The Efficiency Costs of Tax Enforcement:

Evidence from a Panel of Spanish Firms[*]

Miguel Almunia   David Lopez-Rodriguez

(Job Market Paper)

This version: November 17, 2012

Latest version: Available here

Abstract

In modern tax systems, firms remit the majority of tax revenues raised by the government, which gives them opportunities to avoid and evade taxes by misreporting their activities. In this paper, we use a natural policy experiment to analyze how firms respond to different tax enforcement regimes. The Spanish Large Taxpayers' Unit (LTU) monitors and enforces the taxation of companies with operating revenue above €6 million, resulting in more frequent tax audits and more information requirements for those firms. We exploit this discontinuity in enforcement intensity to estimate the impact of low enforcement on tax reporting behavior, using a panel dataset of financial statements for 85% of Spanish firms in the period 1999-2007. We apply two different identification strategies. First, we find an excess mass of firms locating ("bunching") just below the revenue threshold. Based on the number of bunching firms, we estimate that firms reduce reported revenue by 1.4% to 7.5% to avoid falling in the high enforcement regime. Second, we run panel regressions with firm fixed effects and find that reported input costs do not respond significantly to the change in tax enforcement. This suggests that firms react to the enforcement discontinuity mostly by reporting lower revenues, without significantly distorting production. The efficiency costs of tax enforcement are thus likely to be small because tax evasion constitutes a reallocation of income to tax-evading firms. A back-of-the-envelope calculation shows that the loss in tax revenue due to tax evasion by non-LTU firms is, however, significant. In light of these results, we discuss potential reforms to improve tax enforcement policies.

JEL codes: H25, H26, K42, M48.

1 Introduction

Firms remit more than three-quarters of the tax revenues raised by governments in advanced economies.[1] As taxpayers, they remit corporate income tax and their share of payroll tax. As tax collectors, they withhold income and payroll taxes from employees, and in many countries they also remit value added tax (VAT). Despite playing such a crucial role as fiscal intermediaries, the empirical literature on tax evasion has largely neglected firms, focusing instead on individual behavior.[2] The information asymmetry between businesses and tax authorities gives firms incentives to misreport their own income and also third-party incomes in order to evade taxes.

In this paper, we take advantage of a natural policy experiment to analyze how tax enforcement policies (i.e., tax audits and compliance requirements) affect firms' reporting and production decisions. In 1995, the Spanish tax agency established a Large Taxpayers' Unit (LTU) to monitor and enforce the taxation of companies with annual operating revenue above €6 million.[3] Firms assigned to the LTU are subject to more frequent tax audits and information cross-checking by the tax authority, but their tax schedule is unaffected. This discontinuity in tax enforcement intensity gives firms an incentive to remain below the revenue threshold. They can do this either by reducing their output or by underreporting their revenue (or both). We analyze what type of response is more important, which allows us to put an upper bound on the efficiency cost of tax enforcement.

To guide the empirical estimation, we set up a theoretical framework where profit-maximizing firms decide (i) how much to produce and (ii) how much of their revenue to underreport in order to reduce their tax liability. Firms receive an exogenous productivity draw that determines their optimal size in equilibrium. The probability of evasion detection is continuously increasing in firm size and in the amount evaded. This reflects the intuition that larger firms are more visible to the tax authority, and that egregious evasion is easily detectable. We introduce the concept of LTU in the model such that the detection probability jumps up discretely at a fixed level of reported revenue. This generates a "notch" in tax enforcement, meaning that the probability of detection increases for *all* inframarginal units evaded when a firm crosses the threshold. The existence of the notch drives some firms to report lower revenue and bunch at the LTU threshold to avoid high enforcement. We define the "marginal buncher" as the firm with the highest exogenous productivity that chooses to bunch at the cutoff point.

---

[1] In the United States, businesses remit 84 percent of taxes (Christensen, Cline, and Neubig, 2001).

[2] Slemrod (2004, 2008) has repeatedly stressed the relevance of firms in the analysis of tax evasion. In contrast, Andreoni, Erard, and Feinstein (1998) state in the introduction of an authoritative survey of the evasion literature: "[...] Nor do we have the space to discuss corporate or business tax evasion".

[3] Firms in the LTU represent only 2.5% of all registered business, but they employ 50% of private sector workers and report 75% of taxable profits (A.E.A.T., Several years). Most tax agencies in advanced countries, and an increasing number of emerging countries, have some type of LTU to deal with large businesses (see I.M.F., 2002 and O.E.C.D., 2011).

For the empirical analysis, we use financial statements and balance-sheet data from Amadeus. This database compiles information reported by firms to the Commercial Registry, covering more than 85% of registered businesses in Spain with operating revenue above €2 million for the period 1999-2007. The longitudinal structure of the dataset is well suited to the use of panel-data techniques, which helps us deal with sorting issues. An advantage of this data source over administrative tax returns is that it provides an overall picture of the firms' activities, allowing us to observe several dimensions of firm behavior in a single data source.[4] We estimate the response in reported revenues and input costs using two separate identification strategies (described below), and then combine them to provide an estimate of the change in reported profits around the enforcement cutoff.

To estimate the effect of a change in tax enforcement intensity on reported revenue, we analyze the distribution of annual operating revenue. As predicted by the model, a significant number of firms bunch below the LTU threshold. This behavioral response is persistent over time for the entire 1999-2007 period and it is due exclusively to the existence of the LTU, not to other regulations affecting firms in the same size range. We construct a counterfactual revenue distribution and use it to estimate the excess bunching mass in a short interval below the threshold. We then use this excess mass as a sufficient statistic for the revenue response of the marginal buncher. Despite the notch in enforcement, many firms choose to locate just above the LTU threshold, suggesting the existence of optimization frictions. Several reasons could explain this behavior, for example prior exposure to the LTU or the inability to misreport revenue due to the types of clients served (e.g., government contracts). Another factor could be heterogeneous preferences toward evasion, as there might be honest business managers who would not evade taxes under any enforcement level. We use the missing mass in an interval above the threshold (where the bunching firms would have located in the absence of the policy) as a proxy for the degree of optimization frictions. Dividing the original bunching estimator by this proxy, we obtain a treatment-on-the-treated estimator of the total revenue response.

We find that the marginal buncher reduces its reported revenue by €86,000 (1.4% of total revenue) under the assumption of no optimization frictions, and €449,000 (7.5% of total revenue) once frictions are taken into account. This is a sizeable response, considering that average reported profits around the LTU threshold are €290,000 (4.5% of revenue). The effect is somewhat stronger among firms that are small in non-revenue dimensions such as fixed assets or number of employees. Across sectors of activity, there is an inverted-U relationship between the size of the response in a

---

[4]In order to perform our empirical tests using administrative data, we would need tax returns from the corporate income tax, the value added tax, and social security contributions. It is rare for researchers to have access to all these sources of information simultaneously, and especially to be able to link them (since governments provide anonymized data).

given sector and a "scope of evasion" index that takes into account the median number of employees and the share of output sold to final consumers in each sector.

Measuring the response on the input cost side is complicated precisely by the fact that firms sort endogenously below the threshold. A naïve regression discontinuity estimation comparing reported expenditures just below and just above the LTU cutoff would be biased because the running variable (revenue) is manipulated. We exploit the panel structure of the dataset to control for the sorting effect by estimating a model with firm fixed effects. The fixed effects control for all time-invariant characteristics (observable and unobservable) that may drive the sorting process. Identification in this model comes from firms that change their enforcement regime during the period under analysis. We estimate the fixed effects regression separately for labor and nonlabor input costs, because they generate opposite tax incentives. If a firm intends to evade taxes, it should overreport nonlabor inputs (e.g., raw materials used for production) to lower its VAT and corporate income tax liability. The incentives are mixed regarding labor input costs, because underreporting reduces payroll tax liability but that mechanically increases corporate tax liability. Since marginal tax rates are similar (38% and 35%, respectively), it is only marginally profitable to overreport labor input costs. We find point estimates close to zero and precisely estimated for the effect of tax enforcement on reported labor and nonlabor input costs. Hence, we conclude that firms do not report input costs differently under the high and low enforcement regimes. This implies that the observed bunching response must be due almost entirely to revenue underreporting, with little effects of enforcement on actual production.

In order to analyze the efficiency costs of tax enforcement in this context, we make two simplifying assumptions. First, we assume that each firm is owned by one individual, whose total income is given by the profits of the firm. That way, firm profits can enter a standard utilitarian welfare function. Second, we consider an increase in tax enforcement intensity equivalent to an increase in the expected tax rate. This allows us to draw insights from the extensive literature on the deadweight loss of taxation. Feldstein (1999) made an influential contribution to this literature by establishing that the elasticity of taxable income is a sufficient statistic for the deadweight loss of taxation. Besides the standard labor supply response, his framework allowed for the existence of evasion and avoidance responses, assuming that taxpayers equalize the marginal resource cost of evasion to the tax rate (marginal benefit). Chetty (2009) pointed out that Feldstein's result does not hold when the marginal *social* cost of evasion is different from the tax rate. This situation arises when taxpayers face a "transfer cost" for evading, such as a monetary penalty if detected. The penalty represents a private cost to the taxpayer, but has no social cost because it is a transfer to the government. Chetty shows that, when taxpayers face transfer costs, the correct sufficient statistic is the elasticity of total

income, not just (reported) taxable income. Since we do not find a significant change in actual profits, but mostly underreporting, our results indicate that the efficiency costs of tax enforcement are likely to be low in the context we study. In other words, an increase in enforcement would not reduce the firms' true profits significantly, but it would generate a transfer from tax-evading firms to the government.

A related question is whether the observed response is relevant in terms of tax revenue. The lack of significant effects on input costs allows us to use the estimated revenue response to make a back-of-the-envelope calculation of the tax revenue lost due to evasion. Extrapolating these local results to all non-LTU firms, we estimate that lost tax revenues could be in the range of 0.17% of GDP (no frictions estimate) and 0.95% of GDP (frictions estimate). This is equivalent to 2%-10% of the current budget deficit of the Spanish Government. Expanding enforcement policies to a larger set of firms would therefore solve a small part of the country's fiscal troubles, but this could be profitable because the efficiency cost is relatively low. In making this cost-benefit calculations, we would need to incorporate the additional costs of tax enforcement. In the short-run, a marginal expansion of the LTU threshold to including firms currently bunching would yield a substantial amount of tax revenue. Another option is to make the LTU threshold "fuzzy", such that enforcement intensity increases in a continuous rather than discrete way over a reported revenue interval. This would lead to less salient changes in tax enforcement, which according to the literature on taxable income elasticities generates proportionally smaller behavioral responses (e.g., Chetty, Looney, and Kroft, 2009).

The findings in this paper contribute to the empirical literature on business tax evasion by providing a well identified measure of the effects of tax enforcement on firm behavior. Our finding that stronger enforcement does not create large inefficiencies complements the findings of Paula and Scheinkman (2010) and Pomeranz (2011). These papers emphasize the role of information for effective tax enforcement, particularly in the presence of a VAT that uses the invoice-credit system. The self-enforcing mechanisms of VAT lead to the transmission of evasion (or compliance) behavior down the production chain from suppliers to intermediate goods producers and then to retailers. Our results indicate that, since the efficiency cost of tax enforcement is relatively low, an increase in tax enforcement on suppliers could generate positive compliance spillovers down the production chain.

A recurring challenge for tax enforcement policy is how to effectively monitor small businesses. Evasion becomes riskier and more costly to firms as they get bigger, because they need sophisticated accounting systems to carry out complex operations. This facilitates the tax agency's task of ob-

taining information from large firms[5] (Kleven, Kreiner, and Saez, 2009). Such information-related constraints on tax evasion are much weaker in the case of small businesses, which represent the vast majority of firms.[6] This is particularly relevant in the Spanish case, where average firm size is small given the country's level of development, and it is an even more crucial issue in developing countries.[7] Since the expected return from a tax audit grows more than proportionally with firm size, Dharmapala, Slemrod, and Wilson (2011) make the theoretical argument that it may be optimal for tax agencies with limited resources to focus all their enforcement efforts on large firms.

The empirical techniques used in this paper draw on the recent bunching literature in public finance. In the seminal paper, Saez (2010) exploits kink points – income thresholds at which the marginal tax rate jumps – to identify taxable income elasticities.[8] We apply the techniques proposed by Kleven and Waseem (2012) to adapt this estimation method to the case of *notches* – income thresholds at which the *average* tax rate jumps. Our setting has two novel characteristics within this literature. First, the LTU generates a notch in the *probability of evasion detection*, rather than the tax rate (which is unaffected), allowing us to study tax enforcement in isolation. Second, the notch is defined in terms of operating revenue, which is not a measure of taxable income. The latter adds an extra step in the empirical estimation, because it requires a separate estimation of the effects on revenue and input costs, as explained above.

Finally, this paper contributes to the extensive literature on firm size distribution and size-contingent policies. This topic has received a lot of attention in Spain because of reports (e.g., LaCaixa, 2012) showing that Spanish and German firms are equally productive after controlling for firm size (measured by number of employees). The implication is that the entire productivity gap between the two countries is due to differences in the firm size distribution. The findings in this paper suggest that the observed firm size distribution in Spain could be substantially distorted by evasion behavior, raising questions about such productivity calculations.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 provides an overview of tax administration in Spain. Sections 4 and 5 present the empirical analysis. Section 6 discusses the efficiency costs of tax enforcement and provides a rough estimation of tax revenue losses due to evasion. Section 7 concludes.

---

[5]On the other hand, large firms tend to spend more resources to hire top accountants and lawyers to maximize legal tax *avoidance*.

[6]The exact percentage depends on the country and the precise definition of what constitutes a small firm.

[7]The share of small firms seems to be positively correlated with the size of the underground economy. Schneider, Buehn, and Montenegro (2010) estimate that the underground economy accounts for approximately 25% of GDP in Greece, Italy and Spain, where the firm size distribution is skewed towards small family firms. This is high compared to about 15% in France and Germany, and less than 10% in the United States, where firms are larger on average.

[8]Several recent studies (Chetty, Friedman, Olsen, and Pistaferri (2011), Chetty, Friedman, and Saez (2012), Bastani and Selin (2012)) apply Saez's method to derive taxable income elasticities using large administrative datasets from Denmark, Sweden and the United States.

2 Theoretical Framework

We model the problem of a profit-maximizing firm that can choose to evade part of its tax liabilities but faces the risk of getting caught with some probability and punished by the tax agency. The basic setup extends the classic individual tax evasion framework (e.g., Allingham and Sandmo, 1972) to firms. We then enrich this framework by endogenizing the probability of detection, making it depend on firm size and on the amount of evasion.

2.1 Setup

Consider an economy with a continuum of firms of measure one. Firms produce good y using inputs m according to the production function $y = \psi f(m)$, where ψ is an idiosyncratic productivity parameter distributed over the range $[\underline{\psi}, \overline{\psi}]$ with probability density function (pdf) $h_0(\psi)$. The production function exhibits positive but decreasing returns to material inputs ($f_m > 0, f_{mm} < 0$). All markets are competitive, so firms purchase inputs at price c and sell all their output at price p (which we normalize to 1 for simplicity). There is no entry or exit of firms, such that in equilibrium all firms with $\psi > \underline{\psi}$ can sustain positive profits.

The government sets a proportional tax t on profits, so after-tax profits are given by $\Pi = (1-t)[\psi f(m) - cm]$. Assuming that tax evasion is not possible, profit maximization yields the standard condition:

$$\psi f_m \left(m^{NoEv} \right) = c \tag{1}$$

where m^{NoEv} is the optimal input use when there is no evasion. Given the definition of y, this defines optimal true production $y^{NoEv} = \psi_i f \left(m^{NoEv} \right)$. The proportional tax on profits does not distort production efficiency in this simple partial equilibrium setting. Firms optimize production as they would without taxation, but they now transfer part of their profits to the government.

Now assume that firms can evade taxes by underreporting their revenue, which reduces their tax liability. Let $u \equiv y - \overline{y}$ denote the amount of revenue underreported, where $\overline{y}$ is reported revenue. We assume that input costs are always reported truthfully, so reported profits are given by $\overline{\Pi} = (1-t)[\overline{y} - cm]$. The tax agency detects tax evasion with probability $\delta \in (0, 1)$, which is endogenously determined as we explain below. We think of δ as the audit probability, and we make the simplifying assumption that evasion is always detected if there is an audit. When evasion is detected, a penalty rate θ is applied on the total amount evaded, and after-tax profits are given by $\Pi^D = (1-t)\Pi - \theta t[\Pi - \overline{\Pi}]$. If no evasion is detected, after-tax profits are $\Pi^{ND} = \Pi - t\overline{\Pi}$.

We can then write expected after-tax profits as follows:

$$\mathbb{E}\Pi = (1-\delta)\,\Pi^{ND} + \delta\Pi^D$$
$$= (1-t)\,\Pi + tu\,[1 - \delta\,(1+\theta)]. \tag{2}$$

2.2 Benchmark Case

Let the probability of detection $\delta = \delta\,(u,m)$ be a continuous and strictly monotonic function of evasion and true input use. We assume that $\delta_m\,(u,m) > 0$ (which implies $\delta_y\,(u,y) > 0$ because the production function is monotonically increasing), to capture the intuition that larger firms are more visible and hence more likely to be audited by the tax agency.[9] Additionally, we assume that $\delta_u\,(u,m) > 0$, which has two important implications. First, firms face a trade-off between the benefits of evasion (lower tax payments) and the increased probability of detection. Second, the tax agency's enforcement strategy is influenced by the reporting behavior of firms. One way to motivate this assumption is to consider commonly used "relative audit rules", under which tax agencies use aggregate information obtained from firms in similar markets to identify suspicious behavior (Bayer and Cowell, 2009). For example, a company operating in a booming industry that reports negative profits is very likely to be audited because it stands out from its peers.[10]

The probability of detection is common knowledge. To ensure that the probability is bounded, we further assume that $\lim_{u \to 0} \delta\,(u,y) = 0$ and $\lim_{u \to y} \delta\,(u,y) = \frac{1}{1+\theta}$. The latter condition implies that the detection technology is not perfect, because even when a firm reports zero revenue there is no certainty that it will be detected. This assumption is also convenient to rule out corner solutions and ensure that all firms have a positive amount of underreporting in equilibrium. The left panel of Figure 1 depicts the relationship between δ and u graphically. We assume that $\delta\,(u,m)$ is locally convex in the neighborhood of y^{LTU}, i.e. $\delta_{uu}\,(u,m|\bar{y} \approx y^{LTU}) > 0$.

Firms simultaneously make production (m) and reporting (u) decisions to maximize expected profits. Optimal conditions for an interior optimum are given by:

$$\psi f_m\,(m^*) = c + u\left[\frac{t}{1-t}\right][1+\theta]\,\delta_m\,(u,m^*) \tag{3}$$

$$1 = [1+\theta]\,[\delta\,(u^*,m) + u^* \cdot \delta_u\,(u^*,m)] \tag{4}$$

Condition (3) is similar to the standard optimality condition (1), but with an additional positive

---

[9]The idea of an endogenous probability of detection that depends positively on the amount evaded was first introduced by Reinganum and Wilde (1985).

[10]Notice that this type of audit rule provides "good" incentives, because firms are better off reporting higher profits in order to avoid tax audits, holding all else equal.

term on the right-hand side. This term accounts for the fact that higher production increases the probability of detection. Since $u \geq 0$ by definition, in an interior optimum we obtain that $m^* < m^{NoEv}$, which implies $y^* < y^{NoEv}$. In the corner solution where $u^* = 0$, condition (3) reduces to (1). Comparative statics are intuitive: optimal input use m^* is larger when (i) its effect on the detection probability is weaker (i.e., $\delta_m(u,m)$ is smaller), (ii) the tax rate t or the penalty θ are lower, and (iii) the equilibrium amount of underreported revenue u^* is smaller.

Condition (4) equates the expected marginal benefit of an additional unit of evasion to the expected marginal cost. Firms optimally choose to underreport sales as long as $\delta (1 + \theta) < 1$, which we assumed above. Comparative statics show that optimal evasion u^* is higher when (i) the penalty rate θ is lower, and (ii) the probability of detection δ is lower. To show the latter, we apply the implicit function theorem (see Appendix A.1).

The analysis above shows that, when enforcement policies respond endogenously to firms' production and reporting decisions, such policies will in turn affect firm behavior. Compared to the situation with no evasion, firms produce less output and engage in revenue underreporting. These results are qualitatively similar to those obtained by Bayer and Cowell (2009) in a model where they explicitly introduce relative audit rules. Since the production and cost functions are the same for all firms, each firm's optimal size in equilibrium depends uniquely on their idiosyncratic productivity level ψ. It can be shown that if the productivity distribution $h_0(\psi)$ is smoothly decreasing in its full domain $\left[\underline{\psi}, \overline{\psi}\right]$, then there exists a density function $g_0(\cdot)$ such that the distribution of firms' operating revenue, $g_0(\overline{y}^*)$ is also smoothly decreasing in its full domain $\left[\overline{y}^*\left(\underline{\psi}\right), \overline{y}^*\left(\overline{\psi}\right)\right]$.[11]

2.3   Policy Intervention: Large Taxpayers' Unit (LTU)

Assume now that the tax agency sets up a Large Taxpayers' Unit that monitors and enforces the taxation of firms with reported revenue higher than y^{LTU}. Dharmapala, Slemrod, and Wilson (2011) provide a theoretical rationale for the existence of this type of institution when the tax agency's resources are limited. In their model, the trade-off between the tax agency's administrative costs of enforcement and its tax collection goals yields an optimal threshold below which firms should be exempted from taxation.[12] They argue that the full exemption for small businesses exists *de facto* in most developing countries via lenient tax enforcement.

The probability of detection is no longer a continuous function of reported revenue. As shown in the right panel of Figure 1, δ now features a discrete jump at y^{LTU}. The detection probability is

---

[11]The specific mapping between the two density functions depends on the functional forms of the production function $f(m)$ and the probability of detection $\delta(u,m)$.

[12]The threshold in Dharmapala, Slemrod, and Wilson (2011) involves changes in both tax liability and enforcement, whereas in our setting only the enforcement intensity changes.

exactly the same as before for firms below the revenue cutoff, but it is now higher for firms above it. Formally, the new detection probability function can be expressed as follows:

$$
\delta = \begin{cases} \delta\left(u,m\right), & \text{if } \overline{y} \leq y^{LTU} \\ \delta^{LTU} \equiv r \cdot \delta\left(u,m\right), & \text{if } \overline{y} > y^{LTU} \end{cases}
$$

where, $r > 1$. We assume that $\delta\left(\cdot\right)$ is locally convex at y^{LTU} such that the optimal conditions (3) and (4) continue to hold for firms with $\bar{y} \leq y^{LTU}$.

The introduction of the LTU generates a "notch" in δ, meaning that the probability of detection increases for *all* inframarginal units evaded when a firm crosses the (reported) revenue threshold. We assume that firms face no optimization frictions (we relax this assumption later), so they can re-optimize to new levels of production and reporting in response to the new policy. The pre-reform and post-reform revenue distributions are depicted in Figure 2, where they are labeled "counterfactual" and "observed" density, respectively, to be consistent with the terminology of the empirical section.

To study the response of different types of firms to the policy change, we define three distinct groups. First, there are low productivity firms, defined as those that report revenue $\overline{y}^{*} \leq y^{LTU}$ in the benchmark case. Nothing changes for these firms with the new policy because they are not LTU-eligible, so their behavior continues to be defined by optimality conditions (3) and (4). We denote by ψ^{L} the productivity level of the firm that chooses exactly $\overline{y}^{*} = y^{LTU}$ in the benchmark case (without LTU). Hence, all firms with $\psi_{i} \in \left[\underline{\psi}, \psi^{L}\right]$ belong to the "low productivity" group.

Second, there is a group of firms whose pre-reform reported revenue was just above y^{LTU}. These firms react to the reform by reporting lower revenue in order to locate exactly, or "bunch", at the LTU threshold, i.e. $\bar{y}^{**} = y^{LTU}$ (we denote the optimal choices in the LTU case with two stars, to distinguish them from optimal choices in the benchmark case, which had one star). This bunching response is a combination of lower production and higher evasion, where the relative importance of each action depends on the functional forms of $f\left(m\right)$ and $\delta\left(u,m\right)$. We define the "marginal buncher" as the firm with the highest exogenous productivity that chooses $\overline{y}^{**} = y^{LTU}$. We denote by ψ^{MB} the exogenous productivity of the marginal buncher. Formally, ψ^{MB} is the unique value that equalizes expected profits when facing the low probability of detection $\left(\delta\right)$ an expected profits when facing the high probability $\left(\delta^{LTU}\right)$:

$$
\mathbb{E}\Pi\left(u^{**}, m^{**} | \psi^{MB}, \delta\right) = \mathbb{E}\Pi\left(u^{**}, m^{**} | \psi^{MB}, \delta^{LTU}\right) \tag{5}
$$

An important point to notice about expression (5) is that the optimal values $\left(u^{**}, m^{**}\right)$ are different under each probability of detection. Given the above definitions, all firms with productivity $\psi \in$

$(\psi^L, \psi^{MB}]$ belong to the group of "bunching firms".

Third, there is a group of high productivity firms, with $\psi > \psi^{MB}$, which are affected by the introduction of the LTU because they now face a higher probability of detection. For these firms, reducing reporting revenue all the way to y^{LTU} is too costly because it involves either inefficiently low production or too much exposure to being detected by the tax agency (or both). The optimality conditions faced by these firms are equivalent to (3) and (4), but with $\delta^{LTU}(u,m)$ instead of $\delta(u,m)$. Hence, these "high productivity" firms re-optimize and report higher revenue than they did in the benchmark case: $\overline{y}^{**}(\psi > \psi^{MB}) > \overline{y}^*(\psi > \psi^{MB}) > y^{LTU}$.

We can sum up the characterization of these three groups of firms as a function of exogenous productivity levels:

- If $\psi_i \in [\underline{\psi}, \psi^L]$, firm i is a Low Productivity Firm

- If $\psi_i \in (\psi^L, \psi^{MB}]$, firm i is a Bunching Firm

- If $\psi_i \in (\psi^{MB}, \overline{\psi}]$, firm i is a High Productivity Firm

Bunching firms are the most important group for our analysis. We use a first-order approximation to relate the number of bunching firms and the reported revenue response of the marginal buncher. For analytical simplicity, consider the case where the LTU raises the detection probability by an arbitrarily small amount $d\delta \equiv \delta^{LTU}(\cdot) - \delta(\cdot)$. In this case, the range of bunching firms would also be arbitrarily small and we can define $d\psi \equiv \psi^{MB} - \psi^L$, which is the difference in exogenous productivity between the marginal buncher and the largest of the low productivity firms. In the benchmark case, we established that there is a direct mapping from the pdf of the productivity parameter, $h_0(\psi)$, to the pdf of reported revenue, $g_0(\overline{y})$. Hence, we can define the excess mass of bunching firms, B, as follows:

$$B = \int_{y^{LTU}}^{y^{LTU}+d\overline{y}} g_0(\overline{y}) \, d\overline{y} \approx g_0(y^{LTU}) \, d\overline{y}^{MB}, \tag{6}$$

where the approximation assumes that the counterfactual density $g_0(\overline{y})$ is approximately flat in the neighborhood of y^{LTU}. The term $g_0(y^{LTU})$ denotes the height of the density distribution at the LTU threshold (in the benchmark case), while $d\overline{y}^{MB}$ is the change in reported revenue for the marginal buncher in response to the introduction of the LTU.[13] Under the strong assumption that firms face no optimization frictions[14], $d\overline{y}^{MB}$ can also be interpreted as the length (in million Euros)

---

[13]In the benchmark scenario, the marginal buncher reported $\overline{y}_0^{MB} = y^{LYU} + d\overline{y}$, but in presence of the LTU this firm reports $\overline{y}^{**} = y^{LTU}$.

[14]We discuss at length the implications of the existence of optimization frictions in Section 4.2.

of the interval were the density is zero, as shown in Figure 2. To be able to estimate this amount, we use (6) to define b as the ratio of excess bunching over the counterfactual density at the threshold:

$$b \equiv \frac{B}{g_0 \left(y^{LTU} \right)} \approx d\overline{y}^{MB} \tag{7}$$

In Section 4.1, we develop an empirical strategy to build a counterfactual distribution and calculate the excess bunching mass in order to estimate b in the data. We will refer to b as the measure of "bunching intensity". In Section 4.2, we relax the assumption of no optimization frictions and define an alternative estimator of b that takes frictions into account.

2.4 Welfare Implications

As noted above, the introduction of a proportional tax on profits did not generate an inefficiency in this framework, but the distortions created by tax enforcement elicit behavioral responses from firms that could lead to efficiency costs. To set up a social welfare function, we make the simplifying assumption that each firm is owned by one individual, whose total income is the after-tax profit of the firm. That way, we can aggregate these individuals' welfare to the tax revenue raised by the government and make meaningful comparisons. In this theoretical framework, an increase in tax enforcement (summarized by the probability of detection, δ) is equivalent to an increase in the expected tax rate. Therefore, we can evaluate how expected welfare changes in response to an increase in tax enforcement in the same way that the literature on the deadweight loss of taxation evaluates the welfare implications of tax changes.

We define expected welfare as the sum of expected profits and expected tax revenue:

$$\mathbb{E}W \left(\delta \right) = \left\{ (1 - t) \Pi + tu \left[1 - \delta \left(1 + \theta \right) \right] \right\} + t \left[\Pi - u \left[1 - \delta \left(1 + \theta \right) \right] \right]$$

By envelope theorem, we can ignore behavioral responses in the term in curly brackets, because firms are already choosing m and u to maximize expected profits. Hence, an increase in tax enforcement leads to the following change in expected welfare:

$$\frac{d}{d\delta} \mathbb{E}W = t \left[-u \left(1 + \theta \right) + \frac{d\Pi}{d\delta} - \frac{du}{d\delta} \left[1 - \delta \left(1 + \theta \right) \right] + u \left(1 + \theta \right) \right]$$

$$= t\frac{d\Pi}{d\delta} - t \left[\frac{du}{d\delta} \left[1 - \delta \left(1 + \theta \right) \right] \right] \tag{8}$$

$$= t\frac{d\overline{\Pi}}{d\delta} + t\frac{du}{d\delta} \left[\delta \left(1 + \theta \right) \right] \tag{9}$$

We know from comparative statics that $\frac{du}{d\delta}|_{u=u^*} < 0$, so the second term in (9) is negative. This

implies that the change in expected welfare due to an increase in enforcement is neither proportional to the elasticity of reported taxable profits $\left(\overline{\Pi}\right)$ nor to the elasticity of true profits (Π), but to an intermediate amount. Formally,

$$t\frac{d\Pi}{d\delta} \leq \frac{d}{d\delta}\mathbb{E}W \leq t\frac{d\overline{\Pi}}{d\delta}.$$

Hence, the efficiency cost of tax enforcement cannot be calculated based solely on the effects on reported profits. The response of true profits has to be taken into account as well. We return to this discussion in Section 6, after presenting our empirical results.

3 Institutional Context and Data

Tax agencies around the world monitor large taxpayers more closely than small ones. This policy is justified because the expected tax revenue recovered is higher when monitoring large taxpayers, even considering that expected enforcement costs per taxpayer (i.e., the cost of conducting tax audits, requesting and processing information) increase with firm size. Most OECD countries have some type of Large Taxpayers' Unit (LTU) dedicated exclusively to monitoring and enforcing taxes on the largest companies (O.E.C.D., 2011). International institutions like the IMF have supported the establishment of LTUs in developing countries over the last 20 years, arguing that they improve enforcement policies and increase tax revenue (I.M.F., 2002). Below, we summarize the key characteristics of the Spanish LTU, the main source of variation used in the empirical strategy of this paper. We also describe a second policy threshold above which firms are required to hire an external audit firm to . Finally, we provide a brief description of the Spanish tax system and compare to the United States.

3.1 Tax Administration Thresholds in Spain

3.1.1 LTU threshold

The Spanish tax agency established a LTU (*"Unidades de Gestión de Grandes Empresas"*) in 1995 to closely monitor tax compliance by the largest firms operating in the country. The threshold to define a "large firm" was set at €6 million[15] in annual operating revenue and has not been modified since then. When a firm reports revenue above the threshold in a given year, it is automatically added to a 'census' of large firms starting the *following* year. Exporters are always included in the LTU, regardless of their total revenue, because they can potentially claim large VAT reimbursements on their exports.

---

[15]The threshold was originally set at 1 billion Pesetas, the official currency at the time. The exchange rate is 166.386 Pesetas per Euro, so the threshold is exactly €6.010121 million (no rounding was applied).

Firms in the LTU census are subject to stricter monitoring and higher compliance requirements. The LTU performs comprehensive tax audits on approximately 10% of large firms each year, while barely 1% of firms below the threshold are audited.[16] In terms of compliance requirements, firms in the LTU census are required to file their value-added tax declarations on a monthly basis (instead of quarterly) and in electronic form (as opposed to on paper).[17] Moreover, the withholding rate on the corporate income tax is 25%, compared to 18% for small firms.[18] To summarize, (i) firms in the LTU are more likely to be audited, (ii) it is easier to cross-check their individual transactions, and (iii) they may face liquidity constraints due to more frequent and higher tax withholding.

Over time, the composition of the LTU Census has changed because the threshold has remained fixed in nominal terms so inflation has brought some small firms above the cutoff, even if they were not growing in real terms. Combined with a 3-percent average annual GDP growth rate, the number of companies in the LTU census increased from 18,860 (2.4% of all registered firms) in 1999 to 40,571 (2.9%) in 2007. Since firms in the LTU are much larger and more productive on average, they report more than 80% of all profits, although they take advantage of more tax deductions, so their share of *taxable* profits is closer to 75% (A.E.A.T., Several years). In the period under study, overall LTU staff stayed essentially constant, but there were substantial technical improvements, so the net change in enforcement intensity over time was likely to be limited. The LTU has one office in each of the 17 Spanish autonomous regions (*Comunidades Autónomas*).

3.1.2 External Audit threshold

Firms above a certain size are required by law to have their annual accounts audited by a private auditing firm. This external audit requirement applies when two of the following criteria are fulfilled for two consecutive years: (i) annual revenue above €4.7 million; (ii) total assets above €2.4 million[19]; and (iii) more than 50 employees on average during the year. These criteria also determine whether a firm has the option to elaborate an abridged version of its accounts, rather than the standard (long) version.

Despite not being implemented directly by the tax agency, the external audit requirement is related to tax enforcement because official tax audits typically use the private auditor's report as a source of information. Private auditors are required by law to provide a "truthful assessment of

---

[16]As reported by AEAT, the tax Spanish tax agency, in its Annual Reports A.E.A.T. (Several years).

[17]A recent reform extended electronic reporting to all firms since July 1st, 2008. This does not apply to the period under analysis in this study, which is 1999-2007.

[18]To be more precise, the withholding rate for firms in the LTU firms is 5/7 of the statutory rate, yielding $35 * \frac{5}{7} = 25\%$ for most firms. For companies below the threshold, the withholding rate is exactly 18% (BOE, Several years). Post-2007 reforms have modified these rates.

[19]As with the LTU threshold, these amounts were established before the adoption of the Euro. The revenue limit was originally 790 million Pesetas (€4.748 million), and the assets limit was 395 million Pesetas (€2.374 million).

the company's accounting", and they face legal responsibility if any misreporting is found. For this reason, auditors are wary to sign an audit report if they find obvious evidence of tax evasion, which limits the ability to evade for audited firms.[20]

The fee charged by private auditors varies with the size of the business and the complexity of its operations. For a firm with revenue close to €4.7 million, the average costs during the period under study was in the range €10,000 - €20,000, a small but non-negligible expenditure (0.2 to 0.4% of total revenue, but 4 to 8% of *profits* on average).

3.2 Overview of the Spanish Tax System (1999-2007)

The Spanish tax system is similar to that of other European countries, with relatively high tax rates compared to the US. The four most important taxes in terms of tax revenue are the payroll tax (PT) individual income tax (IIT), the value-added tax (VAT) and the corporate income tax (CIT). The payroll tax accounted for a stable 33% of all tax revenues in the period 1999-2007, followed by the IIT with 22%, the VAT with 19%, and the CIT with 13% (although with an upward trend in that period going from 10 at the beginning to 15 at the end). The rest is collected through other indirect taxes and fees (Instituto de Estudios Fiscales 2011).

The top marginal tax rate on the individual income tax (IIT) in Spain was 48% in 1999-2002 and then lowered to 46% in 2003-2007, a substantially higher rate than the 35% (30 for small firms) tax rate of the corporate income tax.[21] Thus, unlike in the US, high-income individuals have an incentive to shift taxable income from the IIT base to the CIT base to lower their tax liability. The general VAT rate in Spain during 1999-2007 was 16%, with reduced rates for some goods and services.[22] The VAT is collected by firms at each production stage and then remitted to the government.

Finally, payroll tax (PT) rates are relatively high in Spain, since they are designed to fund generous unemployment and disability insurance programs and the pay-as-you-go pension system. Combining the rates required by each of these programs, the statutory tax rate 31% for employers and 7% for employees, adding up to an overall 38% payroll tax rate.

Penalties for tax evasion vary depending on the size of the infraction. If the amount evaded is above €120,000 (€90,000 prior to 2004), then the taxpayer faces criminal responsibility, whereas if it is below there can only be administrative penalties. There is a great deal of discretion regarding penalties, which according to law could go from 50 to 600% of the amount evaded, depending on the gravity of the offense. Fiscal crimes legally prescribe after four years, which in some cases limits

---

[20]In private conversations, some auditors admit that they tolerate "small" amounts of misreporting, equivalent to about 2-3% of the firm's total operating revenue.

[21]The revenue threshold to benefit from this tax break has changed over time. Details are provided in the Appendix.

[22]There were two reduced rates of 7 and 4% for items like staple foods, medicine and culture-related goods and services. Education and financial services were fully exempted from VAT.

the tax agency's ability to recover fiscal debts because the legal process is too slow.[23]

3.3 Data

We use firm-level data from Amadeus, a comprehensive database of European businesses put together by Bureau van Dijk, a market research company (www.bvdinfo.com).[24] All firms in Spain are required by law to deposit their annual accounts at the Commercial Registry (*Registro Mercantil Central*). Amadeus compiles all these annual reports into a longitudinal database. The information available for each firm in each year includes: business name, location (5-digit postal code), sector of activity at the 4-digit level, 26 balance sheet items, 26 profit and loss account items, and 32 standard financial ratios. This is therefore a very rich panel of firms, a type of dataset that has seldom been used in the public finance literature.

The main advantage of this dataset is that it allows us to analyze firm behavior along several dimensions, e.g. different tax bases, from a single data source. This is not the case when researchers obtain access to administrative data, because these are often anonymized data that cannot be linked to other source. Another important aspect is the panel structure of the dataset, which permits the application of econometric techniques such as fixed-effects estimation.

The dataset has also some limitations. First, a large number of small firms do not fulfill the reporting requirement because it is costly to them and the associated fines are small. Note, however, that submitting the annual accounts is a requirement to obtain loans from commercial banks and government contracts. Amadeus contains information from approximately 85% of firms with annual revenue between €1.5-€60 million that submitted a corporate tax return to the Spanish tax agency. The percentage is close to 90% for firms larger than €60 million, but just below 50% for firms smaller than €1.5 million. Table 2 shows the comparison between the two data sources. This study focuses on firms with revenue between €2-€12 million, so we treat the available data as the quasi-universe of Spanish firms in that size range, which corresponds to a small-medium enterprise size.[25]

A second limitation, common to the corporate tax literature, is that the financial statements may not provide an accurate measure of actual tax liability, because we do not observe the tax deductions applied by each firm to arrive at fiscal profit.[26] To know the exact tax liability, we would need administrative tax return data for all the major taxes, which is not available to researchers. Aggregate data published by the tax agency (A.E.A.T., Several years) shows that the *effective*

---

[23]On the other hand, the tax inspector can request financial statements from the previous four years during an audit, and the company is legally obliged to provide them.

[24]For the purposes of this paper, we accessed the online version of Amadeus in November-December 2011. Since the dataset is continuously updated, the information currently available in the online version may have suffered changes.

[25]Assuming that missing firms are more likely to be tax evaders than those included in Amadeus, the worst possible scenario is that selection bias would make our estimations of tax evasion a lower bound.

[26]The dataset does include a self-reported estimation of corporate income tax liability.

corporate tax rate paid by small and medium firms is higher than that of very large ones (25% vs. 22%), even though the statutory rate is higher for the latter group (30% vs. 35%). This indicates that tax deductions are of second-order importance for the size range we study. The information submitted to the Commercial Registry is essentially the starting point of the tax return, and the amount must match exactly. With these caveats in mind, we consider these data to be almost as good as administrative data for the small and medium-size firms we study in the following sections.

4 Empirical Analysis: Reported Revenue

4.1 Revenue Distribution

We begin with the analysis of the empirical distribution of firms' reported revenue. In the absence of any regulation, our model predicts a smoothly decreasing density distribution, consistent with standard models of firm size determination (e.g., Lucas, 1978) and empirical regularities from comparable countries confirm its validity (see Cabral and Mata (2003) for the case of Portugal). Any bunching of firms at the revenue thresholds described above indicates a behavioral response to tax administration policies.

Using data from Amadeus, Figure 3 shows the revenue distribution for Spanish firms with operating revenue between €3-€9 million in the period 1999-2007. Since the policy thresholds remained constant in nominal terms during this period, we pool nine annual cross-sections to increase the sample size and obtain smoother histograms. We observe two spikes in this distribution: a large one below the LTU threshold, and a much smaller one below the External Audit threshold. These behavioral responses indicate that firms are willing to incur a cost to *report* lower revenue in order to avoid entering the LTU census (and, to a lesser extent, having their accounts audited by an external firm). The bunching pattern is very similar for each individual year, as shown in the annual histograms in Figure 11.

One concern about drawing causal inference from this graphical evidence is that there could be other policies that simultaneously affect these firms' behavior. In particular, there was a corporate income tax break for small firms that had a revenue-based eligibility threshold. In contrast to the LTU, the cutoff was modified every 1-2 years to account for nominal economic growth.[27] The eligibility cutoff was €1.5 million in 1999-2000 and after several modifications it was set at €6 million for the fiscal year 2004, and then raised to €8 million in 2005. We provide full details in the Appendix. The annual revenue distributions in Figure 11 show no discernible bunching at this tax break threshold in any year other than 2004 (when it overlapped with the LTU threshold).

---

[27]There were also political motivations, because extending a small business tax cut is a popular policy.

The fact that the revenue distributions show absolutely no reaction to a 5 percentage-point (14 *percent*) reduction in the corporate income tax rate but firms respond strongly to a change in the tax enforcement intensity (i.e., LTU) is an indication that the latter policy has a substantial effect on firms' decision making.

4.2 Quantifying the Bunching Response

We can use techniques from the bunching literature (in particular, Chetty et al., 2011 and Kleven and Waseem, 2012) to quantify the size of the response in reported revenue using the amount of bunching observed in the distribution. The main idea is to construct a counterfactual revenue distribution to estimate the excess bunching mass near the tax enforcement notch. To do this, we fit a high-degree polynomial to the observed density. We exclude an interval around the threshold, where manipulation is most likely to occur. We then divide the data in small bins[28] and estimate the following polynomial regression:

$$C_j = \sum_{i=0}^{q} \beta_i \cdot (y_j)^i + \sum_{k=y_{lb}}^{y_{ub}} \gamma_k \cdot \mathbb{1}\,(y_j = k) + \eta_j \tag{10}$$

where C_j is the number of firms in bin j (which has width w), q is the order of the polynomial, y_j is the revenue midpoint of bin j, the interval $[y_{lb}, y_{ub}]$ corresponds to the excluded region, and the γ_k's are intercept shifters for each of the bins in the excluded region.

We obtain an estimate of the counterfactual distribution by calculating predicted values from this regression using only the polynomial coefficients (i.e., excluding the γ_k's), such that:[29]

$$\widehat{C_j} = \sum_{i=0}^{q} \beta_i \cdot (y_j)^i \tag{11}$$

This counterfactual density allows us to estimate the excess bunching mass to the left of the threshold (B), and similarly the missing mass to the right of the threshold (H):[30]

$$\widehat{B} = \sum_{j=y_{lb}}^{y^{LTU}} \left(C_j - \widehat{C_j} \right) \text{ and } \widehat{H} = \sum_{j=y^{LTU}}^{y_{ub}} \left(\widehat{C_j} - C_j \right) \tag{12}$$

Determining the lower and upper bounds of the excluded region in a consistent and sensible way

---

[28]All the results reported in the paper use a bin width of €42,070, which allows us to precisely match each policy threshold to the border between two different bins.

[29]In the application of this technique, we add dummies for a short interval below the External Audit threshold, to prevent it from influencing the estimation of the counterfactual density around the LTU threshold.

[30]We switch signs in the definition of $\widehat{H}$ to ensure that the estimator is defined as a positive number.

is critical for this estimation method to provide credible estimates.[31] The lower bound, y_{lb}, can be determined visually because the point where the empirical density ticks up from its downward trend is fairly clear. However, the "correct" value for upper bound, y_{ub}, is less obvious because bunching firms need not come from just above the threshold, and we do not observe a sharp drop in the density above the cutoff. To deal with this issue, we apply the property that the area under a density function must add up to one. In other words, the excess mass to the left of the cutoff must be equal to the missing mass to the right. Thus, we run equation (10) multiple times, starting with $y_{ub} \approx y^{LTU}$ (which generally yields large estimates of $\widehat{B}$ and small estimates of $\widehat{H}$) and increasing the value of y_{ub} in small steps until we reach a level such that $\widehat{B} = \widehat{H}$.

Consistent with the derivation in Section 2.3, we define our estimator of bunching intensity under the assumption of no frictions $\left(\widehat{b}_{NF}\right)$ as the ratio of excess bunching mass over the average height of the counterfactual density in the interval $\left(y_{lb}, y^{LTU}\right]$:

$$\widehat{b}_{NF} = \frac{\widehat{B}}{\left[\frac{1}{1+(y^{LTU}-y_{lb})/w}\right] \sum_{j=y_{lb}}^{y^{LTU}} \beta_i \cdot (y_j)^i}, \tag{13}$$

where the subscript NF stands for "no frictions" and the term $\left[1 + \left(y^{LTU} - y_{lb}\right)/w\right]$ is the number of excluded bins below the threshold.[32] To interpret the estimator $\widehat{b}_{NF}$, we make two assumptions. First, we assume that firms face no optimization frictions. Second, we assume that the smoothly decreasing counterfactual density defined by (11) is a good approximation of the theoretical revenue distribution in the absence of the LTU threshold. Recall form Section 2.3 that, once the LTU is introduced, every firm with productivity in the range $\psi \in \left(\psi^L, \psi^{MB}\right]$ bunches at the cutoff. Hence, if there are no optimization frictions, the density distribution will feature a large spike at y^{LTU} and have zero mass an interval just above y^{LTU}, as depicted in Figure 2. Since this estimation procedure is applied to the universe of Spanish firms rather than a random sample, there is no sampling error and therefore we cannot construct the usual confidence intervals. To test whether the point estimates are statistically significant, we sample the residuals from regression (10) a large number of times (with replacement) to obtain bootstrapped standard errors.[33]

Applying this technique, we obtain $\widehat{b}_{NF} = .086$ (s.e. .004). This implies that, under the strong assumption of no optimization frictions, the marginal buncher reports revenue €86,000 lower than it would if the LTU did not exist. In proportional terms, the marginal buncher reports approximately

---

[31]We follow Kleven and Waseem's (2012) approach to this problem.

[32]Notice that w is the width of the bins used to build the counterfactual. The fact that there is a finite number of bins means that, in practice, we need to impose the weaker condition that the ratio is "close" to one: $\widehat{B} \in \left[0.9\widehat{H}, 1.1\widehat{H}\right]$.

[33]We thank Michael Best for sharing his Stata code to perform this bootstrapping routine. In all the results shown below, we perform 200 iterations to obtain the standard errors, but using a larger number does not affect our results.

1.4% lower revenue than it would in the counterfactual scenario with no LTU.

Quantifying changes in the probability of detection δ is complicated because enforcement strategies include many elements (audit probabilities, the ability to cross-check transactions, etc.) that are extremely hard to measure. This means that we cannot estimate an elasticity of reported income with respect to tax enforcement. Considering that LTUs share many common characteristics across countries, the external validity of our estimates resides in the fact that they are informative about the expected response to a similar LTU with a revenue threshold.

4.2.1 Optimization Frictions

Contrary to prediction of the stylized model without frictions, we do not observe a hole in the distribution just above the LTU threshold – just a small dip. This indicates that some firms adjust their revenue more easily than others. Thus, the monetary interpretation of the bunching measure $\widehat{b}_{NF}$ is not a precise measure of firms' structural response to a change in enforcement. Optimization frictions have been a persistent issue in the bunching literature, because the cost of not re-optimizing is low for individuals in many contexts.[34] Chetty (2012) shows that an adjustment cost equivalent to 1% of total expenditure makes a high intensive-margin individual elasticity compatible with no bunching response. This issue is likely to be less important in the case of notches because the associated cost of inaction grows at a first-order rate with the size of the policy change (Slemrod, 2010; Chetty, 2012).

Even though businesses have more control over their reported income than individuals (excluding the self-employed), there are several reasons why firms might not respond to the existence of the LTU. First, about half of the firms locating just above the cutoff in any given year had previous exposure to the LTU. That is, their revenue had already been above €6 million for at least one year before the year in which we observe them. Second, some firms may not be planning to misreport their activities regardless of the enforcement regime. This could be due to preferences of the manager against tax evasion or perhaps due to inability to evade given some sector characteristics (e.g., government contracts). For these firms, the only consequence of being in the LTU is the additional compliance costs. Third, firms might be unable to control their revenue with precision due to adjustment costs or unexpected shocks. Fourth, as mentioned in the previous section, exporters are always included in the LTU regardless of their revenue, so they do not have incentives to misreport to avoid the LTU.

We illustrate the importance of the first reason with some evidence for growing and shrinking firms. Recall that firms are added to the LTU census the year after their revenues rise above

---

[34]Especially at kink points, where the marginal tax rate jumps but the average tax rate varies smoothly at the kink.

€6 million, and they are taken out one year after their revenues drop below the cutoff. Despite this apparent symmetry, entering the LTU in practice forces some businesses to make important administrative changes to adapt to the higher enforcement regime. For example, they would have to give up having two sets of accounting books. Once the firm puts an end to the parallel accounting system, it is hard to set it up again after dropping out of the LTU census.[35]

To test whether entering the LTU is seen as a fixed cost, we compare the behavior of firms that are growing to those that are shrinking in size. A growing (respectively, shrinking) firm is defined as having higher (lower) revenues in year t than in $t-1$. Figure 5 shows the striking differences in the revenue distributions for these two groups. Growing firms bunch very strongly at the LTU threshold, but barely react to the External Audit threshold. In contrast, shrinking firms do the exact opposite: they bunch in response to the External Audit requirement, but their response to the LTU cutoff is minimal.[36] We conclude that some growing firms avoid the LTU because they anticipate it will involve paying a one-time adjustment cost and it will reduce their ability to evade taxes in future years. In contrast, shrinking firms with previous LTU exposure have less to gain from bunching just below the threshold because they have already incurred the fixed cost.

Rather than introducing each source of rigidity explicitly into the model, we assess their combined impact to an upper bound of the structural response.[37] We define α as the proportion of firms locating in the interval $\left(y^{LTU}, y_{ub}\right]$, compared to the counterfactual density. This includes all firms that do not bunch even though there are firms similar to them (according to our counterfactual) that do bunch. We use this measure to re-weight the estimates of $\widehat{b}_{NF}$. The frictions estimate can be thought of as treatment-on-the-treated estimator for firms with low-enough adjustment costs:

$$\widehat{b}_F = \frac{\widehat{b}_{NF}}{1 - \alpha}$$

We interpret $\widehat{b}_F$ as an upper bound of the firms' response to a change in tax enforcement, since $\widehat{b}_F \geq \widehat{b}_{NF}$ by definition (because $\alpha \geq 0$). We calculate standard errors for this estimator with the same bootstrapping procedure used above.

The estimate taking frictions into account is $\widehat{b}_F = 0.449$ (s.e. 0.044). To provide a sense of the magnitude of this response, consider that the average profit margin of firms around the LTU

---

[35]This is especially true in small regions where there is only a few hundred large firms that are well-known by the local LTU staff. Anecdotally, officers from the tax agency report that marginal firms in these regions often move their headquarters to a large city (e.g., Madrid, the capital) to blend into a larger group of firms.

[36]In a more disaggregated analysis, we observe that the only subset of shrinking firms that features bunching at the LTU threshold is composed of firms with revenue falling between 0% and -3%. However, firms with a revenue decrease of -3% or beyond show no bunching response. There is always some bunching at the External Audit threshold for these two groups.

[37]Kleven and Waseem (2012) propose a similar method to account for optimization frictions, although in their case there is a well-defined dominated region in which no taxpayer should locate under any preferences.

threshold is 4.4% of revenue, approximately €290,000. Thus, the marginal buncher's reduction is equivalent to 30% of reported *profits* in the frictionless case and 154% according to the frictions estimator. These estimates are large, but we cannot interpret them in isolation because we first need to study how input-cost reporting changes around the enforcement notch. We return to this issue in section 5.

4.3   Heterogeneous Responses

We analyze subsamples of firms with particular characteristics related to their incentives to evade taxes (and their ability to do so). All the results from this subsection are reported in Table 3. We group firms based on their size in non-revenue dimensions, e.g. number of employees or the amount of fixed assets owned. Conditional on being the neighborhood of the LTU threshold, firms with more employees and/or assets tend to have a more complex structure, so they need to have sophisticated accounting systems in place that make misreporting more costly and risky. Holding everything else constant, we expect to see the strongest bunching response among smaller firms along these dimensions. Figures 6 and 7 plot the revenue distributions for groups of firms of different sizes in terms of employees and fixed assets. The density distribution is strongly right-skewed for the smallest firms, while it is almost flat for the largest ones. Bunching at the LTU threshold is strongest among firms below the 50-employee and €2.4-million-in-assets marks.[38] The no frictions bunching estimates are in the same order of magnitude for the very small (less than 40 employees) and large firms (more than 50), with $\hat{b}_{NF} \approx .08$. Similar results are found for firms with less than €5 million in assets, but the bunching estimates are much smaller and only marginally significant for firms with more assets.

Table 3 also reports bunching estimates for firms with different legal forms, *"Sociedad Limitada (SL)"* (comparable to Limited Liability Company in the US) and *"Sociedad Anónima (SA)"* (comparable to a Corporation). The capital requirements to set up a SL are smaller than for a SA, but the latter is the natural legal form for publicly traded companies. SL's are more numerous and smaller on average, but we do not find significantly different bunching responses, as can be seen in Figure 12. This can be explained by the fact that both legal forms are treated equally in terms of taxation. Finally, we document some differences in the bunching response across regions, which we report in detail in the Appendix.

Firms in different sectors of the economy face different constraints on misreporting, depending on their technology, e.g. whether they are labor intensive or not. A restaurant with €6 million in revenue is typically a medium-large company with dozens of employees, and most likely with more

---

[38]Notice that these are two of the three eligibility criteria in the External Audit threshold. There are also a number of labor regulations that apply only to firms with more than 50 employees, for example the obligation to have a Workers' Council where unions are represented and acquire some decision power within the company.

than one location. In contrast, a merchant wholesaler that sells electronic products typically reaches that revenue level about 15 employees. To explore how companies operating in such different markets respond to the same nominal revenue threshold, we define 12 different sectors of economic activity (details on how the sectors are defined can be found in the Appendix). We define a "scope of evasion" index at the sector level with two components. First, we obtain the percentage of a sector's output that is sold to final consumers from the input-output tables of the Spanish economy.[39] Second, we calculate the median number of employees for each sector, using only firms with revenue between €5.5-€6.5 million. The intuition for the first element is that selling to final consumers makes underreporting much easier because the VAT self-enforcing mechanism breaks down at that stage. For the second element, the idea is that it is easier to underreport if the number of employees is small. Formally, we compute:

$$Scope = ShareFinalCons * (1 - employees/100)$$

where we normalize the number of employees so that both terms are smaller than one and $Scope \in (0, 1)$.

Figure 9 shows a scatterplot of bunching responses (measured by $\hat{b}_{F-lb}$, vertical axis) and the scope of evasion index (horizontal axis), to illustrate their relationship for different sectors in the economy. There is not an obvious pattern, but it seems clear that a linear fit is not the best way to approximate the relationship. Instead, we have overlaid a quadratic fit that shows an inverted-U pattern. One way to interpret the relationship is to consider the case of extreme data points. In the sector with lowest scope of evasion (restaurants and hotels), the bunching response is very small. This might be explained by the fact that these are "large" firms in terms of employees, which makes misreporting more difficult. On the other extreme are retailers, which have the highest scope for evasion but do not bunch very strongly. One hypothesis could be that when retailers underreport, it is extremely hard even for the LTU to detect it because 85% of sales go to final consumers (and most of them are untraceable cash transactions). Moreover, the median number of employees of retail firms around the revenue cutoff is just 21. Between these two extremes we observe mostly the industrial sectors: manufacturing, wholesale and construction. Focusing only on these sectors, it seems that the relationship between scope of evasion and bunching intensity could be positive, but the small number of data points limits our ability to draw a strong conclusion.

---

[39]The input-output tables are published by the National Statistics Institute (www.ine.es).

5   Empirical Analysis: Input-Cost Reporting

The effects of tax enforcement on reported revenue documented in the previous section only tell half of the story. To characterize the full response of taxable *profits*, we need to measure the effects on input-cost reporting. Since the LTU threshold is determined in terms of revenue, a naïve regression discontinuity approach would be to compare average input costs reported by firms on either side of the LTU threshold. However, that estimation strategy would violate the key assumption required by regression discontinuity designs because firms can manipulate the running variable – reported revenue.

To deal with this problem, we estimate a model with firm fixed effects to control for time-invariant firm characteristics. By capturing only within-firm variation, this strategy may attenuate the estimated effect of enforcement because it cancels out the effects on firms that remain always below (or above) the threshold.

5.1   Labor and Nonlabor Input Costs

Rather than analyzing the impact on all input costs grouped together, we break them up in two groups that receive different tax treatment and could potentially respond in opposite directions: labor and nonlabor inputs. Nonlabor expenses (e.g., materials and services used for production) are deductible under the value added tax (VAT) and the corporate income tax (CIT), so overreporting them unambiguously lowers the amount of VAT and CIT remitted to the government.[40] In contrast, labor expenses (employee compensation) cannot be deducted from the VAT, and they are taxed through the payroll tax. Underreporting labor inputs lowers the amount of payroll tax remitted, but it increases tax liability on the CIT. The payroll tax rate during the period under study was 38%[41] and the corporate income tax rate was 35%. Therefore, underreporting labor inputs is profitable at the margin, but the incentive is not as strong as in the case of nonlabor inputs.

The left panels of Figure 10 plot average labor and nonlabor input costs in year $t + 1$ against revenue in year t.[42] The graphs show unconditional bin averages with 95% confidence intervals, and pools data for 1999-2007 as before. There is no adjustment for inflation because the outcome variable is a ratio of two nominal amounts. Labor inputs as a share of revenue decrease as firms get bigger, but there is an upward jump exactly at the LTU threshold from about 16 to 17% of revenue.

---

[40]For example, firms could include personal expenditures of the owner into the company books. There are many anecdotal reports of this practice with durable goods like automobiles (intended for personal use) and also with large social events such as weddings.

[41]This tax rate adds up the statutory rates assigned to employers (31%) and employees (7%). Since wage negotiations usually focus on the net-of-tax wage, this is the most natural way of thinking about the overall payroll tax rate, without getting into the question of who bears the tax burden.

[42]Recall that revenue in year t determines LTU eligibility in year $t + 1$.

Nonlabor inputs feature a similar but inverted pattern: their relative importance increases with size, but there is a downward jump at the threshold from 75.5 to 74.5%.

There are two possible interpretations of these findings. First, it could be that firms respond to the differential tax incentives by overreporting nonlabor inputs and underreporting labor expenses. A second interpretation is that labor-intensive firms are less likely to bunch below the LTU threshold, which mechanically yields lower average labor inputs in the bunching interval. In other words, the discontinuities at the threshold could be explained purely by sorting, rather than being a behavioral response to the different enforcement regimes. While the first interpretation is broadly consistent with the incentives for evasion, the second is consistent with the heterogeneous responses documented in section 4.3. The two interpretations are observationally equivalent, so we need an additional test to ascertain which hypothesis is more plausible.

As a first test, we consider reporting behavior in prior years for firms locating near the threshold. The right panel of Figure 10 shows the average reported input costs in years $t-1$, t, and $t+1$ (vertical axis) against reported revenue in year t (horizontal axis). Since reported revenue in t determines LTU eligibility in year $t+1$, the outcome variable in $t+1$ may be influenced by two factors: enforcement intensity faced by the firm (determined by whether $\overline{y}_t \lessgtr y^{LTU}$) and selection effects due to sorting. Meanwhile, the outcomes in year $t-1$ *for those same firms* are only affected by the sorting effect.[43] The patterns for both labor and nonlabor inputs costs in year $t-1$ and $t+1$ are almost identical, and in particular the discontinuities at the threshold are essentially the same. This suggests that the discontinuity in year $t+1$ is most likely due to sorting.

5.2 Fixed-Effects Regressions

A standard approach to overcome the sorting bias described above is to use firm fixed effects to control for time-invariant firm characteristics, both observable and unobservable. We test the effect of a change in enforcement intensity on reported input costs by fitting a fixed-effects model to our nine-year panel dataset. This strategy identifies within-firm effects based on the behavior of firms that switch their enforcement regime at some point. We introduce year dummies to capture year-specific shocks. The model we estimate can be written as follows:

$$z_{it} = \alpha + \phi_i + \lambda_t + \rho \cdot f\left(y_{it} - y^{LTU}\right) + \beta \cdot \mathbb{1}\left[y_{i,t-1} \leq y^{LTU}\right] + \varepsilon_{it}, \tag{14}$$

where z_{it} is the expenditure item (labor or nonlabor input costs as a percentage of revenue) for firm i in year t, ϕ_i denotes firm fixed-effects, λ_t denotes year fixed-effects, $f\left(y_{it} - y^{LTU}\right)$ is a polynomial

---

[43]We avoid using year t as a comparison group for year $t+1$ because in that period is it more plausible that firms could modify their behavior based on anticipating their future enforcement regime.

in the distance to the revenue threshold, and $\mathbb{1}\left[y_{i,t-1} \leq y^{LTU}\right]$ is an indicator for whether firm i will be in the low enforcement regime in year t. We cluster the standard errors at the firm level.

The results are reported in Table 4. The coefficient on the low enforcement dummy is negative and significant (-0.8 percentage points) in column 1, where simple averages are compared. Once we add firm fixed-effects, the coefficient becomes smaller than 0.11 percentage points in absolute value (columns 2-4) regardless of the functional form of the revenue polynomial (none, linear, or quadratic). The largest confidence interval for the effect of low enforcement on reported nonlabor costs is $(-0.4, 0.3)$. Considering that average nonlabor costs are 74.1% of revenue, these effects are both statistically insignificant and economically small. The results for labor inputs are similar, but with opposite sign. The coefficients in columns 7 and 8 are very close to zero (0.03 and 0.08 percentage points, compared to an average of 16.2%) and statistically insignificant.

One limitation of this estimation strategy is that it may attenuate the estimated effects of enforcement on reported input costs. The reason is that all variation due to firms remaining always below (or above) the LTU threshold is netted out by the fixed effects. If reported revenue is persistent or, in other words, if there are many firms staying permanently on one side, then the size of this attenuation bias could be large. We evaluate the persistence of bunching behavior by looking at how many years firms locate in the bunching region. There are 39,424 firms that locate in the bunching interval $\bar{y} \in (5.6, 6.0)$ for at least one year in the 1999-2007 period. Of those, 25% are observed in that interval in two different years, and 9% in three or more years. As a comparison, 23,701 firms are ever observed in the interval $\bar{y} \in (6.0, 6.4)$, of which 17% stay in that interval for two years, and 5% for three or more years. This indicates that there is some persistence in bunching behavior, but not enough to bias our results by an order of magnitude. As a robustness check, we are currently trying to obtain data on electricity consumption for some of the manufacturing firms in our dataset to test whether this conclusion is correct. The idea is to compare the evolution of reported revenue and energy consumption over time. We would expect to see a divergence in these two magnitudes for bunching firms, with energy consumption growing faster than reported revenue as firms approach the LTU threshold (controlling for output prices and other sector-specific characteristics).

6   Efficiency Costs and Tax Revenue Losses

The empirical results obtained in the previous sections indicate that firms respond to the tax enforcement threshold mostly by underreporting their operating revenue, without reducing actual production in a significant way. Drawing from the literature on the deadweight loss of taxation in the presence of evasion and avoidance, we provide an upper bound for the efficiency costs of tax

enforcement. In the final subsection, we perform a back-of-the-envelope calculation of the losses in tax revenue due to evasion in the low enforcement regime, using the estimates from Sections 4 and 5.

6.1 Efficiency Costs of Tax Enforcement

A crucial question for the design of tax administration policies is whether there are large efficiency costs from tax enforcement. We already laid the ground for this estimation in Subsection 2.4, where we derived expressions for the change in expected welfare associated to an increase in tax enforcement. In our theoretical framework, an increase in tax enforcement is equivalent to an increase in the expected tax rate.

In two influential papers, Feldstein (1995, 1999) argued that the elasticity of taxable income with respect to tax rate changes is a sufficient statistic to estimate the excess burden of taxation. This result is useful because it accounts for all the key behavioral responses to taxation (labor supply, avoidance, and evasion), and also because taxable income data is widely available. The key assumption driving Feldstein's result is that tax evaders equate the marginal private cost of evasion (or avoidance) to the marginal cost of reducing true income (by producing less), so that the specific reason why they report lower income does not matter for efficiency.

Chetty (2009) points out that Feldstein's result implicitly assumes that the marginal social cost of evasion and avoidance differs from the tax rate (the marginal benefit). Chetty considers two types of sheltering costs (where sheltering includes both evasion and avoidance). First, there are "resource costs" that make production less efficient when there is evasion. For example, the need to have accountants keeping two different books, or the lost profits for operating in cash. If evaders only incur a resource cost, then Feldstein's result holds. A conceptually different cost is what Chetty calls "transfer costs", for example a monetary penalty to punish evasion behavior. A penalty has a private cost to the evader, but no social cost because the resources are transferred to the government another agent (assuming risk neutrality, as is standard when modeling firm behavior). Chetty shows that the excess burden of taxation in the presence of such transfer costs is directly proportional to the elasticity of total earned income (as opposed to taxable income).

The theoretical framework presented in Section 2 is slightly more complicated than Chetty's because the probability of detection depends on firm size, besides the amount evaded. In spite of this difference, equations (8) and (9) deliver a qualitatively similar result: when firms face only transfer costs of evading, the deadweight loss generated by an increase in tax enforcement is less than proportional to the effect on reported profits. The lack of a real production response to the existence of the LTU implies that the efficiency cost of increasing tax enforcement is not high in this

context. The effect of this enforcement policy is to redistribute resources from tax-compliant firms (or other taxpayers) to tax evaders.

One aspect that we have not addressed so far is the administrative cost of higher enforcement. This could be easily introduced in the model with the function $q(\delta)$, which is increasing in δ. The modified equations (8) and (9) would be:

$$\frac{d}{d\delta}\mathbb{E}W = t\frac{d\Pi}{d\delta} - t\left[\frac{du}{d\delta}\left[1 - \delta\left(1+\theta\right)\right]\right] - \frac{dq}{d\delta} \tag{15}$$

$$= t\frac{d\overline{\Pi}}{d\delta} + t\frac{du}{d\delta}\left[\delta\left(1+\theta\right)\right] - \frac{dq}{d\delta} \tag{16}$$

It is challenging to obtain measures of the marginal increase in administrative costs associated to an increase in tax enforcement. We only have access to the total cost of the tax agency in Spain, which was €1.33 billion in the year 2007, when it raised €250 billion in tax revenue (€188 collected per €1 spent or, equivalently, 0.5 cents of a Euro per each Euro collected).[44] The marginal return to spending an additional Euro on enforcement is likely to be below this average return, but also well above an additional Euro in tax revenue. To guide a cost-benefit calculation to determine what is the socially optimal enforcement intensity, we need to consider the tax revenue lost due to low enforcement.

6.2 Estimates of Lost Tax Revenue

In section 4, we showed that the marginal buncher reduces its reported revenue by 1.4% in the case of no optimization frictions, and between 7.5% once frictions are taken into account. These effects are precisely estimated and statistically significant. In Section 5, we found that the response along the input-cost margin was close to zero both for labor and nonlabor inputs. The estimates are less precise, but the confidence intervals bound the effects at ±0.4 percentage points of revenue at most. Since the effects on input costs are very small compared to those on revenue, we can make a back-of-the-envelope calculation of total tax revenue lost using only the revenue response.

To perform these calculations, we make a number of assumptions. First, we extrapolate our local estimates for firms near €6 million in revenue to all firms with smaller reported revenue. This implies assuming that firms below the LTU threshold conceal the same percentage of their revenue as the marginal buncher. Concealing revenue is in theory easier for small firms because they have simpler operations and fewer employees than large firms. Thus, we think this assumption makes the extrapolation of local results acceptable. Second, we assume that the marginal corporate income

---

[44]The Internal Revenue Service (IRS) of the United States, considered one of the most efficient tax agencies in the world, collected $2.4 trillion in 2007, with an administrative cost of $10.7 billion. Hence, the IRS collects $224 per $1 spent, higher than the Spanish tax agency, but in the same order of magnitude. (Source: www.irs.gov).

tax rate is 30% for all non-LTU firms (even though the actual rate was 35% for some of these firms in the early years of the 1999-2007 period).

The thought experiment we propose is the following: what would be the tax liability of non-LTU firms if they reported their true revenue? We define true revenue as actual reported revenue plus the percentage underreported, according to our estimates from section 4. We use data from Amadeus for the year 2005 to perform the calculations (the results are similar for other years). In that year, the corporate income tax raised €38 billion (4.2% of GDP), of which €8.3 billion came from non-LTU firms, according to official statistics from A.E.A.T. (Several years). There are 553,956 firms in Amadeus with revenue below €6 million, of which 64.8% report positive profits. In the official statistics, there are 1.1 million firms in that revenue range, of which 42% report positive profits. Applying the statutory 30-percent tax rate on the firms with positive profits in Amadeus and summing over firms yields a total tax liability of €7.98 billion. The number is quite close to the official statistics, despite the fact that Amadeus has only half the number of firms. This indicates that the firms missing from Amadeus are mostly tiny firms with little or no declared profits, which favors the validity of our results.

We calculate the "true" revenue of each firm using our three estimates (no frictions and lower/upper bound with frictions). Given our results from Section 5, we keep input costs constant. The results are reported in Table 5. The first point to notice is that the proportion of firms reporting positive profits prices to 75.1% (in the case of no frictions) and up to 83.4 (in the upper bound with frictions). The increase in overall tax revenue from the corporate income tax is €1.5 billion (0.17% of GDP) using the no frictions estimate, and €8.5 billion (0.95% of GDP) using the estimate that accounts for frictions.

These numbers are reasonably large considering that they almost double (in the upper-bound case) the total tax revenue raised from non-LTU firms in practice. To put them in perspective, the budget deficit of the Spanish government in the year 2011 was 9.3% of GDP, so eliminating corporate tax evasion and avoidance among non-LTU firms could reduce the deficit by up to 10%.

7   Concluding Remarks

The findings of this paper indicate that firms respond strongly to the change in tax enforcement induced by the Spanish LTU, reducing their reported revenue by between 1.4% and 7.5%. This sizeable behavioral response is driven mostly by underreporting rather than a reduction in output, implying that the efficiency losses associated with tax enforcement policies are likely to be modest. In parallel, a rough estimation indicates that the losses in tax revenue due to this behavior could

be substantial. A natural question to ask is whether the current tax enforcement policy is optimal from the perspective of social. A proper cost-benefit calculation would involve putting weights on different economic agents, so we leave that normative exercise in the hands of policy makers.

We have abstracted from evaluating whether having a clear-cut revenue threshold to delimit two enforcement regimes is an optimal policy from the mechanism design point of view. An intuitive alternative would be to make the threshold "fuzzy". That is, the transition to the high enforcement regime would happen over a revenue interval, rather than a fixed level. In practical terms, this would apply mostly to audit rates, but not to other administrative LTU elements such as extra compliance requirements. With a fuzzy threshold, the firms' decision would be more uncertain and, at the same time, less salient because there would be no point at which the perceived probability of detection changes dramatically.

A second option is to expand the LTU by lowering the threshold so that more firms fall under its jurisdiction. This would have high short-term returns, because the LTU would now target bunching firms, which are the most likely to be evading taxes. The fact that the threshold has not been updated at least to account for inflation in more than 15 years is the probably the worst possible option. On the one hand, it gives firms prone to evading strong and clear incentives to manipulate their reporting to remain in the low enforcement regime. On the other hand, it makes the number of firms in the LTU increase mechanically over time due to nominal growth (while the human resources available to the LTU remain essentially constant). We hope that the results from this paper can be used to improve the design of the LTU in Spain and other countries.

References

A.E.A.T. (Several years): "Memoria de Administración Tributaria," Discussion paper, Agencia Estatal de Administración Tributaria (Spanish Tax Agency).

ALLINGHAM, M., AND A. SANDMO (1972): "Income tax evasion: A theoretical analysis," *Journal of Public Economics*, 1, 323–338.

ANDREONI, J., B. ERARD, AND J. FEINSTEIN (1998): "Tax Compliance," *Journal of Economic Literature*, 36, 818–860.

BASTANI, S., AND H. SELIN (2012): "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule," CESifo Working Paper No. 3865.

BAYER, R., AND F. COWELL (2009): "Tax compliance and firms' strategic interdependence," *Journal of Public Economics*, 93, 1131–1143.

BOE (Several years): "Ley de Presupuestos Generales del Estado," Boletín Oficial del Estado.

CABRAL, L., AND J. MATA (2003): "On the Evolution of the Firm Size Distribution: Facts and Theory," *American Economic Review*, 93(4), 1075–1090.

C.B.O. (2011): "Budget and Economic Outlook: Historical Budget Data," Congressional Budget Office.

CHETTY, R. (2009): "Is the Taxable Income Elasticity a Sufficient Statistic to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy*, 1(2), 31–52.

——— (2012): "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80(3), 969–1018.

CHETTY, R., J. FRIEDMAN, T. OLSEN, AND L. PISTAFERRI (2011): "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 126, 749–804.

CHETTY, R., J. FRIEDMAN, AND E. SAEZ (2012): "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings," NBER Working Paper 18232.

CHETTY, R., A. LOONEY, AND K. KROFT (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99(4), 1145–1177.

CHRISTENSEN, K., R. CLINE, AND T. NEUBIG (2001): "Total Corporate Taxation: "Hidden", Above-the-Line, Non-Income Taxes," *National Tax Journal*, LIV(3), 495–506.

DHARMAPALA, D., J. SLEMROD, AND J. D. WILSON (2011): "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs," *Journal of Public Economics*, 95(9-10), 1036–1047.

FELDSTEIN, M. (1995): "The Effect of Marginal Tax Rates on Taxable Income: A Panely Study of the 1986 Tax Reform Act," *Journal of Political Economy*, 103(3), 551–572.

——— (1999): "Tax Avoidance and The Deadweight Loss of Income Taxation," *Review of Economics and Statistics*, 81(4), 674–680.

I.E.F. (2011): "Evolución del Sistema Fiscal Español: 1978-2010," Instituto de Estudios Fiscales.

I.M.F. (2002): "Improving Large Taxpayers' Compliance: A Review of Country Experience," Team led by Katherine Baer.

KLEVEN, H., K. KREINER, AND E. SAEZ (2009): "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries," NBER Working Paper 15218.

KLEVEN, H., AND M. WASEEM (2012): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," Working Paper, LSE.

LACAIXA (2012): "Â¿Por Qué No Aumenta la Dimension de la Empresa Española?," .

LUCAS, R. (1978): "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 9(2), 508–523.

O.E.C.D. (2011): "Tax Administration in OECD and Selected non-OECD Countries: Comparative Information Series (2010)," Discussion paper, Organisation for Economic Co-Operation and Development.

PAULA, Á. D., AND J. SCHEINKMAN (2010): "Value-Added Taxes, Chain Effects, and Informality," *American Economic Journal: Macroeconomics*, 2, 195–221.

POMERANZ, D. (2011): "No Taxation Without Information: Deterrence and Self-Enforcement in the Value-Added Tax," Harvard University.

REINGANUM, J., AND L. WILDE (1985): "Income Tax Compliance in a Principal-Agent Framework," *Journal of Public Economics*, 26, 1–18.

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

SCHNEIDER, F., A. BUEHN, AND C. MONTENEGRO (2010): "Shadow Economies all over the World: New Estimates for 162 Countries from 1999 to 2007," Working Paper 322, Universidad de Chile.

SLEMROD, J. (2004): "The economics of corporate tax selfishness," *National Tax Journal*, LVII(4), 877–900.

———— (2008): "Does it Matter Who Writes the Check to the Government? The Economics of Tax Remittance," *National Tax Journal*, LXI(2), 251–275.

———— (2010): "Buenas Notches: Lines and Notches in Tax System Design," Working Paper, University of Michigan.

Tables

Table 1: Comparison of Tax Systems: Spain vs. United States

	Spain		United States	
	top tax rate	*share of revenue*	*top tax rate*	*share of revenue*
Social Security Contributions (SSC)	38%	33%	15.3%	36%
Individual Income Tax (IIT)	48% (46%)	22%	38.6% (35%)	46%
Value-Added Tax (VAT)	16%	19%	-	-
Corporate Income Tax (CIT)	35%	13%	35%	11%
Other indirect taxes and fees	-	13%	-	7%
Federal Tax Revenue / GDP		34%		18%

Sources: for Spain, Instituto de Estudios Fiscales (I.E.F., 2011). For the United States, CBO Historical Tables (C.B.O., 2011). The data for the United States does not include State nor Local-level tax revenues (for example, retail sales taxes are not included), whereas the Spanish data do include regional-level revenues, which partly explains the large difference in the share of Federal tax revenue to GDP.

Table 2: Amadeus Dataset Compared to Official Statistics

	€0 - €1.5 million	€1.5 - €6 million	€6 - €60 million	€60+ million
Official Statistics	994,644	92,261	29,075	2908
Amadeus	440,089	78,474	24,240	2618
%	44.2%	85.06%	83.37%	90.02%

Source: the official statistics on the number of firms that submitted a corporate income tax return in the year 2004 are taken from annual reports of the Spanish tax agency (A.E.A.T., Several years). The data from Amadeus also refer exclusively to the year 2004. The Amadeus dataset is described in detailed in section 3.3.

Table 3: Estimates of Bunching on Reported Revenue

	No Frictions $\left(\widehat{b}_{NF}\right)$	Frictions $\left(\widehat{b}_{F}\right)$
Aggregate Results		
All firms	.086	.449
	(.004)	(.044)
By Number of Employees		
$1-25$ *employees*	.078	.309
	(.006)	(.040)
$26-40$ *employees*	.084	.462
	(.006)	(.069)
$41-50$ *employees*	.128	1.167
	(.013)	(.808)
More than 50 *employees*	.078	.745
	(004)	(.116)
By Fixed Assets		
$0-0.6$ *million Euros*	.089	.352
	(.006)	(.042)
$0.6-2.4$ *million Euros*	.097	.638
	(.005)	(.084)
$2.4-5$ *million Euros*	.084	.528
	(.006)	(.075)
More than 5 *million Euros*	.019	.209
	(.006)	(.098)
By Organizational Form		
SA (Corporation)	.084	.585
	(.004)	(.066)
SL (L.L.C.)	.088	.381
	(.005)	(.039)
By Growing Trend		
Growing	.106	.524
$\left(\overline{y}_t > \overline{y}_{t-1}\right)$	(.004)	(.040)
Shrinking	.015	.940
$\left(\overline{y}_t < \overline{y}_{t-1}\right)$	(.014)	(1.674)

Note: this table reports the bunching intensity estimates for different subsamples of firms: by number of employees, by fixed assets, by organization form, and by growing trends. In the latter, $\overline{y}_t$ stands for reported revenue in year t.

Table 4: Input-Cost Reporting: Fixed-Effects Estimates

	Nonlabor Inputs (% of Revenue), year t				Labor Inputs (% of Revenue), year t			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Low Enforcement	-0.803***	0.094	-0.035	-0.106	1.117***	0.296***	0.027	0.078
$\left(y_{i,t-1} \leq y^{LTU}\right)$	(0.161)	(0.088)	(0.110)	(0.155)	(0.120)	(0.061)	(0.075)	(0.108)
Mean Dep. Var.	74.08	74.08	74.08	74.08	16.21	16.21	16.21	16.21
Revenue Polynomial	none	none	linear	quadratic	none	none	linear	quadratic
Firm Fixed-Effects	no	yes	yes	yes	no	yes	yes	yes
Year Fixed-Effects	no	no	yes	yes	no	no	yes	yes
Observations	176,649	176,649	176,649	176,649	200,062	200,062	200,062	200,062
Clusters	52,170	52,170	52,170	52,170	58,397	58,397	58,397	58,397
R-squared	0.00	0.91	0.91	0.91	0.00	0.92	0.93	0.93

Note: *** = 1%, ** = 5%, and * = 10% significance level. Standard errors are clustered at the firm level. The table shows the results from estimating the following fixed-effects model

$$z_{it} = \alpha + \phi_i + \lambda_t + \rho \cdot f\left(y_{it} - y^{LTU}\right) + \beta \cdot \mathbb{1}\left[y_{i,t-1} \leq y^{LTU}\right] + \varepsilon_{it},$$

where z_{it} is the expenditure item (labor or nonlabor inputs as a share of total revenue) for firm i in year t, ϕ_i denotes firm fixed-effects, λ_t denotes year fixed-effects, $f\left(y_{it} - y^{LTU}\right)$ is a polynomial of the distance to the threshold, and $\mathbb{1}\left[y_{i,t-1} \leq y^{LTU}\right]$ is an indicator for whether firm i will be in the low enforcement regime in year t.

Table 5: Lost Tax Revenue Calculations

Revenue Measure:	Amadeus As reported	No frictions Reported + 1.4%	Frictions Reported + 7.5%
Observations	553,956	553,956	553,956
% with Positive Profits	64.8%	75.1%	83.4%
Average Tax Liability (million €)	0.014	0.017	0.030
Total Tax Liability (million €)	7,988.5	9,508.5	16,578.9
Difference (million €)	–	1,520.0	8,590.4
Difference (% of tax revenue)	–	0.48%	2.65%
Difference (% of GDP)	–	0.17%	0.95%

Note: this table summarizes the calculations of lost tax revenue in the low enforcement regime for the year 2005. The first column shows the actual observations from Amadeus. The next two columns present the results of creating a new reported revenue measure equal to actual reported revenue plus a percentage based on the bunching estimates: 1.4% (no frictions estimate), and 7.5% (frictions estimate).

Figures

Figure 1: Theoretical Probability of Detection

(a) Benchmark Case

(b) Introduction of the LTU
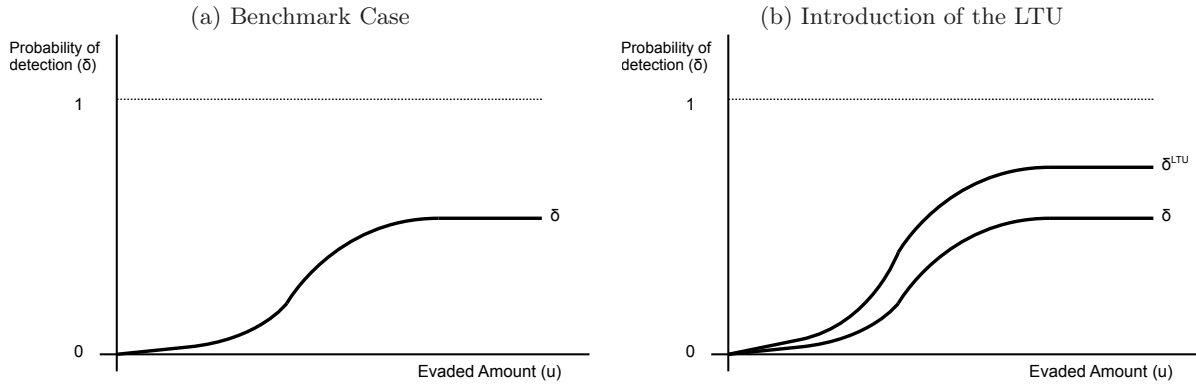
Note: Panel (a) shows the benchmark case where the probability of evasion detection grows with the amount evaded, but does not vary for different levels of *reported* revenue, so all firms face the detection probability $\delta\left(u, m\right)$ regardless of their reported revenue $\overline{y}$. Panel (b) depicts the situation after the LTU is introduced, when firms that report revenue $\overline{y} > y^{LTU}$ face the probability of detection to $\delta^{LTU}\left(u, m\right) \equiv r \cdot \delta\left(u, m\right)$, where $r > 1$. Firms with reported revenue $\overline{y} \leq y^{LTU}$ face $\delta\left(u, m\right)$ as in the benchmark case.

Figure 2: Theoretical Revenue Distributions
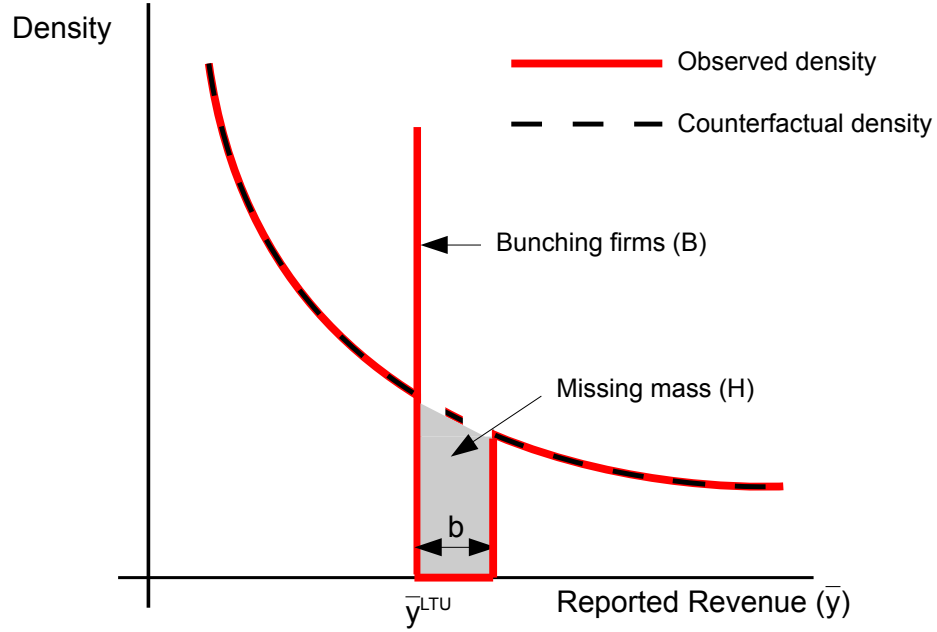
Note: this figures depicts the theoretical revenue distributions, with and without the discontinuity in tax enforcement at the revenue cutoff y^{LTU}. The dashed (black) line shows the counterfactual distribution of revenue in the benchmark scenario when the probability of detection is the same for all levels of reported income. The solid (red) line shows how the theoretical distribution of reported revenue changes when the LTU is introduced in the absence of optimization frictions. A group of firms in an interval above y^{LTU} respond to the new policy by underreporting more of their revenue to report exactly $\bar{y} = y^{LTU}$. This generates a spike at the threshold (with excess mass B), and an area of missing mass (H) to the right of the threshold. The length of the interval where bunching firms originated from equals the bunching parameter $\widehat{b}_{NF}$, defined in Section 4.1.

Figure 3: Operating Revenue Distribution

Note: pooled 1999–2007 data
N=288111
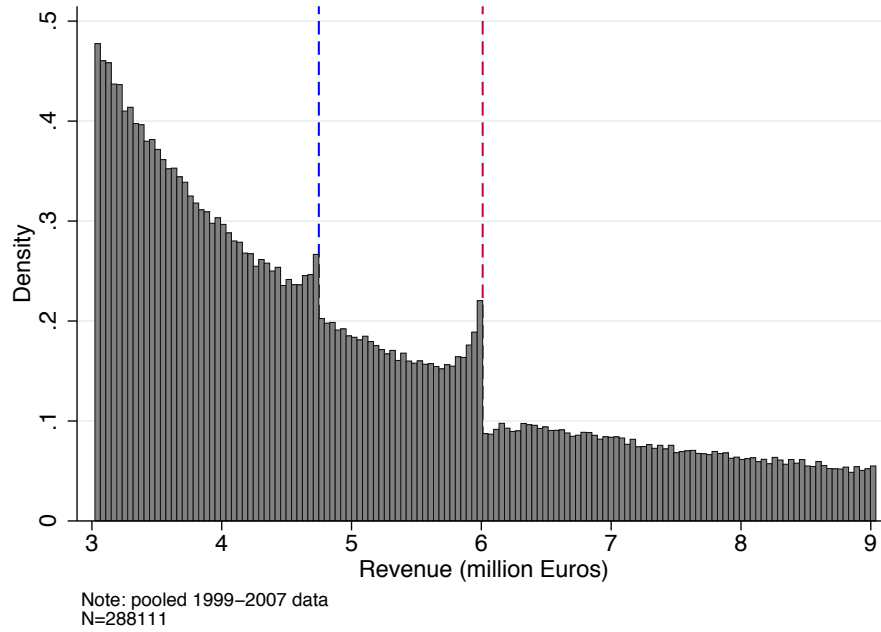
Note: this histogram pools data from the nine years between 1999-2007. The first dashed line from the left (blue) indicates the External Audit threshold, set at €4.7 million in nominal terms. The second dashed line (red) indicates the LTU threshold, set at $\overline{y}$ =€6 million in nominal terms. The bins are €42,000 wide and defined in such a way that both thresholds match bin limits precisely.

Figure 4: Counterfactual Revenue Distribution
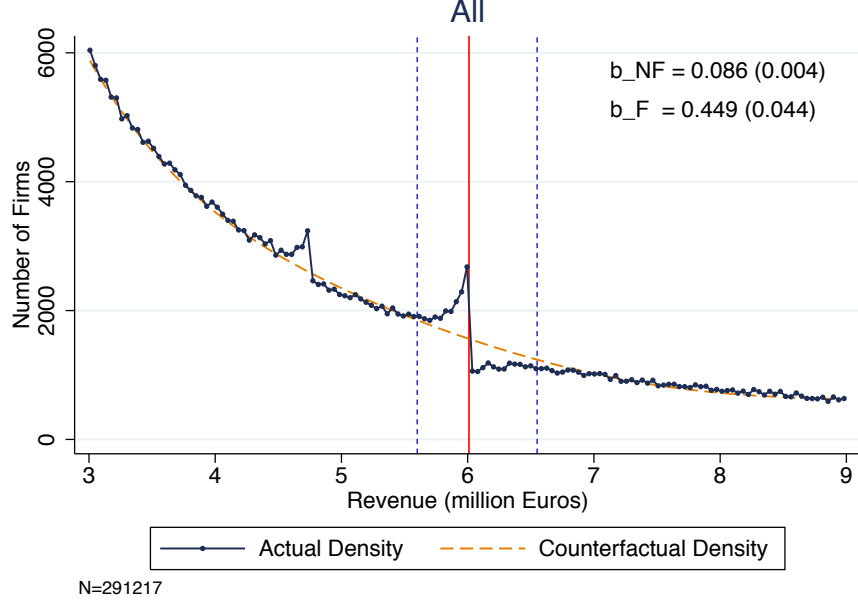
Note: the data used in the actual density plot (dark blue line connecting dots) are the same as those used for the histogram in Figure 3. "b_NF" denotes the estimate of bunching intensity derived under the assumption of no optimization frictions (b_{NF}). "b_F" denotes the estimate that takes into account the existence of frictions (b_F). The orange dashed line is a counterfactual density distribution estimated by fitting a polynomial regression to the density points, excluding the interval between the two vertical dotted (blue) lines. The first dotted blue line from the left denotes y_{lb} and is fixed at €5.6 million, while the second denotes y_{ub}. To determine the value of y_{ub}, we fit the polynomial regression multiple times, starting with $y_{ub} \approx y^{LTU}$ and then increasing the value in small increments until we reach a point where the bunching mass (B) equals the missing mass (H). This way, the area under the counterfactual density adds up to one.

Figure 5: Growing vs. Shrinking Firms
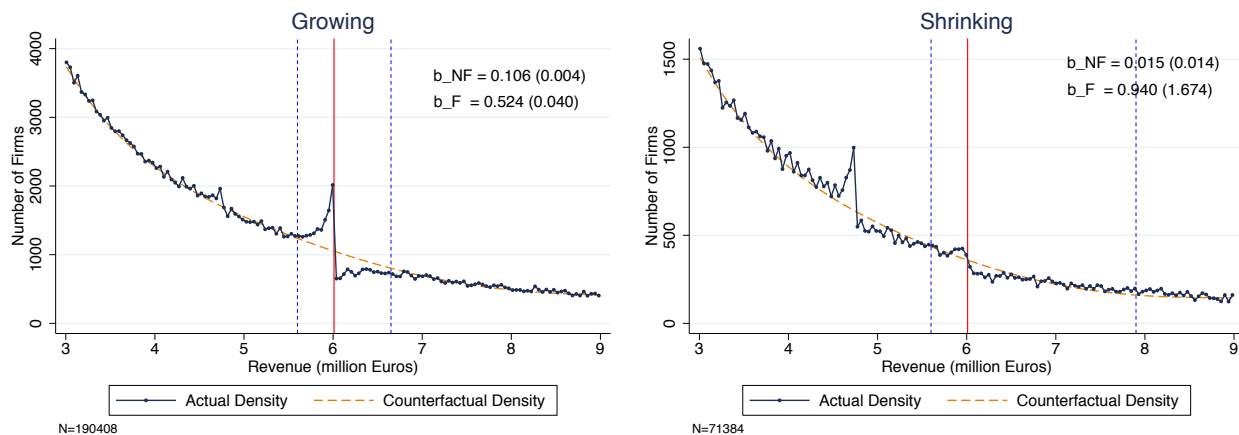
Note: these graphs show the actual and counterfactual revenue distributions for subsamples of firms: those that are growing and those that are shrinking. A firm is defined as growing if its reported revenue in year t is higher than in year $t-1$, i.e. $\overline{y}_t > \overline{y}_{t-1}$. A firm is defined as shrinking if its reported revenue in year t is lower than in year $t-1$, i.e. $\overline{y}_t >< \overline{y}$. The counterfactual distribution is constructed as explained in the note to Figure 4.

Figure 6: Revenue Distribution by Number of Employees
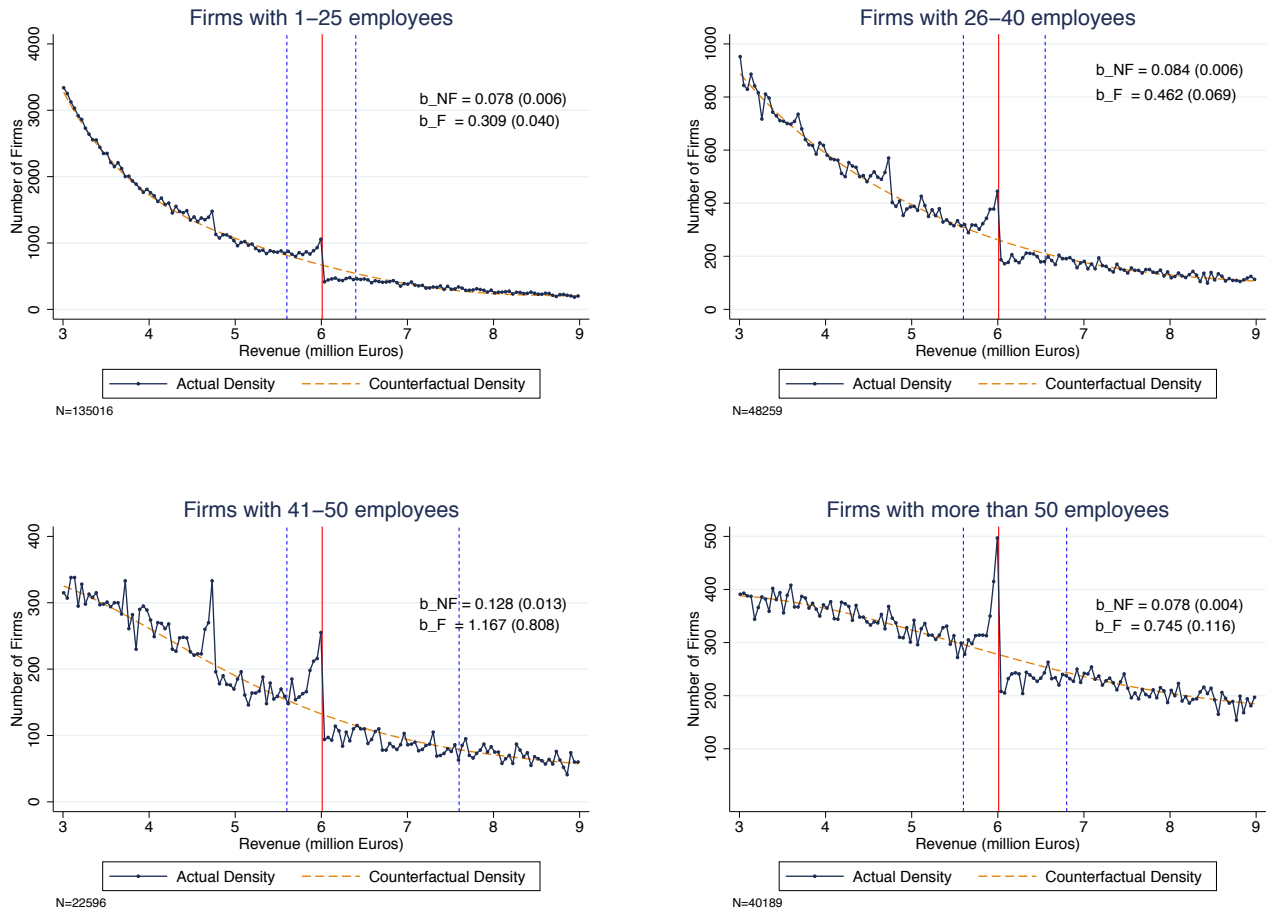
Note: these graphs show the actual and counterfactual revenue distributions for subsamples of firms with a given number of employees. The counterfactual distributions are constructed as explained in the note to Figure 4.

Figure 7: Revenue Distribution by Fixed Assets
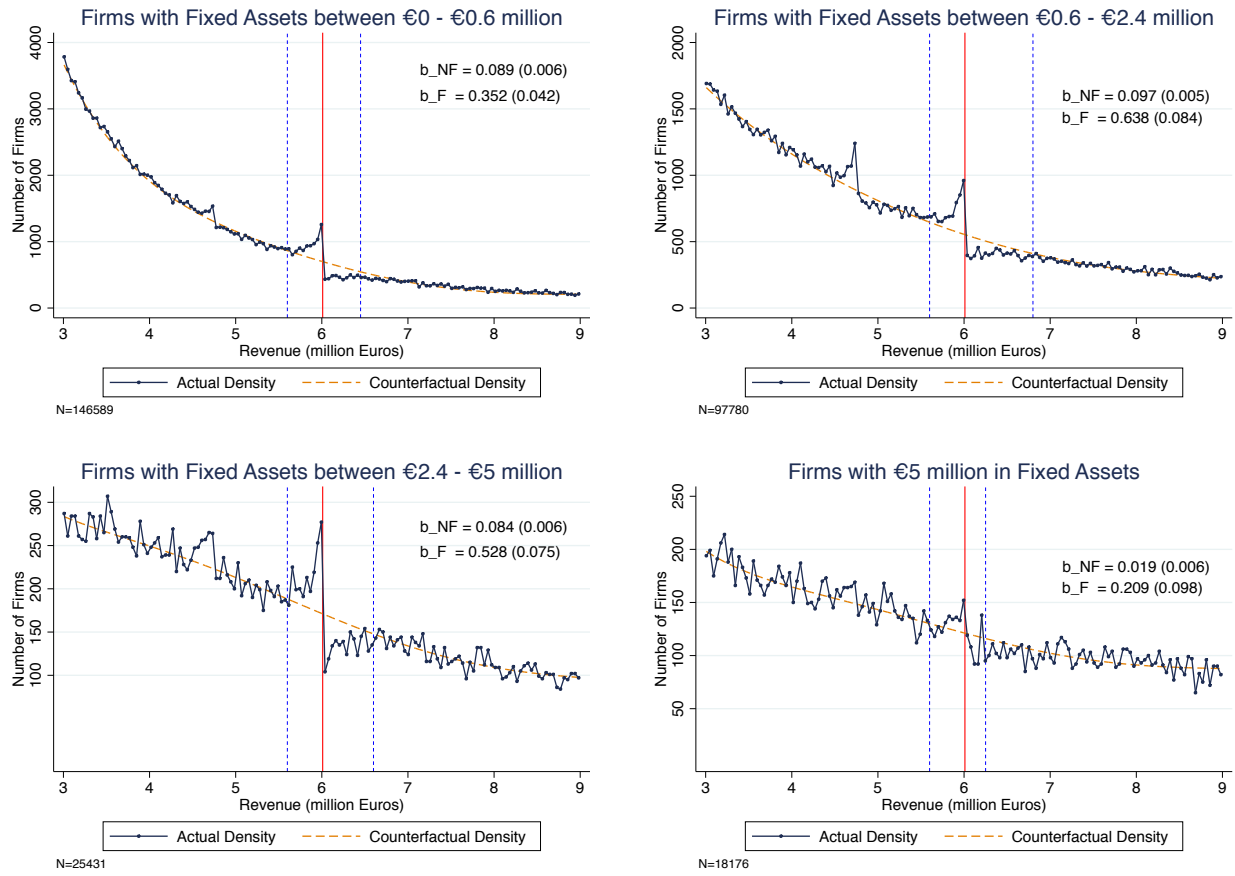
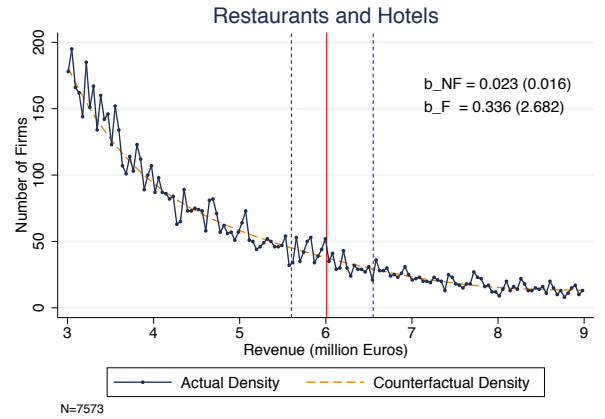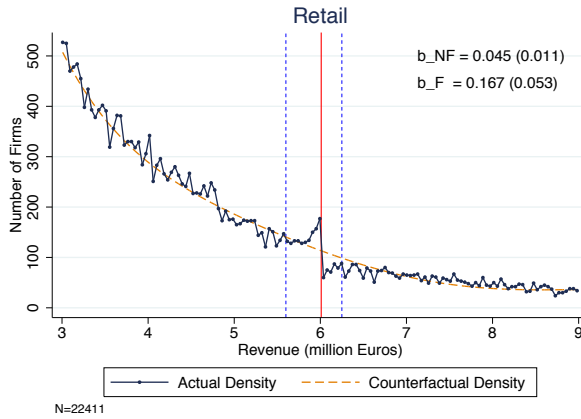Note: these graphs show the actual and counterfactual revenue distributions for subsamples of firms with a given level of fixed assets (measured in million Euros). The counterfactual distributions are constructed as explained in the note to Figure 4.

Figure 8: Revenue Distribution by Sector of Activity

(a) High Bunching Sectors
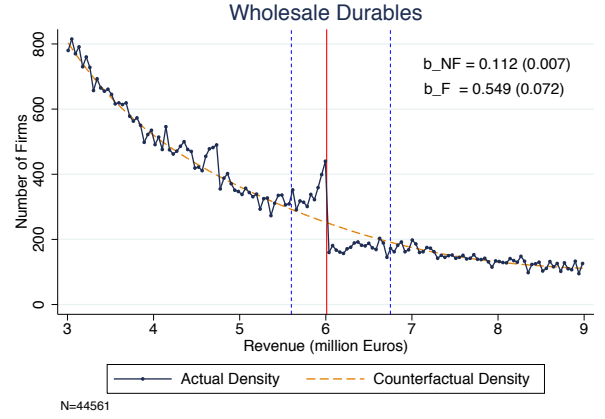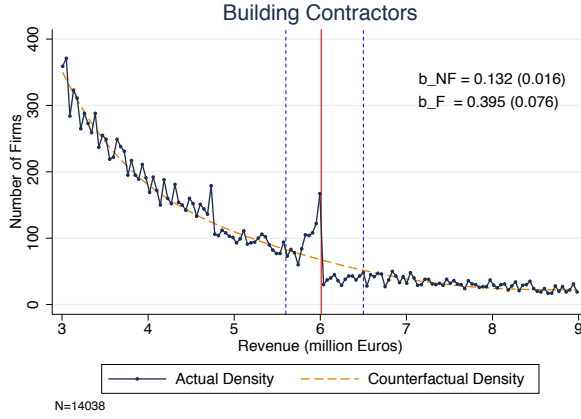
(b) Medium Bunching Sectors
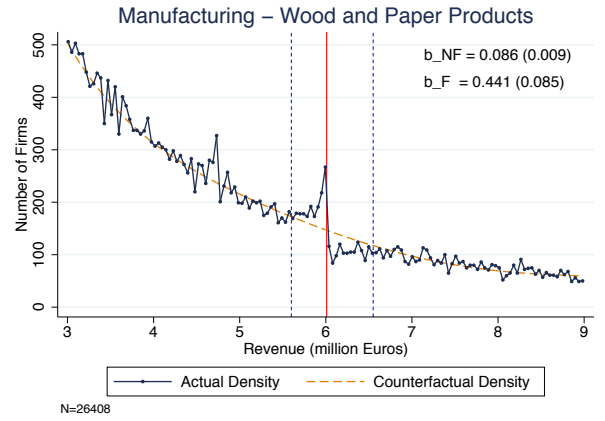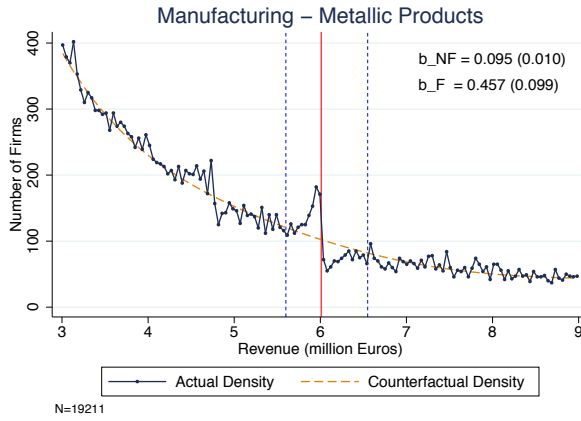
(c) Low Bunching Sectors

Note: these graphs show the actual and counterfactual revenue distributions for selected sectors (six out of a total of 12 sectors defined). The counterfactual distribution is constructed in each case as explained in the note to Figure 4.

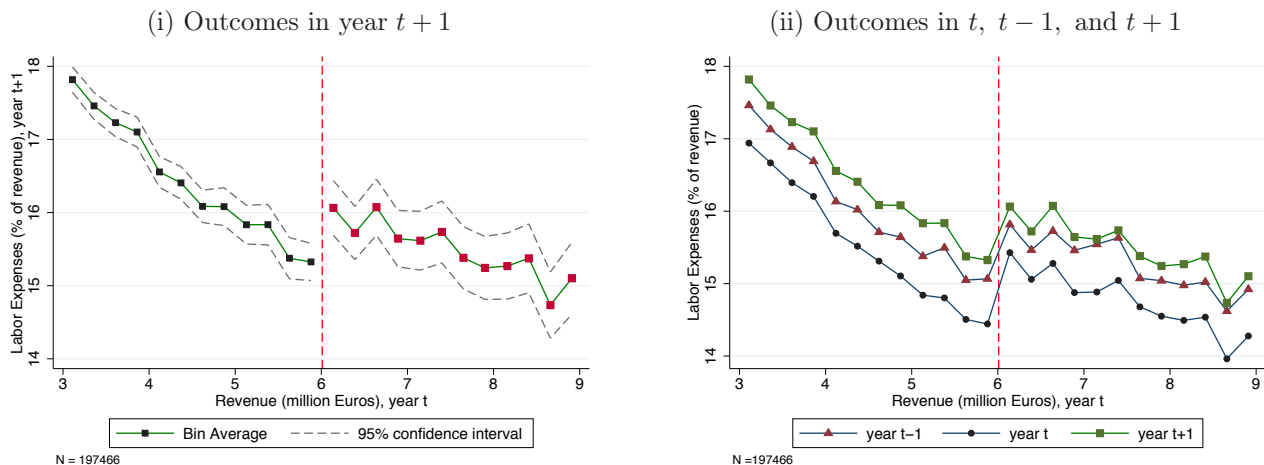Figure 9: Bunching Response by Scope of Evasion
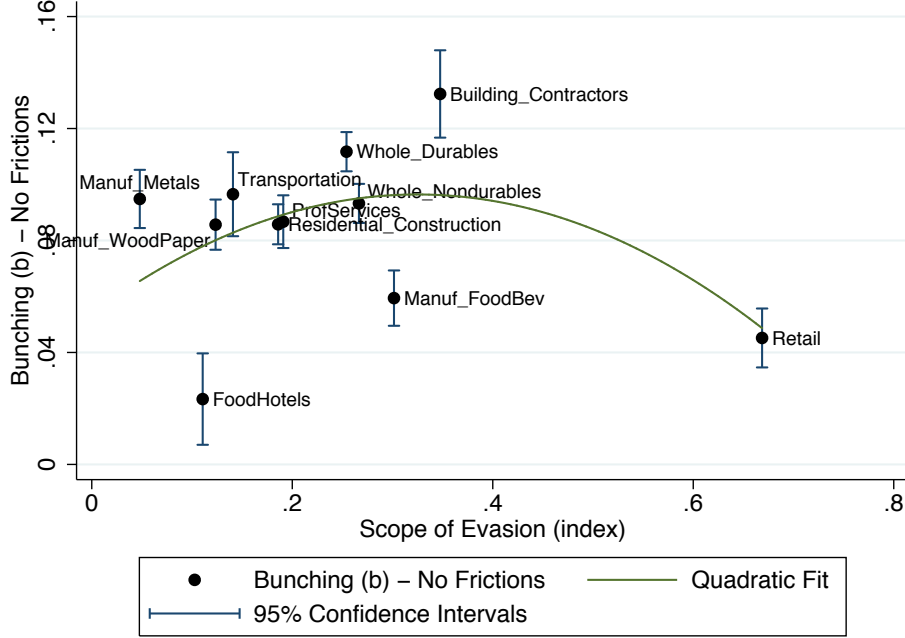
Note: the bunching measure $\widehat{b}_{NF}$ is calculated for each sector as explained in Section 4.1 in the main text. The scope of evasion index is the product of two elements: first, the percentage of a sector's output that is sold to final consumers. Second, the median number of employees that firms with revenue between €5.5-€6.5 million have in a given sector. The intuition for the first element is that selling to final consumers makes underreporting much easier because there VAT self-enforcing mechanism breaks down. For the second element, the idea is that it is easier to underreport if the number of employees is small. Specifically, we compute:
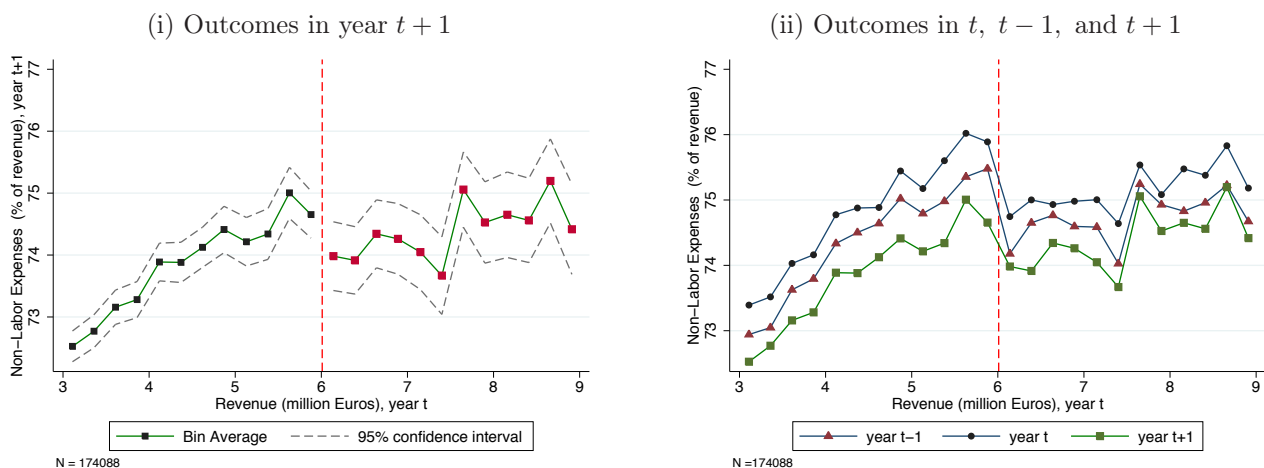
$$Scope = ShareFinalCons * (1 - employees/100)$$

where we divide the number of employees by 100 so that both numbers are smaller than one and $scope \in (0, 1)$.

Figure 10: Input-Cost Reporting

(a) Labor Input Costs

(i) Outcomes in year $t+1$ — (ii) Outcomes in t, $t-1$, and $t+1$

(b) Nonlabor Input Costs

(i) Outcomes in year $t+1$ — (ii) Outcomes in t, $t-1$, and $t+1$

Note: panel (a) shows average reported labor input costs (as a percentage of revenue) for year $t+1$ in (left panel). with 95% confidence intervals around each bin average. The average values for the same outcome variable in years $t-1$, t, and $t+1$ are shown in the right panel. Panel (b) shows the same graphs for average reported nonlabor input costs (materials and services used for production), also expressed as a percentage of revenue.

Appendix

A.1 Comparative Statics of an Increase in Enforcement

We apply the Implicit Function Theorem to do the comparative statics of an increase in the probability of detection. Let $F(u, \delta) \equiv \frac{d}{du}\mathbb{E}\Pi = 1 - [1 + \theta][\delta + u^* \cdot \delta_u(u^*, m)]$. Then:

$$\frac{du}{d\delta}\Big|_{u=u^*} = -\frac{dF/d\delta}{dF/du}\Big|_{u=u^*}$$

$$= -\frac{1 + \theta}{[1 + \theta][\delta_u + \delta_u + u^*\delta_{uu}]}\Big|_{u=u^*}$$

$$= -\frac{1}{2\delta_u + u^*\delta_{uu}}\Big|_{u=u^*}$$

$$< 0, \text{ since } \delta_u, \delta_{uu} > 0.$$

A.2 Corporate Tax Break for Small Firms

In parallel to the enforcement policies and regulations presented in Section 3.1, there are tax incentives for small firms in the corporate income tax. For the period we study, the incentive consisted of a lower marginal tax rate (30% instead of 35) for the first €90,151 of taxable profits, a quantity then raised to €120,202 in 2005. Notice how this introduces a *kink* (only the marginal tax rate changes) in the budget set of firms that are below the threshold, but a *notch* (both the average and the marginal tax rate change) at the revenue threshold for firms with low profits. Eligibility for this tax break is established solely based on an annual revenue criterion. Contrary to the case of the LTU, this threshold has been revised over time to account for nominal growth, from €1.5 million in 1999-2000 to €3 M. in 2001, €5 M. in 2002-03, €6 M. in 2004, and €8 M. in 2005-09.

The fact that this threshold overlaps with the LTU threshold in 2004 poses a potential problem for our analysis, because in that year we cannot distinguish whether any observed response is due to the LTU or the corporate tax break. This is not a concern, because 11 show that there is no behavioral response to this tax break in any year other than 2004 (when it overlaps with the LTU threshold).

A.3 Revenue Distribution by Region

Given that the LTU is organized in regional offices, there might be variation in the enforcement intensity change experienced when crossing the threshold in each region. Figure 13 shows a map with

the 17 Autonomous Regions in Spain. We use a color scale to show the different bunching intensity observed in the revenue distribution in each region. Lighter (yellow) tones apply to low bunching regions, while darker tones (red) denote high-bunching regions. The lowest bunching is observed in Navarra and País Vasco, the two regions in the North-Center where the Large Taxpayers' Unit (LTU) only applies to firms those that operate extensively in the rest of Spain. There is relatively (but statistically significant) low bunching in the Northern and Eastern regions of Cataluña, Aragón, Valencia and Baleares. Meanwhile, bunching is relatively high in the South, Center and North-West. The top bunchers are Extremadura, a relatively poor region in the Center-West, and Cantabria, a middle-income small region in the North.

Appendix Figures

Figure 11: Revenue Distribution, Year by Year
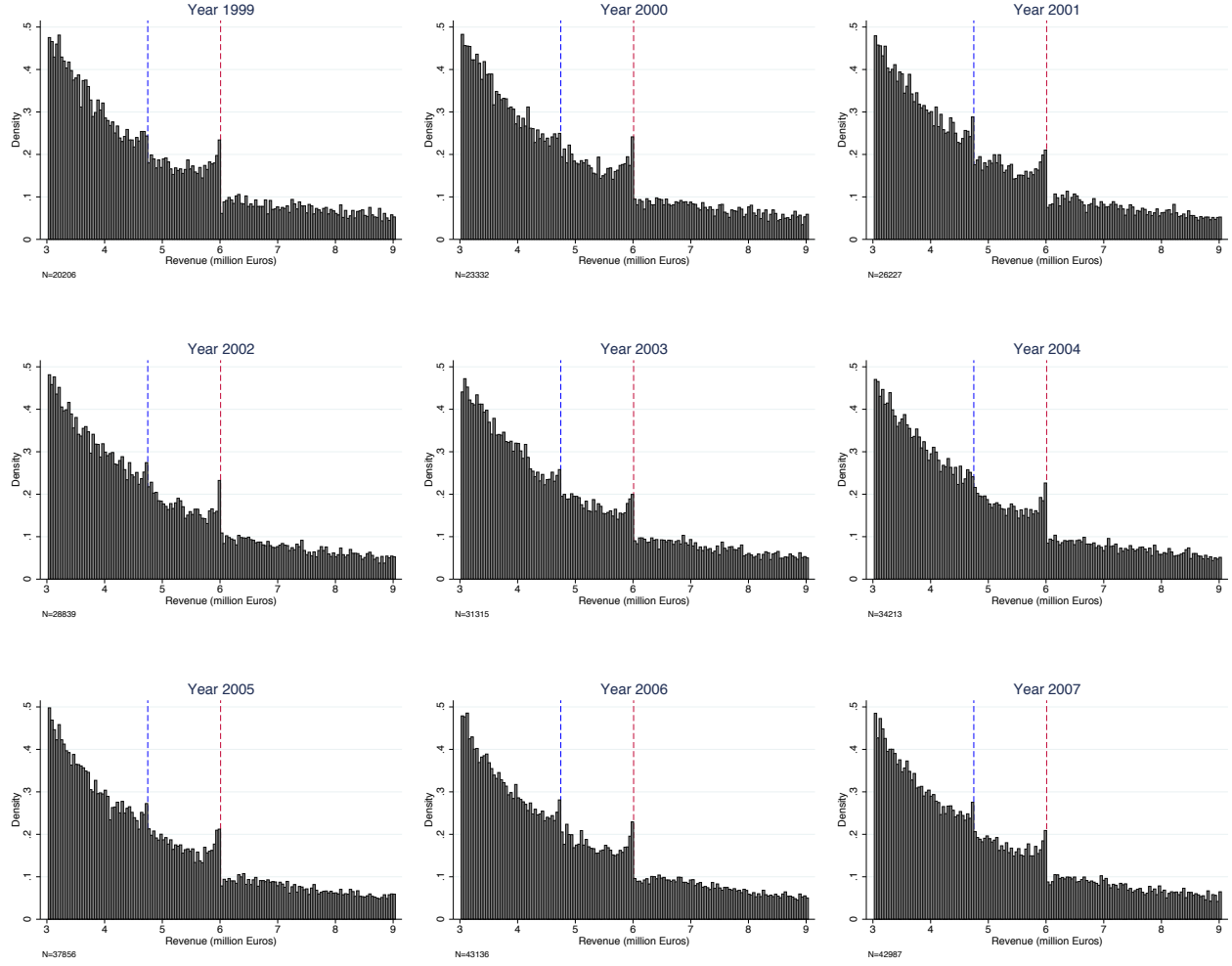
Note: this figure shows the annual histograms of reported revenue for each year in the period 1999-2007. The density distribution is essentially identical in every year, with some noise due to the fact that these subsamples are relatively small. There is no visible bunching at the threshold for the corporate tax break for small firms. This threshold was set at €1.5 million in 1999-2000, then raised to €3 million in 2001, to €5 million in 2002-03, to €6 million in 2004, and to €8 million for the period in 2005-2009.

Figure 12: Revenue Distribution by Organizational Form
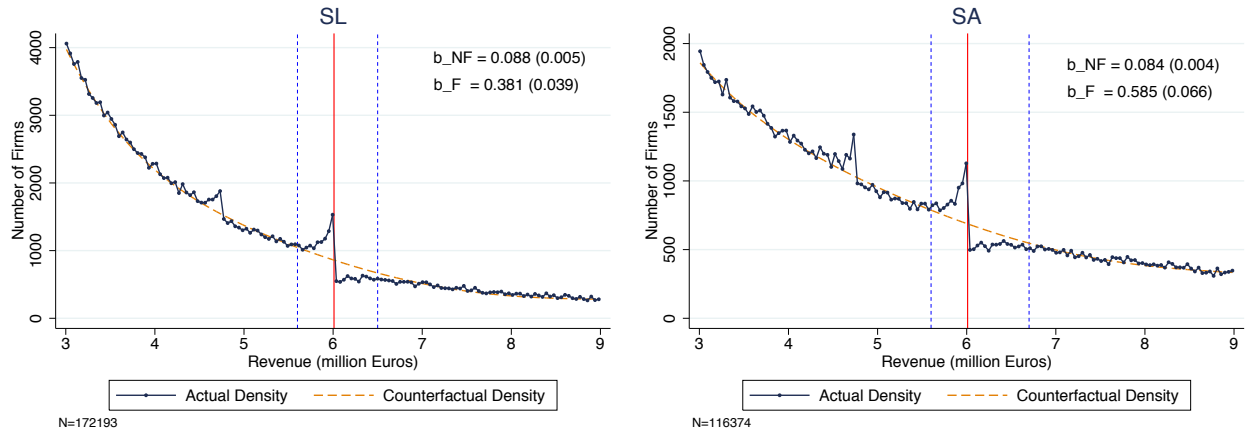
Note: these graphs show the actual and counterfactual revenue distributions for firms with different organizational forms. SL stands for *Sociedad Limitada*, equivalent to a Limited Liability Company. SA stands for *Sociedad Anónima*, equivalent to a Corporation. The counterfactual distribution is constructed in each case as explained in the note to Figure 4.

Figure 13: Bunching Intensity by Region
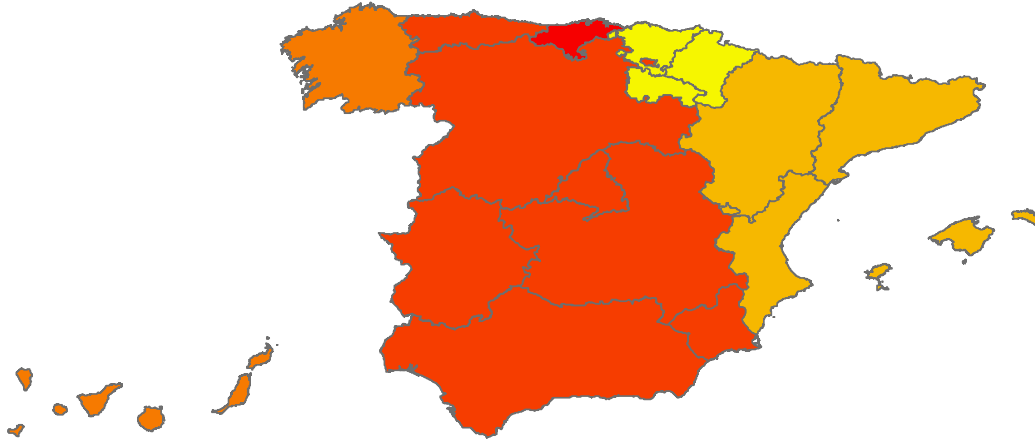
Note: this maps represents the 17 Autonomous Regions of Spain. We use a color scale to show the different bunching intensity observed in the revenue distribution in each region. Lighter (yellow) tones apply to low bunching regions, while darker tones (red) denote high-bunching regions. The lowest bunching is observed in Navarra and País Vasco, the two regions in the North-Center where the Large Taxpayers' Unit (LTU) only applies to a some firms (those that operate extensively in the rest of the country). For the other regions, the pattern is: relatively low bunching in the Northern and Eastern regions (Cataluña, Aragón, Valencia and Baleares) and relatively high bunching in the South, Center and North-West.