\underline{z_0} = N.\left(\frac{w}{\nu}\right)^{\frac{1}{1-\nu}}$ is the productivity index of the theoretical lowest buncher. Note that labor is simply a linear function of the unobserved productivity. The productivity index of the theoretical top buncher $\bar{z_0}$, is defined by a profit indifference condition. For this marginal firm, the profit is the same whether the firm's manager decides to bunch at N or to pay the regulatory cost and pick the unconstrained optimal employment above the threshold $L(\bar{z_0})$. To see this, notice that the equilibrium profit function is

$$\Pi(z) = \begin{cases} (1-\tau)(1-\nu)\, z\left(\frac{\nu}{w}\right)^{\frac{\nu}{1-\nu}} & \text{if } z \leq \underline{z_0} \\ (1-\tau)\left[z^{1-\nu}N^{\nu} - wN\right] & \text{if } z \in \left[\underline{z_0}, \bar{z_0}\right] \\ (1-\tau-\Delta\tau)(1-\nu)\, z\left(\frac{\nu}{w}\right)^{\frac{\nu}{1-\nu}} & \text{if } z \geq \bar{z_0} \end{cases} \tag{5}$$

Hence, the profit indifference condition for the top marginal buncher with productivity index z_0 is as follows

$$\Pi\left(L(\bar{z_0})\,; \bar{z_0}\right) = \Pi\left(N; \bar{z_0}\right)$$

Using the expression for profits, this indifference condition is

$$(1-\tau-\Delta\tau)\,\bar{z_0}\left(\frac{\nu}{w}\right)^{\frac{\nu}{1-\nu}} = (1-\tau)\left[\bar{z_0}^{1-\nu}N^{\nu} - wN\right] \tag{6}$$

which, after substituting out the threshold $N = \underline{z_0}\left(\frac{\nu}{w}\right)^{\frac{1}{1-\nu}}$ and rearranging, yields

$$\frac{\bar{z_0}}{\underline{z_0}} = \left[\frac{1-\tau}{1-\tau-\Delta\tau}\right]\left(\frac{1}{1-\nu}\right)\left[\left(\frac{\bar{z_0}}{\underline{z_0}}\right)^{1-\nu} - \nu\right] \tag{7}$$

Because of the notch in the profit function, the marginal top buncher hires $\bar{N} = L(\bar{z_0}) > N$ workers and leaves a gap in the firm size distribution. It is important to realize that the size of the theoretical hole is independent of any assumption on the unobserved productivity distribution. It is therefore useful to express equation (7) solely in terms of the size of the theoretical gap in the distribution with the cost of the regulation. In keeping with the bunching literature I write $\bar{N} = N + \Delta N$, where ΔN is the size of the hole after the threshold.

Proposition 1. *The size of the gap is strictly increasing with the policy cost and we have*

$$\frac{\Delta\tau}{1-\tau} = 1 - \left(\frac{1}{1-\nu}\right)\left[\left(1+\frac{\Delta N}{N}\right)^{-\nu} - \nu\left(1+\frac{\Delta N}{N}\right)^{-1}\right]. \tag{8}$$

Because N is known and $\bar{N}$ is in principle observable, the hassle cost can be inferred from this simple formula given a value for the returns to scale ν.[17]If the regulations creates no distortion at all, i.e. if $\Delta\tau = 0$ then there is no bunching, $\Delta N = 0$ and $\bar{z_0} = \underline{z_0}$. Finally, the labor schedule 4, together with the productivity distribution $H(.)$, defines the firm size distribution $S_0(.)$. If the unobserved productivity follows a distribution with $H(.)$, then the firm size density is given by

$$S_0(n) = \begin{cases} \left(\frac{w}{\nu}\right)^{\frac{1}{1-\nu}} h\left(n\left(\frac{w}{\nu}\right)^{\frac{1}{1-\nu}}\right) & \text{if } n \leq N-2 \\ H(\bar{z_0}) - H(\underline{z_0}) & n = N-1 \\ 0 & N \leq n \leq \bar{N} \\ \left(\frac{w}{\nu}\right)^{\frac{1}{1-\nu}} h\left(n\left(\frac{w}{\nu}\right)^{\frac{1}{1-\nu}}\right) & \text{if } n \geq \bar{N} \end{cases} \tag{9}$$

In practice, I estimate h non-parametrically from the observed firm size distribution. As mentioned previously, this basic model with a notch has stark predictions: a mass point just before the threshold and a hole with no firms at the right of the threshold. This predicted hole, as in many bunching designs with notches (e.g. Kleven and Waseem (2013)), is clearly counterfactual as can be seen from Figure . In Section 2.3 I show how the mass of firms in that region can be understood as the consequence of optimization frictions. Before that, I analyze how capital/labor substitution in the production function increases the amount of bunching and can be measured from distortions in input choices.[18]

Capital Labor Substitution. Now suppose that the production function F allows for capital-labor substitution. I show in this section that the amount of firms bunching at the threshold and the structural parameters are linked by a transparent formula. The revenue production function of a firm is given by

$$pY = z^{1-\nu} \cdot F(K,L)^{\nu} \tag{10}$$

where the productivity index z follows an exogenously given distribution, with a cdf $H(.)$ that will be estimated nonparametrically. The elasticity of substitution between capital and labor is defined by

$$\sigma(K,L) \equiv -\frac{d\ln {}^{K}\!/_{L}}{d\ln {}^{F_K(K,L)}\!/_{F_L(K,L)}}$$

[17]This formula is thus the counterpart to equation 5 in Kleven and Waseem 2013.

[18]Assuming that the underlying productivity follows a Pareto distribution leads to counterfactual results and substantial misspecification erros. If $H(.)$, then $h(x) \propto x^{-\eta-1}$ and $\ln h(x) \propto \ln x$. It follows that the logarithm of the cdf, $\ln S_0(n)$, would be proportional to $\ln n$ with the same slopes on both sides of the threshold. This prediction is rejected in the data (see Appendix Figure **??**). This observation leads us to reject the usual Pareto assumption to a more flexible approach that measure the bunching mass more accurately.

With only two production factors, it corresponds to the inverse of the slope of the isoquant of the production function. When the input markets are perfectly competitive, the factors of production are paid their marginal costs and the elasticity simply relates the relative demand of factors to their relative prices, that is

$$\sigma(K,L) = -\frac{d\ln {}^{K}\!/_{L}}{d\ln {}^{r}\!/_{w}} \tag{11}$$

In this article I focus on the case where firms belonging to the same sector have access to the same production function F. I suppose that F is defined by the following CES production function [19].

$$F(K,L) = \left[\alpha K^{\frac{\sigma-1}{\sigma}} + (1-\alpha) L^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

with $\alpha \in (0,1)$. In order to simplify the exposition, I have assumed here that firms belonging to the same sector operate with the same capital-labor ratio. When I take the model to the data, I let αbe a function of $z$such that the capital intensity in my model increases with size as in the data. The CES production function nests three special cases. The familiar Leontieff production obtains when $\sigma \to 0$. In this case there is no substitution possibilities between capital and labor and the production function writes $F(K,L) = \min\{\alpha K, (1-\alpha)L\}$. On the other hand, taking the limit $\sigma \to 1$ yields the Cobb Douglas production function $F(K,L) = K^{\alpha}L^{1-\alpha}$. Finally as $\sigma \to \infty$ the production function becomes linear and capital and labor are perfect substitutes.

Except at $L = N$, where the firms are bunching, the solution to the firm's maximization problem is interior and the first order conditions are satisfied. It follows that the capital-labor ratios implied by this model are constant and lie in an interval at $L = N$:

$$\begin{cases} \frac{K}{L} = k^* \equiv \left(\frac{\alpha}{1-\alpha}\frac{w}{r}\right)^{\sigma} & \text{if } L \neq N \\ \frac{K}{L} = [k^*, k^* + \Delta k] & \text{if } L = N \end{cases} \tag{12}$$

To simplify notations, let $f(k)$ denote the intensive-form production function, that is, $f(k) \equiv F(k,1)$. For all the firms that are not bunching at the threshold, the labor demand is linear in the productivity index

$$L(z) = z \cdot L_1 \tag{13}$$

where $L_1 = \left(\frac{\nu f(k^*)^{\nu}}{rk^*+w}\right)^{\frac{1}{1-\nu}}$. In the bunching region, where $L = N$,[20] the first order condition in capital is satisfied and yields

$$k(z)^{\frac{1}{\sigma}}\left[\alpha k(z)^{\frac{\sigma-1}{\sigma}} + 1 - \alpha\right]^{\frac{(1-\nu)\sigma-1}{\sigma-1}} = z^{1-\nu}\frac{r\nu N^{-\nu}}{\alpha} \tag{14}$$

This equation implicitly defines the equilibrium capital intensity, $k(z)$, as a function that is strictly

[19]More generally, I show in the appendix that the first order conditions provides a nonparamerical local identification of F.

[20]More precisely, it is equal to $N-1$ since this value is reported as an integer. I abstract from the integer problem in the stylized model but return to this issue in more details in the empirical section.

increasing in z.[21] Intuitively, with an elasticity close to 0, the shape of $k(z)$ is flat as the production function exhibit almost little substitution whereas in the Cobb Douglas case, the relationship is exact and we have $d\log k = \frac{1-\nu}{1-\alpha\nu} d\log z$. The relative productivity gap between the top and bottom marginal bunchers is $\frac{\Delta z}{z} = \frac{\Delta N}{N}$. Therefore, if the hole in the distribution has size 10 ($\frac{\Delta N}{N} = 0.2$), a distortion in the capital ratio of about 5% is expected in the Cobb-Douglas case. For the more general CES case I solve this equation numerically and provide a first order approximation.

Proposition 2. *The amount of distortion in the capital intensity for the firms in the bunching region increases with the elasticity of substitution if $\sigma < 1$. At the first order approximation, we have*

$$\frac{\Delta k}{k} \approx \sigma \frac{1-\nu}{1-\alpha\nu} \frac{\Delta z}{z}$$

This relationship captures the link between the unobserved productivity and the distortion. However, the productivity is not an object that is observed. Instead, the capital distortions in terms of observable revenues.

The profit functions are also linear in the productivity index, except in the bunching region (see Figure **??**), we have

$$\begin{cases} \Pi^*(z) = (1-\tau)\, z \cdot \Pi & z < \underline{z_1} \\ \Pi^*(z) = (1-\tau)\left[z^{1-\nu} N^{\nu} f(k(z))^{\nu} - rk(z)\, N - wN\right] & \text{if } z \in \left[\underline{z_1}, \bar{z}_1\right] \\ \Pi^*(z) = (1-\tau-\Delta\tau)\, z \cdot \Pi & z > \bar{z}_1 \end{cases} \tag{15}$$

where $\Pi \equiv \left(\frac{\nu f(k^*)^{\nu}}{rk^*+w}\right)^{\frac{\nu}{1-\nu}}$ is a constant that corresponds to the before-tax profit of a firm with productivity index $z = 1$. The index of the lowest buncher, $\underline{z_1}$, is determined by $L\left(\underline{z_1}\right) = \underline{z_1}\left(\frac{\nu f(k^*)^{\nu}}{rk^*+w}\right)^{\frac{1}{1-\nu}} = N$ and the unobserved productivity index $\bar{z}_1$ of the top buncher is determined by the following indifference condition

$$(1-\tau-\Delta\tau)\, \bar{z}_1 \cdot \Pi = (1-\tau)\left[\bar{z}_1^{1-\nu} N^{\nu} f(k(\bar{z}_1))^{\nu} - rk(\bar{z}_1)\, N - wN\right] \tag{16}$$

To obtain a formula similar to the one derived in the previous section, I take a first order approximation of equation16 . The capital intensity is increasing in the bunching region and I write $k(\bar{z}_1) = k^* + \Delta k$. Using the first order condition in capital taken at $z = \underline{z_1}$ and the fact that, in equilibrium, the before-tax profit is equal to a fraction $1-\nu$ of output, we have

$$\frac{\bar{z}_1}{\underline{z_1}} = \frac{1-\tau}{1-\tau-\Delta\tau} \frac{1}{1-\nu}\left[\left(\frac{\bar{z}_1}{\underline{z_1}}\right)^{1-\nu} - \nu + \frac{\Delta k}{k^*} s_K \left[\left(\frac{\bar{z}_1}{\underline{z_1}}\right)^{1-\nu} - 1\right]\right] \tag{17}$$

where s_K is the revenue share of capital. The bunching formula (17) is the counterpart of (7) when the manager can substitute between capital and labor. Firms in the bunching region have

[21] This monotonicity holds in general for a larger class of productions F and is not specific to the CES.

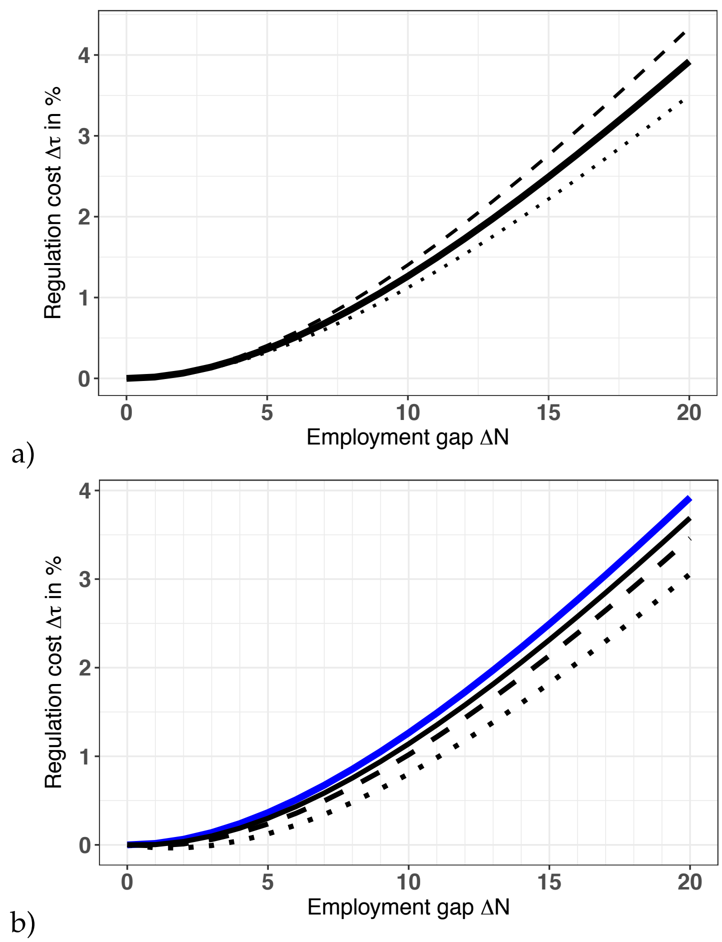
an incentive to boost their capital intensity instead of crossing the threshold. The extra profits generated by this increase in capital intensity is captured by the last term in the brackets. After substitution, this formula yields a relationship between the size of the hole after the threshold, the policy hassle cost and and the distortion.

Proposition 3. *The amount of bunching at the threshold increases with the policy cost and with the elasticity of subsitution, and the three are related by the following equation*

$$\frac{\Delta\tau}{1-\tau} = 1-\left(\frac{1}{1-\nu}\right)\left(\left(1+\frac{\Delta N}{N}\right)^{-\nu} - \nu\left(1+\frac{\Delta N}{N}\right)^{-1} + \frac{\Delta k}{k^*}s_K\left[\left(1+\frac{\Delta N}{N}\right)^{-\nu} - \left(1+\frac{\Delta N}{N}\right)^{-1}\right]\right)$$

where s_K is the capital share.

Figure 1: Relationship between elasticity, regulation costs, and the theoretical employment gap.



Panel a) plots the relationship between the regulation cost and the size of the hole in the firm size distribution in the baseline one-factor model. Panel b) represents the formula computed for $\sigma = 0$ (solid), $\sigma = 0.1$ (dash), and $\sigma = 0.3$ (dotted).

This formula generalizes the formula in Proposition 2 to the case where capital and labor are substituable. In the extreme case where there is no substitution at all, then the capital intensity stays the same in the bunching region and $\frac{\Delta k}{k^*} = 0$. In this case the formula reduces to the perfectly rigid case covered by Proposition 2. This formula maps three observables, the labor share, the capital intensity distortion and the gap in the size distribution ΔN into the policy cost. Different combinations of policycosts and elasticities can account for a given level of excess mass. Figure (1) plots the combinations of elasticities and policy costs that can jointly generate the observed excess mass.

The firm size density, $s_1(.)$ is a simple linear transformation of the productivity density $h(.)$, which directly follows from the labor demand 4. It is defined piecewise as follows

$$s_1(n) = \begin{cases} \frac{1}{L_0} h\left(\frac{n}{L_0}\right) & \text{if } n \leq N-2 \\ H\left(\frac{\bar{N}}{L_0}\right) - H\left(\frac{N}{L_0}\right) & n = N-1 \\ 0 & N \leq n < \bar{N} \\ \frac{1}{L_0} h\left(\frac{n}{L_0}\right) & \text{if } z \geq \bar{z}_1 \end{cases}$$

That the firm-size density adopts such a simple expression facilitates the empirical work, as a simple rescaling delivers the unobserved productivity distribution from the firm size distribution. This allows me recover the unobserved productivity distribution $h()$ non-parametrically instead of having to assume a strong functional form. Relaxing these functional forms is important as it removes the dependency of the policy cost and elasticity estimates on the unobserved productivity distribution. This problem is even more acute if one views these regulations as a combination of a fixed cost and a variable cost. In this case, these policy costs are not separetly identified and the identification ultimately rests on the choice of functional for for the unobserved productivity (see appendix B for a detailed proof).

The (gross) welfare cost of these policies in this model is transparent can be broken down into two pieces. The first part simply corresponds to the hassle cost $\Delta\tau$ multiplied by the before-tax profits. Given the decreasing returns, this cost corresponds to a fraction $1-\nu$ of output produced. The second component of the welfare cost is due to the local distortions at the threshold. By refusing to grow and distorting their capital intensity, firms uses resources inefficiently. This welfare cost can be viewed as an upper-bound on the effective welfare cost of these regulations since, in the model, the cost of the policies is captured exclusively by a hassle cost for the manager. To the extent that some of the regulations achieve their intended goal of providing workers with more safety (the compulsory safety committee), less job uncertainty (firing rules), or better incentives (the profit-sharing schemes), the *net* welfare cost might be smaller. Another concern is that parts of the payments made by the firm are compensated by a decrease in the wages. However, decreasing the employees' wages to implement a profit-sharing scheme is illegal[22] so this cancelling effect

[22] Ordonnance n° 86-1134 du 21 octobre 1986 relative à l'intéressement et à la participation des salariés aux résultats de l'entreprise et à l'actionnariat des salariés.

could only happen slowly overtime in terms of dampened wage growths.

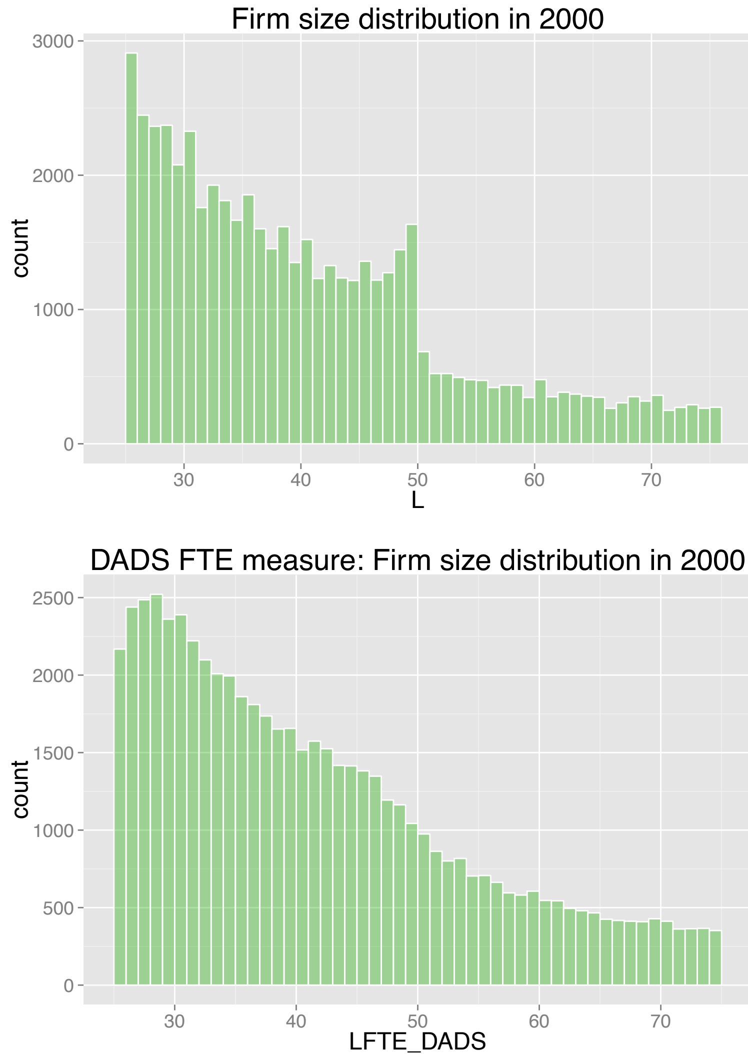
4 Data and Institutional Background

This section presents the data and the institutional background. I test the theory using two comprehensive administrative datasets. Merging these two datasets allows me to use several measures of employment and look for direct evidence of evasion. Evasion is by nature hard to measure, by using the two datasets jointly I can capture evasion behavior better precision than with the traditional approach of using some proxies.[23] Both datasets are generated by the administration and cover the universe of all French firms that have at least one employee . We merge these two datasets using the unique firm and establishment identifiers created by the administration. The first dataset is generated from the tax retunrs while the second one comes from payroll data. Since these datasets are not new to the macro/labor literature I focus only on the aspects that are the most relevant for this paper and refer to Bagger et al. (2014) for instance for details.

Regulations. A complete list of the regulations can be found in Ceci-Renaud and Chevalier (2010). The most important policies that are activated at the 50-employee threshold are: (i) an elected workplace committee must meet once every other month. Elections must be held (unless there is no candidate); (ii) information must be transmitted quarterly and yearly about the firm's financial situation to the workplace committee; (iii) external "representative" trade unions can designated a union representative; (iv) a workplace safety committee must be elected, with up to 3 members, which 2 paid hours every month to each member in order to run the committee; (v) compulsory annual negociation about wages, work duration, etc; (vi) if the firm increases dividends with respect to the previous two years then a bonus must also be paid to employees; (vii) a specific procedure ("Plan de Sauvegarde de l'Emploi") must be put in place a mass layoff is considered (where mass layoff is defined as the layoff of at least 10 employees over less than 30 days); (viii) workplace negociations must be held about prevention of accident injury and work penibility; (ix) a compulsory monthly reporting of job turnover must be sent to the administration

[23] For instance, Chen et al. (2017), Bachas and Soto (2018), or Hurst et al. 2014 look at distortions or breaks in variables potentially related to their main variable interest.

Figure 2: Firm Size Distribution around 50-employees in the two administrative datasets



The top panel shows the firm size distribution using the employment reported on the tax form. The bottom shows the firm size distribution using the measure checked by the administration. Sources: FICUS and DADS 2000

Tax return dataset. The first dataset is created by the Tax administration from mandatory corporate tax returns. The tax form[24] contains the detailed balance-sheet and income statement information: assets, broken into different types, as well as investment, liabilities, wage bill and materials. Some of these variables are retreated by the administration to reduce errors. The methodologies used to construct and cross-validate some of the variables have evolved over time and this is reflected in several versions of the datasets. I follow Garicano et al. (2016) and use FICUS, which is the most comprehensive one. I observe several measures of capital, gross and net values. The net value correspond to the gross value minus the accumulated depreciation, which is also observed.

[24]The official form is CERFA 2050-9 . It is the equivalent of the IRS Form 1120 for Corporate Income Tax Return.

Capital is broken down into several categories and it is in principle possible to look separately at tangible and intangible capital. Intangible capital represent on average 20% of the gross non financial assets.. It is notoriously difficult to measure capital precisely (see Becker et al. (2006) for an extensive discussion of the numerous challenges). The book value, which is most commonly available in micro data, can differ substantially from the economic value, the relevant concept for the production function. Available measures of capital therefore tend to exhibit substantial measurement error[25] and differ from the concept of capital used in production function. I use alternatively several measures. First, I use the net value of plant, property, and equipment capital, the closest counterpart in the French data of the variable called PPENT in Compustat[26], used in most U.S. studies. For consistency reasons, I focus on the period 2000-2007. My panel is not long enough to make it reliable to use a perpetual inventory method to construct a measure of capital. For some of the robustness checks, I use different measures of investment that distinguish different kinds of tangible and intangible investments. Some of these variables are only available in the BRN dataset, also generated by the administration from the tax returns, but ith a coverage that excludes self-employed companies that report under the simplified tax regime.

The FICUS dataset also contains the number of employee. More precisely, firms are asked to report the arithmetic average of the employees headcount at the end of each quarter. Importantly, this number is reported in the appendix of the corporate income tax form and, importantly, is *not* checked at all by the tax administration, which indeed warns of the "mediocre quality" of this measure. [27]This can partly be blamed on the official form[28] containing the explaination on how to fill up the tax return.

Social Security dataset. The second main dataset I use is a matched employer-employee dataset generated from Social Security data, called DADS (Donnees Annuelles des Declarations Sociales). Since it has been used several times in the macro labor, see for instance Bagger et al. (2014), I will only emphasize the points that are most important for the purpose of this paper. The data are based upon mandatory employer reports of the gross earnings of each employee subject to French payroll taxes. These taxes apply to all "declared" employees and to all self-employed persons, essentially all employed persons in the economy. The data comes from mandatory reports by firms to the administration that are used to compute various social security contributions as well as rights to unemployment benefits, health insurance, work accident benefits and retirement. It contains information about each job spells, the number of hours worked, the type of job, the net and gross salary etc. These can be aggregated at the establishment or the firm levels. The social security administration computes three employment measures at the firm level: 1) a Full-Time

[25]See Collard-Wexler and De Loecker (2016): "There is reason to believe that, more than any other input in the production process, there are severe errors in the recording of capital stock." See also Kim et al. (2016) for a two-step estimation method in presence of measurement error.

[26]PPENT stands for Property, Plant, and Equipment Net Total.

[27]"Il existe également des informations sur l'emploi dans les liasses fiscales, mais de médiocre qualité (beaucoup d'informations manquantes en particulier)" Béguin and Haag (2017), p.20.

[28]Form N° 2032-NOT-SD specifies how to fill in forms n° 2050 to 2059-G that constitute the corporate income tax returns for firms under the general regime.

Equivalent of the total number of employees par firms 2) a headcount on December 31, 3) an arithmetic average of the headcounts at the end of each quarter. As detailed below, the concept that is the closest to the one used in the regulations is the Full-Time Equivalent count. Notice that it is also the one that is the closest in spirit to the one that should enter the production function.

Legal workforce concepts. In the French Labor Code, the relevant concept for the workforce is the Full-Time Equivalent employment. Almost all the policies coming from the Labor Code hinge on the FTE measure. There can be, however, a slight ambiguity as to whether this FTE count must be made at the firm of at establishment level. Article L1111-2 of the French Labor Code lays down the definition of the workforce to be used for all the Labor Code regulations: (i) full-time employees with open-ended contracts count for one unit; (ii) employees on a fixed-duration contract, on intermittent contract, employees "lended" by another company that work in the company premises since at least a year, and temporary workers are to be counted in porportion of their presence over the last twelve months; (iii) part-time employees, regardless of their working contract, are counted in proportion of the number of hours written in their contracts divided by the legal or conventional worktime duration. There is one exception: employees lended by another company who replace an employee on maternal/paternal leave are not counted. The regulations contained in the other Codes mostly follow the same definition.

Do firms report labor accurately ? Evasion is a serious concern in many bunching designs, since evasion that can substantially bias the estimates. Direct evasion is by nature often hard to observe in administrative datasets. An advantage of my setting is that by combining two measures of employment coming from different administrative source, I am able to provide direct evidence that firms under-report their workforce around the thresholds. This evidence suggests that evasion drives the majority of the observable response. To make this point more rigorously I explore in details the discrepancies between the two labor measures.

I start by looking for traces of misreporting by comparing the two firm size distributions that can be computed from the two sources. If One reason to proceed that way is that the coverage of the datasets do not perfectly coincide. In this preliminary steps I do not throw out any firms and instead analyze the relative changes in the shape of the distribution. The first observation is that the two distributions look roughly similar. However they differ significantly by the amount of distortion around the threshold. To see this more clearly, I compute the log-ratios of the density based on self-reported employment divided by the density based on the measured FTE employment (Figure). A clear and characteristic spike appears right before the threshold. The density of firms reporting 49 employees is about 1.9 times the density of firms that have 49-employees according to the administration. The number of firms declaring 9 employees is 1.45 times larger. There is no small discernable bump around the 20 threshold but it is not significant. This is exactly the pattern we would expect if the incentives to misreports were large around the 50 and 10 threshold and weak around the 20 threshold. I repeat the same exercise in each 2-digit sector. Figure 9 shows the log ratios of the density for manufacturing , retail, wholesale and transportation and communication sectors . The same characteristic pattern is visible there.

Next I focus on the individual discrepancies by merging the two datasets using unique the firm identifiers and compute the difference between the two measures. I group firms by the amount of labor they report. Figure (10) reports the mean and the standard deviation of the gap. The significant bump in the gap just before the threshold is a clear evidence of misreporting. The gap between the two measures of labor could also be due to differences in the concept of workforce. On the tax form, firms are instructed to compute the arithmetic mean of the workforce at the end of each quarter, while on the other dataset, the Full Time Equivalent is computed. This is the relevant measure for the majorities o the policies according to the Labor Code.

There are several reasons, both in theory and in practice why the two measures do could not coincide. Counting ETP is not entirely straightforward. The methodology that the INSEE itself uses consists in several steps. Indeed, despite following their method carefully I am still left we some residual discrepancy. I thus use directly their measure. As mentioned above, the fiscal dataset contains an employee count. This count is the arithmetic average of the employees headcount at the end of each quarter. This number is reported in the appendix of the corporate tax form, and is not checked thoroughly by the tax administration.

Several reasons why the DADS measure is more reliable. First, while the FICUS headcounts are not verified by the administration (Béguin and Haag, 2017), the DADS measures are the subject of detailed verifications. The reason for it is that employee benefits and retirement claims are based on the information submitted by the firms. The original information is generally generated from the firm's accounting and payroll software. It seems difficult to under-declare the number of workers without harming directly one or several workers individually. On the contrary, nothing really prevents firms from misreporting there employment count on their corporate tax form. After receiving the raw data, the statistical agency performs several rounds of checks. This is therefore the highest quality measure of employment according to the French National Statistical Agency.[29]

Also note that the employment count reported in the fiscal source is rounded at the nearest integer. In contrast, using the social security data and aggregating at the firm level, I obtain non-integer numbers. I compute the gap before rounding but to in order to compute the firm size distributions I round these employment counts at the nearest integer. I do not measure significant rounding at multiples of 5 or 10. Figure 10 shows the gap between the two measures for each firm size. The self-reported number tends to be always larger than the Full-Time equivalent measure in the DADS.

Primary and secondary jobs. There is no such thing as a perfect measure of employment. In the social security data, the statistical agency computes for each worker a full-time equivalent .

[29] Béguin and Haag, 2017 strongly recommend the use of the DADS employment count instead of one in the fiscal data: "La source administrative la plus naturelle (car à la fois la plus exhaustive et la plus fiable) pour estimer les données d'emploi dans Ésane était donc la source DADS-U2" p.150. See also Ceci-Renaud and Chevalier (2010).

But this count excludes those job spells deemed "secondary" ("emplois annexes" in French). The dataset documentation contains the precise definition of these secondary jobs as welle as detailed explanations of the implementation procedure. Excluding secondary jobs is actually innocuous. While they account for about a fifth of the *observations* over the period they represent a negligeable share of *hours* worked.[30] A potential concern would be that these secondary jobs could be heavily concentrated in the firms bunching below the threshold. If that were the case then we would need to accomodate for a channel through which "secondary jobs" could substitute for "primary jobs", which could create a downward bias on our estimate.[31]The data shows that this is not the case. In Figure (13), I add up all the hours worked in secondary jobs and express them as a fraction of total hours worked. This fraction is stable around 1% and does not vary around the thresholds. Moreover, they are heavily concentrated in some specific activies among the service sectors such as: cleaning, restaurants, hotels, and various recreational and sports activities, associations.[32] In addition, some of these spells are created artificially either by the employer to register variable salary components such as bonuses.[33] The more important point is that the employment measure obtained by summing up full-time equivalent spells at the firm-level slightly underestimate the true employment. This is because the construction of FTE measure makes it top-censored at 1. [34] This underestimation

Compliance and Enforcement. Are these policies really binding? Do we have evidence that firms are compliant? Given the large number of policies, it is difficult to give a definitive answer. The enforcement of the Labor Code is the responsibility of the Corps des Inspecteurs du Travail (Workplace Inspectors Unit), a specialized group of civil servant. Its function is similar to the Office of Inspector General in the US Department of Labor. A public report is produced every year at the intention of the International Labor Organization. The impression that emerges is that while these controls are resented by the firm managers, the audits' frequency is low and monetary sanctions for failing to comply with these regulations are uncommon.

First, the average effective number of firms to be controlled by each controller is large. In the2001, they were on average $1,159$ establishments and $11,521$ employees per agent of the follow a rigid division of labor following the seniority of the public servant. Junior employees are called "controlers" and are in charge of establishments with less than 50 employees while the more senior

[30]In 2002 the share of secondary jobs was 24.4% but these jobs represent only 1.2% of the aggregate share of hours worked. Moreover, this share is constant across firms sizes, see appendix and Béguin and Haag (2017) p.154.

[31]Garicano et al. (2016) use that figure to argue in favor of using the employment variable in FICUS and not the one in the DADS : "unfortunately, full-time equivalent information in DADS misses almost one-quarter of jobs" p.3455.

[32]Temporary workers count towards the legal employment threshold, in proportion of their time spent in the company. However they are counted neither in the fiscal dataset or in the social security dataset. Temp workers have therefore no impact on the gap between these two measures. I am not aware of comprehensive firm-level data on temp workers, but according to surveys they typically amount to a few percents, See Havez (2018) for details on those sectors who employ a lot of temporary workers.

[33]For instance, starting from 2004, unemployment benefits received in year n appears as job spells if the individual is employed either in year n+1, or n-1 . This addition was made to allow researcher to track movements in and out of the labor force.

[34]"légère sous-estimation du volume d'emploi total tel qu'il est pris en compte dans la masse salariale" p.154 Béguin and Haag (2017)

"inspectors" are in charge of those above the threshold. Breaking up the workload according to this line, there were on average $1,667$ establishments and $8,964$ employees per controler and 110 establishments with $16,262$ salariés per inspector.

However, the regulations activated at the thresholds are only a tiny fraction of the numerous provisions in the Labor Code. Workplace inspectors must also investigate[35]: infraction to workplace safety and especially workplace injuries, illegal hirings of undocumented workers, the respect of work duration laws, timely payment of wages, compliance with collective wage bargaining, discriminations, sexual and moral harassment, human trafficking, smoking interdiction. Given the large scope of their duties, priorities are defined by the Minister in charge of Labor relations [36] and other potential violations – among which some of the regulations this study focuses on – receive far less attention.[37] For instance, the construction is over-represented. Because a status of limitation of three years applies to many of these infractions, in principle establishments must be controlled at least once every three years. No more than 2% of the cases opened lead to formal indicments or convictions. Most of the times, controlled firms are given a warning and ordered to comply. In theory, penalties for infractions to the Safety Code can go up to $3750.$ euros.[38]

For some of the regulations, direct evidence of compliance can be gathered from surveys run by the administration. I consider two of these surveys: one called REPONSE and one called ACEMO. The first one, REPONSE, is a survey conducted every five years on a sample of about 2% of the firms. The aim of this survey is to better understand the working conditions in the firms and the interactions between workers, workers' representative and management.[39] The survey is addressed to each of these three parties. From this survey, I can test whether some of these policies are followed. If a firm fails to put in place a profit-sharing scheme, employees can ask the juge to impose a default one specified by law.[40]

Finally, empirical studies have found that individuals can overestimate the cost of evasion (Andreoni et al. 1998). It is moreover difficult to estimate the cost of the penalties given that an infraction has been detected. Moreover, the great majority of investigations for financial fraud do not give rise to prosecution or end in confidential settlements with the administration.

Evasion is a often serious concern in bunching designs, and third-party sources are held more accurate (see Fack and Landais, 2016 or Carrillo et al. 2017). The true impact of regulatory distortions depends on the extend of the economic response, purged of the evasion behavior. In order to separate the two, I enriched the basic model with an evasion margin, allowing firms can divert resources to engage in misreporting and operate above the threshold without complying with

[35] Article L8112-2 of the Labor Code

[36] Article L8112-1 of the Labor Code

[37] In theory an inspector that notices the unlawful absence of a workplace safety committee is mandated to order the firm to institute one.

[38] Perhaps not surprisingly, the 50-employee threshold also serves for the division the labor among labor inspectors. It is a sensible division since different sets of policies apply below and above that threshold, however it is not particularly helpful to catch misreporting firms. Firms with 52 employees who misreport and declare 48 employees fall under the purview of a more junior controller.

[39] See Appendix for details

[40] see decision Cass. Soc. n° 03-03-10502, 13/09/2005

the regulations. This section provides evidence that many firms bunching around the threshold are in fact misreporting and then shows how to leverage the two datasets to infer the perceived cost of the regulation. Since firm owners have a incentive to stay below the threshold for tax purpose, they could be tempted to misreport the number of workers or hire undeclared workers, in which case the sharp discontinuity exagerates the true allocative distortion. To assess the extend of misreporting, we confront two datasets: a fiscal dataset where firm owners (or, for that purpose, their accountant) report the number of employees on their corporate tax form, and the one generated from payroll data, used to compute payroll taxes as well as entitlements (unemployment and retirement benefits).

Another reasonable concern raised in previous studies that attempted to estimate the impact of the threshold relates to the timing of the labor measures. In principle whether the labor variable in the Social Security data .[41] Payroll data employment figures are the ones coming from other sources and this is true as well in France (Ceci-Renaud and Chevalier, 2010). Failure to account for this misreporting channel leads to serious biases. For instance, in an unpublished working paper Smagghue (2014) estimates an elasticity of 0.6 based on the fiscal data. Similarly, Garicano et al. (2016)'s analysis is based on the tax data. What they interpret as investment per capita spikes before the regulatory threshold are actually a spurrious consequences of the denominator lowered by the under-reporting behavior of the firms before the threshold.

Missing data. I also verified that the reporting gap is not due to missing data. About 0.4% of all the firms in the tax dataset lack a firm identifier. For this reason we are unable to match them with their corresponding Social Security data and cannot ascertain whether they are misreporting. This identifiers are missing despite numerous procedures that the administrationto undertakes to check the data and retrieve partly missing or erroneous firm identifiers in a separately maintained exhaustive list of firms called "Fichier des Redevables Professionnels de la DGI"(Béguin and Haag 2017). I check that these missing observations have very little impact on our measures of firm size distribution. For instance, in 2004, there is exactly 1 firm with 49 that could not be identified, and 3 firms with 48 employees. Around the 10-employee threshold, the absolute numbers are respectively 168, 174, and 103 firms of $8, 9$, or 10 employees. Only at the 20 threshold do we observe some bunching before the threshold: 17 (145), 18 (252), 19 (456), 20 (25).

5 Estimation of Firm-level elasticities

This section details the estimation of the sufficient statistics needed to estimate the micro-elasticities and present the baseline results.

Bunching and Counterfactual distribution. I construct a counterfactual distribution by fitting a 6th order polynomial outside of the bunching region, following the standard approach in the

[41] "recalé sur la date de fin d'exercice et transformé, en ce qui concerne l'emploi en nombre de postes, en un concept identique à celui des liasses fiscales"

bunching literature (Saez, 2010). The magnitude of the bunching – the so-called "missing mass"– is then computed as the deviation from the polynomial fit in an interval around the threshold.

I proceed in a similar fashion and obtain B using by regressing the observed distribution on a polynomial of degree 6, and excluding the region around the threshold. This corresponds to running the following regression

$$h(n) \quad = \quad \sum_{p=1}^{6} a_p \cdot n^p + \quad \sum_{j=N-\delta^-}^{N+\delta^+} d_j \cdot 1_{\{n=j\}} \quad + \xi_n \tag{18}$$

where the dummies d_j are just here to absorb the firm counts in the region of the firm size distribution that is distorted by the policy. The counterfactual size distribution is is then estimated by

$$\hat{h}(n) = \sum_{p=1}^{6} \hat{a}_p \cdot n^p$$

Following the methodology that is now standard for the bunching literature (cf. Kleven, 2016), I set the lower bound of the bunching region, δ^- , at the point where the distribution starts increasing. To obtain the upper bound δ^+ I conduct an iterative procedure that ensures that the two distributions – the actual one and the counterfactual one – add up to the same number of firms in the interval considered. The upper bound is such that the displaced mass of firms before the threshold and after the threshold exactly compensates. Typically because the deviation is much larger before than after the threshold, this method implies an asymetry in the response region: the upper bound δ^+ is further from the threshold than the lower bound δ^- is. In practice, I cannot equalized the excess mass and the missing mass exactly perfectly because of the round numbers of employees. In practice, I pick the closest value and conduct robustness checks with the other value.

The detailed estimation goes as follow. First, I apply the standard bunching procedure to recover the firm size distribution, the bunching mass $\hat{B}$, the missing mass $\hat{M}$, and the boundaries of the excluded region δ^- and δ^+. As explained above, this step involves iterating until δ^+ is such that the bunching mass and the missing mass coincide, that is, such that the counterfactual firm distribution and the empirical firm distribution integrate to the same number of firms. For the nonparametric estimation of the firm size distribution, I select the best fit among the space of 6^{th} order polynomials as is standard in the bunching literature. This gives me the distribution of productivity up to a scale factor. I then compute the distortions

The standard errors are obtained by boostrapping. More precisely, I resample the residuals $\{\xi_i\}$ in equation (18) 200 times . The standard errors I then report are the standard deviation of the distribution of estimates. It is important to recognize that because our data consists in the whole universe of firms, – up to at tiny fraction of missing data – these standard errors do not correspond to sampling errors per se but are a inform us about misspecification errors.

The bunching mass is the difference between the empirical distribution and the counterfactual

$$\hat{B} = \sum_{j=N-\delta^{-}}^{N-1} d_j - \hat{c}_j \tag{19}$$

Similarly the missing mass is sum of the difference between the counterfactual number of firms after the threshold and the observerd number of firms.

$$\hat{M} = \sum_{j=N}^{N+\delta^{+}} \hat{c}_j - d_j \tag{20}$$

It is convenient to write these quantities as ratios

$$\hat{b} = \frac{\hat{B}}{\sum_{j=N-\delta^{-}}^{N-1} \hat{c}_j}\ ,\ \hat{m} = \frac{\hat{M}}{\sum_{j=N}^{N+\delta^{+}} \hat{c}_j} \tag{21}$$

The fraction of firms that do not adjust (Kleven and Waseem, 2013) is $1 - \hat{m}$. In a version of the model with fixed cost of adjustment in capital, the interpration is that a fraction $1 - \hat{m}$ of firms are constrained. At the first order, the bunching mass is $\hat{B} \approx s(N)\,\Delta N$ where $s(N)$ is the empirical firm size distribution. Therefore an estimate for ΔN that can be used in the formulas is $\Delta N \approx \frac{\hat{B}}{s(N)}$. With the reduced-form frictions, $\hat{B} \approx \hat{m}.\, s(N)\,\Delta N$. Therefore an estimate for ΔN that can be used in the formulas is $\Delta N \approx \frac{\hat{B}}{\hat{m}.s(N)}$. The presence of frictions implies that the underlying structural elasticity could be higher than the one estimated from the response.

Capital distortions. I now turn to the measures of capital distortions. To control for measurement errors in capital and potential downward biases, I consider several measures of the capital distortions. In theory, one should look at the top buncher, for which the capital distortion reaches its maximum. In practice I implement three estimation strategies, one that uses the average capital distortion in the bunching region, one that exploits the cross-sectional distortion in capital for the bunchers and one that uses investment rates. The model presented in the theoretical framework is extended to accomodate a capital-labor ratio that is increasing with firms' sizes, as it is the case in the data. The firm-level production can be written instead

$$F(K,L) = \left[\alpha(z)\,K^{\frac{\sigma-1}{\sigma}} + (1-\alpha(z))\,L^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

Let $\kappa(z) \equiv \frac{\alpha(z)}{1-\alpha(z)}$, then the capital labor ratio outside of the bunching region is $k(z) = \left(\kappa(z)\,\frac{w}{r}\right)^{\sigma}$. And $\kappa(\cdot)$ can be recovered from the empirical relation between capital intensity and size.

The average capital distortion $\mathbb{E}[\Delta k]$ can be readily measured as the deviation in the capital intensity observed in the bunching region. I test for the presence of capital intensity distortions by running the following regression at the sector level

$$k_i = \beta_0 + \beta_1.L_i + \beta_2.(L_i)^2 + \sum_{j=N-\delta^-}^{N} \delta_j \cdot 1_{\{L_i=j\}} + \epsilon_i \qquad (22)$$

I use the first two moments of labor to absord trends in the capital intensity. The dummies interacted with size capture the deviation in the capital intensity. There are several reasons to think that this measure understates the top buncher's distortion. First, each coefficient δ_j in the bunching region is an average response weighted over an interval of firms, as

$$\delta_j \equiv \mathbb{E}[\Delta k \,|L=j] = \frac{1}{H(\bar{z}) - H(\underline{z})} \int_{\underline{z}}^{\bar{z}} \Delta k(z)\, h(z)\, dz$$

At the first order approximation, we have if the distribution is taken to be roughly constant over the interval $[\underline{z}, \bar{z}]$, in which case $\mathbb{E}[\Delta k]$ corresponds to about half of the maximum capital distortion. Such an approximation creates a downward bias since $h(\cdot)$ is decreasing. To control for the downard bias, I can therefore use the the counterfactual density $\hat{h}$ to better estimate the downard bias created by the bias. Second, measurement errors in capital can create a attenuation bias in the measure of the capital gap. As a robustness check, I estimate how the elasticity would change if the distortions were higher. The results are displayed in Table 2 and 3. . The first table

The second approach exploits the static first order condition in capital and the relationship it predicts in the cross-section of the firms bunching around the threshold. The intution is that as firms over-accumulate capital and deviate from the non-distorted capital-labor ratio, their efficiency decreases. The main specification I run to capture the relationship between the distorted output and capital is

$$\log k_i = \gamma_0 + \gamma_1.\log y_i + \gamma_2.(\log y_i)^2 + \gamma_3.(\log y_i)^3 + X'\beta + \epsilon_i$$

The third approach addresses potential measurement issues in capital. I use the information contained in the investment variable, which is known to be more accurately measured than the net value of the firm's capital stock(e.g. Collard-Wexler and De Loecker, 2016). In order to do so, I map the parameters of the models into investment distortions using a dynamic version of my baseline model, the details are relagated to the Appendix. The basic intuition is the following. Suppose a one-period time-to-build such has investing I_t units of good at time t deliver an equivalent of installed capital in the next period. The Bellman equation for the firm's manager takes the following form

$$V_t(K;z) = \max_{I,L} (1-\tau)[\Pi(K,L) - I] + \beta \cdot \mathbb{E}\left[V_{t+1}(K';z')\right]$$

where the law of motion for capital is $K' = (1-\delta)K + I$ and the law of motion for the productivity index is exogenously specified, $z' = G(z)$. The first order condition in investment is therefore simply

$$I\left(K,z\right)=\frac{\beta}{1-\tau}\mathbb{E}_{t}\left[\frac{\partial V_{t+1}\left(K';z'\right)}{\partial K}\,|z\right]$$

As long as the mean growth rate of productivity is not too large, the local deviation around the threshold in investment per capita, $\frac{I}{L}$, reveals how much capital distortion firms are willing to operate on when they get stuck at the threshold. Consistent with the previous approaches, I find little evidence of distortion in investment. Figure plots invesment per capita around the threshold for various measure of investment per capita: investment in all assets and investment in PPE capital. The two series are remarkably similar. As a robustness check, I also compare with the series obtained after replacing the numerator with the log changes in PPE capital and the one obtained after replacing the denominator with the self-reported measure of labor instead of FTE labor. The log changes in capital do not display any significant uptick. In contrast – and as expected if self-reported labor under-estimate the true quantity of labor – the last meausre display a slight bump before the 50-employee threshold.

Estimation. I implement this procedure at each of the regulatory thresholds, i.e. at 10, 20, and 50 employees. This allow me to estimate the elasticity at different points of the distribution. Moreover, the number of firms at the the 10 and 20 thresholds are much bigger which considerably increases the power of our estimation. It turns out that the 20-employee threshold create little distortion. I then assume that the 10 annd 50- employee thresholds are well-separated so that we can analyze them separately. I first consider the 50-employee where the impact of the regulations seems the most salient.

Bunching evidence and Firm-level Elasticities. Looking at the aggregate size distributiion I find a share of non optimizers is 89% in the social security data. This high magnitude is similar to what has been found in the literature. Kleven and Waseem 2013 estimate that about 90% of workers do not adjust labor supply due to frictions and Liu and Lockwood (2015) find that a similar amount of firms do not adjust turnover in the presence of VAT notches.

I first estimate the elasticity around the 50 threshold and implement the method at the sector level, to allow for potentially large technological differences across sectors for each of the two datasets.At the 50-employee threshold, the elasticities range from 0.04 in Retail to 0.21 in the Wholesale and Trade sector. These

Aggregate elasticities. The micro estimates obtained in the previous sections characterize the *intensive* margin of capital-labor substitution at the firm-level. I how do these micro-level elasticities translate into an aggregate elasticity of substitution for the economy? First, this framework requires the estimation of several other elasticities which measurement might be uncertain. Second the model abstract from This is why we view this exercise mainly as a way to generate reasonable bounds on the aggregate elasticity. I try different alternatives, setting the markups equal to a standard value or computing the markups from the ratio of sales over costs. The results are presented in table Most of the sectoral markups lie between 0.2 and 0.5. Second, the definition of sectors is likely to matter In our data, we consider all 2-digit industries. And we find as a consequence an elasticity around 0.15. Consequently, these results rule out quite clearly the explanation based on falling capital prices.

6 A Structural Model with Frictions in Labor Choice and Misreporting

The baseline model presented in the previous section relates the structural parameters and the policy costs to sufficient statistics. This tractable approach yields closed-form formulas at the expense of abstracting from some potentially important margins. I therefore complement this estimation strategy with a more structural appraoch. However, as is common with notch designs, the model delivers predicts that there should be a hole in the firm-size distribution right after the threshold, since the decision to locate in this region is strictly dominated. Empirically, the missing mass of firms in the dominated region is between 5 and 10 times smaller than predicted by the simple model. In order to explain the absence of such a "hole", I develop a more structural approach. I then add two margins: optimization frictions and the possibility of misreporting the amount of employees. The optimization friction margin accounts for the mass of firms located after the threshold while misreporting is necessary to account for the gap between the labor measures coming from the two datasets. From revealed preferences, what matters for the input choices is the perceived cost of the policy by the firm manager. The large response in the dataset where the labor measure is subject to manipulation clearly exclude the possibility that firms are unaware of the policies. This framework is built to provide an estimate of the elasticity of substitution that is as reliable as possible and tries to limit misspecification error by placing only limited parametric assumptions. The drawback of the approch is that we have to resort to numerical method and cannot obtain the closed-form expressions or maximum likelihood estimators that one could obtain with stronger parametric assumptions.[42]

Substantial biases in the estimates of the elasticity and of the policy cost are created by failing to take into account that firms might misreport their labor size. Under-reported labor results in an overstated economic response, which creates an upward bias on the elasticity. The investment bump reported in Garicano et al. (2016) and in Smagghue (2014) appears as an artefact of firms misreporting their true number of employees. Investment per capita looks higher in the bunching because the denominator is under-reported in this region. The reason is that the presence of firms in the "hole" could be due to frictions that interfere with manager's decisions to bunch. If this is the case, only part of the theoretical, structural response is observed. It is widely recognized that firms face frictions in employment choices (Cooper et al. 2007). In the French labor market, often presented as one of the least flexible (Blanchard et al. 1997) these frictions are more prevalent than in most developped countries. I therefore recognize in the model that managers can only choose

[42] For instance Garicano et al., 2016 assume that the unobserved distribution of firms's productivities follows a Pareto distribution but their likelihood estimator is uninformative about the elasticity parameter as it is completely flat on this parameter's dimension (p.8 of their Appendix).

imperfectly their desired employment level and face optimization frictions. I introduce noise in the optimization program of the manager as in Chetty (2009).[43] This is just one parcimonious way of capturing the effect of several labor frictions that have been examined in the literature such as search frictions or economic uncertainty. If L is the targeted employment level, then the realized employment is given by $\tilde{L} = L \cdot \exp(\epsilon)$, where $\epsilon \sim N(0, \theta)$. As a consequence, the annual Full Time Employment count differs from the equilibrium value chosen *ex ante*.

The timing of the manager's decision is the following: at the beginning of the period, the manager chooses the optimal amount of capital, then optimal chooses a target labor, then the uncertainty is realized. Given a choice of capital, the firm's objective function becomes

$$
\begin{aligned}
\Pi_K(L; z, K) &= \mathbb{E}_\epsilon\left[\left(1 - \tau - \Delta\tau.1_{\{L\exp\{\epsilon\} \geq N\}}\right)\left(z^{1-\nu}F(K, L\exp\{\epsilon\})^\nu - wL\exp\{\epsilon\} - rK - c(s)\right)\right] \\
&= (1-\tau)\mathbb{E}_\epsilon\left[z^{1-\nu}F(K, L\exp\{\epsilon\})^\nu - wL\exp\{\epsilon\} - rK\right] \\
&\quad - C(L; \theta, \tau) \\
&= (1-\tau)z^{1-\nu}\mathbb{E}_\epsilon\left[F(K, L\exp\{\epsilon\})^\nu\right] \\
&\quad - wL\exp\{\theta^2/2\} - rK - C(L; \theta, \tau)
\end{aligned} \tag{23}
$$

where $C(L; \theta, \tau)$ is the expected cost for the firm of crossing the threshold and having to implement the regulations. It is defined as

$$
\begin{aligned}
C(L; \theta, \tau) &= \Delta\tau.\mathbb{E}_\epsilon\left[z^{1-\nu}F(K, L\exp\{\epsilon\})^\nu - wL\exp\{\epsilon\} - rK \mid L\exp\{\epsilon\} \geq N\right] \\
&\quad \times \Pr\{L\exp\{\epsilon\} \geq N\} \\
&= \Delta\tau.z^{1-\nu}\int_{\ln\frac{N}{L}}^{\infty} F(K, L\exp\{\epsilon\})^\nu \phi(\epsilon)\, d\epsilon \\
&\quad - \Delta\tau w - \Delta\tau.rK\left[1 - \Phi\left(\ln\frac{N}{L}\right)\right]
\end{aligned} \tag{24}
$$

where $\Phi(.)$ is the cdf of a the standard normal distribution. The labor friction has several implications on the profit function and on the firm's behavior in equilibrium. The main distinction with the baseline model is that the profit function is no longer discontinuous, as the labor friction smoothes out the discrete cost of crossing the threshold. Moreover, the uncertainty costs increase with the amount of noise.

Proposition 4. *The expected cost of the regulation increases with the amount of unvertainty, i.e.* $\frac{\partial C}{\partial \theta} > 0$.

The intuition is that as the labor friction because larger, firms'managers have an incentive to target an employment level further below the threshold to reduce the likelihood of paying the regulation cost.

[43] An alternative hypothesis would be to suppose that managers can be of several types: "good" or "honest" managers who would never cheat and "dishonest" managers who are prone to cheating and misreporting their size.

Misreporting. In the data, many firms located around the threshold seem to be misreport their true amount of labor. I therefore add an evasion margin in the model. As in Chetty (2009), I suppose that hiding a fraction s of their labor force is costly and that firms must divert an amount $c(s)$ of resources . One interpretation of that cost is that to bypass the regulation manager need to divert time or pay an accountant to help with the dissimulation. Finding ways to evade the regulation might create all sorts of inefficiences. Managers must also balance the benefits of evading with the expected cost of being caugt. In that case one can write $c(s) = p(s) \cdot t(s)$, where $p(s)$ is the perceived probability of being caught and $t(s)$ the associated fine. I follow Chetty et al. (2011) in assuming that the evasion cost $c(.)$ is nondecreasing and strictly convex. A parsimonious parametrization is to consider the special case: $c(s) = \frac{s^\gamma}{\gamma}$, following Chen et al. (2017). More general shapes of misreporting cost could be considered. Ultimately what matters for the purpose of this article is that the convexity ensures that misreporting arises as local phenomenon around the threshold. The realized employment is again given by $\tilde{L} = L \cdot \exp\{\epsilon\}$, where $\epsilon \sim N(0, \theta)$ however the reported labor is now different from the actual labor and depends on the amount of misreporting $\tilde{L}_{rep} = L \cdot \exp\{\epsilon - c(s)\}$. The cost of the regulation still applies to the actual workforce but its cost is borne only if the reported labor force exceeds the threshold. [44]The timing of the decisions is the same as above, after having chosen K, the profit is

$$\begin{aligned}
\Pi_K(L; z, K) &= \mathbb{E}_\epsilon\left[\left(1 - \tau - \Delta\tau.1_{\{Le^\epsilon \geq N\}}\right)\left(z^{1-\nu}F(K, Le^\epsilon)^\nu - wLe^\epsilon - rK - c(s)\right)\right] \\
&= (1-\tau)\, z^{1-\nu}\mathbb{E}_\epsilon\left[F(K, L\exp(\epsilon))^\nu\right] - c(s) \\
&\quad -wL\exp\left(\theta^2/2\right) - rK - C(L; \theta, \tau)
\end{aligned} \tag{25}$$

where the expected cost of the regulation is

$$\begin{aligned}
C(L, s; \theta, \tau) &= \Delta\tau.\mathbb{E}_\epsilon\left[z^{1-\nu}F(K, Le^\epsilon)^\nu - wLe^\epsilon - rK - c(s) \mid Le^{\epsilon - s} \geq N\right] \\
&\quad \times \left[1 - \Phi\left(\ln\frac{N}{L} + s\right)\right]
\end{aligned} \tag{26}$$

It corresponds to the expected fraction of the profit that will be dissipated by the hassle cost when the labor count crosses the threshold. In equilibrium, the incentives to misreport balance the expected reduction in the burden created by the regulation and the first order condition in the amount of misreporting is

$$c'(s) = -\frac{1}{(1-\tau)}\frac{\partial C(s, L; \theta, \tau, \phi)}{\partial s}$$

Misreporting workers, reduces the expected burden because it correspond to a downward shift in the distribution of the labor shock. Alternatively, the effect of misreporting is equivalent to raising

[44]There could be at least two sources for these costs: a penalty after a control by the administration and deteriorated working relations with the employees. Section 3 provides details regarding the enforcement. It is also possible that managers overestimate the costs of non-compliance (Andreoni et al. 1998) either by overestimating the detection probabilities or the penalties associated with tax evasion (see Scholz and Pinney 1995; Chetty 2009).

the threshold above which the regulations must be enforced, whis reduces the probability of being subjected to the regulation. Firms that report a workforce above the threshold do not misreport neither do those that are far below the threshold. The possibility of misreporting creates an additional distortionary effect on the marginal product of labor. This distortion is negligible away from the threshold as the uncertainty of whether the threshold is crossed the threshold or not vanishes. However in the region before the threshold, the uncertainty has a "shading effect" on the choice of labor as firms cut down on labor to limit the risk of going above the threshold. In this structural model, the amount of bunching firms is influenced by three effects: i) the substitution between capital and labor ii) the misreporting of the true amount amount of employee iii) the "shading" effect to prevent against crossing the threshold because of the noise. Because the model is not linear, these effects interacts and it is not possible to separate them in closed-form expressions. The misreporting channel exagerates the amount of bunching in the self-reported data and the shading turn the sharp bunching more into a smooth bump, which better corresponds to the pattern in the matched employer-employee data.

I estimate the structural model using a simulated method of moments. I simulate the model and require the simulated moments to match the size and capital distortion. More formally, my estimator is chosen to minimize the vector of moments

$$\hat{\beta} \in \arg\min_{\beta \in B} \left(\mathbf{m} - m\left(\beta\right)\right)' W \left(\mathbf{m} - m\left(\beta\right)\right)$$

where m is the vector of moment and W is a weighting matrix that is set initially to be the identity matrix. I then compute the optimal weighting matrix W_n by taking the inverse of the variances of the empirical moments at the minimum. Like in Einav et al. 2015 I include in those moments the histogram of the distribution around the threshold , the excess mass, and the gap between the two measures. In effect, I use the bunching equations as an auxiliary model to provide further moments and finds the parameters that minimize the moments. For robustness checks, I use alternative measures of capital: total net assets, net tangible capital and gross tangible capital. Using the two datasets for the estimation increases the power of our estimates as it provides cross-equations restrictions. For instance the profile of the gap between the reported measure depends on both the regulation cost and the misreporting cost. The moments need to be carefully chosen in order to minimize misspecification errors while capturing the essential economic features. Some of the earlier works mentioned previously lead to very wide standard errors for the elasticity estimate. I do not make parametric restrictions on the underlying, unobservable distribution of productivities and instead choose a flexible polynomial approach.

Besides the elasticity of substitution between capital and labor, σ, we have four other parameters that we need to estimate: the variance of the optimization error (θ), the convexity of the misreporting cost function (γ), the returns to scale (ν), and the capital share coefficient in the CES (α). The convexity of $c(.)$ is obtained from matching the average gap between the reported employment and the social security measure. More precisely, we reproduce the icnrease in the gap around the threshold. The magnitude of the returns to scale calibrated to range commonly found in the literature. I pick a baseline value of $\nu = 0.85$ as in Atkeson and Kehoe (2005) and conduct sensitivity analyses around this value, in the range suggested by Basu and Fernald (1997). An alternative estimation approach would be to use directly the first order conditions and derive their implications on factor ratios. Such an approach raises at least three concerns: i) it first implies to make assumptions about r , the rental rate of capital, for instance following Hsieh and Klenow, 2009, and ii) measurement errors in capital can create severe bias (see Collard-Wexler and De Loecker 2016). For the baseline tax rate τ, I use the prevailing statutory corporate income tax rate in each year. During the period considered, the smallest firms can qualify for a "reduced tax rate". To qualify for this reduced rate, the firm's revenues must be below a certain threshold (about 1 million euros[45]). The reduced rate then applies to the first bracket of the revenues. For all the revenues above this bracket, the normal rate applies. Because this bracket is very low – less than $50k$ euros – it is not directly relevant for this quantitative exercise and I do not model it explicitely.

The results of this structural estimation are in line with the results based on the reduced-form formulas. In the structural estimation, I consider only the 50-employee threshold and I assume, as in Kleven and Waseem (2013) that the other thresholds are far enough that they do impact the behavior of the firms around the 50-employee threshold. In Table 4 I present the results I obtain for the manufacturing sector. The cost of regulation that is abouth 0.12% of the profits and the elasticity is 0.18, somewhat higher than what estimated with the first approach.

Discussion and implications. How do these estimates help us understand the evolution of factor shares? Between 1980 and 2010, the French labor share fell by 7.3 percent points, a fall of similar magnitude than the decline of the U.S. labor share's over the same period (7.5 percentage points). However, a closer look at their evolutions over the three decades reveal strikingly different very patterns. The U.S. labor share was flat roughly until 2000 and then dropped sharply (Figure (14)).[46] In France, the pattern in almost the complete opposite: a brutal drop from 1983 to 1989 and then a flat if somewhat upward trajectory. A low elasticity leaves little scope for factor prices to affect the labor share. Even if one takes the view that my estimates are mostly reflecting the short-run response. It is highly unlikely that such a sharp drop can be accounted for by entry and exit patterns. [47]

A low elasticity has several other important implications. For instance, in a static input-output model as in Baqaee and Farhi 2017 one can show the impact of a productivity shock in sector $k$on aggregate consumption depend on a combination of firm-level elasticities and markups . With an elasticity smaller than one, the shocks to productivity are *attenuated*. With elasticities close to 0,

[45]See General Tax Code (CGI) article 219, I-b

[46]The U.S. labor might have fallen by 2 percentage points more if one include the fact that U.S. multinational firms have increasingly shifted part of their domestic profits abroad. See Guvenen et al. (2017) for the impact of offshore profit shifting on U.S. national accounting.

[47]I measure the labor share as the "corporate labor share", using the definition and data provided by **?**. The corporate labor share excludes government activities as well as the labor and capital income earned by unincorporated businesses or sole proprietors, for which the labor share needs to be imputed. Thus the corporate labor share is a metric that is less sensible to these imputations methods and more robust for international comparisons. The French corporate labor share fell from 73.4 % to 66.1% while the U.S. labor share fell from 64.7% to 57.2%.

these shocks barely propagate. Therefore low elasticity limit some of the propagation channels of production-network macroeconomic models.

How would adding adjustment costs to capital affect my elasticity estimates? While labor frictions are present in my structural framework and help account for the sizeable amount of firms lying in the dominated region, inputs in my models are static and I abstracted from the full dynamics of firms' input decisions. However, because some firms in the bunching region have been here for several years, my estimate captures also this medium-run response. For instance, about 20% of the firms that were bunching at 50 in 2000 where still bunching there 7 years later (see Figure 11). Quadratic adjustment costs in capital of the type popularized by Sargent, 1978 create second order disturbances for the relatively small investment deviations considered here, and are therefore unlikely to generate any substantial downward bias. More critical are fixed cost of adjustment, which generate the large inaction bands characteristic of S-s models (Caballero and Engel, 1999). An important body of literature argue that fixed costs of adjustment are necessary to match the moments of the investment distribution at the plant-level (Cooper and Haltiwanger, 2006). However the magnitude of these fixed costs is hard to pin down precisely and seems sensitive to modeling choices. For instance, Khan and Thomas (2008) and Bachmann et al. (2013)'s estimates are one order of magnitude apart and, moreover, Khan and Thomas (2008) show that the plant-level investment cross-section is consistent in general equilibrium with negligible fixed costs. If their view is correct, then one can reasonably expect my elasticity estimates to be largely unaffected by adding these fixed costs. Otherwise, my elasticity estimates need to be interpreted as the *effective* elasticities of substitution, corresponding to the firms' reactions in the face of lumpy adjustment.

Could a more explicitly dynamic model of the bunching behavior deliver substantially higher elasticity estimates? There are good reasons why this would not necessarily the case, as several forces play against a high elasticity. If anything, my static design suffers from an overestimation bias. Because static designs pool together cross-sections from different years, firms who bunch in consecutive years will be counted several times in the excess mass. Those "repeat" bunchers who stay at the threshold several years in a row tend therefore to inflate the excess mass in the bunching region. By overweighting those repeat bunchers, the estimate produced will overestimate the true elasticity. To exploit the panel dimension in my setting I can replace my identification assumption that the counterfactual employment distribution is smooth by the assumption that the counterfactual employment *growth rate* distribution is smooth. By correctly reweighting the contribution of repeat bunchers, Marx (2015) shows that the overestimation bias can be up to an order of magnitude. Because my baseline estimate for the elasticity is already low, this potential upward bias does not present a fundamental challenge to my conclusions.

Finally, this paper has so far concentrated on the *intensive* margin of substitution between capital and labor. Indeed, the static framework I develop abstracts from entry and exit of firms and from technological change. An active extensive margin would provide a mechanism through with the long-run elasticity could diverge from the short-run elasticity. For instance, if technological change is embodied into capital, then the entry of firms with capital intensity consistently higher than the incumbent's capital intensity would progressively raise the long-run aggregate capital labor ratio. This mechanism is best analyzed in a model with putty-clay technology.[48] When the production function is putty-clay, capital is malleable *ex ante* but once the capital intensity is chosen, it must stay fixed forever. For instance the production function can be CES *ex ante*, but becomes Leontief *ex post*. The speed at which the short-term elasticity converges to the long-run elasticity hinges on the amount of churn in the economy.[49] With entry rate in France slightly below than 9%, and concentrated among the smallest firms (less than 10 employees) this channel is quantitatively limited in my setting. Moreover, since my estimation pools firms across years – and some of these firms have been staying at the threshold for several years – this adjustment effect is already partly reflected in my estimates.

7 Concluding Remarks

This paper argues that the aggregate elasticity of substitution between capital and labor might be lower than conventionally assumed. The main reason is that the firm-level elasticities are close to zero, putting it differently, the production structure of firms is sticky or rigid, similar to a Leontieff production. Thes finding is robust across sectors. I develop a new method to estimate this micro elasticity using the distortions generated by size-dependent policies in France. I implement this method on two comprehensive administrative datasets. Despite the striking discontinuity visible in the French firm size distribution, most of the bunching response is in fact due to misreporting by firms. I identified this misreporting behavior using the gaps between the self-reported measure in one dataset with a measure verified by the administration using employe-employee data. These regulations are salient enough that they elicit a significant response at the threshold – even when misreporting is taken into account –, but they generate little distortions in the use of production factors. In particular, distortions in the capital intensity at the threshold or in investment behavior are about 4 to 5 times smaller than what would be implied by a cobb-douglas production function. I then aggregate these firm-level elasticities in a framework that allows for substitituion across firms and sectors.

More general lessons can be gleaned from this exercise. A low elasticity of substitution between capital and labor has important consequences for investment policies and tax incidence. My findings imply that is that little aggregate capital-labor substitution is the result of the intensive margin. If one takes the view that the aggregate elasticity is large, then the a stronger emphasis should be put on the extensive margin, that is, the entry of new firms and the exit of failing firms.

[48] The seminal contribution by Johansen (1959) launched a large literature investigating the property of growth models with putty-clay technology (Solow, 1962; Phelps, 1963; Cass and Stiglitz, 1969; Calvo, 1976). More recently, putty-clay model have been used to investigate the evolution of factor shares (Caballero and Hammour, 1998), energy use (Atkeson and Kehoe, 1999), business cycle dynamics (Gilchrist and Williams, 2000), or stock market volatility (Gourio, 2011).

[49] Sorkin (2015) develops a model along these lines in order to study how employment respond over time to changes in the minimum wage.

If new firms embody technological changes, then a dynamic entry margin can alter the capital labor mix in the medium run. Alternative explanations might account for the drastic fall in the labor share observed in some countries, such as drastic institutional changes or macroeconomic shocks. My empirical design, which by its very nature focuses on small and local changes, is silent about the role of these institutional changes. Examining in greater depth and through the lens of firm-level data how new entrants contribute empirically to the aggregate elasticiy of the economy is an interesting question for left for future research.

References

Aghion, Philippe, Ufuk Akcigit, Matthieu Lequien, and Stefanie Stantcheva, "Tax simplicity and heterogeneous learning," Technical Report, National Bureau of Economic Research 2017.

Andreoni, James, Brian Erard, and Jonathan Feinstein, "Tax compliance," *Journal of economic literature*, 1998, *36* (2), 818–860.

Antras, Pol, "Is the US aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution," *Contributions in Macroeconomics*, 2004, *4* (1).

Arrow, Kenneth J, Hollis B Chenery, Bagicha S Minhas, and Robert M Solow, "Capital-labor substitution and economic efficiency," *The review of Economics and Statistics*, 1961, pp. 225–250.

Atalay, Enghin, "How important are sectoral shocks?," *American Economic Journal: Macroeconomics*, 2017, *9* (4), 254–80.

Atkeson, Andrew and Patrick J Kehoe, "Models of energy use: Putty-putty versus putty-clay," *American Economic Review*, 1999, *89* (4), 1028–1043.

_ **and Patrick J. Kehoe**, "Modeling and Measuring Organization Capital," *Journal of Political Economy*, 2005, *113* (5), 1026–1053.

Bachas, Pierre and Mauricio Soto, "Not (ch) your average tax system: corporate taxation under weak enforcement," 2018.

Bachmann, Rüdiger, Ricardo J. Caballero, and Eduardo M. R. A. Engel, "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model," *American Economic Journal: Macroeconomics*, October 2013, *5* (4), 29–67.

Bagger, Jesper, Francois Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin, "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics," *American Economic Review*, 2014, *104* (6), 1551–96.

Baqaee, David Rezza and Emmanuel Farhi, "Productivity and Misallocation in General Equilibrium.," Technical Report, National Bureau of Economic Research 2017.

_ **and** _ , "The Microeconomic Foundations of Aggregate Production Functions," Technical Report 2018.

Barrot, Jean-Noël and Julien Sauvagnat, "Input specificity and the propagation of idiosyncratic shocks in production networks," *The Quarterly Journal of Economics*, 2016, *131* (3), 1543–1592.

Basu, Susanto and John G Fernald, "Returns to scale in US production: Estimates and implications," *Journal of political economy*, 1997, *105* (2), 249–283.

Becker, Randy A, John Haltiwanger, Ron S Jarmin, Shawn D Klimek, and Daniel J Wilson, "Micro and macro data integration: The case of capital," in "A new architecture for the US national accounts," University of Chicago Press, 2006, pp. 541–610.

Benzarti, Youssef and Jarkko Harju, "Are Taxes turning human into machines: Using Payroll variation to estimate capital-labor elasticity of substitution," 2018.

Blanchard, Olivier J, William D Nordhaus, and Edmund S Phelps, "The medium run," *Brookings Papers on Economic Activity*, 1997, *1997* (2), 89–158.

Boehm, Johannes, Swati Dhingra, John Morrow et al., "Swimming upstream: input-output linkages and the direction of product adoption," *CEP Discussion Paper*, 2016, (1407).

Béguin, Jean-Marc and Olivier Haag, "Méthodologie de la statistique annuelle d'entreprises, Description du système Esane," 2017.

Caballero, Ricardo J and Eduardo MRA Engel, "Explaining investment dynamics in US manufacturing: a generalized (S, s) approach," *Econometrica*, 1999, *67* (4), 783–826.

_ **and Mohamad L Hammour**, "Jobless growth: appropriability, factor substitution, and unemployment," in "Carnegie-Rochester Conference Series on Public Policy," Vol. 48 Elsevier 1998, pp. 51–94.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg, "The anatomy of French production hierarchies," *Journal of Political Economy*, 2015, *123* (4), 809–852.

Calvo, Guillermo A, "Optimal growth in a putty-clay model," *Econometrica: Journal of the Econometric Society*, 1976, pp. 867–878.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal, "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement," *American Economic Journal: Applied Economics*, April 2017, *9* (2), 144–64.

Cass, David and Joseph E Stiglitz, "The implications of alternative saving and expectations hypotheses for choices of technique and patterns of growth," *Journal of Political Economy*, 1969, *77* (4, Part 2), 586–627.

Ceci-Renaud, Nila and Paul-Antoine Chevalier, "L'impact des seuils de 10, 20 et 50 salariés sur la taille des entreprises françaises," *Economie et Statistique*, 2010, *437.*

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D Investment with Corporate Income Tax Cuts in China," 2017.

Chetty, Raj, "Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance," *American Economic Journal: Economic Policy*, 2009, *1* (2), 31–52.

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

—, John N Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The quarterly journal of economics*, 2011, *126* (2), 749–804.

Chirinko, Robert S, "σ: The long and short of it," *Journal of Macroeconomics*, 2008, *30* (2), 671–686.

Collard-Wexler, Allan and Jan De Loecker, "Production function estimation with measurement error in inputs," Technical Report, National Bureau of Economic Research 2016.

Cooper, Russell, John Haltiwanger, and Jonathan L Willis, "Search frictions: Matching aggregate and establishment observations," *Journal of Monetary Economics*, 2007, *54*, 56–78.

Cooper, Russell W and John C Haltiwanger, "On the nature of capital adjustment costs," *The Review of Economic Studies*, 2006, *73* (3), 611–633.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "Declining dynamism, allocative efficiency, and the productivity slowdown," *American Economic Review*, 2017, *107* (5), 322–26.

Doraszelski, Ulrich and Jordi Jaumandreu, "Measuring the Bias of Technological Change," *Journal of Political Economy*, 2018, *126* (3), 1027–1084.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf, "The response of drug expenditure to nonlinear contract design: evidence from medicare part D," *The quarterly journal of economics*, 2015, *130* (2), 841–899.

Fack, Gabrielle and Camille Landais, "The effect of tax enforcement on tax elasticities: Evidence from charitable contributions in France," *Journal of Public Economics*, 2016, *133*, 23–40.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, November 2016, *106* (11), 3439–79.

Gattaz, Yvon, *La fin des patrons*, Robert Laffont, 1979.

Gilchrist, Simon and John C Williams, "Putty-clay and investment: a business cycle analysis," *Journal of Political Economy*, 2000, *108* (5), 928–960.

Gourio, Francois, "Putty–clay technology and stock market volatility," *Journal of Monetary Economics*, 2011, *58* (2), 117–131.

Gourio, François and Nicolas Roys, "Size-dependent regulations, firm size distribution, and reallocation," *Quantitative Economics*, 2014, *5* (2), 377–416.

Guvenen, Fatih, Raymond J Mataloni Jr, Dylan G Rassier, and Kim J Ruhl, "Offshore profit shifting and domestic productivity measurement," Technical Report, National Bureau of Economic Research 2017.

Havez, Pierre, "Les seuils sociaux en France, Quel impact sur la démographie des PME?," *Master Thesis, Paris-I Sorbonne*, 2018.

Hicks, John, "The theory of wages," 1932.

Hopenhayn, Hugo A, "Firms, misallocation, and aggregate productivity: A review," *Annu. Rev. Econ.*, 2014, *6* (1), 735–770.

Houthakker, Hendrik S, "The Pareto distribution and the Cobb-Douglas production function in activity analysis," *The Review of Economic Studies*, 1955, *23* (1), 27–31.

Hsieh, Chang-Tai and Benjamin A Olken, "The Missing" Missing Middle"," *Journal of Economic Perspectives*, 2014, *28* (3), 89–108.

_ **and Peter J Klenow**, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403–1448.

Hurst, Erik, Geng Li, and Benjamin Pugsley, "Are household surveys like tax forms? Evidence from income underreporting of the self-employed," *Review of economics and statistics*, 2014, *96* (1), 19–33.

Johansen, Leif, "Substitution versus fixed production coefficients in the theory of economic growth: a synthesis," *Econometrica: Journal of the Econometric Society*, 1959, pp. 157–176.

Karabarbounis, Loukas and Brent Neiman, "The global decline of the labor share," *The Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Khan, Aubhik and Julia K Thomas, "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics," *Econometrica*, 2008, *76* (2), 395–436.

Kim, Kyoo, Amil Petrin, and Suyong Song, "Estimating production functions with control functions when capital is measured with error," *Journal of Econometrics*, 2016, *190* (2), 267–279.

Kleven, Henrik J and Mazhar Waseem, "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan," *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Kleven, Henrik Jacobsen, "Bunching," *Annual Review of Economics*, 2016, *8* (1), 435–464.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

León-Ledesma, Miguel A, Peter McAdam, and Alpo Willman, "Identifying the elasticity of substitution with biased technical change," *American Economic Review*, 2010, *100* (4), 1330–57.

Levhari, David, "A note on Houthakker's aggregate production function in a multifirm industry," *Econometrica: journal of the Econometric Society*, 1968, pp. 151–154.

Liu, Li and Ben Lockwood, "VAT notches," 2015.

Lucas, Robert E, "Labor-capital substitution in US manufacturing," *The taxation of income from capital*, 1969, pp. 223–274.

_ , "On the size distribution of business firms," *The Bell Journal of Economics*, 1978, pp. 508–523.

Marx, Benjamin M, "Dynamic bunching estimation and the cost of reporting regulations for charities," *MPRA Paper #88647*, 2015.

Miyagiwa, Kaz and Chris Papageorgiou, "Endogenous aggregate elasticity of substitution," *Journal of Economic Dynamics and Control*, 2007, *31* (9), 2899–2919.

Oberfield, Ezra and Devesh Raval, "Micro data and macro technology," Technical Report, National Bureau of Economic Research 2014.

Phelps, Edmund S, "Substitution, fixed proportions, growth and distribution," *International Economic Review*, 1963, *4* (3), 265–288.

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic dynamics*, 2008, *11* (4), 707–720.

Ricardo, David, *On the principles of political economy and taxation*, London: John Murray, 1817.

Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, August 2010, *2* (3), 180–212.

Sargent, Thomas J, "Estimation of dynamic labor demand schedules under rational expectations," *Journal of Political Economy*, 1978, *86* (6), 1009–1044.

Satō, Kazuo, *Production functions and aggregation*, Vol. 90, North-Holland, 1975.

Scholz, John T and Neil Pinney, "Duty, fear, and tax compliance: The heuristic basis of citizenship behavior," *American Journal of Political Science*, 1995, pp. 490–512.

Smagghue, Gabriel, "Size-Dependent Regulation and Factor Income Distribution," 2014.

Solow, Robert M, "Technical change and the aggregate production function," *The review of Economics and Statistics*, 1957, pp. 312–320.

_ , "Substitution and fixed proportions in the theory of capital," *The Review of Economic Studies*, 1962, *29* (3), 207–218.

Sorkin, Isaac, "Are there long-run effects of the minimum wage?," *Review of economic dynamics*, 2015, *18* (2), 306–333.

Appendix A: Mathematical Appendix

A1. Alternative models of the cost of regulation

In order to capture the large set of regulations that kick in at the regulatory thresholds, the main model analyzes their impact if they operate as a wedge on the tax rate. In order to asses the robustness of this assumption, this section analyzes alternative specifications of the policies – as a "pure" notch, as a combination of a labor wedge and a fixed cost, and as a wedge on output. I derive the bunching formulas in each case. These formulas follow from the profit indifference condition and capture the trade-off between the regulation cost or the cost of distortions around the threshold.

"Pure" Notch

Instead of modelling the cost of the regulations as an increment in the tax rate, $\Delta\tau$, one can view it instead as a fixed cost ΔT that creates a "pure" notch following the terminology of Kleven and Waseem (2013). The profit function for the baseline model then becomes

$$\Pi(L;z) = \begin{cases} (1-\tau)\left[z^{1-\nu}L^{\nu} - wL\right] & \text{if } L < N \\ (1-\tau)\left[z^{1-\nu}L^{\nu} - wL\right] - \Delta T & \text{if } L \geq N \end{cases}$$

For this case with no substitution, the bunching formula is

$$\Delta T = (1-\tau)\left(\frac{\nu}{w}\right)^{\frac{\nu}{1-\nu}} z\left(1+\frac{\Delta z}{z}\right)\left[1-\left(\frac{1}{1+\Delta z/z}\right)^{\nu}\right] \tag{27}$$

Labor Wedge and a Fixed Cost

Suppose that instead of incurring a hassle cost, the firm's faces a wedge on the marginal labor cost τ_ℓ and a fixed cost ϕ, as in Garicano et al. (2016). The profit function for the case with no substitution, which is the one they consider, is

$$\Pi(L;z) = \begin{cases} z^{1-\nu}L^{\nu} - wL & \text{if } L < N \\ z^{1-\nu}L^{\nu} - w(1+\tau_\ell)L - \phi & \text{if } L \geq N \end{cases}$$

The profit indifference condition delivers the following bunching formula

$$\left(1+\frac{\Delta N}{N}\right)^{1-\nu} - (1-\nu)\left(1+\frac{\Delta N}{N}\right) = \nu \cdot \frac{1}{1+\tau_\ell}\left(1-\frac{\phi}{wN}\right) \tag{28}$$

First, it is worth stressing that this formula is independent from assumptions on the underlying productivity distribution. Second, either an increase in τ_ℓ or in ϕ increases the size of the gap, as the left-hand side is strictly decreasing in $\bar{N}$, declining from ν to 0, when the ratio $\frac{\bar{N}}{N}$ increases from 1 to $\left(\frac{1}{1-\nu}\right)^{\frac{1}{\nu}}$ (≈ 9.3 if $\nu = 0.85$).

Proposition. *The labor wedge and the fixed cost are not separately identified without further assumptions on the unobserved productivity distribution.*

This can be seen directly in (28). As long as the pair of costs (τ_ℓ, ϕ) is such that the quantity $\frac{1}{1+\tau_\ell}\left(1-\frac{\phi}{wN}\right)$ stays constant, then the size of the hole stays the same. Indeed this predicted linear relationship holds almost perfectly for the estimates in Table 1 in Garicano et al. (2016). Any other moments of the size distribution – the amount of bunching or shifts in the slope of the size distribution – are directly dependent on assumptions on the unobserved productivity. For instance, the amount of bunching depends on the distribution between N and $\bar{N}$ and any shift in the distribution above $\bar{N}$ can be produced either by increasing τ_ℓ or by a rescaling of $h(.)$, which is not observed. Indeed the firm's labor demand

$$L(z) = \begin{cases} z\left(\frac{\nu}{w}\right)^{\frac{1}{1-\nu}} & \text{if } z \leq \underline{z}_0 \\ N & z \in \left[\underline{z}_0, \bar{z}_0\right] \\ z\left(\frac{\nu}{(1+\tau)w}\right)^{\frac{1}{1-\nu}} & \text{if } z \geq \bar{z}_0 \end{cases}$$

implies the following size distribution, in terms of the underlying productivity distribution $H(.)$,

$$s(n) = \begin{cases} \frac{1}{L_0} h\left(\frac{n}{L_0}\right) & \text{if } n \leq N-2 \\ H(\bar{z}) - H(\underline{z}) & n = N-1 \\ 0 & N \leq n < \bar{N} \\ (1+\tau)^{\frac{1}{1-\nu}} h\left(\frac{n}{L_0}(1+\tau)^{\frac{1}{1-\nu}}\right) & \bar{N} \leq n \end{cases}$$

where $L_0 = \left(\frac{\nu}{w}\right)^{\frac{1}{1-\nu}}$. With a wedge on labor, the capital-labor ratios are different on both sides of the thresholds, and the relative gap is $\sigma\Delta\tau$. This provides an additional identification equation.

Output Wedge

Suppose that instead of incurring a hassle cost, the regulations act as an output wedge τ_y. One interpretation is that the manager must now divert some of his time to handle the implementation of these regulations. For the sake of notation, I omit the baseline tax rate and the profit function is

$$\Pi(L;z) = \begin{cases} z^{1-\nu}L^\nu - wL & \text{if } L < N \\ \left([1-\tau_y]\,z\right)^{1-\nu} L^\nu - wL & \text{if } L \geq N \end{cases}$$

The profit indifference condition delivers the following bunching formula

$$\left(1+\frac{\Delta N}{N}\right)^{1-\nu}-(1-\nu)\left(1-\tau_y\right)\left(1+\frac{\Delta N}{N}\right)=\nu \tag{29}$$

A2. Investment and Capital distortions in a Dynamic Model

Investment is measured with greater accuracy than capital, which is know to be measured with substantial margin of error [50]. Using the distortions in per capita investment for firms around the thresholds to detect distortionary impacts of the policy is a good way of checking the robustness of the results that rely on capital measures.[51] I therefore extend my baseline model to a dynamic setting in order to make explicit the connection between investment distortions and capital distortions as a reaction to the policies. Investment takes one period to be fully operational and the Bellman equation for the firm's manager takes the following generic form

$$V_t(K;z)=\max_{I,L}(1-\tau)\left[\Pi(K,L)-I\right]+\beta\cdot\mathbb{E}\left[V_{t+1}\left(K';z'\right)\right]$$

where the law of motion for capital is $K'=(1-\delta)K+I$ and the law of motion for the productivity index is exogenously specified, $z'=G(z)$. The first order condition in investment is therfore

$$I(K,z)=\frac{\beta}{1-\tau}\mathbb{E}\left[\frac{\partial V_{t+1}\left(K';z'\right)}{\partial K}|z\right]$$

relationship between investment per capita, $\frac{I}{L}$ and firm size L depends on the specifications of the law of motion of the productivity index z. To build intuition, lets' first consider some simple cases and assume $\beta=1$. If z grows deterministically at rate g then investment per capita is equal to $\frac{I}{L}=\delta k^*+k^*\cdot g$ for all the firms except for the ones bunching before the threshold. For the firms who expect to bunch in the following period it is optimal to choose a higher investment rate $\frac{I}{L}=\delta k^*+k^*\cdot g+\Delta k$ in order to operate at a higher capital intensity when they are bunching. In this case the distortion in the capital labor ratio is exactly informative of the capital distortion Δk, as in the case of sharp bunching.

Starting from 2000, there are two direct measures of investment in FICUS, one that sum tangible and intangible investments (INVAVAP) and one that measure tangible investments only (INVCORP). In order to better understand investment patterns, I also look at a more granular version of the datasets only available after the 2009 great recession. I consider three broad investment categories: investment in tangible assets (property, plant, equipment, and land), investment in intangible(intellectual property, R&D, licenses), and financial (mostly security deposits and guarantees).

[50] See Becker et al. (2006) for survey of the methods to measure capital.

[51] Collard-Wexler and De Loecker (2016) develop a control function approach that uses investment lags as an instrments for capital.

Capital-Skill complementarity The firm-level production functions can be extended in order to allow for a more than two factors. For instance one can think of several categories of labor as in Krusell et al. (2000). The French matched-employee data seems specially appealing since in the dataset, each job spells receives an occupation code classifying this job in a 500+ occupation category. This occupations are used in Caliendo et al. (2015) to analyze the hierarchical organization of French manufacturing firms. I compute at the firm-level the share of each category. After removing the unemploment and retirement spells, jobs are organized into five broad categories, each of them being itself with subdivisions: CEO, firm owners or firm directors; senior staff or top managers; employees at the supervisor level which includes quality control technicians, technical, accounting, and sales supervisor; qualified and nonqualified clerical employees (secretaries, human resources or accounting employees, telephone operators, and sales employees), and blue-collar qualified and nonqualified workers (e.g. welders, assemblers, machine operators, maintenance workers). I group the first three together to form a high-skill category and the last two in a low-skill category. The share of each category in the wages and hours worked is depicted in Figure . There is suggestive evidence that the composition of labor changes slightly around the threshold. Unfortunately, there are many missing occupation labels and, moreover, an abrupt change in the percentage of missing values happens concurrently around the 50-employee threshold. One interpretation is that the extra administrative requirements generated by the regulatory threshold also forces some employer to be more diligent. To measure carefully the changes in labor composition, one would therefore need a deeper investigation to precisely understand how this two effects interact. In a chapter of I leave this for future research.

Appendix B: Data and Institutional Background

Size-dependent Regulations

This section provides details about the list of the regulations activated when a firm crosses the 10-, 20- and 50-employee thresholds. The specifics of these policies have been modified several times and calls for heir suppression or smoothing out have been frequent in policy and business circles. After some changes in 2008 (Loi de modernisation de l'économie du 4 août 2008), and in 2015, some of the threshold-based regulations have been further relaxed in a bill voted in the French National Assembly this Fall.). I focus on the regulations that were in place during the period 2000-2008. These policy thresholds are commonly referred to, in the public debates, as "Social and Fiscal Thresholds" ("Seuils sociaux et fiscaux"). The regulations are scattered across several pieces of legislation: the Labor Code (Code du Travail, CT), the Social Security Code (Code de Securite Sociale, CSS), the Territorial Authorities Code (Code General des Collectivites Territoriales, CGCT), the General Tax Code (Code General des Impôts, CGI). These regulations have been altered over the last decades, but not in the period I consider. Calls to eliminate these thresholds go as far back as 1984 (Gattaz 1979). A last change happened in 2015 when it was decided to smooth over time the effect of these thresholds. Firms crossing the threshold are given 2 years to comply.

Regulations starting at the 50-employee threshold A complete list of the regulations can be found in Ceci-Renaud and Chevalier (2010). The most important policies that are activated at the 50-employee threshold are: i) an elected workplace committee must meet once every other month. Elections must be held (unless there is no candidate; ii) information must be transmitted quarterly and yearly about the firm's financial situation to the workplace committee; iii) external "representative" trade unions can designated a union representative; iv) a workplace safety committee must be elected, with up to 3 members, which 2 paid hours every month to each member in order to run the committee; v) compulsory annual negociation about wages, work duration, etc; vi) if the firm increases dividends with respect to the previous two years then a bonus must also be paid to employees; vii) a specific procedure ("Plan de Sauvegarde de l'Emploi") must be put in place a mass layoff is considered (where mass layoff is defined as the layoff of at least 10 employees over less than 30 days); viii) workplace negociations must be held to discuss about prevention of accident injury and work penibility; ix) a compulsory monthly report of job turnover must be sent to the administration.

Regulations starting at the 10-employee threshold Very small businesses are exempted from many administrative regulations and many exemptions stop when the firm reaches 10 employees. When they cross that threshold, these firms are subject to: i) a professional training contribution amounting to at least 1.6% of the wage bill (Labor Code - Article L6331-9) ii) a transportation contribution (Code général des collectivités territoriales - Article L2333-64); iii) extra administrative requirements: transmission of national unemployment agency of proof of employment termination (Labor Code - Article R1234-9) and rules regarding dismissal are strenghtened.

Regulations starting at the 20-employee threshold. At the 20-employee threshold, the main regulations are the following: i) an increase in the professional training contribution from 1.05% to 1.60% of the wage bill, ii) the obligation that employees with disabilities represent at least 6% of the workforce,[52] iii) two extra contributions each amounting to 0.45% of the wage bill; iv) extra cost for overtime work;

[52]Typically, a firm with just 20 employees which did not have any employee with disabilities, has to hire one when crossing the threshold.

Dataset coverage and the universe of French firms.

They are several available administrative firm datasets generated from tax data, following evolutions in their collection and in their statistical treatment. In some years, these datasets overlap. The datasets whose access is given to researchers is not raw. The administration has developped procedures to remove some inconsistencies and errors. Efforts are made to cross-check some of the information (for instance firm identifiers). These procedures have evolved over time. The last version of the cleaning procedure is explained in painstaking details in Béguin and Haag, 2017. The dataset initially covered only firms subject to corporate income tax. In FICUS, the coverage has been extended to firms operating under different, usually simplified, tax regimes, such as pass-through businesses, partnerships or medical practices. Because the tax forms differ, some variables for firms reporting under the other tax regimes can only be estimated. In particular, the only balance sheet items reported by firms under the "Non commercial profits" regime (BNC for "Benefices non-commerciaux", about $500,000$ legal entities report under this tax regime in 2005) are gross capital and operating income.[53] To make my work more comparable to previous studies I conduct my main analysis using FICUS, that is the whole universe of French firms combining all tax regimes despite potential heterogeneity concerns. As a robustness check I performed the analysis excluding these other firms.

Evidence for Evasion in France

The gap between reported and measured employment – and its implication in terms of misreporting and regulation avoidance – might at first be surprising. However my empirical finding are in line with Ceci-Renaud and Chevalier (2010) and confirmed by Havez (2018).Fack and Landais (2016) finds other evidence regarding lack of enforcement regarding charitable contributions in France, while Spire (2017) marshalls detailed evidence that tax evasion is not strongly prosecuted in France. Overall tax evasion rates in France is found to be in line with the tax evasion rates in the US, estimated at around 16% for individual income (Internal Revenue Service 2008) .

Corporate Income Taxation

The French corporate income tax rate has been remarkably stable over the period I study. From 1993 to 2016 the standard tax rate stood at 33,33%. There is also a reduced tax rate that applies for profits up to 38,120 Euros. All profits exceeding that threshold are taxed at the standard rate. To qualify for the reduced rate, the firm's revenues must not exceed 7,63 million euros. Because such a threshold could be easily taken advantage by large firms splitting themselves into several several units, an extra requirement is that more than 75% of the paid-in capital of the firm must be detained by natural persons. The 7,63 million euros threshold is therefore a kink point. Consistent with many studies which find little bunching at kink points, I find that there is no significant bunching at the 7,63 million euros threshold. As a result the distribution of the average effective

[53] see Béguin and Haag (2017) p.126

tax rate paid by firms has a peak at 33¹/₃% and a peak at 0% for firms who are not turning profits. There is a small mass of firms lying between 15%, and 33¹/₃%, and no significant bunching at 15%, the reduced-tax rate.

Boundary of the Firm

I adopt the same unit of observation as the one used in the FICUS dataset, that is, a firm is defined as a legal entity, associated with a unique SIREN identifier. This by far the most frequent choice, although there are at least three possible levels of observations, at the establishment level, at the firm leve, and at the group level. For the manufacturing sector, the establishment level would correspond to the plant level, which is also often used. In the administrative dataset based on the tax returns, firms are uniquely identified by their 9-digit SIREN number. But firms can locate their activities across several establishments. Establishments are uniquely identified by a SIRET number, which corresponds to the SIREN prefixed by an block of 5 digits. In principle, the social security forms are filled at the establishment level[54]. In addition to that, European Law[55] provided in 2008 a definition of an "enterprise", as "the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities at one or more locations. An enterprise may be a sole legal unit." This definition has been transposed into French Law in 2008 by the "Loi de Modernisation Economique". The French statistical agency is in the process of constructing comprehensive datasets that reflects this new definition. This is a difficult process that requires additional and detailed information about intra-group transactions. Only 120 such "enterprises" have been constructed in the 2015 datasets out of about 80,000 enterprises. Therefore I leave for future research a detailed exploration of capital-labor substitution within the boundaries of these enterprises.

[54]see Béguin and Haag (2017).

[55]Council Regulation 696/93, https://publications.europa.eu/en/publication-detail/-/publication/1ea18a1a-95c2-4922-935c-116d8694cc40/language-en

Additional Figures

Figure 3: Effects of the Policies on Labor, Capital intensity

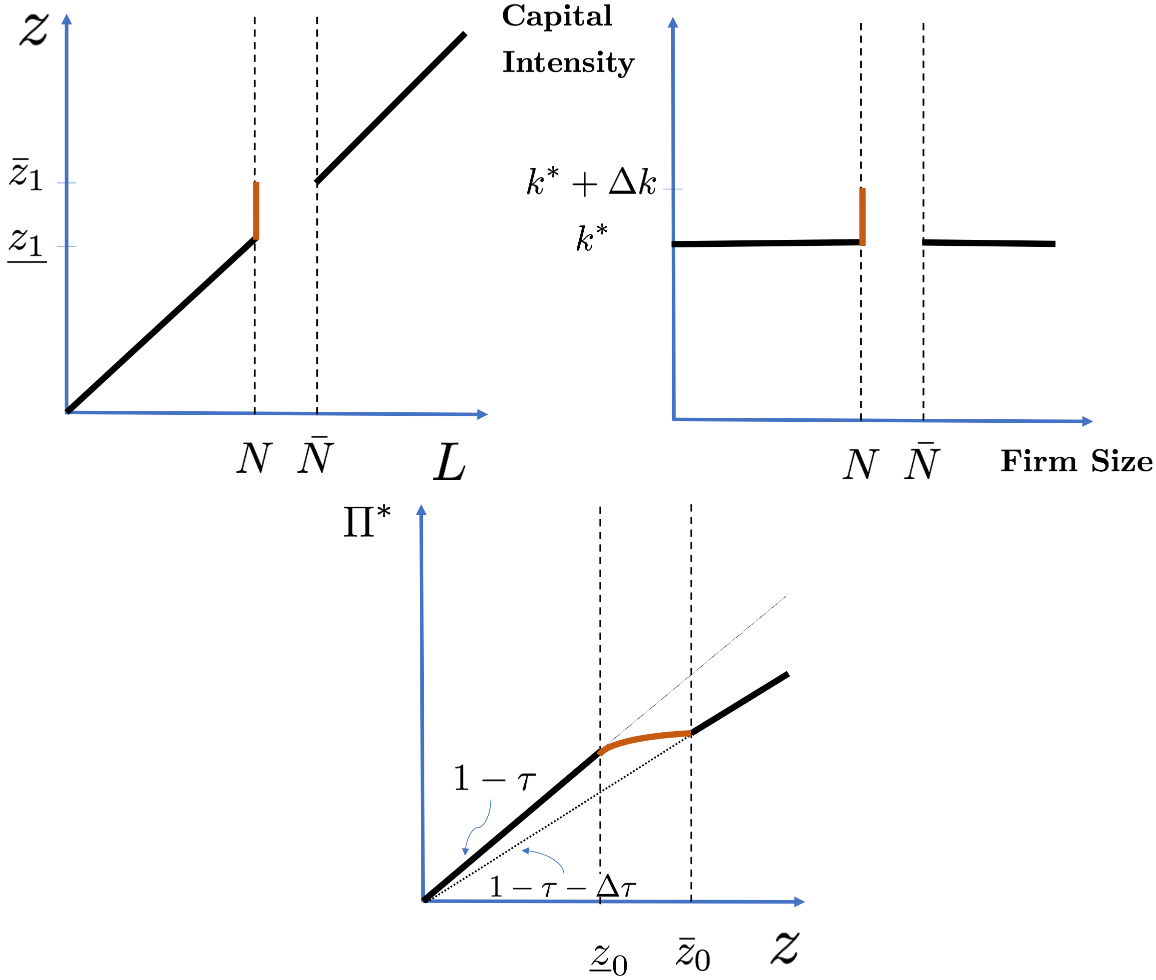
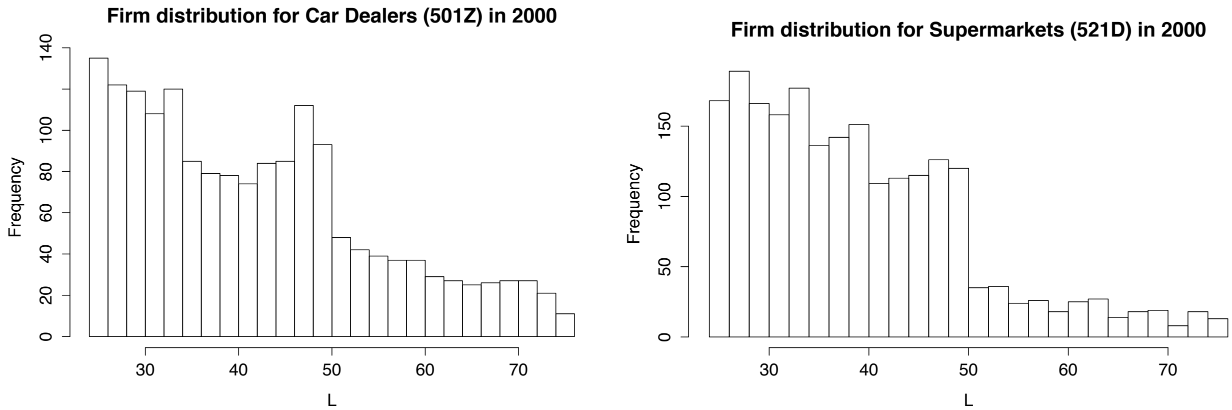


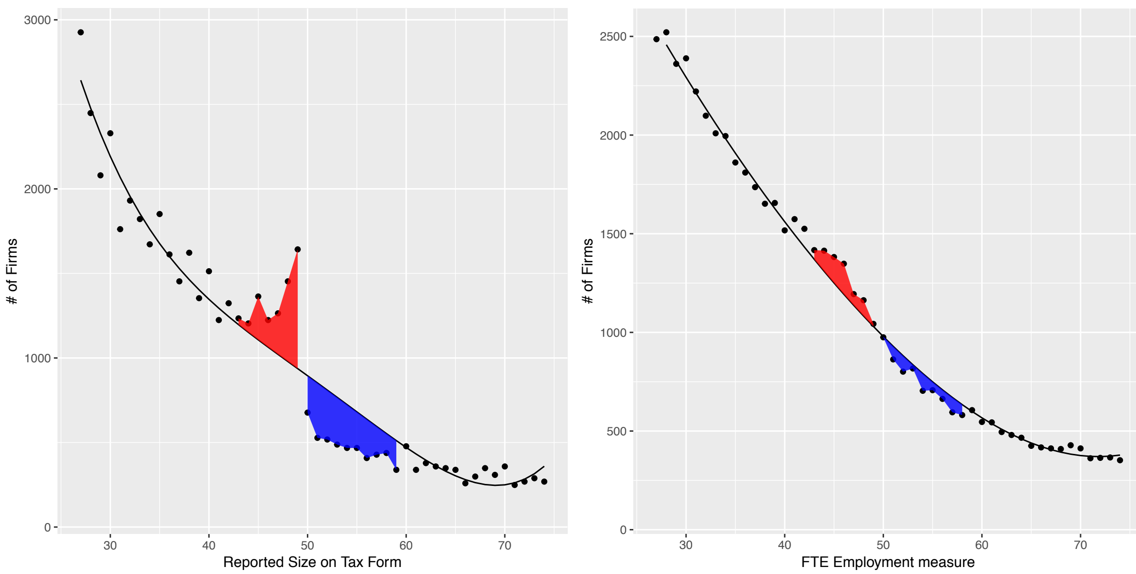
Figure 4: Firm Size distributions at selected 4-digit sectors



Source: FICUS

Figure 5: Excess Mass in the Two Datasets

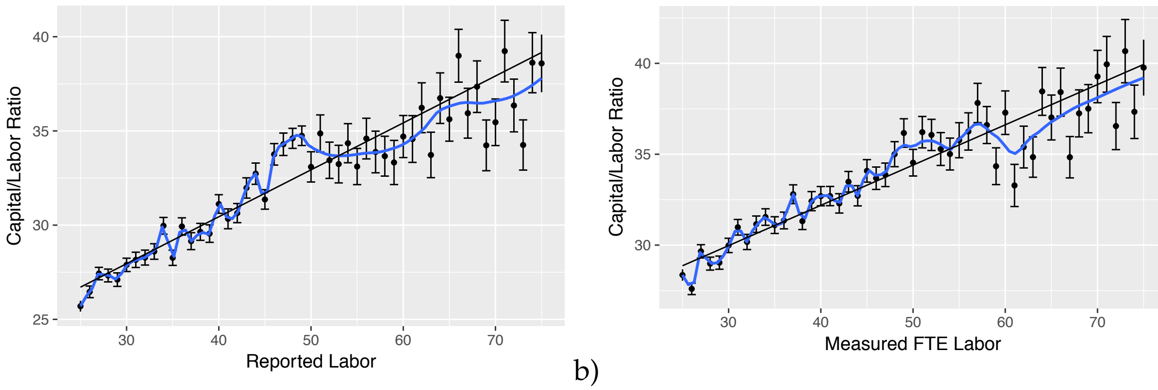
a) At the 50-employee threshold



b) At the 10-employee

The excess mass are computed in each dataset by fitting a 6th order polynomial The left panel uses the self-reported data while the right panel uses the measured data. Source: FICUS and DADS

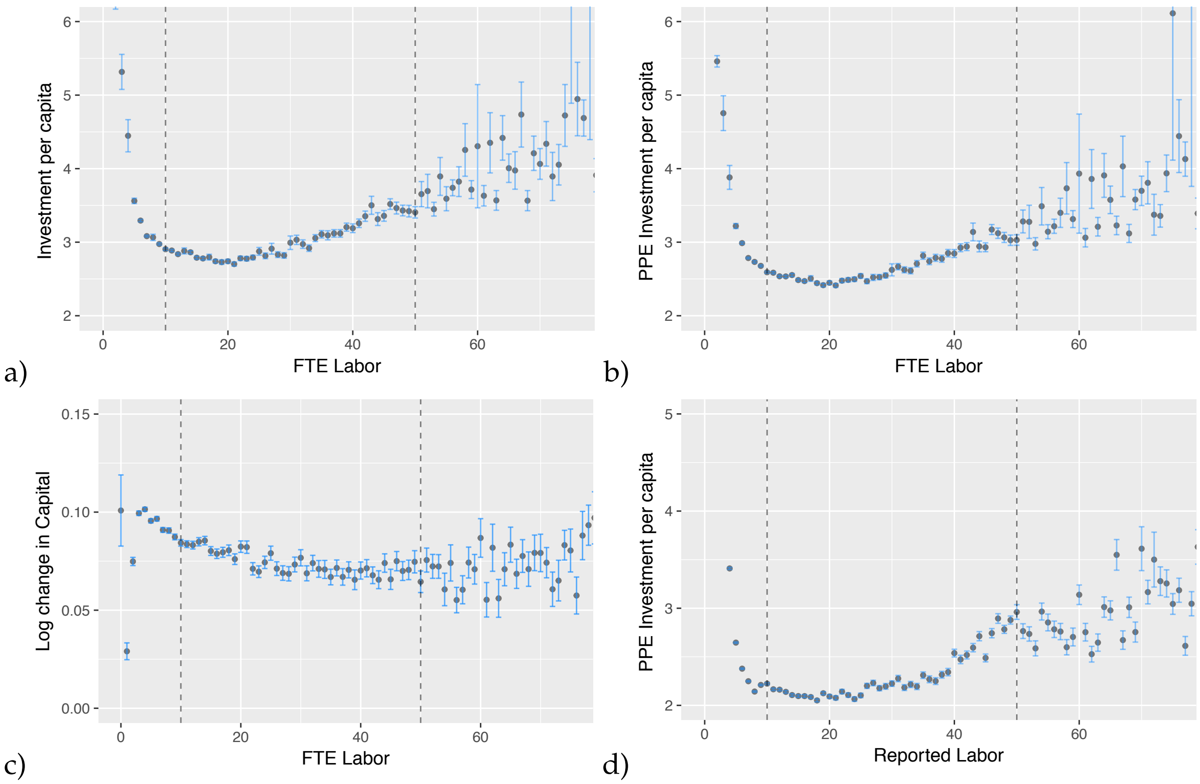
Figure 6: Distortions in Capital-Labor ratios



a) b)

Note: Panel a) shows the distortion in capital-labor ratio where the reported labor variable is used. Panel b) shows the capital-labor for the same firms but using the FTE measure computed by the administration instead. Source: FICUS and DADS.

Figure 7: Distortions in Investment per capita



Note: I report the investment response around the threshold for three measures of investment: a) total gross investment, b) gross investment in property, plant, and equipment, and c) the log change in PPE capital. In all panels except for d) the x-axis is the FTE measured Labor. Only when the x axis is the self-reported labor does investment seem abnormally high before the cutoff.

Figure 8: Notch with Frictions

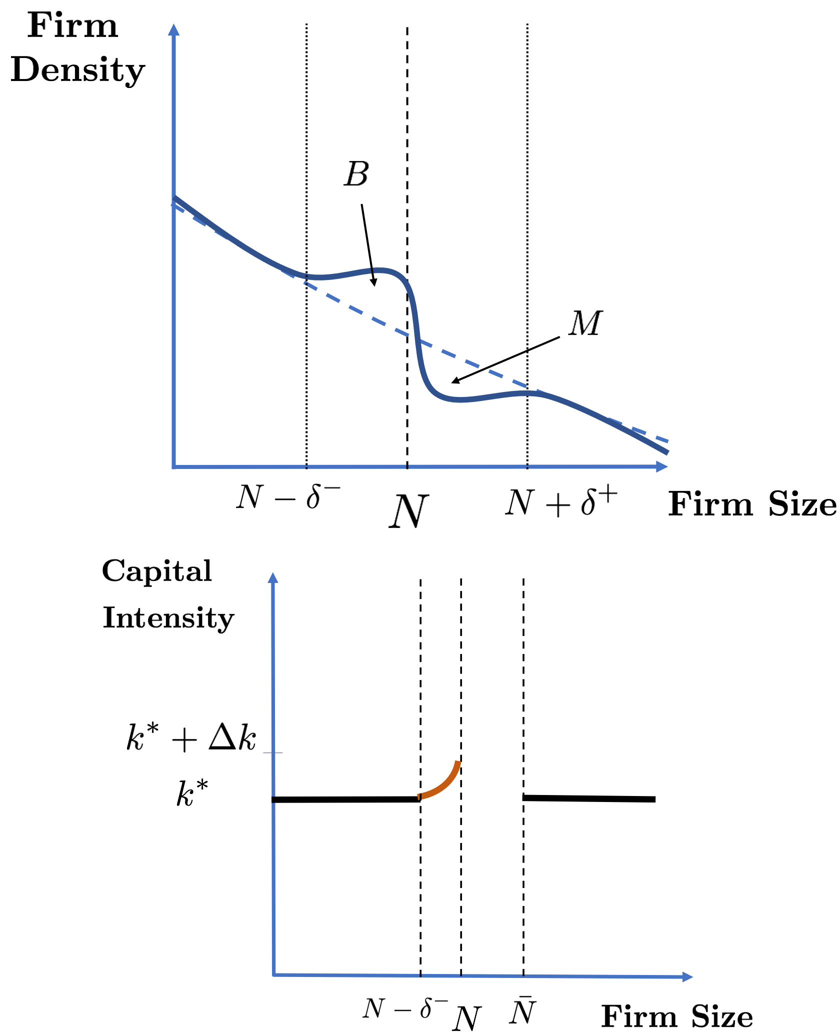
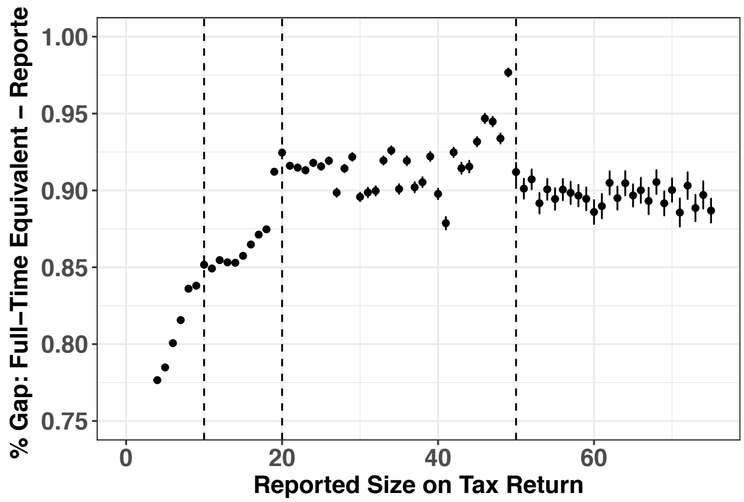


Figure 9: Misreporting Evidence: Ratios of Firm-Size Densities at the Sector Level

Note: the firm size densities are computed for each sector using alternatively reported and FTE-measured employment. Then I compute the ratio of these two densities (y axis). The plots shows these density ratios for a) the manufacturing sector, b) the retail sector, c) the transportation and communication sector, d) Business services. For instance, there are 75% more firms with 49 reported employees than with 49 FTE-measured employees. The dotted line is the ratios' average, excluding the bunching regions. Source: FICUS 2004-2007

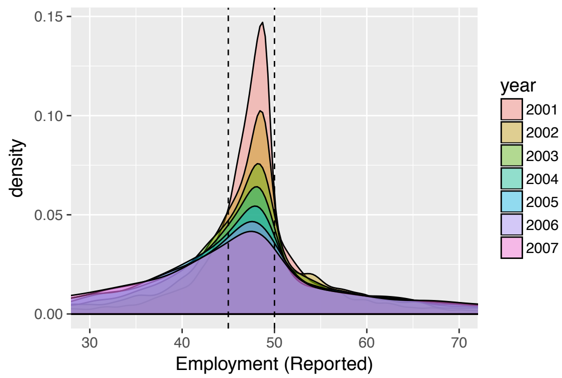
Figure 10: Misreporting Evidence: Self-Reported vs Administration's Full-Time Employment measure.



b)

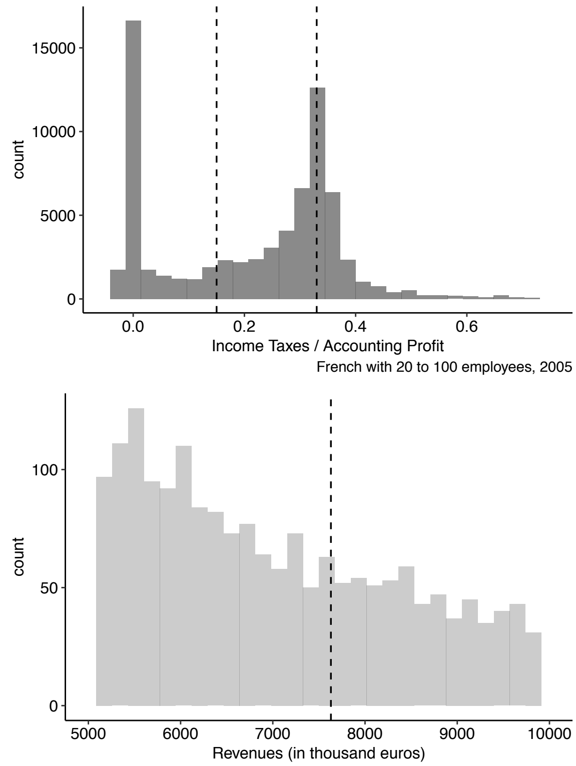
Note: the figure plots the percentage gap between the Social Security Labor measure and the self-reported labor measure. The formula used by the administration to compute the FTE measure is top-censored at 1, which creates a downard bias of 10% on average. Because o f misreporting this gap shrinks abruptly at the 20- and 50- employe threshold. Under 50-employee, the standard errors are tiny and hardly distinguishable from the point estimate. Source: Autor's calculation from FICUS and DADS.

Figure 11: Bunching Persistence around the 50-employee threshold



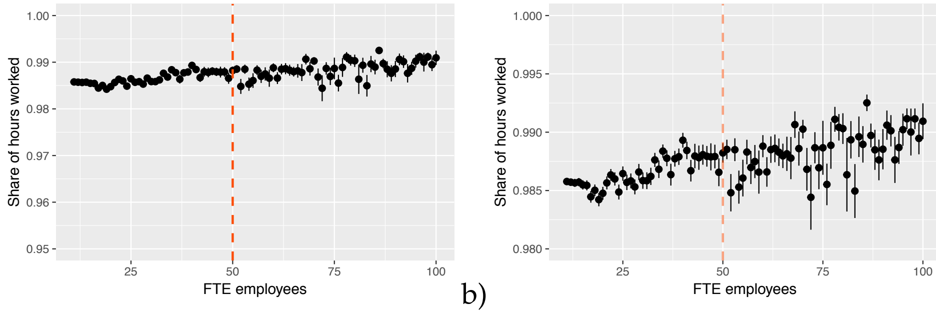
Note: This is the employment density over time of the firms who were bunching at the 50-employee threshold in 2000. About half of them are bunching the next period as well and 20% of them are still bunching 7 years later. Source: Author's calculation from FICUS.

Figure 12: Effective tax rates



Note: The top panel shows the average effective income tax rate paid by firms with 20 to 100 employees. It implies that for almost all of these firms turning profits, the relevant marginal corporate income tax rate is the standard rate at $33, 1/3\%$. The bottom panel shows that there is no significant bunching at the kink point between the reduced and standard rates. Source: Author's calculation from FICUS.

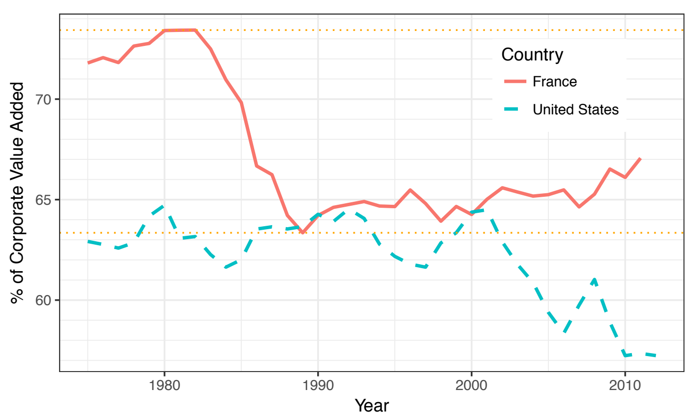
Figure 13: Primary jobs account for 99% of the hours worked



b)

Note: these figures show that the Full-Time Employment measure in the Social Security Data is a very realiable measure of employment at the firm level. It is unproblematic to leave aside "secondary jobs" as they account for less than 1% of the hours worked, and this fraction is stable with firm size. For each firm, I use the matched employer-employee data to sum up the total number of hours worked in primary jobs and in secondary jobs down. I then compute the share of hours corresponding to primary jobs. Finally I bin firms per FTE size (rounded) and compute the mean and standard error. Source: DADS 2004

Figure 14: Evolution of the Labor Shares in France and in the United States



The corporate labor share is the share of labor compensation in the value added of the corporate sector, that is, exluding the government sector, sole proprietors and unincorporated enterprises. This concept arguably makes cross-country comparisons more transparent (Karabarbounis and Neiman, 2014). Source: author's calculations from OECD data.

Table 1: Descriptive statistics: Comparing the Firm Size Distributions

	FICUS	DADS (FTE over whole year)	DADS (FTE on 12/31)
# of firms with	2,504,173	1,573,610	1,573,594
0 employees	1,215,064	123,673	236,816
>0 employees	1,289,109	1,449,937	1,336,778
1	210,564	245,395	224,118
9	36,263	28,289	28,449
10	18,098	21,837	21,890
18	12,273	7,276	7,434
19	11,783	6,540	6,623
20	9,661	5,974	5,870
47	2,547	1,292	1,304
48	3,021	1,222	1,225
49	1,647	1,047	1,097
50	678	1,018	1,012

Note: The difference in the total number of firms is mainly due to the firms with no employee as the DADS (Social Security data) registers payroll data. Starting from ... FICUS records income of partnerships, medical practices, entrepreneurs etc. Source: FICUS and DADS 2004 .

Table 2: Capital-Labor Distortions: Reported Employment

	Dependent variable:					
	log(Capital-Intensity)					
	OLS					*Sector FE*
	(1)	(2)	(3)	(4)	(5)	(6)
L	0.005***	0.008***	0.004***	0.005***		0.006***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)		(0.0001)
L^2		−0.00003***				
		(0.00000)				
Bunching			0.184***	0.095***	0.083***	0.088***
			(0.009)	(0.028)	(0.028)	(0.026)
L<45				−0.058**	−0.274***	−0.005
				(0.026)	(0.026)	(0.025)
L>55				−0.145***	0.035	−0.125***
				(0.027)	(0.027)	(0.025)
Constant	2.538***	2.518***	2.539***	2.591***	2.853***	
	(0.001)	(0.001)	(0.001)	(0.027)	(0.026)	
Obs.	2,360,558	2,360,558	2,360,558	2,360,558	2,360,558	2,340,215
R^2	0.004	0.004	0.004	0.004	0.002	0.115
Adj. R^2	0.004	0.004	0.004	0.004	0.002	0.115
Res. S.E.	1.286	1.286	1.286	1.286	1.287	1.207
F Stat.	8,591.332***	4,629.088***	4,501.677***	2,281.012***	1,930.439***	

*p<0.1; **p<0.05; ***p<0.01

Table 3: Capital-Labor Distortions: FTE-measured Employment

Table 4:

	Dependent variable:						
	log(Capital-Intensity)						
	OLS					*Sector FE*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
etpr	0.004***	0.004***	0.004***	0.001***			
	(0.0001)	(0.0001)	(0.0001)	(0.0001)			
L						0.003***	0.001***
						(0.0001)	(0.0001)
bunching			0.147***	0.024	0.022	0.051**	0.040*
			(0.009)	(0.027)	(0.027)	(0.026)	(0.024)
L<45				−0.221***	−0.268***	−0.129***	−0.133***
				(0.026)	(0.026)	(0.025)	(0.023)
L>55				0.054**	0.075***	−0.073***	−0.049**
				(0.026)	(0.026)	(0.025)	(0.023)
Constant	2.714***	2.714***	2.716***	2.955***	3.013***		
	(0.001)	(0.001)	(0.001)	(0.026)	(0.026)		
Obs.	1,957,932	1,957,932	1,957,932	1,957,932	1,957,932	1,957,932	1,957,932
R^2	0.002	0.002	0.002	0.003	0.003	0.120	0.228
Adj. R^2	0.002	0.002	0.002	0.003	0.003	0.120	0.228
Res. S.E.	1.248	1.248	1.247	1.247	1.247	1.181	1.106
F Stat.	3,673.795***	3,673.795***	1,970.334***	1,306.335***	1,692.167***		

*p<0.1; **p<0.05; ***p<0.01

Table 5: Dispersion in Income Shares and Elasticity by Sectors

NAF code	Sector	Dispersion (χ_{ni})	Firm-level Elasticity (σ_i)	Demand Elasticity (Benchmark)	Sectoral Elasticity
70	Real Estate	0.77	0.12	0.3	0.26
74	Business Services	0.37	0.14	0.3	0.19
52	Retail	0.10	0.04	0.3	0.06
55	Hotel and restaurants	0.18	0.06	0.3	0.10
45	Construction	0.19	0.11	0.3	0.15
50	Wholesale and Trade	0.07	0.21	0.3	0.22
24	Chemical Industry	0.27	0.18	0.3	0.21

Note: Elasticities measured using the 50-employee threshold. The dispersion coefficient is measured using net assets. NAF is the acronym of the French classification of economic activity. Activities are ranked in decreasing order of the amount of capital they command. Source: FICUS 2004.

Table 6: Sectoral Dispersion in Income Shares and Elasticities

NAF code	Sector	Firm-level Elasticity σ	Share α_n
74	Business Services	0.19	0.37
52	Retail	0.06	0.10
55	Hotel and restaurants	0.10	0.18
45	Construction	0.15	0.19
50	Wholesale and Trade	0.22	0.07
15-35	Manufacturing	0.12	0.21

Note: elasticitities are measured at the 50-employee threshold. Dispersion is measured using net assets. NAF is the acronym of the French classification of economic activity. Activities are ranked in decreasing order of the amount of capital they command. Source: FICUS 2004.

Table 7: Structural Estimates

	Parameter	Value	S.E
$\hat{b}$	Bunching Mass	0.09%	0.002
$\Delta\hat{k}$	Capital distortion	0.011	0.017
τ	Regulation Cost	0.12 %	0.0023
σ	Elasticity of Substitution	0.18	0.13
θ	Variance of the Optimization Error	0.05	0.17
γ	Convexity of misreporting cost	3	.
ν	Returns to Scale	0.85	.

Note: Firms in manufacturing Sector (NAF 15-35) 50-employee threshold. Source: Ficus 2004-7.