Oct 2015 No.245

Under the Radar: The Effects of Monitoring Firms on Tax Compliance

Miguel Almunia and David Lopez-Rodriguez

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

Under the Radar:

The Effects of Monitoring Firms on Tax Compliance*

Miguel Almunia
University of Warwick and CAGE

David Lopez-Rodriguez
Banco de España

August 31, 2015

Abstract

This paper analyzes the effects on tax compliance of monitoring the information trails generated by firms' activities. We exploit quasi-experimental variation generated by a Large Taxpayers Unit (LTU) in Spain, which monitors firms with more than 6 million euros in reported revenue. Firms strategically bunch below this threshold in order to avoid stricter tax enforcement. This response is stronger in sectors where transactions leave more paper trail, implying that monitoring effort and the traceability of information reported by firms are complements. We calculate that there would be substantial welfare gains from extending stricter tax monitoring to smaller businesses.

Keywords: tax enforcement, firms, bunching, Spain, Large Taxpayers Unit (LTU).

JEL codes: H26, H32.

*Almunia: m.almunia@warwick.ac.uk, University of Warwick Department of Economics and Centre for Competitive Advantage in the Global Economy (CAGE). Lopez-Rodriguez: david.lopezr@bde.es, Banco de España. A previous version of this paper circulated under the title *"Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Frims"*. We thank Emmanuel Saez, Alan Auerbach, Fred Finan and Ted Miguel for constant support and encouragement throughout this project. We gratefully acknowledge many useful comments and suggestions from Juan Pablo Atal, Pierre Bachas, Henrique Basso, Michael Best, David Card, Lorenzo Casaburi, Raj Chetty, Francisco de la Torre, François Gerard, Jonas Hjort, Simon Jäger, Anders Jensen, Juan F. Jimeno, Henrik Kleven, Attila Lindner, Justin McCrary, Craig McIntosh, Adair Morse, Gautam Rao, Daniel Reck, Ana Rocca, Michel Serafinelli, Monica Singhal, Juan Carlos Suárez Serrato, Joel Slemrod, Victoria Vanasco, Andrea Weber, Danny Yagan, Owen Zidar and numerous seminar participants. Letizia Borgomeo and Gonzalo Gaete provided excellent research assistance. Almunia gratefully acknowledges financial support from Fundación Rafael del Pino and the Burch Center for Tax Policy and Public Finance. Any views expressed in this paper are only those of the authors and should not be attributed to the Banco de España.

1 Introduction

In contrast to the predictions of the classical deterrence model of tax evasion (Allingham and Sandmo, 1972), modern tax systems in advanced economies feature high levels of tax compliance despite low audit rates. Recent theoretical studies argue that this can be explained by the additional deterrence effect implied by third-party information reporting (Kopczuk and Slemrod, 2006; Gordon and Li, 2009; Kleven, Kreiner, and Saez, 2015). Indeed, experimental evidence shows that the tax compliance rate on third-party reported income is much higher than on self-reported income (Kleven et al., 2011; IRS, 2012) and that threat-of-audit letters have a significant effect on self-reported tax bases (Slemrod, Blumenthal, and Christian, 2001; Kleven et al., 2011; Pomeranz, 2015). While these studies focus on individuals and small businesses, there is little micro evidence on the determinants of tax compliance by larger firms. Even with third-party reporting, monitoring large firms is more complex due to the large amount and variety of transactions that need to be verified jointly to determine the correct tax liabilities in several tax bases.[1]

This paper investigates to what extent stricter monitoring interacts with third-party information to increase firms' tax compliance. We first derive theoretical predictions on how firms respond to an increase in the monitoring resources used by the tax authority to examine the information trails generated by business activities. We then test these predictions using quasi-experimental variation in monitoring effort, generated by the existence of a Large Taxpayers Unit (LTU) in Spain.[2] The LTU, a special unit within the tax authority, exclusively monitors firms with more that €6 million in annual operating revenue and has more auditors per taxpayer than the rest of the tax authority. Thus, while all firms face the same tax schedule and procedures for tax compliance, the monitoring effort exerted by tax auditors changes discretely at an arbitrary revenue threshold, which allows us to study how the traceability of information trails and monitoring by tax auditors jointly affect firms' tax compliance.

In our baseline theoretical framework, firms make production and tax reporting decisions given a tax rate on reported profits. Underreporting revenue lowers tax liabilities, but implies some resource costs (e.g., keeping two sets of accounting books or foregoing business opportunities). The detection probability increases endogenously with the amount evaded by firms, depending on the interaction between (i) the resources devoted

[1] As taxpayers, firms remit corporate income tax and a share of payroll tax. As tax collectors, they withhold income and payroll tax from employees. In a broad range of countries, firms also remit value added tax (VAT) payments (Bird and Gendron, 2007; Keen and Lockwood, 2010).

[2] Most tax authorities in advanced countries, and an increasing number of emerging countries encouraged by multilateral institutions, have some type of LTU to increase tax compliance by large businesses, which concentrate a large share of tax-collection capacity (IMF, 2002; OECD, 2011).

by the tax authority to analyze tax returns and other paper trail generated by business activities ("monitoring effort"); and (ii) the legal reporting requirements and the available technology to process the information generated by business transactions ("enforcement technology"). In line with the quasi-experimental variation created by the LTU, we assume that monitoring effort jumps up discretely at a given level of reported revenue, while reporting requirements and the enforcement technology remain constant. This strengthens the effectiveness of the enforcement technology above the threshold, creating an incentive to bunch below the threshold in order to avoid stricter tax enforcement. We introduce heterogeneity across firms in the resource costs of evasion and in the effectiveness of monitoring effort, which depends on the traceability of firms' transactions. Given that intermediate input sales are easier to trace because they leave more paper trail, the model predicts a stronger bunching response by upstream firms compared to downstream firms (which sell mostly to final consumers), implying that monitoring effort and the traceability of information trails are complementary.

In the empirical analysis, we use data from financial statements that Spanish firms must submit by law to the Commercial Registry (*Registro Mercantil*) and also report to the tax authority. The *Banco de España* compiles and digitalizes this information since 1995, creating a confidential administrative dataset. This dataset contains firm-level information on annual net operating revenue, input expenditures, number of employees, payroll taxes, total value added, and the tax base and liability in the corporate income tax, making it possible to analyze multiple margins of firms' responses to the tax enforcement threshold. The dataset covers more than 80% of registered businesses in Spain with operating revenue between 3 and 9 million euros (the relevant range in our analysis) for the period 1995-2007, during which the LTU threshold remained constant at €6 million.

We find that a significant number of firms report revenues just below the LTU revenue threshold to avoid stricter monitoring. Our bunching estimates indicate that firms reduce their reported revenue by €94,000 on average (about 1.6% of their total revenue). Taking into account adjustment costs (resource costs of evasion and other frictions), we estimate that, conditional on bunching, the average reported revenue response is bounded between €276,000 and €520,000 (4.6% to 8.7% of total reported revenue). Robustness checks show that the bunching response is not due to other size-dependent policies, and the estimates are robust to different parametric assumptions to estimate the counterfactual distribution.

Regarding the relationship between monitoring effort and information trails, we distinguish 16 sectors of activity and use the percentage of revenue obtained from sales to final consumers as a proxy for the traceability of transactions made by firms in each sector. We find that the average bunching response is significantly higher in upstream sectors (e.g., heavy manufacturers and wholesalers) than in downstream sectors (e.g., retailers, restaurants and hotels), suggesting that the effectiveness of additional monitoring effort is higher the easier it is to trace firms' transactions. This finding indicates that information trails and monitoring effort by tax auditors are complements, because it is the interaction of these two elements that deters firms from evading taxes. We show that this complementarity cannot be attributed to a systematic relation between other firm characteristics, such as number of employees or fixed assets, and the size of the bunching response.

Finally, we examine the mechanisms used by firms to avoid stricter tax enforcement. In the presence of multiple taxes (corporate income tax, CIT, value-added tax, VAT, and payroll tax, PRT), firms may evade taxes not only by underreporting revenue but also by overreporting material inputs and underreporting labor costs. To study the nature of the bunching response, we derive predictions on the average relative use of inputs at the threshold depending on whether the bunching response is due to lower output (real response) or underreported revenue (evasion response). Our results show that the average ratio of material expenditures over total revenue shifts down sharply at the LTU threshold. In contrast, the average ratio of the wage bill (net of payroll taxes) as a share of total revenue shifts upward at the threshold. While these patterns do not identify causal effects, they are not compatible with only a real response, which would lead to more productive bunching firms using less of both inputs and create an upward shift of both ratios at the threshold. We also discard composition effects due to relative lower labor intensity of bunchers as the source of those patterns.

Our findings suggest that stricter monitoring creates a tax compliance effect on large firms that is equivalent to a broadening of reported tax bases. We document a remarkably stable one percentage-point gap (5% vs. 6%) in the average reported taxable profit (the tax base of the CIT) and significant jumps in the VAT and PRT tax bases between the two tax-enforcement regimes. These gaps are persistent well beyond the interval around the LTU threshold, suggesting that they cannot be explained only by selection due to the bunching response. Even in the presence of third-party reporting and detailed information requirements, medium and small firms take advantage of their role as fiscal intermediaries to appropriate part of the VAT that they should remit to the tax authority, and evade the corporate income and payroll taxes. Overall, the significant tax-reporting regime shift associated to the LTU suggests that this policy is an effective tool to reduce tax evasion. Moreover, the effectiveness of the policy on reporting behavior implies that significant reductions of the total resource costs of evasion can be achieved by incurring a relatively small additional administrative cost. We extend our theoretical framework to analyze the net welfare change associated with extending stricter tax enforcement marginally and redistributing the additional revenue lump-sum to taxpayers. Applying this framework to the Spanish case, we estimate that there would be substantial net welfare gains from extending the LTU to cover more medium-sized and small firms.

The sizeable bunching response associated to the stricter monitoring of large firms is consistent with the insights provided by the extensive literature on the effects of size-dependent policies on firm behavior. One strand of this literature has focused on evasion and avoidance responses at thresholds that determine eligibility to pay certain taxes, in particular the VAT (Keen and Mintz, 2004; Onji, 2009; Kanbur and Keen, 2014; Liu and Lockwood, 2015). Other studies have instead analyzed the impact of size-dependent regulations on aggregate productivity, given the pervasive incentives for firms to remain inefficiently small (Guner, Ventura, and Xu, 2008; Restuccia and Rogerson, 2008; Garicano, LeLarge, and van Reenen, 2013). Our results indicate that firms under low tax monitoring may look smaller in the data than they are in reality because of revenue under-reporting, which could have important implications for the estimates of the relationship between productivity and firm size.

The empirical techniques used in this paper draw on a growing literature in public finance that analyzes agents' responses to thresholds in taxes and regulations.[3] Our estimation strategy is very similar to the one proposed by Kleven and Waseem (2013), who exploit notches[4]—i.e., income thresholds at which the average tax rate jumps—to identify structural elasticities in the presence of optimization frictions. The novel feature of our setting is that the Spanish LTU generates a notch in enforcement intensity, rather than the tax rate, allowing us to study the effects of tax enforcement policies in isolation.

Our findings are mostly related to the empirical literature on the relevance of information trails to reduce tax evasion. The bunching estimates provide a well-identified measure of the impact of effectively using this information on firms' tax compliance. The complementarity result reveals the positive impact of information flows through the production chain on tax compliance in line with the results obtained by De Paula and Scheinkman (2010) and Pomeranz (2015) in developing economies. Our results also show the limitations, even in advanced economies, of third-party reporting to eliminate tax evasion given the complexity and multiplicity of business activities that must be effectively monitored.[5] This evidence contrasts with the essentially full effectiveness of third-party

[3] In the seminal paper of this literature, Saez (2010) exploits kinks—income thresholds at which the marginal tax rate jumps—to identify taxable income elasticities. Several recent studies apply Saez's method using large administrative datasets (Chetty et al., 2011; Bastani and Selin, 2014; Chetty, Friedman, and Saez, 2013). Devereux, Liu, and Loretz (2014) also use bunching techniques to estimate the elasticity of corporate taxable income in the United Kingdom.

[4] Slemrod (2010) provides a general description of notches in tax and regulatory systems. Two recent working papers, Best and Kleven (2015) and Kopczuk and Munroe (2015) study notches generated by property transaction taxes.

[5] Carrillo, Pomeranz, and Singhal (2014) show the limitations of third-party reporting to tax the full corporate income in Ecuador. Firms respond to better monitoring of revenue through third-party reporting by substituting evasion into less verifiable margins, such as input expenditures. Similarly,

reporting on individual tax compliance stressed by Kleven et al. (2011), but at the same time confirms the critical role of information trails as a necessary condition for effective taxation.

The creation of special tax units to monitor the largest taxpayers has been broadly implemented worldwide (OECD, 2011) and has been encouraged by multilateral institutions (IMF, 2002) as an effective way to raise the tax collection capacity in both advanced and developing countries. This paper provides novel evidence on the positive effects of the stricter monitoring implemented by a LTU on firms' tax compliance. This evidence demonstrates that there is no full self-enforcement of medium-sized businesses even in advanced country settings, indicating the potential welfare gains of spending additional resources to monitor those firms, taking advantage of the available information trails. The complementarity between monitoring and third-party information suggests that LTUs may be more effective in economies with a higher proportion of upstream firms (e.g., sectoral composition biased to industry), because their misreported transactions are most likely to be detected by monitoring their paper trail.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 describes the empirical strategy and derives the bunching estimators. Section 4 describes the institutional context and the data. Section 5 presents the estimation results. Section 6 discusses welfare and policy implications. Section 7 concludes.

2 Theoretical Framework

We extend the Lucas (1978) model to examine the problem of profit-maximizing firms that can evade taxes and face the risk of being detected and punished by the tax authority. In the baseline setting, firms make production decisions and can underreport their revenue (i.e., evade taxes), which implies incurring a resource cost. The probability of detection depends on the interaction between (i) the *monitoring effort* devoted by the tax authority to examine tax returns and other paper trails generated by firms' transactions and (ii) the existing *enforcement technology* used in this process, including the degree of detail of information-reporting requirements. We use this framework to analyze how firms respond to a discontinuous increase in monitoring effort at an arbitrary revenue threshold. We then extend the model to incorporate heterogeneity across firms in the effectiveness of monitoring effort and in the resource costs of evasion. The model with heterogeneity yields testable predictions on the expected effects of stricter tax enforcement on firms' decisions.

Slemrod et al. (2015) show that small businesses in the US offset the introduction of third-party reporting of credit card sales by increasing their reported expenses to leave their reported tax base essentially unchanged.

2.1 Corporate Taxation with Risky and Costly Evasion

Consider an economy with a continuum of firms of measure one. Firms produce good y combining tax-deductible inputs x and nondeductible inputs z according to the production function $y = \psi f(x, z)$, where ψ is a productivity parameter and $f(\cdot, \cdot)$ is strictly continuous, increasing and concave in both arguments. Productivity ψ is exogenously distributed over the range $[\underline{\psi}, \overline{\psi}]$ with a smoothly decreasing and convex density $d_0(\psi)$ in the population of firms. Firms purchase deductible and nondeductible inputs in competitive markets at unit cost w and q, respectively, and sell their output at the market price p, which is normalized to unity.

The government levies a proportional tax t on taxable profits $P = y - wx$, so net-of-tax profits with truthful reporting are given by $\Pi = (1-t)P - qz$. Since the tax authority does not perfectly observe all transactions in the economy, firms may attempt to evade taxes by misreporting taxable profits. In the baseline case, firms can underreport their revenue by an amount $u \equiv y - \overline{y} \geq 0$, where $\overline{y}$ is reported revenue[6] and, therefore, reported taxable profits are given by $\overline{P} = (1-t)[\overline{y} - wx]$. The direct and indirect resource costs of evasion are captured with the reduced form $\kappa(u)$, which is an increasing and convex function of concealed revenue.[7]

The tax authority detects evasion with probability $\delta = \phi h(u)$, where $\phi > 0$ is the monitoring effort parameter, and $h(\cdot)$ is a continuous, increasing and convex function of underreported revenue.[8] The endogenous component, $h(u)$, represents the enforcement technology used to match tax returns among trading partners and to review the paper trail created by information-reporting requirements. The shape of the function captures the intuition that a larger amount of unreported activity increases the probability of detection more than proportionally, because each inconsistency in reported transactions leaves a paper trail that can be examined (e.g. discrepancies in the monetary value of sales reported by firms and the purchases claimed as tax credits by their clients). The monitoring effort parameter, ϕ, represents the amount of resources devoted by the tax authority to the audit process. It is an exogenous policy choice that encompasses both quantity (e.g., the total number of auditors hired by the tax authority) and quality (e.g., the auditors' skill level and the actual effort exerted during audits).

We model the probability of detection as the interaction between monitoring effort and

[6] In subsection 5.3, we discuss the predictions of an extended model in which firms can also evade taxes by misreporting their input costs. We fully derive the extended model in the online appendix.

[7] One example of these resource costs of evasion is the need to maintain parallel accounting books to keep track of "black" payments made in cash. Tax evading firms may also forego business opportunities by not accepting credit cards or bank payments, given that it is much easier to conceal cash transactions. See Chetty (2009b) for a detailed discussion on the economic nature of these resource costs.

[8] For simplicity, we assume that when discrepancies between firms' reported transactions are detected, the authorities uncover the full amount evaded.

the available technology to capture the intuition that the two elements are complementary to determine the deterrence effect of tax enforcement policy. Finally, we assume that, when evasion is detected, the tax authority imposes a fine with a penalty rate θ over the amount of tax evaded, on top of the true tax liability.[9]

Firms make production (i.e., demand of inputs x and z) and reporting (i.e., under-reported revenue u) decisions in order to maximize expected after-tax profit, given by $\mathbb{E}\Pi = (1-t)[\psi f(x,z) - wx] - qz - \kappa(u) + tu[1 - \phi h(u)(1+\theta)]$. An interior optimum satisfies the following system of first-order conditions:[10]

$$\begin{aligned} \psi f_x(x,z) &= w \quad &(1)\\ \psi f_z(x,z) &= q/(1-t) \quad &(2)\\ t[1-\phi h(u)(1+\theta)] &= \kappa_u(u) + tu(1+\theta)\phi h_u(u) \quad &(3) \end{aligned}$$

where the term $[1-\phi h(u)(1+\theta)] \equiv r$ is the expected rate of return of evasion. This system of equations indicates that a positive tax rate has two effects. First, it distorts the choice of inputs, reducing production below the zero-tax optimum. Second, it creates incentives to evade taxes, thereby reducing reported revenue for all firms in equilibrium. Simple comparative statics show that an increase in monitoring effort ϕ leads to a decrease in concealed revenue u without affecting production decisions.

To provide more intuition on firms' incentives to evade taxes, we define the elasticity of detection probability with respect to concealed income as $\varepsilon_{\delta,u} \equiv \phi h_u \cdot u/\delta$, and rewrite the optimal evasion condition (3) as follows[11]

$$1 = \frac{\kappa_u(u)}{t} + (1+\theta)\delta(u)\left[1 + \varepsilon_{\delta,u}\right]. \quad (4)$$

The right-hand side of (4) identifies the two mechanisms that contribute to raising tax compliance by firms. The first term shows the disincentive effect created by the presence of marginal resource costs (relative to the marginal benefit of evasion, i.e., the tax rate). The second term represents the deterrence effect generated by the interaction between the tax authority's monitoring effort and the enforcement technology to review the paper trail generated by misreporting behavior.

Given that the production and resource cost functions, $f(\cdot)$ and $\kappa(\cdot)$, are the same for

[9] The canonical Allingham and Sandmo (1972) model of income tax evasion assumes that the penalty applies to the total amount evaded, but Yitzhaki (1974) points out that the common practice in most countries is to make the penalty proportional to the *amount of tax* evaded. We follow Yitzhaki's approach.

[10] The assumption of a convex detection probability is sufficient to ensure the second-order condition for an interior optimum is satisfied.

[11] This equation is similar to the one derived by Kleven et al. (2011), but obtained from the choice problem of firms, with an additional term to capture the impact of resource costs of evasion.

all firms, all the variation in reported revenue $\overline{y}$ is due to differences in productivity ψ. Assuming that monitoring effort ϕ is constant across all firms, there exists a density function of reported revenue $g_0(\overline{y})$ which is smoothly decreasing and convex in its full domain $[\overline{y}_{\min}(\underline{\psi}), \overline{y}_{\max}(\overline{\psi})]$.[12] The theoretical revenue distribution $G_0(\overline{y})$ is smoothly decreasing and convex in firms' productivity, as depicted by the dashed curve in Figure 1a.

Large Taxpayers Unit (LTU): A Tax Enforcement Notch

Assume now that the government provides additional resources to the tax authority in order to create a Large Taxpayers Unit (LTU). This reform raises the monitoring effort without affecting the enforcement technology. Formally, the LTU increases monitoring effort from ϕ_0 to $\phi_1 \equiv \phi_0 + d\phi$ (where $d\phi > 0$) *only* for firms with reported revenue $\overline{y} > y^L$, where y^L denotes the threshold for LTU eligibility. The probability of detection is then given by $\delta = \left[\phi_0 + d\phi \cdot \mathbf{1}\left(\overline{y} > y^L\right)\right] \cdot h(u)$, where $\mathbf{1}(\cdot)$ is the indicator function. The introduction of the LTU creates a tax enforcement notch, meaning that monitoring effort and, consequently, the probability of detection increase discretely at the arbitrary revenue level y^L.

The predicted behavioral response to the tax enforcement notch allows us to classify firms in three groups depending on their exogenous productivity draw: low productivity, bunchers, and high productivity firms. First, consider a firm with productivity ψ^L such that its optimal pre-LTU reported revenue is exactly the enforcement threshold, y^L. This firm determines the upper bound of the "low productivity" group of firms with $\psi \in [\underline{\psi}, \psi^L]$ that are not LTU-eligible. The production and reporting decisions of these firms are not affected by the introduction of the LTU.

Second, consider a firm with productivity ψ^M and pre-LTU reported revenue $y^M > y^L$. We assume that, for this firm, the costs of adjusting reported revenue to y^L in order to avoid the stricter enforcement of the LTU are equal to the benefits of such action. Formally, expected profits are equal in both cases, that is $\mathbb{E}\Pi_0\left(x, z, u|\phi_0, \psi^M\right) = \mathbb{E}\Pi_1\left(x', z', u'|\phi_1, \psi^M\right)$. We denote this firm the "marginal buncher", i.e., the firm with the highest productivity level that bunches at the threshold. All firms with $\psi \in \left(\psi^L, \psi^M\right]$ react to the introduction of a LTU by reporting lower revenue in order to bunch exactly at the LTU threshold, so we label this group the "bunchers". The bunching response generates a mass point at y^L in the distribution of reported revenue, and a hole with zero mass in an interval just above y^L.

Third, there is a group of "high productivity" firms with $\psi \in (\psi^M, \overline{\psi}]$. After the reform, these firms are monitored by the LTU because the costs of bunching (resource

[12] The specific mapping between the productivity and reported revenue density functions depends on the functional forms of the production function $f(\cdot)$ and the enforcement technology $h(u)$.

costs of evasion plus the increased probability of detection) outweigh the benefits. When monitored by the LTU, these firms report higher revenue than they did before the reform, which creates a rightward shift in the distribution of reported revenue above y^L. This shift also contributes marginally to extend the hole in the post-LTU distribution. The pre- and post-LTU distributions for this baseline model are depicted by the solid blue curve in Figure 1a.

To obtain a quantifiable measure of the behavioral response to the tax enforcement notch, we use a first-order approximation to relate the number of bunching firms to the change in the marginal buncher's reported revenue, following the method first proposed by Saez (2010). For analytical simplicity, consider the case in which the LTU raises monitoring effort by a small amount $d\phi = \phi_1 - \phi_0 > 0$, such that bunching firms adjust their reported revenue by $d\overline{y}^M$. The adjustment is proportional to $d\psi = \psi^M - \psi^L$, the difference in productivities between the marginal buncher and the firm that locates at the notch before the LTU is introduced. Since there is a direct mapping between the productivity distribution $d_0(\psi)$ and the pre-LTU reported revenue distribution $g_0(\overline{y})$, we can define the number of bunching firms at the threshold as

$$B = \int_{y^L}^{y^L + d\overline{y}^M} g_0(\overline{y}) d\overline{y} \approx g_0(\overline{y}^L) d\overline{y}^M, \tag{5}$$

where $g_0(\overline{y}^L)$ denotes the height of the pre-LTU density distribution at the threshold.[13] The number of bunching firms that respond to the notch depends positively on the increase in monitoring effort at the notch and negatively on the extent of resource costs associated to tax evasion. Using the approximation in (5), we define the bunching estimator b as the ratio of excess bunching over the height of the counterfactual density at the LTU threshold, that is,

$$b \equiv \frac{B}{g_0(y^L)} \approx d\overline{y}^M. \tag{6}$$

This is equivalent to the standard estimator used in the existing bunching literature, (e.g., Chetty et al., 2011). Geometrically, the estimator approximates the length (in euros) of the interval of reported revenue where bunching firms would have located in the counterfactual scenario without the LTU.

2.2 Heterogeneous Firms

We extend the baseline model to introduce firm heterogeneity along two dimensions: the effectiveness of the tax enforcement policy across firms, and the importance of resource

[13] The approximation in (5) assumes that the pre-LTU density $g_0(\overline{y})$ is approximately flat in the neighborhood of the enforcement threshold y^L.

costs of evasion. Considering this heterogeneity, the LTU creates different incentives to bunch for firms with the same productivity level, and the model predicts differential increases in reported revenue for high productivity firms (i.e., those monitored by the LTU). As a result, the extended model no longer predicts a hole with zero mass in an interval above the LTU threshold, but only a triangle of missing mass in that interval. In the presence of heterogeneous agents, the bunching method applied to notches allows us to derive estimators of both the average revenue response to the LTU policy attenuated by the costs of reaction (short-run response), and the average structural (or long-run) response accounting for the existence of such costs.

Heterogeneous Effectiveness of Monitoring Effort

We assume now that the effectiveness of tax enforcement policy depends on the traceability of misreported business transactions. An increase in the resources devoted to monitoring activities is more effective at uncovering evasion when firms sell to other businesses, because these transactions generate a paper trail,[14] compared to situations in which firms sell directly to final consumers, where there is little or no paper trail. Hence, for a given productivity level, the LTU is more effective at monitoring upstream firms than downstream firms (whose transactions are harder to trace). As a consequence, upstream firms have stronger incentives to bunch at the LTU threshold, and they also report higher revenue when they belong to the "high productivit" group.

To capture this intuition formally, assume a joint distribution of productivities and monitoring effort with density $\widetilde{g}(\psi, \phi)$ on the domain $(\underline{\psi}, \overline{\psi})$ x $(\phi_0, \overline{\phi})$. The behavioral response at each monitoring effort level is characterized by the set of conditions (1)-(3), taking into account the discrete change at the cut-off productivity level ψ^M that separates bunchers and high productivity firms. At each monitoring effort ϕ, firms in the pre-LTU density interval $(y^L, y^L + d\overline{y}^M_\phi)$ have incentives to bunch if they expect tax savings from that action. Note that the revenue response of the marginal buncher, $d\overline{y}^M_\phi$, is increasing in the effective monitoring effort that it faces. We assume that the distribution of monitoring effort effectiveness is such that there exists a dominated region $\left(y^L, y^D\right)$, where $y^D = d\overline{y}^M_{\phi_{min}}$ is the pre-LTU revenue level of the marginal buncher associated to the lowest effectiveness change. No firm has an incentive to locate in this region, because they could simultaneously reduce reported revenue and increase expected profits. The rest of the bunching interval is given by $\left(y^D, y_{ub}\right)$, where $y_{ub} = d\overline{y}^M_{\phi_{max}}$ is the pre-LTU revenue level of the marginal buncher associated to the highest monitoring effort change. Bunching provides expected tax savings only if the change in monitoring effort at the LTU threshold is large enough to overcome differential resource costs $d\kappa(u)$, so only a

[14] Especially in a context with a value added tax (VAT) that is applied using the credit-invoice method.

share of firms in the interval $\left(y^D, y_{ub}\right)$ have incentives to bunch. Since downstream firms experience a smaller increase in effective monitoring effort, they have lower incentives to bunch (for a given productivity level and resource cost function).

With this setup, we can quantify the average reported revenue response from the observed bunching at the LTU threshold. Let $\widetilde{g}_0(\overline{y}, \phi)$ be the joint distribution of reported revenue and monitoring effort when tax enforcement is smooth with constant ϕ, and denote by $g_0(\overline{y}) \equiv \int_\phi \widetilde{g}_0(\overline{y}, \phi) d\phi$ the unconditional reported revenue distribution absent the tax enforcement notch. Assuming that the counterfactual density is roughly flat around the LTU threshold, we can write the excess bunching mass at the threshold as

$$B = \int_\phi \int_{y^L}^{y^L + d\overline{y}_\phi^M} \widetilde{g}_0(\overline{y}, \phi) d\overline{y} d\phi \approx g_0(y^L) \cdot E[d\overline{y}_\phi^M], \tag{7}$$

where $E[d\overline{y}_\phi^M]$ is the average response in reported revenue. This term is a weighted average of the marginal buncher's response at each tax enforcement change implied by the LTU. We denote by b_{av} the estimator of the average bunching response in the population, which is the ratio of excess bunching over the counterfactual reported revenue density at the threshold,

$$b_{av} \equiv \frac{B}{g_0(y^L)} \approx E[d\overline{y}_\phi^M]. \tag{8}$$

This estimator measures the average revenue response to the LTU policy attenuated by the costs of adjustment. That is, the estimator averages the response of the bunchers and the lack of response of those that do not bunch (*nonbunchers*).

Heterogeneous Resource Costs of Evasion

We now let resource costs of evasion $\kappa(u)$ differ across firms with the same productivity. For instance, the costs of evading a certain amount of taxes might vary depending on the complexity of firms' operations or the number of employees with whom the employers would need to collude in order to evade Kleven, Kreiner, and Saez (2015). These costs can also depend on managers' preferences (e.g. risk-aversion and honesty), or on their attention and effort devoted to misreport transactions rather than allocating their skills to productive activities. Besides, evader firms bear the opportunity costs of foregoing business opportunities with trading partners that do not engage in misreporting (Chetty, 2009a). Through any of these channels, firms might face different resource cost functions and, therefore, firms with the same productivity have heterogeneous costs of misreporting. In the LTU context, this implies that firms with the same incentives to react face differential costs of adjustment. For some firms with incentives to bunch given their productivity draw ψ and the change in monitoring effort effectiveness $d\phi$, adjustment costs can be so large ("prohibitive") that they exceed the benefits of bunching to avoid the higher monitoring effectiveness of the LTU.

Following Kleven and Waseem (2013), we leverage the strong incentives generated by the notch to quantify the structural revenue response, accounting for these costs of adjustment. Let $\alpha(\overline{y}, \phi)$ denote the proportion of firms that have incentives to bunch but face prohibitive adjustment costs at each level of reported revenue and monitoring effort. For analytical simplicity, we assume that this proportion is constant in the bunching interval associated to the LTU, such that $\alpha(\overline{y}, \phi) = \alpha$ for $\overline{y} \in (y^L, y^L + d\overline{y}^M_\phi)$. Assuming that the pre-LTU reported revenue density is locally flat in the neighborhood of the threshold, the excess bunching mass at the LTU threshold is now given by

$$B_{adj} = \int_\phi \int_{y^L}^{y^L + d\overline{y}^M_\phi} [1 - \alpha(\overline{y}, \phi)] \cdot \widetilde{g}_0(\overline{y}, \phi) d\overline{y} d\phi \approx g_0(y^L) \cdot (1 - \alpha) \cdot E[d\overline{y}^M_\phi], \tag{9}$$

where $E[d\overline{y}^M_\phi]$ is the average reported revenue response at the LTU threshold.

Given (9), the bunching parameter that measures the structural response (i.e., the response that we would observe in the absence of adjustment costs) can be written as

$$b_{adj} \equiv \frac{B}{g_0(y^L) \cdot (1 - \alpha)} \approx E[d\overline{y}^M_\phi]. \tag{10}$$

Expression (10) indicates that the larger the number of bunching firms and the smaller the hole in the bunching region (i.e. the higher the share of nonbunchers) the larger is the structural response to stricter tax enforcement. We discuss in subsection 3.1 how we estimate the lower and upper bounds of this structural response.

Summary of Predictions with Heterogeneous Firms

This extended model with heterogeneity in (a) the effectiveness of monitoring effort, and (b) the magnitude of resource costs of evasion, yields a new set of predictions for the behavioral responses to the introduction of a LTU, which are depicted in Figure 1b. Compared to the baseline model, there is still bunching at the LTU threshold, but there is no longer a hole in the distribution just above it. Instead, the model predicts a triangle of missing mass, because the resource costs of evasion dominate the benefits of bunching for a fraction of firms. According to this model, firms monitored by the LTU (high productivity firms) do not modify their production decisions, but stricter tax enforcement induces them to declare a higher proportion of their true tax base.

3 Empirical Strategy: Bunching Estimators

This section presents the empirical procedure to estimate the reported revenue response of firms to a tax enforcement notch created by the existence of a LTU. To quantify this response, we adapt the techniques from the bunching literature in individual taxation (Saez, 2010; Chetty et al., 2011; Kleven and Waseem, 2013) to estimate the parameters derived in the previous section.

The basic procedure to estimate the reaction of firms to a LTU relies on constructing a counterfactual distribution of reported revenue in the absence of a tax enforcement notch, and comparing it with the observed distribution. To build the counterfactual, we fit a high-degree polynomial to the observed density, excluding an interval around the threshold. We discuss below how the excluded interval is determined. Dividing the data in small bins of width w, we estimate the polynomial regression

$$F_j = \sum_{i=0}^{q} \beta_i \cdot (y_j)^i + \sum_{k=y_{lb}}^{y_{ub}} \gamma_k \cdot \mathbb{1}\left(y_j = k\right) + \eta_j, \tag{11}$$

where F_j is the number of firms in bin j, q is the order of the polynomial, y_j is the revenue midpoint of bin j, y_{lb} and y_{ub} are the lower and upper bound of the excluded interval (respectively), and the γ_k's are intercept shifters for each of the bins in the excluded interval. Then, using the estimated coefficients from regression (11), we estimate the counterfactual distribution of reported revenue, that is, $\widehat{F_j} = \sum_{i=0}^{q} \widehat{\beta}_i \cdot (y_j)^i$. This expression excludes the γ_k shifters to ensure that the counterfactual density is smooth around the threshold.

Comparing this counterfactual density to the observed distribution we can estimate the excess bunching mass to the left of the threshold (B), and similarly the missing mass to the right of the threshold (H), given by

$$\widehat{B} = \sum_{j=y_{lb}}^{y^L} \left(F_j - \widehat{F_j}\right) \geq 0 \qquad \textit{and} \qquad \widehat{H} = \sum_{j=y^L}^{y_{ub}} \left(\widehat{F_j} - F_j\right) \geq 0.$$

Determining the lower and upper bounds of the excluded region in a consistent way is critical for this estimation method to provide credible estimates. We follow the approach proposed by Kleven and Waseem (2013) to determine these bounds. This procedure imposes that the areas under both the counterfactual and the observed density have to be equal, and thus the missing area (H) has to be equal to the excess mass (B). Implicitly, this is equivalent to assuming that all responses to the tax enforcement notch are on the intensive margin (i.e., firms don't go out of business due to the introduction of the LTU, they only adjust their reported revenue). One potential concern with this method to estimate the counterfactual distribution is that high productivity firms may increase their reported revenue in response to the LTU. This intensive margin response shifts the distribution to the right for firms above the LTU threshold, increasing the missing mass in the interval $\left(y^L, y_{ub}\right)$. However, this rightward shift in the distribution is of second-order importance for the estimation of the counterfactual as long as the distribution is approximately flat in the bunching segment.[15]

To obtain consistent bunching estimates, we first fix the lower bound y_{lb} at the bin where the decreasing trend of the reported revenue density reverts to an increasing rate due to the bunching response. Second, we set the upper bound at $y_{ub} \approx y^L$ and then run regression (11) multiple times, increasing the value of y_{ub} by a small amount after each iteration. When bunching is substantial, the first few iterations yield large estimates of $\widehat{B}$ and small estimates of $\widehat{H}$. This estimation procedure iterates until reaching a value of y_{ub} such that missing and bunching areas converge, i.e $\widehat{B} = \widehat{H}$.[16]

Once we have an estimate for the excess number of bunching firms B and the counterfactual density at the threshold $g_0(\overline{y})$, we can estimate the average bunching parameter b_{av} defined in equation (8). The explicit formula for the estimator is given by

$$\widehat{b}_{av} = \frac{\widehat{B}}{\left(\frac{1}{y^L - y_{lb}}\right) \sum_{j=y_{lb}}^{y^L} \widehat{F_j}} \tag{12}$$

where $\left(y^L - y_{lb}\right)$ is the length, in euros, of the excluded interval to the left of the threshold. This estimator measures the average reported revenue response of firms in the bunching interval to the LTU.[17] The magnitude of the response depends on the effectiveness of the policy across firms and the costs of their reaction. We estimate the standard errors using bootstrapping.[18]

[15] This point is also noted in Kleven and Waseem (2013) and Best and Kleven (2015).

[16] In the empirical application there is a finite number of bins, so we impose the weaker condition that the ratio be "close" to one, i.e. $\hat{H}/\hat{B} \in [0.9, 1.1]$. Even though bunching strategy imposes some discretion in the choice of the lower bound of the excluded region, sensitivity analysis presented section 5 and the appendix shows that the bunching estimates are robust to the precise choice of the excluded region.

[17] In earlier bunching papers (e.g., Chetty et al., 2011), the b estimator is defined as the ratio of excess bunching mass over the height of the counterfactual density at the threshold, and then the elasticity is calculated separately. In the results presented below, we multiply the estimator of b_{av} times the bin width w in order to obtain a money metric directly.

[18] We thank Michael Best for sharing his Stata code to perform the bootstrapping routine. In all the results shown below, we perform 200 iterations to obtain the standard errors. Using a larger number does not affect any of the results.

3.1 Adjusted Bunching Estimator: Bounding the Structural Response to Tax Enforcement

In subsection 2.2, we derived the parameter b_{adj} to quantify the structural response that would be observed in the absence of adjustment costs. To estimate this structural response empirically, we need to estimate the proportion of firms (α) that do not react because their adjustment costs are higher than the expected tax savings of bunching. According to the model with heterogeneity, there are two factors that may lead firms to locate in the interval (y^L, y_{ub}): the LTU may be relatively ineffective (i.e., small change in monitoring effort at the threshold) or adjustment costs may be too large. We exploit the fact that the only factor preventing a bunching response in the dominated region (y^L, y^D) is the existence of adjustment costs.[19] We thus propose an empirical procedure to estimate α in the dominated region, which allows us to provide a lower bound for the structural response.

The length of the dominated region depends on the change in the effectiveness of monitoring effort, $d\phi$, which is heterogeneous across firms. Therefore, the length of the dominated region is not uniquely defined. To obtain an empirical estimate of α, we assume that the dominated region is given by a short interval of length $(y^D - y^L) > 0$ above the LTU threshold, and then conduct a sensitivity analysis using different values for y^D. The empirical mass in this interval can be used to estimate the proportion of firms, α, whose reaction costs are larger than the expected tax savings. Specifically, we define the estimator $\widehat{\alpha} \equiv \frac{\int_{y^L}^{y^D} g(\overline{y})d\overline{y}}{\int_{y^L}^{y^D} g_0(\overline{y})d\overline{y}}$. To be conservative, we assume that the proportion of firms that do not react because of prohibitive adjustment costs is constant in the whole bunching interval. Using the estimate of α, we reweigh the average bunching estimator in (12) to obtain the adjusted bunching estimator

$$\widehat{b}_{adj}^{lb} = \frac{\widehat{b}_{av}}{1 - \widehat{\alpha}}, \tag{13}$$

which measures the average reported revenue response of bunchers and identifies the structural behavioral response to stricter tax enforcement. This estimator is a lower bound of the structural response because, in practice, adjustment costs increase with the distance to the threshold (and, therefore, the proportion of firms along the bunching interval that do not react due to the adjustment costs also increases).

The upper bound of the structural response is determined by the convergence point between the counterfactual and the empirical densities, y_{ub}. Assuming that the effect of

[19] In the rest of the bunching interval, (y^D, y_{ub}), the costs of bunching may outweigh the benefits due to a combination of the two factors, but it is not possible to empirically disentangle the weight of each of them.

the LTU is homogeneous across firms, the convergence point identifies a unique marginal buncher with the whole mass in the bunching interval due to adjustment costs. Instead, when the effectiveness of the LTU is heterogeneous across firms, a share of the mass is due to the differential impact of the policy. The convergence point identifies the behavioral response of the firm on which the LTU is most effective. Therefore, we interpret the estimator $\widehat{b}^{ub}_{adj} = y_{ub} - y_L$ as an upper bound of the structural reaction to stricter tax enforcement.[20]

4 Institutional Context and Data

4.1 Tax Administration Thresholds: the Spanish LTU

The Spanish tax authority established a Large Taxpayers Unit (LTU) in 1995 to increase monitoring effort on the largest taxpayers, defined as firms with annual operating revenue above €6 million.[21] This eligibility threshold has not been modified since then, and the number of firms in the LTU census (excluding public companies) more than tripled from 11,107 in 1995 to 34,923 in 2007, mainly due to nominal economic growth.[22] In the same period, the overall staff of the tax authority, and the LTU in particular, stagnated.[23] Firms in the LTU represent only 2% of those that submit a corporate income tax return, but they report about 80% of all taxable profits, 65% of total sales subject to VAT, and they employ 40% of private sector wage-earners (AEAT, 1995-2007).

Businesses just above and below the LTU threshold face the same tax schedule on all the major taxes (corporate income, value added and payroll taxes), and the same administrative requirements related to invoicing, accounting and information reporting. The monitoring process on taxpayers works in several stages: first, the tax authority's highly-sophisticated electronic system processes all detailed the information reported in tax returns to detect inconsistencies or mistakes. Second, the system automatically cross-checks

[20] Kleven and Waseem (2013) call this the "convergence method".

[21] In Spanish, *Unidades Regionales de Gestion de Grandes Empresas*. The LTU has one office in each of the 17 Spanish regions, although auditors from one region can monitor a firm based elsewhere if it operates in that region. The threshold was originally set at 1 billion pesetas (€6.010121 million at the official exchange rate).

[22] There are two exceptions to the eligibility rule: (i) exporters that claim VAT refunds are included in the LTU census regardless of their operating revenue, and (ii) firms based in two small regions with independent tax authorities (Navarra and País Vasco) are only included if they obtain a large proportion of revenue from sales in the rest of Spain. Due to data limitations, we cannot precisely identify which of those firms are in the LTU census, so we drop Navarra and País Vasco from the analysis but we keep all exporters (large and small). The latter, if anything, would bias our bunching estimates downward. We provide more details in the online appendix.

[23] Note that the tax authority's enforcement technology, including the introduction of electronic reporting and modern data management, has improved steadily over the years (AEAT, 1995-2007). Thus, it is difficult to assess whether the enforcement capacity of the LTU has increased or decreased over this period.

tax returns with other available information such as financial statements or transactions reported by third parties (e.g., trading partners or banks) to detect any discrepancies.[24] Up to this point in the process, the enforcement technology is the same for firms above and below the LTU threshold. Once the system detects an inconsistency on a specific tax return, tax auditors analyze the paper trail available for that firm and requests additional clarifications or tax payments using notification letters. If large discrepancies or inconsistencies are detected, the staff may start a full tax audit.

The key difference we exploit in our empirical strategy is the fact that the LTU has more auditors per taxpayer than the rest of the tax authority, and those auditors have on average higher qualifications and experience. This implies that the LTU makes heavier use of the available information and enforcement technology to detect tax evasion than the rest of the tax authority. This policy setting provides quasi-experimental variation in the monitoring effort (captured by ϕ in the model) faced by large firms, while keeping the information requirements and enforcement technology (captured by $h(u)$) equal for firms around the LTU threshold. Therefore, it allows us to examine firms' responses to a more effective tax enforcement system.

4.2 Data

In the empirical analysis, we use data from financial statements that all Spanish firms must submit by law to the Commercial Registry (*Registro Mercantil*).[25] Firms are also required to reproduce the financial statement on their annual corporate income tax return, so there are no incentives to report different information on the two documents. In fact, auditors from the tax authority routinely cross-check tax returns against published financial statements to detect discrepancies (for example, to ensure that the annual operating revenue figure is the same in both).

The *Banco de España* has compiled and digitalized all the financial statements submitted for the period 1995-2007 to construct a confidential administrative dataset. This micro-dataset has a panel structure and includes the following information for each firm: business name, fiscal identifier, sector of activity (4-digit CNAE-2009 code), Balance Sheet, Profit & Loss Account, and number of employees.[26]

[24] The tax authority introduced in the 1980's a mandatory information form (*Modelo 347*) in which firms must provide detailed information on the monetary value of their transactions with all of their suppliers and clients. The information from these forms is processed electronically and regularly used by tax auditors to cross-check tax returns and detect discrepancies between trading partners.

[25] Financial statements are an official public document, and information for a particular firm can be accessed by anyone paying a small fee.

[26] In an earlier version of this paper (Almunia and Lopez-Rodriguez, 2014) we used a similar dataset from Amadeus, which is also constructed using financial statements from the Spanish Commercial Registry, obtaining almost identical results. The *Banco de España* dataset has two advantages: first, it is an administrative dataset with a higher level of disaggregation in some variables than the Amadeus version,

Comparing the number of firms in the *Banco de España* dataset to official statistics of corporate income tax filers for the period 1995-2007, we observe that our dataset contains almost 80% of all firms with reported revenue between €3 and €9 million, the relevant range in the empirical analysis.[27] The main reason for missing observations is that some firms submit their statements late or on paper form, in which case they may not have been digitalized. The analysis performed by the *Banco de España* statistical division concludes that missing firms are on average less transparent, which may be positively correlated with misreporting behavior. Among reporting firms, the *Banco de España* dataset identifies some firms as having "unreliable data", often because of rounding issues or inconsistent reporting. We drop these firms (about 7% of the total) and obtain a final dataset of 285,570 firms with reported revenue between €3 and €9 million in the period 1995-2007. We expect any bias due to excluding missing and unreliable firms from our sample to go against finding a behavioral response to avoid being in the LTU.

The main variables used in our empirical analysis are: (i) annual net operating revenue, which is used to determine whether firms are eligible to the LTU; (ii) material expenditures, which accounts for the cost of all raw materials and services purchased by the firm in the production process generating tax credits in both the corporate income tax (CIT) and the VAT; (iii) labor expenditures, which accounts for the total wage bill of a firm, excluding social security contributions; (iv) average number of employees reported during the fiscal year; (v) accounting profit, i.e. the gross profit reported in the CIT; (vi) actual tax liability in the CIT, and (vii) taxable profit, which we calculate by applying the CIT schedule for firms that report a positive CIT liability.[28]

5 Results

First, we document and quantify the reported revenue response of Spanish firms to the notch in effective tax enforcement created by the existence of a LTU. Second, we examine the heterogeneity of this response across sectors of activity to test the hypothesis that monitoring effort and the enforcement technology are complements. To do this, we use the relative position of firms in the value chain as a proxy for the effectiveness of monitoring effort in each sector. Finally, we analyze the expected effect of the LTU on tax compliance by analyzing reported tax bases, intentifying manipulation in reported input expenditures as an additional channel of tax evasion.

which allows a more detailed analysis. Second, the *Banco de España* has digitalized some of the files submitted in paper form to the Commercial Registry that were not included in Amadeus.

[27] See Table A.3 in the online appendix for details. Official statistics publish the number of firms in the range between €3 and €10 million, so that is the benchmark we use in the analysis of the representativeness of our dataset.

[28] We provide details of the calculation of taxable profit in the online appendix.

5.1 Bunching Estimates: Reported Revenue Response

Figure 2 shows a histogram of the distribution of reported revenue for Spanish firms in the period 1995-2007, using administrative micro-data on financial statements from *Banco de España*. We focus on firms in the range between €3 and €9 million, such that the LTU threshold is in the center of the graph. There is substantial bunching just below the LTU threshold, indicating that a significant number of firms attempt to avoid stricter tax enforcement by keeping their reported revenue below €6 million.[29]

In order to quantify the excess bunching mass, we estimate the counterfactual distribution by fitting a flexible polynomial, as explained in section 3. Figure 3 shows the counterfactual density (dashed line) and the observed density (solid line), overlaid. For the average bunching estimator, we obtain $\widehat{b}_{av} = 0.094$ (*s.e.* 0.005) which is statistically different from zero at the 1% level. This point estimate implies that firms originally in the bunching interval reduce their reported revenue by about €94,000 *on average* (approximately 1.6% of their total reported revenue) in response to the tax enforcement notch. The interpretation of b_{av} as an average response derives from the assumption that firms are heterogeneous along two dimensions: the effectiveness of monitoring effort (ϕ) and the resource costs of evasion, $\kappa(u)$. The estimate is a weighted average of the bunchers' response and the lack of reaction of nonbunchers.

Two features of the distribution of reported revenue deserve additional discussion. First, bunching below the LTU threshold is somewhat diffuse, rather than a single spike as predicted by theory. This could be due to the indivisibility of some large transactions, which need to be either reported in full or not reported at all. Another explanation could be the artificial division of firms into multiple entities to avoid crossing the threshold. There is some anecdotal evidence of such behavior in Spain and also in other countries.[30] Notice that diffuse bunching biases our estimates *downwards*, because it pushes the estimated counterfactual above its true level. We check the sensitivity of bunching estimates to different choices of y^{lb} in the robustness checks below.

Second, we do not observe a "hole" with zero mass in the distribution to the right of the LTU threshold. Instead, there is a substantial number of firms reporting revenue just above y^L. The model with heterogeneous firms predicts the existence of mass in

[29] There is another modest spike in the distribution at about €4.75 million, associated to a requirement to perform an external audit of the accounts. This spike is smaller in magnitude and more difficult to interpret because the criteria to determine eligibility involve other variables besides operating revenue. Since this bunching response is very localized, it does not affect our main estimates. We provide additional details in the appendix, and for the rest of the paper we focus on the response to the LTU threshold.

[30] Firms could react by creating a new legal entity to transfer part of their activity (e.g. revenue). This division of the activity of one firm in multiple entities has been documented by the Spanish Tax Agency (*Plan General de Control Tributario*, several years). For Japan, Onji (2009) finds evidence that firms masquerade as multiple smaill firms in order to avoid compulsory VAT registration.

the bunching interval for two reasons: heterogeneity in monitoring effectiveness (ϕ), and variation in the resource costs function $\kappa(u)$ across firms. Taken together, these two features lead to the existence of a triangle of missing mass to the right of the LTU threshold, which can be seen empirically comparing the observed and counterfactual distributions of reported revenue in Figure 3. In addition to these two dimensions of heterogeneity, at the empirical level other optimization frictions may prevent firms from bunching. Such frictions, besides the resource costs of evasion, contribute to increasing the number of firms locating in the dominated region.

The empirical procedure presented in subsection 3.1 considers all adjustment costs (i.e., resource costs plus other frictions) together, allowing us to estimate bounds on the structural response to the policy. Using the adjusted bunching estimator, we obtain $\widehat{b}^{lb}_{adj} = 0.276$ ($s.e.$ 0.023). Under the assumption that resource costs are heterogeneous across firms and increasing with the distance to the threshold, this represents a lower bound of the structural response. The upper bound is given by the point of convergence between the observed and counterfactual densities. Convergence occurs at €6.53 million, so we express the upper bound as $\widehat{b}^{ub}_{adj} = 0.52$. These point estimates imply that, *conditional on bunching*, the average firm reduces its reported revenue by an amount between €276,000 to €520,000 (that is, 4.6% to 8.7% of total reported revenue) in order to avoid being monitored by the LTU. The gap between the upper and lower bounds of the structural response suggests that there is more uncertainty in the estimation of this parameter than in the case of the average bunching estimator (b_{av}).

Robustness. We examine the robustness of the bunching estimates reported above in several ways. First, the observed response could be affected by other size-dependent policies, such as a corporate income tax benefit that offers a 5 percentage-point lower tax rate to small firms.[31] We do not find any evidence of bunching in response to this tax break over time, even when we only include firms with positive taxable profits in the estimation. The lack of reaction to such a large reduction in the corporate income tax rate is remarkable in a context where firms respond strongly to a discontinuity in tax enforcement intensity. This evidence indicates that a relevant share of Spanish firms perceives that being under the LTU has a potentially large cost.[32]

[31] The eligibility threshold for this tax benefit changed over time, from €1.5 million in 1995 to €8 million in 2007, as shown in Table A.1 in the online appendix. The distribution of reported revenue under each of the thresholds is shown in Figure A.1.

[32] Devereux, Liu, and Loretz (2014) study firms' responses to corporate tax kinks in the UK and find significant bunching, implying taxable income elasticities between 0.13 and 0.56 for British firms. In their setting, kinks are set in terms of taxable profits, whereas in the Spanish case the tax incentive depends on operating revenue, which is considered less easily manipulable (Best et al., 2014; Carrillo, Pomeranz, and Singhal, 2014). This makes the strong bunching in reported revenue that we observe at the LTU threshold even more remarkable.

Second, pooling several annual cross-sections together increases the effective sample size allowing us to obtain very precise estimates, but it could mask differences in the response across years. Table 2 shows that bunching estimates are all significant and of similar magnitude in every year for the period 1995-2007.[33] Given that the bunching pattern is stable over time and that the LTU threshold remained fixed in nominal terms throughout the entire period, we use the pooled 1995-2007 dataset for our main analysis.

Third, in the results reported above, we set the lower bound of the excluded region equal to $y_{lb} = 5.7$ and the order of the polynomial to construct the counterfactual equal to 5. We estimate the proportion (α) of firms in the dominated region assuming the interval $(6, 6.1)$ as the baseline case. An exhaustive sensitivity analysis shows that the results are robust to different parameter choices.[34] In the empirical exercises in the next subsection, we focus on the average bunching estimator ($\widehat{b}_{av}$) for two reasons: first, the actual attenuated response is the most relevant for policy and welfare discussions, because adjustment costs do not seem to decline over time in this context. Second, the estimator $\widehat{b}_{av}$ is more robust to alternative parameter choices than $\widehat{b}_{adj}$.

5.2 Heterogeneous Responses: Complementarity Result

In the model with heterogeneous firms, we assume that the effectiveness of monitoring effort (ϕ) depends on the traceability of a firm's transactions. According to the theoretical predictions, we expect a larger response to the LTU threshold for upstream firms (which sell mostly to other firms) than for downstream firms (which sell mostly to final consumers). In practice, it is relatively easy to cross-check tax returns to detect misreported intermediate input sales because the buying firm has an incentive to record its expenses to claim tax credits. In the case of final sales, the consumer has no incentive to keep a receipt[35] and therefore it is significantly harder to cross-check those transactions

[33] The annual histograms are shown in Figure A.2 in the online appendix. In a previous version of this paper (Almunia and Lopez-Rodriguez, 2014), we studied the degree of persistence in bunching behavior over time, i.e. the possibility that some firms stay in the bunching interval for many consecutive years. We do not present those results here for space constraints. The conclusion from that analysis was that the static bunching estimates are no systematically biased due to persistent bunching behavior, because we only found evidence of short-term persistence (at most two or three years in the bunching interval).

[34] For the lower bound of the excluded region, we use values in the interval $y_{lb} \in \{5.3, 5.8\}$, and for the order of the polynomial we use values in the interval $q \in (4, 5)$. Table A.4 reports the results for the pooled 1995-2007 sample. The resulting upper bound y_{ub} (determined by the integration constraint $\widehat{B} = \widehat{H}$) is fairly stable between €6.38 and €6.71 million. Similarly, point estimates for $\widehat{b}_{av}$ are all in the interval $(0.089, 0.108)$ and those of $\widehat{b}_{adj}$ are in the interval $(0.256, 0.313)$. Finally, the polynomials of order 4 and 5 to fit the counterfactual distribution yield very similar results. Overall, we conclude that any sensible choice of the parameters yields similar estimates of the aggregate bunching response.

[35] One exception would be self-employed individuals, who have incentives to report part of their personal consumption as intermediate inputs to evade the VAT and the personal income tax. These misreported transactions appear as intermediate inputs in national accounts, rather than final consumer sales.

against other information sources, especially when they are made in cash.[36]

To study heterogeneous responses to the LTU threshold empirically, we define 16 sectors of activity.[37] Since we do not have transaction-level data for each firm, we use the percentage of sales made to final consumers in each sector (from input-output tables[38]) as a proxy for the traceability of sales made by firms in that sector. Figure 4 plots the percentage of final consumer sales (in the horizontal axis) against the average bunching estimates by sector (vertical axis). The relationship is downward-sloping, suggesting that the incentive to remain under the LTU threshold is stronger in sectors where a low percentage of sales is made to final consumers. On the top-left corner we observe sectors such as specialized construction activities, transportation, and metal and equipment manufacturing. Firms in these sectors sell less than 10% of their output to final consumers, and they all present high average bunching estimates, between 0.11 and 0.15. On the bottom-right corner, we find retailers and restaurants and hotels. These businesses obtain more than 80% of their revenue from sales to final consumers and they feature a much lower bunching response, between 0.06 and 0.07.[39]

The negative correlation between a high share of final consumer sales and the size of the bunching response at the enforcement notch is consistent with the predictions of our theoretical framework. Holding information requirements constant, the same LTU policy yields a different increase in the effectiveness of monitoring effort across firms depending on the traceability of their paper trail. The empirical result suggests that the deterrence effect associated to higher monitoring resources is most effective for firms whose misreported transactions are easier to detect. In contrast, the increase in monitoring resources is less binding for firms that sell mostly to final consumers. This does not tell us anything about the level of tax evasion in different sectors. It could well be that the actual level of evasion is higher for restaurants and hotels than for metal manufacturers. The key point is that in sectors with a high share of final sales, even a highly-skilled team of LTU auditors may be unable to detect evasion using standard methods such as information cross-checks. In conclusion, the empirical evidence suggests that there is complementarity between monitoring effort and the existence of a paper trail. The deterrence component that contributes to higher tax compliance derives from the interaction of these two elements, rather than from each of them independently.

[36] In a recent paper, Naritomi (2015) studies an innovative policy implemented in the state of Sao Paulo (Brazil), where consumers were incentivized to keep receipts from all their purchases (by entering a lottery and receiving a partial income tax rebate). She finds that the program had a much stronger effect on the sales reported by retail sector firms compared to wholesalers, consistent with our framework.

[37] Details about how we define each of the sectors can be found in the online appendix.

[38] We use the input-output tables for the Spanish economy in year 2000 published by the Institute of National Statistics (INE).

[39] The counterfactual and empirical distributions of revenue in all sectors are shown in the online appendix Figure A.3, and all the point estimates are reported in Table 3.

Robustness. The empirical pattern shown above is robust to several controls for potential omitted variables. One potential concern is that other firm attributes, such as the number of employees or the complexity of operations, might prevent firms from engaging in tax evasion. Kleven, Kreiner, and Saez (2015) make this point theoretically, arguing that large and more complex firms have more difficulties to reach a colluding agreement to evade taxes, as there is a higher chance that one of the employees may act as a "whistleblower".

Figure 5 shows a scatterplot of the median number of employees in each sector and the share of final sales. The relationship is weakly positive (i.e., sectors where firms have more employees feature *stronger* bunching), but it is not statistically significant. Table 4a shows a set of OLS regressions of average bunching estimates by sector on the share of final sales. In the simple regression (unconditional) case, the coefficient is -0.052 (s.e. 0.018), which implies that a 10 percentage-point (pp) increase in the share of final sales is associated with an average reported revenue response to the LTU threshold €5,200 lower in that sector. Controlling for the median number of employees, the ratio of tangible fixed assets to total revenue (a proxy for complexity) or both variables together does not affect the size nor the statistical significance of the coefficient of interest. Table 4b reports the results of weighted least squares (WLS) regressions, using the inverse of the variance of the bunching estimates as weights. The coefficient on the share of final sales decreases in absolute value to around -0.038, but it remains strongly significant and robust to using different controls in all columns.

These tests suggest that the complementarity result is not due to a systematic relation between firm characteristics such as firm size or complexity. The share of final sales is a robust determinant of the size of the bunching response in each sector, supporting the complementarity between monitoring effort and enforcement technology that we postulate in our theoretical framework.

5.3 LTU Effectiveness and the Effect on Tax Compliance

In subsections 5.1 and 5.2, we have documented a significant reduction in firms' reported revenue to avoid being monitored by the LTU, suggesting that this policy may be an effective tool to reduce tax evasion. In this subsection, we examine the nature of the behavioral response around the LTU threshold and also the impact of the policy on the behavior of firms monitored by the LTU. This is a first step towards analyzing the welfare implications of the policy, but it is challenging from an empirical point of view because bunching is an endogenous response, so we cannot apply a standard regression discontinuity design.

To examine the effect of the LTU on tax compliance, we extend our theoretical framework to allow firms to evade taxes by misreporting their input expenditures (besides misreporting their revenue) in the presence of multiple taxes, namely corporate income tax (CIT), value-added tax (VAT) and payroll tax (PRT). Enabling firms to misreport their inputs yields testable predictions on whether the bunching response is on average due to a change in real production or simply to tax evasion. We summarize the predictions of this extended model below and test them empirically using information on reported input expenditures and tax bases. We use these tests to rule out mechanisms that are inconsistent with theory, rather than to identify causal effects.

Theoretical Predictions with Multiple Taxes. The extended model with multiple taxes shows that firms have incentives to underreport revenue and overreport material expenditures in order to evade both VAT and CIT. Similarly, there is an incentive to underreport labor expenditures to evade PRT and avoid regulatory costs of labor.[40] Since the probability of detection is higher under the LTU, firms above the threshold have lower scope to evade taxes and they report larger amounts in each of the three tax bases.

To study the nature of the bunching response, we consider as outcome variables the average reported ratios of tax-deductible input expenditures (materials and labor) as a fraction of operating revenue. We consider the predictions of the model under three scenarios.[41]

(a) *Real production response*: bunching firms that reduce production have a higher producitivity draw (ψ) than those originally located below the threshold, so they need fewer inputs to produce the same amount of output. As a result, average material and labor expenditures as a fraction of revenue follow a downward slope in the interval (y_{lb}, y^L) due to the increasing share of bunchers, and both ratios shift upward at the threshold.

(b) *Evasion via revenue underreporting*: bunchers demand the same inputs (and produce the same output) as they would without an LTU, but they do not report a fraction of their revenue. This response mechanically increases both reported input ratios, yielding an upward trend in both ratios in the interval (y_{lb}, y^L) and a downward shift at the threshold.

(c) *Evasion via revenue and input misreporting*: firms with a larger scope to misreport inputs to minimize their tax liability (i.e., overreporting materials and underreporting

[40] Underreporting labor expenditures increases corporate tax liabilities, but this can be compensated by the tax savings on the payroll tax. During the period under study, the statutory payroll tax in Spain was 38% (including both the employer's and the employee's shares), compared to a corporate income tax rate that declined from 35% to 30%. Moreover, keeping reported salaries low and paying part under the table protects firms against future negative shocks, because there is downward nominal wage rigidity. There is theoretical support for wage underreporting via collusion agreements between employers and employees in Yaniv (1988, 1992), and pervasive empirical evidence of this practice in developing countries (Kumler, Verhoogen, and Frias, 2015; Best, 2013) and even in the US (Slemrod and Gillitzer, 2014).

[41] See section B in the online appendix for a mathematical derivation of these predictions.

labor costs) have stronger incentives to bunch to avoid being monitored by the LTU, because they obtain larger tax savings in expectation. Combined with revenue underreporting, this leads to an upward trend in the materials ratio in the interval (y_{lb}, y^L), and a downward shift at the threshold. The prediction is exactly the opposite for the ratio of labor expenses over reported revenue.

Empirical Evidence: Reported Input Expenditures. The top-left panel of Figure 6 plots the average reported ratio of material expenditures to reported revenue (vertical axis) against reported revenue (horizontal axis) for the period 1995-2007.[42] The ratio follows an upward slope in the reported revenue range between €3 and €6 million with a concave shape, indicating that firms with higher revenue use an increasingly larger proportion of material inputs. The relative use of material inputs shifts sharply downwards at the LTU threshold by almost two percentage points (from almost 80% to 78%). The bottom panel of Figure 6 shows a similar plot for the reported ratio of the total wage bill (net of payroll taxes) over revenue. The pattern in this case is approximately the reverse: the ratio slopes down smoothly in reported revenue with an upward jump of about one percentage point (from 10% to 11%) at the threshold.

Overall, this evidence is consistent with the scenario of evasion response with both revenue and input misreporting. Under this hypothesis, bunching firms are characterized by having (on average) a large scope to misreport their input expenditures, and they hide part of their revenue to avoid the stricter tax enforcement of the LTU. One potential alternative explanation for the observed patterns could be that labor-intensive firms are less likely to bunch due to prohibitive resource costs of evasion. Indeed, if such costs were systematically correlated with the number of employees, the bunching response would create a composition effect and there would be discontinuities in the input ratios compatible with the evidence. However, we can discard this possibility because there is essentially zero correlation between average bunching at the sector level and the median number employees in each sector, as we showed at the end of section 5.2.

Empirical Evidence: Reported Tax Bases. We study the implications for tax compliance by comparing reported tax bases by firms monitored by the LTU and those below the eligibility threshold. To do this, we exploit the richness of our dataset, which includes the reported tax bases and tax liabilities of the CIT and PRT, and a proxy for the tax base of the VAT.

[42] We trim the top and bottom 1% of all outcome variables in each €1-million interval in the range $\overline{y} \in (3, 9)$ million to prevent undue influence from outliers, and we use wider bins than in the histograms shown before (the bin width here is €120,202 instead of €60,101) to reduce the amount of noise in the bin averages. We do not adjust for inflation because the outcome variables are ratios of two nominal amounts.

Figure 7 shows average taxable profits (the tax base of the CIT) as a percentage of reported revenue.[43] We observe an upward shift in the average reported taxable profit margin from 5% below the threshold to 6% above. The average taxable profit margin is remarkably stable on both sides of the threshold, even far away from the interval affected by selection due to bunching (around €5.5-€6.5 million). This suggests that there is a "reporting regime shift" at the point where firms become eligible for the LTU, which cannot be driven only by selection. This 1pp gap in taxable profit margin is broadly consistent with the overreporting of materials and the underreporting of labor expenditures documented above. Indeed, Figure 8 documents the upward shift of both the average reported tax base and the average CIT liability. This is consistent with the reporting effect associated to the LTU, as there is no change in tax rate schedule associated to the threshold.

Figure 9a shows the average value added reported in financial statements (a proxy for the tax base of the VAT)[44] as a percentage of revenue. Again, we find a break at the LTU threshold, with an upward shift from approximately 19% to 22%, indicating that firms monitored by the LTU report a substantially higher value-added tax base. This structural break suggests that firms below the threshold are illegally appropriating part of the difference between the VAT they charged on their sales and the VAT they paid on their purchases. Failing to report the correct tax base, these firms are taking advantage of their role as fiscal intermediaries to keep part of the tax revenue that they should remit to the tax authority.

The tax base of the payroll tax is best approximated by the net wage bill. In Spain, the payroll tax consists of a combination of employer contributions (31% of the net wage in the period we study) and employee contributions (7% of the net wage). In the data, we separately observe the net wage bill and the employers' part of the payroll tax. As discussed above, Figure 6 shows the average net wage bill as a fraction of reported revenue. The ratio has a downward slope in the interval just below €6 million, and it shifts up from 10% to 11% at the LTU threshold. The structural break separates two significantly different patterns on either side of the tax enforcement regime, not just around the threshold. This evidence suggests that firms monitored by the LTU have

[43] We can only calculate the taxable profit margin for firms with positive taxable profits, because firms with negative taxable profits report a zero tax liability. This implies that the sample on which this figure is calculated is about 20% smaller than in the previous figures, which also included firms with zero tax liability. All the results presented in the paper are robust to using only the subsample of firms with positive tax liability.

[44] This measure of value added is calculated by the Bank of Spain as the difference between operating revenue and operating expenses. As some of the expenses are not considered in the VAT base, we construct an alternative measure of value added as the difference between operating revenue and material input expenditures, which are fully taxable under the VAT. The patterns are qualitatively similar suggesting a uniform distribution of VAT- exempt expenditures among firms.

lower scope to misreport their payroll tax base than firms below the threshold.

Taking the evidence on the three tax bases together, the empirical patterns depict a persistent tax regime shift associated to an arbitrary nominal threshold over a long period of time (1995-2007). This evidence suggests that the LTU effectively increases tax compliance on the subset of firms that it monitors. Firms below the LTU threshold, including both bunchers and low productivity firms, seem to have larger scope to evade taxes. On average, firms under stricter tax enforcement engage in less misreporting and the LTU policy is therefore equivalent to a broadening of reported tax bases. At a general level, we conclude that the deterrence channel is more effective at reducing evasion when the tax authority has more resources to monitor the information trails making full use of the enforcement technology.

We observe similar tax-reporting patterns when we disaggregate the analysis by five broad sectors of activity, as shown in Figure A.4 (for the corporate income tax base) and Figure A.5 (for the value-added tax base). These results indicate that the tax-reporting regime shift observed in the aggregate data is not due to a systematic correlation between the strength of the bunching response and the sectoral composition of that response.

6 Welfare Analysis and Policy Implications

We use the baseline model to examine analytically the welfare effects of increasing tax enforcement on firms. First, we analyze the net welfare change from increasing monitoring effort across firms and returning the additional revenue lump sum to all taxpayers. Second, we examine the mechanisms and conditions under which a firm size-dependent tax enforcement is a welfare improving policy. Finally, we use the analytical insights and the empirical results on the Spanish LTU to illustrate the potential welfare gains of extending stricter tax enforcement to medium-sized and small firms.

6.1 Welfare Effects of Increasing Tax Enforcement

Consider the baseline economy introduced in section 2. Without loss of generality, assume that each firm is owned by one individual whose total income is the net of tax profit of the firm. The government devotes fixed costs F to create both a tax system with information-reporting requirements and a tax authority in charge of tax enforcement with a given technology. The tax authority spends $c(\phi)$ resources per taxpayer to employ qualified staff to examine the discrepancies detected by the enforcement technology and undertake tax audits to uncover tax evasion. The administrative costs of tax enforcement, $AC(\phi) = F + c(\phi)$, are increasing, $c_\phi > 0$, and convex, $c_{\phi\phi} > 0$, in monitoring effort ϕ (which is assumed constant across firms in this baseline model). The social welfare function that aggregates expected profits by firms and expected tax revenue by the government, net of administrative costs of enforcement, is given by

$$\begin{aligned}\mathbb{E}W &\equiv \int_{\overline{y}_{\min}}^{\overline{y}_{\max}} \{(1-t)P - qz - \kappa(u) + tur\} \cdot g_0(\overline{y})d\overline{y} \\ &+ \int_{\overline{y}_{\min}}^{\overline{y}_{\max}} \{t(P-u) + \phi h(u)tu(1+\theta)\} \cdot g_0(\overline{y})d\overline{y} - AC(\phi). \end{aligned} \quad (14)$$

By the envelope condition, the increase in monitoring effort has only a mechanical first-order effect on the firms' expected profits, because firms have already made their real and reporting choices. The marginal change in welfare in response to an increase in monitoring effort, is, therefore,

$$\begin{aligned}\frac{d\mathbb{E}W}{d\phi} &= \int_{\overline{y}_{\min}}^{\overline{y}_{\max}} \left[t\frac{d\overline{P}}{d\phi} + t\phi(1+\theta)\left(\frac{\partial h}{\partial u}\frac{du}{d\phi}u + h(u)\frac{du}{d\phi}\right)\right] \cdot g_0(\overline{y})d\overline{y} - c_{\phi}(\phi) \\ &= \int_{\overline{y}_{\min}}^{\overline{y}_{\max}} \left[t\frac{dP}{d\phi} - \kappa_u(u)\frac{du}{d\phi}\right] \cdot g_0(\overline{y})d\overline{y} - c_{\phi}(\phi), \end{aligned} \quad (15)$$

where the second equality follows from firms' first order condition of net-of-tax profit maximization with respect to concealed revenue. Comparative statics indicates that the increase in monitoring effort without changing the relative cost of inputs does not modify the demand of inputs and thus it does not affect real production, $dP/d\phi = 0$.[45] Arranging terms, we obtain a condition showing that the increase of tax enforcement is a welfare-improving policy when the marginal welfare gains from a lower level of tax evasion are larger than the marginal administrative costs of achieving that level, such that

$$\int_{\overline{y}_{min}}^{\overline{y}_{max}} \left[-\kappa_u(u) \cdot \frac{du}{d\phi}\right] \cdot g_0(\overline{y})\, d\overline{y} > c_{\phi}(\phi). \quad (16)$$

This expression indicates that the welfare gains of increasing tax enforcement depend positively on i) the marginal resource costs of evasion, $\kappa_u(u)$, because the more costly evasion is the larger the gains of reducing it; and ii) the reduction of misreported income due to the increase in monitoring effort. The welfare gains are negatively related with the initial monitoring effort, ϕ, due to convex administrative costs, indicating that there are decreasing returns to increasing tax enforcement.

Condition (16) is derived according to the procedure followed in Chetty (2009a) to obtain the generalization of Feldstein (1999)'s condition to examine the excess burden of taxation. In contrast with the welfare change induced by distortionary taxes, increasing

[45] Our empirical results for the Spanish LTU support this prediction as we do not find any evidence of real responses (i.e. fall of production) to the change in monitoring effort.

monitoring effort does not distort relative input prices and thus it broadens reported tax bases without creating production distortions. Indeed, raising monitoring effort reduces firms' incentives to misreport tax bases creating both direct welfare gains through the reduction of resource costs and a redistributive effect by transferring income among taxpayers. Notice that the welfare gains of raising tax enforcement could be even larger if the extra revenue obtained by the higher tax base reporting was devoted to fund valuable public spending and/or to decrease corporate income tax rates. The latter possibility would reduce the excess burden of taxation justifying an expansion of the tax agency's resources as discussed in Slemrod and Yitzhaki (1987).

The conditions derived in the extended model with heterogeneous firms (discussed in subsection 2.2) show that firms could respond differently to increases of tax enforcement because i) the monitoring effort to detect tax evasion is more effective on firms whose transactions are easier to trace with paper trail (i.e. higher sensitivity of the tax base to monitoring effort); and ii) the costs of misreporting transactions can be heterogeneous across firms, and even so high that a proportion of taxpayers do not react to avoid lager enforcement but report larger tax bases. Given these insights, according to condition (16), welfare gains of increasing tax enforcement would be higher in economies with a higher proportion of upstream firms that sell mostly to other businesses because their misreported transactions are more likely to be detected. Welfare gains would also be higher when tax enforcement policy reduces misreported transactions by evaders that bear high resource costs, because reducing evasion by these taxpayers creates larger reductions in the social costs of evasion.

6.2 Size-Dependent Tax Enforcement

Most governments around the world have established LTUs within their tax agencies to strengthen tax administration and improve tax compliance by the largest taxpayers, often by recommendation of the IMF (IMF, 2002). Despite the widespread adoption of this policy in recent decades, there have been very few attempts to evaluate its welfare effects beyond noting that it has been generally successful at increasing total tax revenues. In practice, there is variation across countries in the specific eligibility criteria, which usually include either total revenue, number of employees, or both. However, the overarching feature of LTUs is that they lead to size-dependent tax enforcement intensity by applying higher monitoring effort per taxpayer to a subset of large firms. We develop a framework to analyze the welfare implications of the policy, taking into account the local distortions generated by bunching to avoid the LTU and also the behavioral responses within the LTU. Using this framework and the estimates from our empirical application, we do a simple welfare calculation of the effects of the Spanish LTU to illustrate the potential welfare gains associated to extending stricter tax enforcement to smaller firms.

Given the social welfare function defined by (14), the welfare change induced by the LTU with an eligibility threshold y^L is given by

$$\begin{aligned}
\Delta \mathbb{E} W = & \int_{y^L}^{y^L+d\overline{y}^M} \left[\Delta \Pi^R + \Delta u \cdot \kappa_u(u)\right] \cdot (1-\alpha(\overline{y})) \cdot g_0(\overline{y}) d\overline{y} + \\
& + \int_{y^L}^{y^L+d\overline{y}^M} \left[\Delta u \cdot \kappa_u(u)\right] \cdot \alpha(\overline{y}) \cdot g_0(\overline{y}) d\overline{y} + \\
& + \int_{y^L+d\overline{y}^M}^{\overline{y}^{\max}} \left[\Delta u \cdot \kappa_u(u)\right] \cdot g_0(\overline{y}) d\overline{y} + \\
& + \int_{y^L+d\overline{y}^M}^{\overline{y}^{\max}} \left[c(\phi_1) - c(\phi_0)\right] \cdot g_0(\overline{y}) d\overline{y}
\end{aligned} \tag{17}$$

where $\Delta \Pi^R$ is the change in gross profits (i.e. $\Pi^R = P - qz$) and Δu is the change in reported tax base, both changes induced by the increase in monitoring effort; and $\alpha(\overline{y})$ is the proportion of nonbuncher firms at each level of pre-LTU reported revenue.

The welfare change associated to a size-dependent tax enforcement policy is thus the result of four components. The first one captures the local welfare loss created by the bunchers' reaction. Bunchers reduce their reported revenue creating welfare losses through either the additional resource costs (evasion channel) and/or the reduction in output (real channel). The second component measures the welfare gains due to the existence of adjustment cost that prevent some firms from bunching. These nonbunchers report larger tax bases, reducing their resource cost of evasion. Similarly, the third component captures the positive welfare gains of the LTU's deterrence on high productivity firms, which report larger tax bases and hence incur lower resource costs. Although firms monitored by the LTU are worse off because of the higher tax liability that they pay, this income transferred to the government without affecting aggregate welfare (i.e. a pure redistributive effect). The main mechanism to create welfare gains is due to the reduction in the resource costs associated to evasion. The final component is the additional administrative costs devoted to increase monitoring effort for firms within the LTU. Overall, the policy is thus welfare improving when the welfare gains created by the reduction of evasion on firms monitored by the LTU, net of the administrative costs, are larger than the local welfare costs created by bunchers' reaction.

6.3 Welfare Calculations: Application to the Spanish LTU

We evaluate now the welfare effect of expanding the scope of the LTU to one additional firm, using the insights from equations (16) and (17) and the empirical results from our application for the Spanish LTU. Framing the question as a marginal reform from the status quo is a simple way to explore the existence of potential marginal net welfare gains of the policy.[46] Given our analysis, extending the LTU monitoring effort to an additional firm would be welfare improving if the reduction of resource costs due to lower tax evasion is higher than the additional administrative costs required to reduce that evasion, given some local welfare losses associated to bunching. The existence of substantial net welfare gains at the margin would indicate the efficiency of spending additional monitoring resources on smaller firms, and also suggest the presence of sizeable welfare gains in the interior of the LTU.

The marginal resource cost of evasion is a critical measure to evaluate the welfare gains of extending the policy, but as dicussed by Chetty (2009a) it is quite challenging to estimate it. We propose a method to put bounds on this measure by looking at tax reporting behavior by firms in the dominated region. We defined the dominated region as the revenue interval where firms affected by higher monitoring effort could save taxes by misreporting revenue, but instead report larger tax bases because of adjustment costs. Assuming that these costs are pure resource costs of evasion,[47] we estimate that nonbunchers in the dominated region report on average an additional one percentage point in taxable profit margin than firms under lower monitoring effort.[48] Over the period of analysis, firms in the dominated region report on average an additional €60,000 of taxable profits which, at a 32.5% tax rate, result in €19,500 of taxes that they could avoid paying by simply reporting €1 less in total revenue. This suggests that these firms bear resource costs of at least 5.9% of their gross income (note that average gross taxable profit for firms around the LTU threshold is €330,000), which prevent them from reacting. Given the shift in the average taxable profits, the inclusion of an additional firm in the LTU yields a welfare gain of €3,300, due to the reduction of resource costs of evasion. As an alternative, taking also into account the potential tax savings in multiple taxes such as the PRT (1pp of tax base shift under a 38% tax rate representing €22,800 of tax liability) and the VAT (3pp of tax base shift at 11.5% of effective tax rate over the period

[46] Conducting a full welfare evaluation of the LTU policy as a whole would require making strong assumptions about how the local estimates obtained at the threshold extrapolate to larger firms. That exercise would incorporate a much larger degree of uncertainty than the one we develop in this subsection. See Kanbur and Keen (2014) for an analytical discussion on the complexity of determining the optimality of tax thresholds.

[47] If firms do not react because of misinfomation about the tax enforcement strategies, the proposed bounds would be biased upwards. However, the proposed measure captures mosts of the costs such as managerial effort devoted to collecting information and understanding the effects of the policy, the resources devoted to both respond to the policy and keeping double accounting, the violation of ethical principles, the risk aversion of taxpayers, or the costs of maintaining an inefficiently small firm (e.g., foregoing business opportunities).

[48] As discussed in section 5.3, the parallel reporting regime shift of 1% in the corporate income tax base is documented well beyond the revenue region affected by selection. Besides, evidence suggests that on average bunchers mimic tax reporting by small firms, making the 1% gap stable in the whole revenue interval we study.

resulting in an additional €20,700 tax remittance), the marginal resource costs could be as large as 19% of their gross income. Given the average increase in reported tax bases of approximately €300,000, the welfare gain generated by marginally extending the LTU would be up to €57,000 in the most optimistic scenario. The bounds on the marginal resource costs of evasion (5.5%-19%) put our estimates in the same order of magnitude as the ones obtained by Gorodnichenko, Martinez-Vazquez, and Peter (2009) in a study about the Russian flat tax reform.

To quantify the marginal administrative cost of including one additional firm in the LTU, we need to estimate the average additional cost per taxpayer associated to the monitoring effort of the LTU (i.e., $c(\phi_1) - c(\phi_0)$). According to official data published by the Spanish tax agency (AEAT, 1995-2007; TdC, 2008), over the period of analysis the LTU has been endowed with 125 highly-qualified tax auditors earning an average annual compensation between €60,000 and €80,000.[49] This group of auditors is in charge of monitoring the approximately 30,000 firms in the LTU Census. Hence, we can approximate the additional cost per taxpayer in the LTU to be in the range of €250-€333.33.[50] The small magnitude of the marginal cost per taxpayer illustrates the potential gains of exploiting the fixed-cost structure of the tax agency by extending higher monitoring effort to smaller firms below the current LTU eligibility threshold.[51]

Indeed, our simple calculations show that broadening the LTU scope at the margin creates substantial net welfare gains by reducing resource costs of evasion in a magnitude much larger than the additional administrative costs required. This welfare gap is due to the non-exhausted complementarity between monitoring effort and enforcement technology that leads to higher tax compliance induced by the thread of audit. Once the fixed-cost investment of creating a tax system with information trails and third-party reporting is made, the low marginal cost of increasing human resources to exploit the existing information trails creates complementarities and decreases the average administrative costs of tax enforcement.

The existence of significant welfare gains at the margin also indicates the presence of sizeable welfare gains in the interior of the LTU. As discussed in section 5.3, the LTU

[49] These figures correspond to, the last in the period under analysis. The average annual compensation of LTU tax auditors was lower in nominal terms in the early years of the period, but we take a the largest numbers to obtain a conservative estimate of the administrative cost of extending the LTU.

[50] Again, this is a conservative estimate because we are assuming that the marginal cost of including one additional taxpayer with reported revenue just below €6 million in the LTU census is equal to the *average* cost of monitoring taxpayers with reported revenue above that level.

[51] Official reports (AEAT, 1995-2007) show that tax auditors in the LTU perform intensive verifications on approximately 11% of LTU firms each year (i.e., roughly over 3,000 firms), mainly through electronic notifications from cross-checkings of information sources. As an example of the potential tax revenue gains of the policy, official statistics indicate that in 2004 the direct effect of the LTU monitoring through verifications yielded more than €800 million in additional tax revenue, to which we should add the indirect impact on other non-audited taxpayers' behavior due to the threat of audit effect.

creates a tax compliance effect on large firms well-above the threshold implying significant increases of tax bases (in euros) and thus creating substantial welfare gains. Given the small magnitude of the local welfare costs and the low additional administrative costs of creating the LTU, the evidence suggests that the Spanish LTU has not been only effective to increase tax revenue capacity but also to raise welfare.[52]

7 Concluding Remarks

In this paper, we have investigated the effectiveness of exploiting the information trails generated by business activities to enforce tax compliance. We have first derived theoretical predictions on how firms respond to increases in the resources used by the tax authority to verify the transactions reported by firms. We have then tested the predictions of the model using quasi-experimental variation in monitoring effort provided by the Large Taxpayers Unit (LTU) in Spain.

The empirical results show that firms react to avoid being under stricter tax enforcement by reducing their reported revenue just below the LTU eligibility threshold. This reaction is heterogeneous among firms depending on the traceability of their transactions, indicating the complementarity between monitoring effort and information requirements to increase tax compliance. In particular, we have found a larger bunching response in sectors that sell intermediate goods, where information trails are easier to verify by using more monitoring resources. Finally, we have documented that firms monitored by the LTU report broader tax bases, indicating that the policy is effective to reduce tax evasion. In contrast, small firms not monitored by the LTU have larger scope to misreport both labor and material expenditures to evade multiple taxes because of the low monitoring effort applied to them. Incorporating these findings into a welfare calculation, we conclude that devoting additional resources to extend stricter tax monitoring to smaller firms would generate net welfare gains.

The results of the paper highlight the relevance of monitoring the information trails created by firms to ensure tax compliance. Firms play a very important role to prevent individual income tax evasion by submitting a third-party report of employees' salaries. However, tax authorities must devote adequate resources to verify business activities, which are notably more complex than those of individuals, to achieve overall effective tax enforcement.

[52] According to official data (AEAT, 1995-2007), the additional administrative costs devoted to the Spanish LTU are on average €7.5-€10 million per year. Assuming a full evasion response of the bunchers (on average 250 firms per year) that on average misreport revenue by €300,000, implies that the local welfare costs are in the range of €4.42-€14.25 million, given the discussed bounds on the marginal resource costs of evasion. The observed tax base reporting regime shift above the treshold more than compensates for these losses.

References

AEAT. 1995-2007. "Memorias de Administración Tributaria." Tech. rep., Agencia Estatal de Administración Tributaria (Spanish Tax Authority).

Allingham, Michael and Agnar Sandmo. 1972. "Income tax evasion: A theoretical analysis." *Journal of Public Economics* 1:323–338.

Almunia, Miguel and David Lopez-Rodriguez. 2014. "Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Firms." Working Paper 1419, Banco de España.

Bastani, Spencer and Hakan Selin. 2014. "Bunching and Non-Bunching at Kink Points of the Swedish Tax Schedule." *Journal of Public Economics* 109:36–49.

Bentolila, Samuel, Juan J. Dolado, and Juan F. Jimeno. 2012. "Reforming an insider-outsider labor market: the Spanish experience." *IZA Journal of European Labor Studies* 1 (4).

Best, Michael C. 2013. "Salary Misreporting and the Role of Firms in Workers' Response to Taxes: Evidence from Pakistan." Working paper, London School of Economics.

Best, Michael C., Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2014. "Production vs Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* forthcoming.

Best, Michael C. and Henrik J. Kleven. 2015. "Housing Market Responses to Transaction Taxes: Evidence from Notches and Stimulus in the UK." London School of Economics.

Bird, Richard and Pierre-Pascal Gendron. 2007. *The VAT in Developing and Transitional Countries*. Cambridge University Press.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal. 2014. "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement." Working paper, Harvard Kennedy School.

Chetty, Raj. 2009a. "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.

———. 2009b. "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Analysis." *Annual Review of Economics* 1:451–488.

Chetty, Raj, John Friedman, Tore Olsen, and Luigi 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, Raj, John Friedman, and Emmanuel Saez. 2013. "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings." *American Economic Review* 103 (7):2683–2721.

Devereux, Michael, Li Liu, and Simon Loretz. 2014. "The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records." *American Economic Journal: Economic Policy* 6 (2):19–53.

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

Garicano, Luis, Claire LeLarge, and John van Reenen. 2013. "Firm Size Distortions and the Productivity Distribution: Evidence from France." NBER Working Paper 18841.

Gordon, Roger and Wei Li. 2009. "Tax Structures in Developing Countries, Many Puzzles and a Possible Explanation." *Journal of Public Economics* 93:855–866.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter. 2009. "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effect in Russia." *Journal of Political Economy* 117 (3):504–554.

Guner, Nezih, Gustavo Ventura, and Yi Xu. 2008. "Macroeconomic Implications of Size-Dependent Policies." *Review of Economic Dynamics* 11:721–744.

IEF. 2011. "Evolución del Sistema Fiscal Español: 1978-2010." Instituto de Estudios Fiscales.

IMF. 2002. "Improving Large Taxpayers' Compliance: A Review of Country Experience." Tech. rep., International Monetary Fund, Washington DC. Team led by Katherine Baer.

IRS. 2012. "Tax Gap for Year 2006." URL `www.irs.gov/file_source/pub/newsroom/overview_tax_gap_2006.pdf`. Internal Revenue Service.

Kanbur, Ravi and Michael Keen. 2014. "Thresholds, informality, and partitions of compliance." *International Tax and Public Finance Tax Public Finance* 21:536–559.

Keen, Michael and Benjamin Lockwood. 2010. "The value added tax: Its causes and consequences." *Journal of Development Economics* 92 (2):138–151.

Keen, Michael and Jack Mintz. 2004. "The Optimal Threshold for a Value-Added Tax." *Journal of Public Economics* 88:559–576.

Kleven, Henrik, Martin Knudsen, Claus Kreiner, Soren Pedersen, and Emmanuel Saez. 2011. "Unable or Unwilling to Cheat? Evidence from a Tax Audit Experiment in Denmark." *Econometrica* 79 (3):651–692.

Kleven, Henrik, Klaus Kreiner, and Emmanuel Saez. 2015. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries." Working paper, UC Berkeley.

Kleven, Henrik and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics* 128 (2):669–723.

Kopczuk, Wojciech and David J. Munroe. 2015. "Mansion Tax: The Effect of Transfer Taxes on the Residential Real Estate Market." *American Economic Journal: Economic Policy* 7 (2):214–57.

Kopczuk, Wojciech and Joel Slemrod. 2006. "Putting Firms into Optimal Tax Theory." *American Economic Review, Papers and Proceedings* 96 (2):130–134.

Kumler, Todd, Eric Verhoogen, and Judith Frias. 2015. "Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico." Working paper, Columbia University.

Liu, Li and Benjamin Lockwood. 2015. "VAT Notches." CESifo Working Paper No. 5371.

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

Naritomi, Joana. 2015. "Consumers as Tax Auditors." Working paper, London School of Economics.

OECD. 2011. "Tax Administration in OECD and Selected non-OECD Countries: Comparative Information Series (2010)." Tech. rep., Organisation for Economic Cooperation and Development.

Onji, Kazuki. 2009. "The response of firms to eligibility thresholds: Evidence from the Japanese value-added tax." *Journal of Public Economics* 93:766–775.

Paula, Áureo De and José Scheinkman. 2010. "Value-Added Taxes, Chain Effects, and Informality." *American Economic Journal: Macroeconomics* 2:195–221.

Pomeranz, Dina. 2015. "No Taxation Without Information: Deterrence and Self-Enforcement in the Value-Added Tax." *American Economic Review* 105 (8):2539–69.

Restuccia, Diego and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics* 11:707–720.

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

Slemrod, Joel. 2010. "Buenas Notches: Lines and Notches in Tax System Design." Working paper, University of Michigan.

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79:455–483.

Slemrod, Joel, Brett Collins, Jeffrey Hoopes, Daniel Reck, and Michael Sebastiani. 2015. "Does credit-card information reporting improve small-business tax compliance?" Working paper, University of Michigan.

Slemrod, Joel and Christian Gillitzer. 2014. *Tax Systems.* Cambridge: MIT Press.

Slemrod, Joel and Shlomo Yitzhaki. 1987. "The Optimal Size of a Tax Collection Agency." *The Scandinavian Journal of Economics* 89 (2):183–192.

TdC. 2008. "Informe de Fiscalización de las Unidades de Gestión de Grandes Empresas de la AEAT." Tech. rep., Tribunal de Cuentas.

Yaniv, Gideon. 1988. "Withholding and Non-Withheld Tax Evasion." *Journal of Public Economics* 35:183–204.

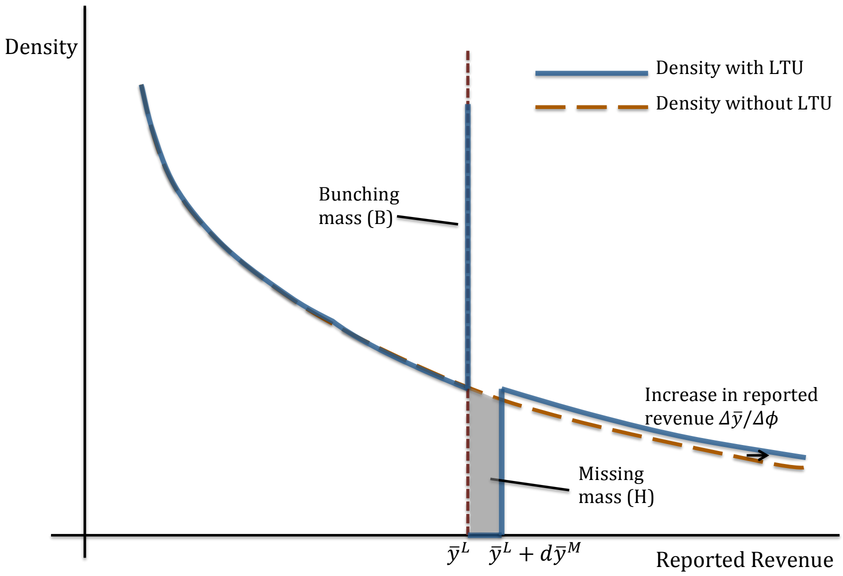
———. 1992. "Collaborated Employee-Employer Tax Evasion." *Public Finance* 47 (2):312–321.

Yitzhaki, Shlomo. 1974. "A Note on "Income Tax Evasion: A Theoretical Analysis"." *Journal of Public Economics* 3 (2):201–202.

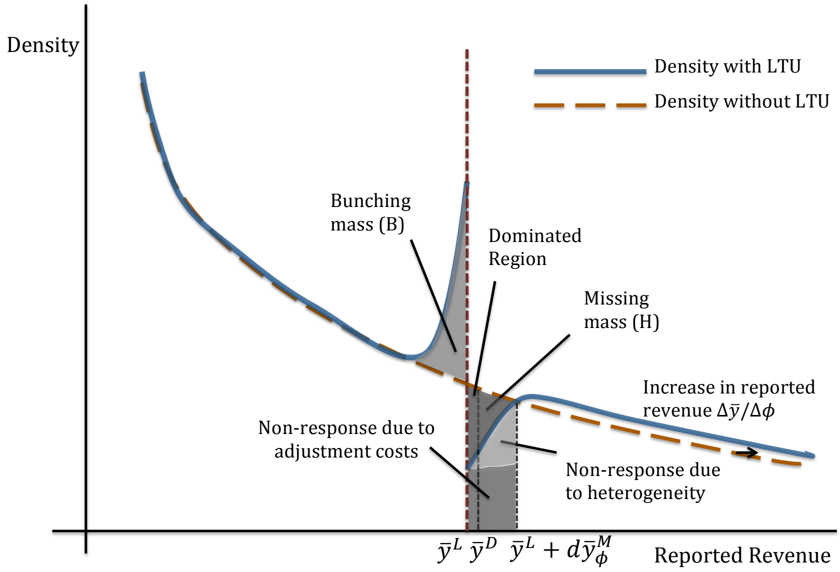
Figures

Figure 1: Theoretical Distribution of Reported Revenue

(a) Baseline Model

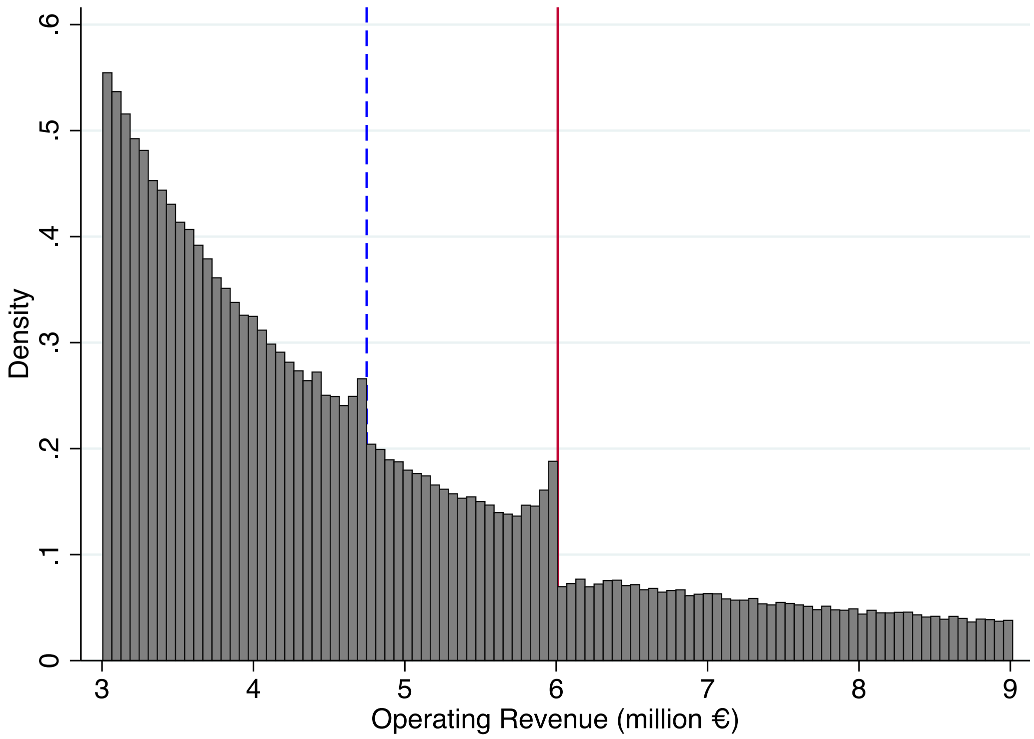


(b) Model with Heterogeneous Firms



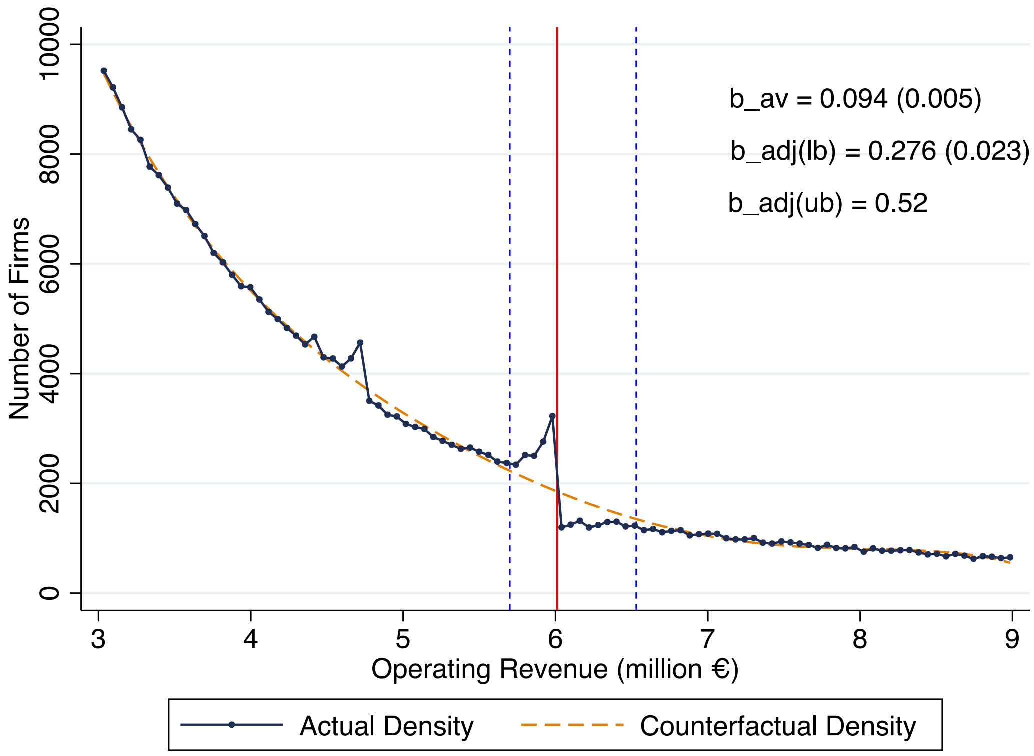
Notes: this figure depicts the theoretical revenue distribution before and after the introduction of the Large Taxpayers Unit (LTU). Without the LTU, all firms face the same monitoring effort and the distribution of revenue is smoothly decreasing as depicted by the dashed (brown) line. When the LTU is introduced, firms reporting revenue above y^L face a higher enforcement intensity. A group of firms in an interval above y^L respond to the new policy by underreporting more of their revenue to report exactly $\bar{y} = y^L$. This generates a spike at the threshold (with excess mass B), and an area of missing mass (H) to the right of the threshold, as depicted by the solid (blue) line. Panel (a) corresponds to the baseline model, where monitoring effort and resource costs of evasion are the same for all firms, so all firms with the same productivity draw respond identically to fiscal incentives. Thus, no firm locates in the interval of length $d\bar{y}^M$ to the right of the LTU threshold. Panel (b) depicts the equilibrium in the model with heterogeneity, where a fraction of firms does not respond to the incentives due to different changes in monitoring effort at the threshold or due to other adjustment costs.

Figure 2: Operating Revenue Distribution, 1995-2007



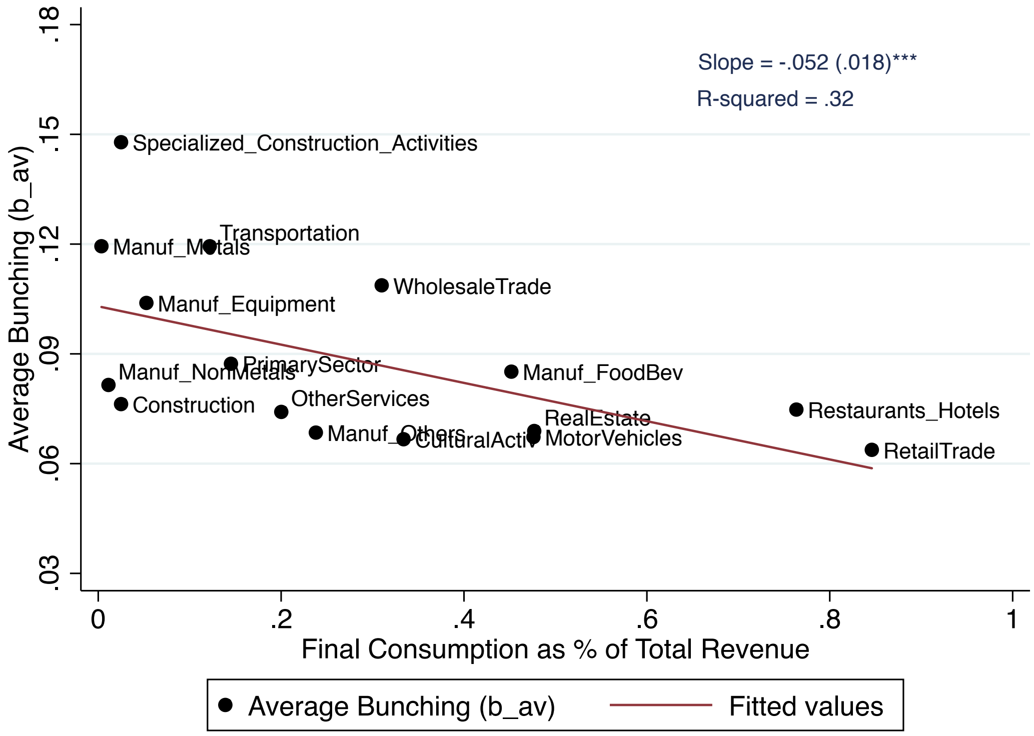
Notes: this graph shows the distribution of operating revenue reported by Spanish firms, pooling annual observations from the period 1995-2007. The solid (red) line indicates the Large Taxpayers Unit (LTU) threshold (€6 million); the dashed (blue) line indicates the revenue threshold (€4.75 million) for the External Audit requirement (see online appendix for more details on this regulatory threshold). The bins are €60,101 wide, delimited such that no bin contains data both to the left and to the right of the relevant policy thresholds. The total number of observations is 285,570.

Figure 3: Counterfactual vs. Observed Operating Revenue, 1995-2007



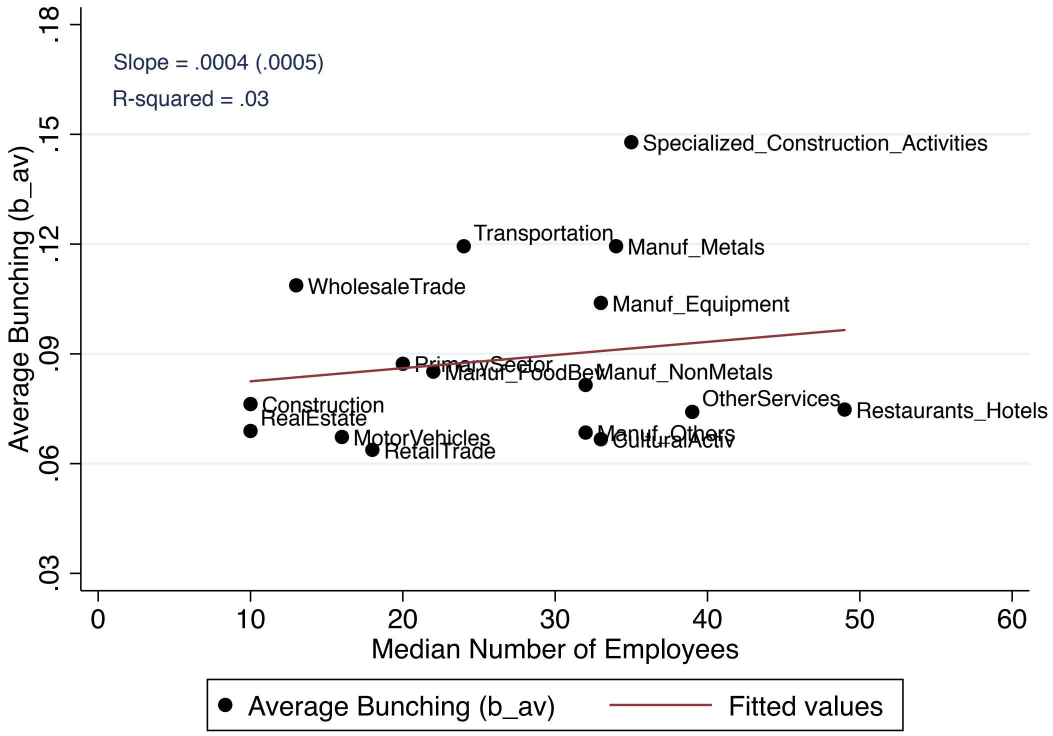
Notes: this graph shows the reported distribution of revenue (dots connected by solid blue line) and the estimated counterfactual (orange dashed curve) for the period 1995-2007. The data for the true distribution are exactly the same as those used to construct the histogram in figure 2. The vertical dotted blue lines indicated the bounds of the excluded region (y_{lb} and y_{ub}) chosen for the estimation of the counterfactual. To determine the value of y_{ub}, we fit a 5th-degree polynomial to the true density in multiple iterations, starting with $y_{ub} \approx y^L$ and increasing the value in small steps until we reach a point where the bunching mass (B) equals the missing mass (H), so that the integration constraint is satisfied. The average bunching parameter (b_{av}) estimates the adjustment in reported revenue for the average firm above the threshold, while (b_{adj}) estimates the adjustment for the marginal bunching firm, accounting for the existence of resource costs of evasion that prevent some firms from responding to the notch. (N = 285,570. Bin width = €60,101).

Figure 4: Average Bunching by Share of Final Consumer Sales



Notes: the bunching measure $\widehat{b}_{av}$ is calculated for each sector as explained in section 3 in the main text. Final consumption as a share of total sales in each sector is calculated using the year 2000 input-output tables for the Spanish economy, published by the National Statistics Institute (INE). (More details about this data source are provided in the online appendix.) The figure shows a negative relationship between the average bunching response (b_{av}) and the percentage of sales made to final consumers in each sector. This suggests that firms with easily traceable revenue are more concerned about crossing the LTU threshold, because their ability to evade taxes by misreporting revenue is substantially reduced in the LTU. The slope coefficient reported is obtained by estimating a simple linear regression of the bunching estimates on final consumption share in each sector, using robust standard errors.

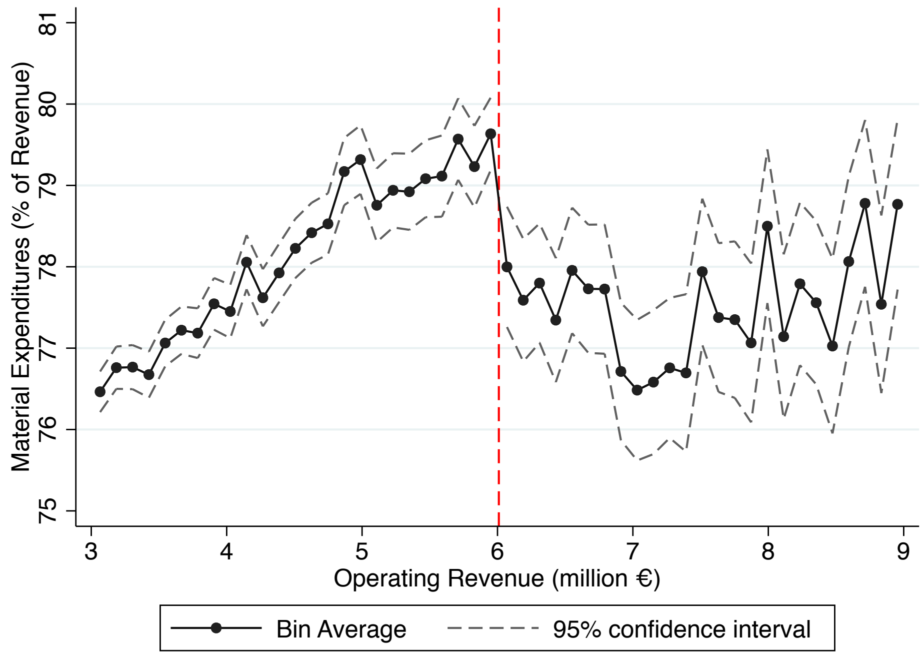
Figure 5: Average Bunching by Median Number of Employees



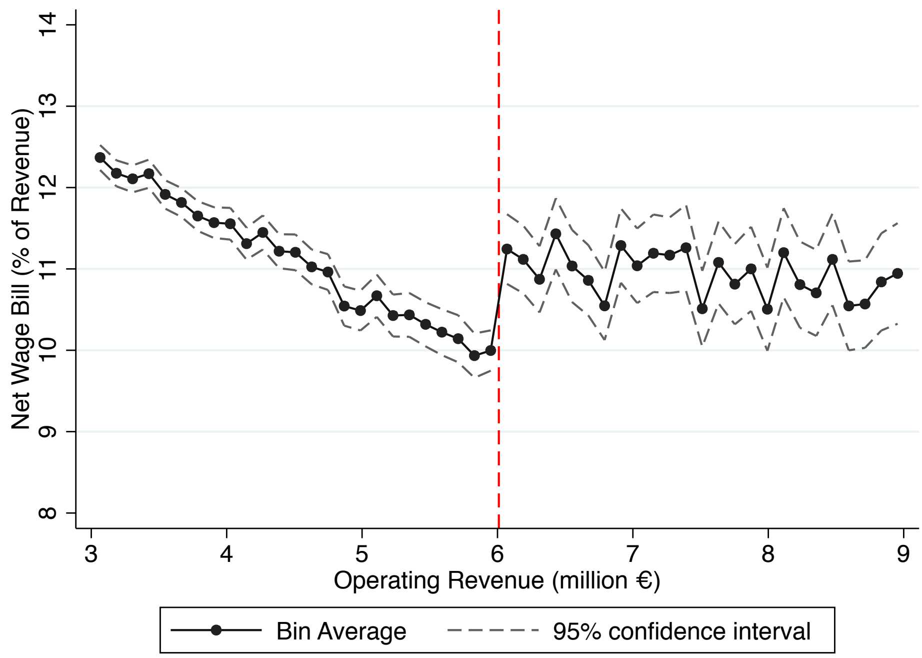
Notes: this graph is similar to figure 4, but in this case the horizontal axis shows the median number of employees by firms in each sector, calculated for firms in the revenue range €3-€9 million. The figure shows a weak relationship between the average bunching report and the median number of employees in each sector. The coefficient of the linear fit is not statistically different from zero. This suggests that employment size is not related to bunching behavior at the sector level.

Figure 6: Input Expenditures

(a) Material Expenditures

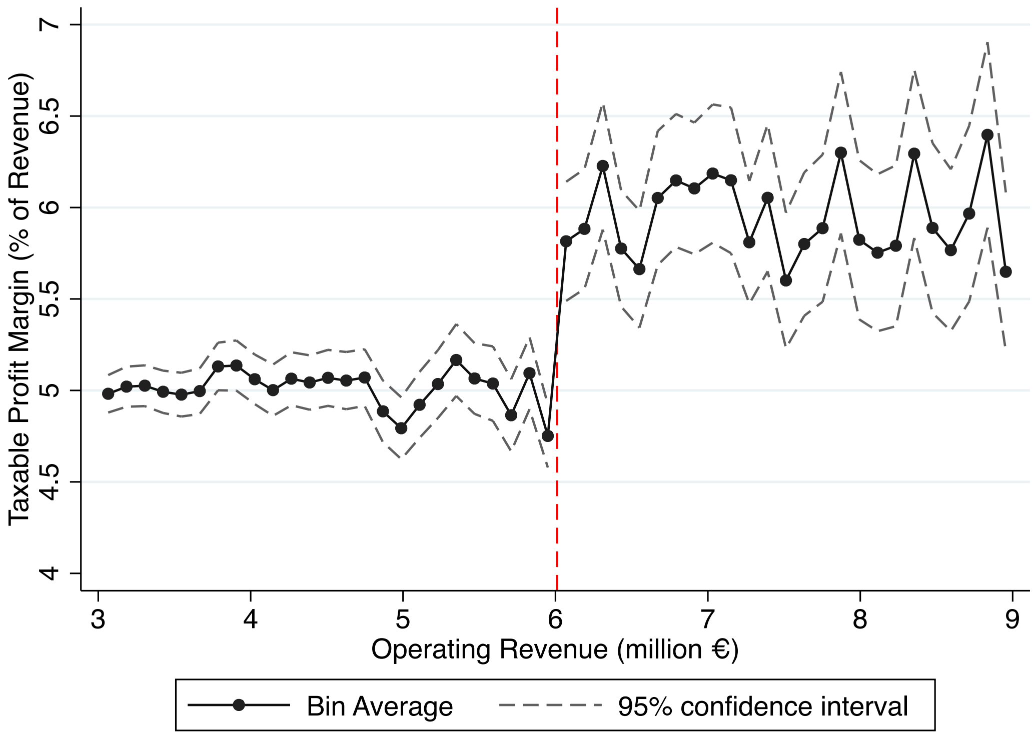


(b) Labor Expenditures



Notes: these graphs show the average ratio of input expenditures over revenue (vertical axis) against reported revenue (horizontal axis), for the period 1995-2007. The dashed (red) vertical line indicates the LTU threshold. The dotted lines denote bin averages and the grey dashed lines show 95% confidence intervals for each bin average. We trim outliers in the data by dropping the observations in the top and bottom 1% of the outcome variable in each €1-million interval in the range $\overline{y} \in (3, 9)$ million. "Material input expenditures" includes all the intermediate inputs used by the firm for production. "Labor input expenditures" is the total wage bill of the firm, excluding employee-contributed payroll taxes (social security contributions). (Bin width = €120,202).

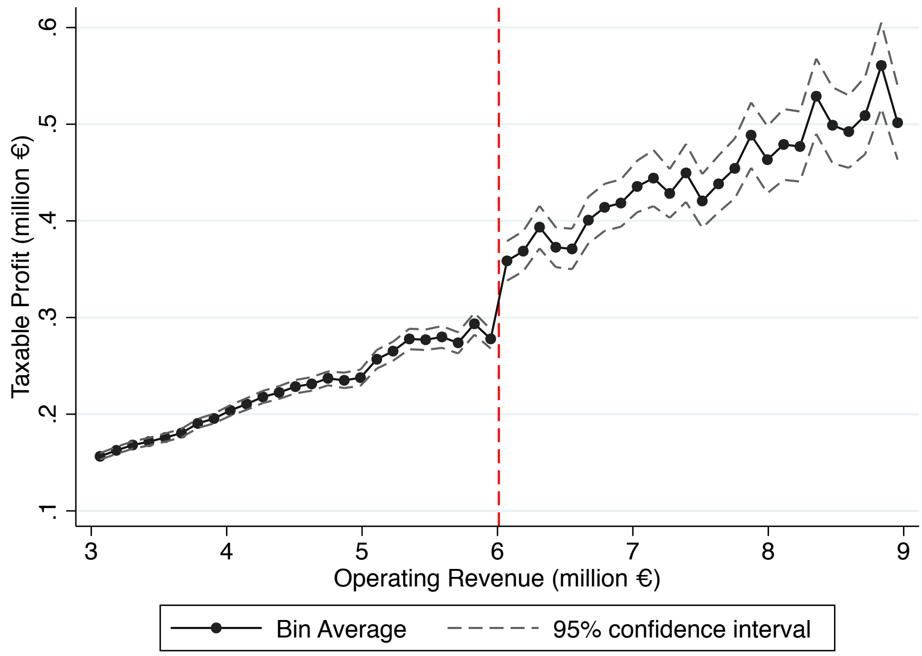
Figure 7: Taxable Profit Margin (% of Revenue)



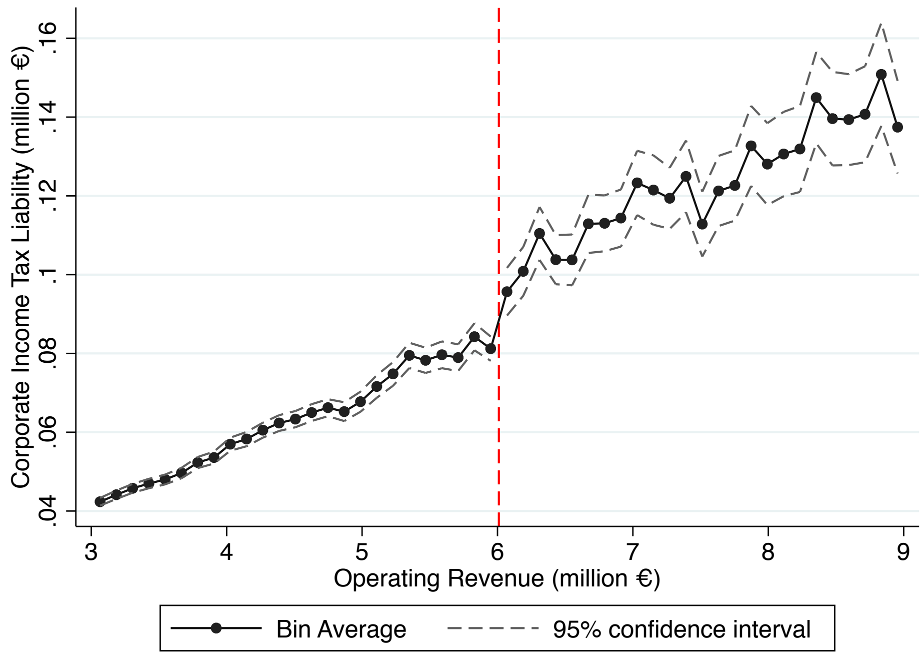
Notes: this graph shows the average taxable profit margin as a percentage of total revenue (vertical axis) against total reported revenue (horizontal axis), for the period 1995-2007. The dashed (red) vertical line indicates the LTU threshold. The black dotted lines denote bin averages and the grey dashed lines show 95% confidence intervals for each bin average. We trim outliers in the data by dropping the observations in the top and bottom 1% of the outcome variable in each €1-million interval in the range $\overline{y} \in (3, 9)$ million. Since taxable profits are not reported directly in the data, we back them out using the reported corporate income tax liability and applying the tax schedule. Note that this calculation can only be made for firms with a positive tax liability, which in this period were more than 80% of all firms in the sample. We then divide taxable profits by total reported revenue to obtain the "Taxable profit margin". (Bin width= €120,202).

Figure 8: Corporate Income: Tax Base and Tax Liability

(a) CIT Base



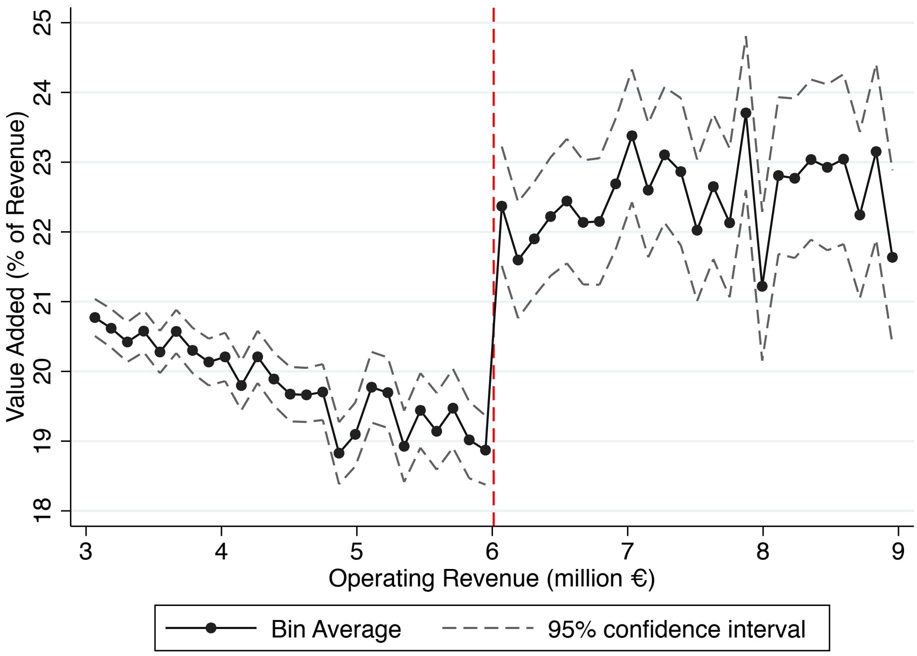
(b) CIT Liability



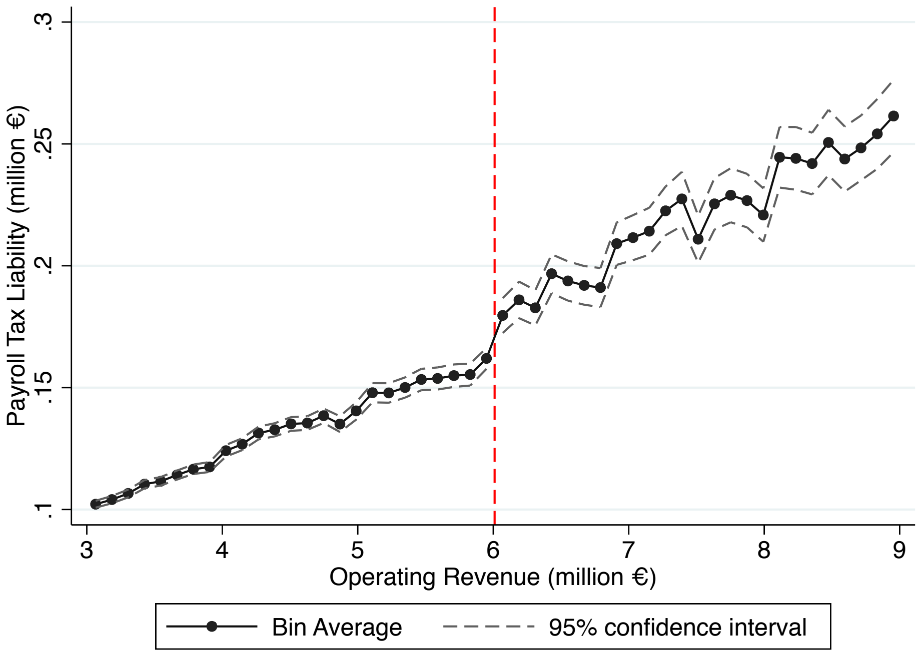
Notes: these graphs show the average CIT tax base (taxable profit) and the CIT tax liability (as reported by the firm) against reported revenue (horizontal axis), for the period 1995-2007. The dashed (red) vertical line indicates the LTU threshold. The black dotted lines denote bin averages and the grey dashed lines show 95% confidence intervals for each bin average. We trim outliers in the data by dropping the observations in the top and bottom 1% of the outcome variable in each €1-million interval in the range $\overline{y} \in (3, 9)$ million. "CIT tax base" is estimated as explained in the note to figure 7. "CIT tax liability" is reported by firms directly on their annual statement. (Bin width= €120,202).

Figure 9: VAT and Payroll: Tax Base and Tax Liability

(a) VAT Base as % of Total Revenue



(b) Payroll Tax Liability



Notes: these graphs show the average VAT tax base (i.e., value added) and the payroll tax liability (as reported by the employer) against reported revenue (horizontal axis), for the period 1995-2007. The dashed (red) vertical line indicates the LTU threshold. The dotted lines denote bin averages and the grey dashed lines show 95% confidence intervals for each bin average. We trim outliers in the data by dropping the observations in the top and bottom 1% of the outcome variable in each €1-million interval in the range $\overline{y} \in (3, 9)$ million. "Value added" is a proxy for the VAT tax liability calculated by the Statistics Divison of *Banco de España* taking into account that some goods could be zero-rated or VAT-exempted. "Payroll tax liability" is reported by firms in their annual statement, and includes only the employee-contributed part of the payroll tax (social security contributions). (Bin width= €120,202).

Tables

Table 1: Summary Statistics

	Mean	SD	Median	Min.	Max.	Obs.
Operating Revenue (million €)	4.669	1.447	4.253	3.005	9.015	285,580
Material Expenditures (million €)	3.630	1.449	3.347	0.000	28.698	279,878
Net Wage Bill (million €)	0.520	0.530	0.369	0.000	11.017	260,884
Taxable Profits (million €)	0.245	0.356	0.116	0.002	5.295	237,180
CIT Liability (million €)	0.068	0.116	0.027	-0.644	1.826	279,879
Value Added (million €)	0.959	1.037	0.692	-6.325	33.579	280,371
Tangible Fixed Assets (million €)	1.041	1.979	0.455	0.000	138.412	282,477
Number of Employees (FTE)	27.8	28.1	20	0	429	247,884
Material Expenditures (% of Revenue)	77.7%	17.8%	82.0%	0.0%	358.7%	279,878
Net Wage Bill (% of Revenue)	11.2%	10.4%	8.3%	0.0%	122.4%	260,885
Taxable Profit Margin (% of Revenue)	5.17%	6.8%	2.63%	0.0%	86.6%	237,184
Value Added (% of Revenue)	20.4%	19.2%	15.9%	-70.2%	419.7%	280,374

Notes: this table shows summary statistics for firms in the final dataset used for analysis, which is restricted to firms with reported revenue $y \in$(€3.01,€9.01) million. The top and bottom 1% of the variables "Materials as % of revenue", "Labor as % of revenue", "Fixed assets as % of revenue" and "Average gross wage" were dropped from the initial dataset to prevent outliers (and potentially incorrect data entries) from biasing the empirical estimations. The number of observations is different for each variable due to missing values, an issue especially relevant for the number of employees variable, which is not reported by about 20% of the firms.
Source: annual data from the Banco de España CBB dataset for Spanish firms in the period 1995-2007, built using administrative data from *Registro Mercantil*. More details about the dataset are given in section D in the online appendix.

Table 2: Bunching Estimations, by Year

	Bunching Estimators		# of Firms		Excl. Interval		Obs.
	b_{av}	b_{adj}	B	H	y_{lb}	y_{ub}	N
Pooled data							
1995-2007	0.094	0.276	3191	2965	5.70	6.53	285,570
	(0.005)***	(0.023)***					
Annual data							
1995	0.069	0.315	82	77	5.70	6.56	10,207
	(0.015)***	(0.149)**					
1996	0.080	0.243	98	91	5.70	6.38	11,014
	(0.012)***	(0.057)***					
1997	0.131	0.413	186	183	5.70	6.59	12,516
	(0.018)***	(0.103)***					
1998	0.124	0.252	198	185	5.70	6.53	14,310
	(0.014)***	(0.039)***					
1999	0.128	0.241	218	203	5.70	6.68	15,835
	(0.016)***	(0.041)***					
2000	0.138	0.429	273	251	5.70	6.80	17,923
	(0.015)***	(0.087)***					
2001	0.098	0.236	229	214	5.70	6.59	21,030
	(0.011)***	(0.041)***					
2002	0.062	0.344	168	151	5.70	6.38	23,977
	(0.008)***	(0.087)***					
2003	0.082	0.242	249	254	5.70	6.41	26,759
	(0.010)***	(0.043)***					
2004	0.109	0.368	382	356	5.70	6.86	29,779
	(0.013)***	(0.080)***					
2005	0.090	0.218	382	376	5.70	6.47	33,342
	(0.008)***	(0.027)***					
2006	0.080	0.247	373	342	5.70	6.35	35,788
	(0.007)***	(0.034)***					
2007	0.080	0.244	336	310	5.70	6.50	33,090
	(0.009)***	(0.040)***					

Notes: b_{av} is the average bunching response and b_{adj} is the marginal buncher's response taking into account resource costs of evasion (both measured in million euros). Bootstrapped standard errors are shown below each estimate in parenthesis. B is the number of firms above the counterfactual density of revenue in the range $y \in \left(y_{lb}, y^L\right)$, where y is revenue, y_{lb} is the lower bound of the excluded region (used to construct the counterfactual) and y^L is the LTU threshold of €6 million. H is the missing number of firms below the counterfactual density in the range $y \in \left(y^L, y_{ub}\right)$, where y_{ub} is the upper bound of the excluded region. The upper and lower bounds of the excluded interval, (y_{lb}, y_{ub}) are also reported. For all years, the counterfactual density is estimated using a 5th-degree polynomial. Finally, N is the number of observations included in the estimations, i.e. the number of firms with revenue $y \in$(€3.01,€9.01) million in each year. Significance levels: *** = 1%, ** = 5%, and * = 10%.

Table 3: Bunching Estimations, by Sector of Activity

	Bunching Estimators		# of Firms		Excl. Interval		Obs.
	b_{av}	b_{adj}	B	H	y_{lb}	y_{ub}	N
Primary Sector	0.087	0.962	40	37	5.70	6.53	4,120
	(0.020)***	(28.453)					
Manuf. Food and Beverages	0.085	0.254	105	108	5.70	6.89	10,100
	(0.017)***	(0.083)***					
Manuf. Non Metals	0.081	0.203	97	89	5.70	6.29	9,415
	(0.011)***	(0.037)***					
Manuf. Metals	0.119	0.337	115	103	5.70	6.47	8,262
	(0.017)***	(0.077)***					
Manuf. Equipment	0.104	0.215	73	67	5.70	6.68	6,137
	(0.020)***	(0.060)***					
Manuf. Others	0.068	0.256	142	128	5.70	6.53	18,818
	(0.010)***	(0.060)***					
Construction of Buildings	0.076	0.282	272	251	5.70	6.62	29,070
	(0.008)***	(0.052)***					
Specialized Constr. Activ.	0.148	0.301	180	170	5.70	6.74	11,838
	(0.018)***	(0.051)***					
Motor Vehicles	0.067	0.323	118	107	5.70	6.41	13,118
	(0.008)***	(0.075)***					
Wholesale (exc. Motor V.)	0.109	0.260	986	939	5.70	6.47	72,865
	(0.006)***	(0.020)***					
Transportation	0.119	0.308	177	161	5.70	6.80	13,025
	(0.014)***	(0.057)***					
Retail Trade	0.064	0.189	155	146	5.70	6.29	22,053
	(0.007)***	(0.030)***					
Restaurants and Hotels	0.074	0.436	64	59	5.70	6.74	7,868
	(0.023)***	(56.070)					
Cultural Activities	0.067	0.190	38	37	5.70	6.38	5,397
	(0.019)***	(0.083)**					
Real Estate	0.069	4.515	38	36	5.70	6.38	4,829
	(0.016)***	(19.349)					
Other Services	0.074	0.223	146	147	5.70	6.53	16,799
	(0.009)***	(0.039)***					

Notes: b_{av} is the average bunching response and b_{adj} is the marginal buncher's response taking into account resource costs of evasion (both measured in million euros). Bootstrapped standard errors are shown below each estimate in parenthesis. B is the number of firms above the counterfactual density of revenue in the range $y \in \left(y_{lb}, y^{LTU}\right)$, where y is revenue, y_{lb} is the lower bound of the excluded region (used to construct the counterfactual) and y^L is the LTU threshold of €6 million. H is the missing number of firms below the counterfactual density in the range $y \in \left(y^L, y_{ub}\right)$, where y_{ub} is the upper bound of the excluded region. The upper and lower bounds of the excluded interval, (y_{lb}, y_{ub}) are also reported. For all years, the counterfactual density is estimated using a 5th-degree polynomial. Finally, N is the number of observations included in the estimations, i.e. the number of firms with revenue $y \in$(€3.01,€9.01) million in each year. Significance levels: *** = 1%, ** = 5%, and * = 10%.

Table 4: Determinants of Average Bunching Response

(a) Ordinary Least Squares Regressions

	Dependent variable: Average Bunching (b_{av})			
	(1)	(2)	(3)	(4)
Share of Final Consumer Sales	-0.052	-0.052	-0.049	-0.046
	(0.018)***	(0.016)***	(0.017)**	(0.015)***
Median Number of Employees		0.000		0.000
		(0.000)		(0.000)
Average Tangible Assets/Revenue			-0.040	-0.054
			(0.049)	(0.050)
Constant	0.103	0.095	0.110	0.101
	(0.009)***	(0.012)***	(0.016)***	(0.016)***
Observations	16	16	16	16
R-squared	0.318	0.334	0.340	0.373

Notes: All regressions are estimated by OLS. The unit of observation is the sector of activity (for details on the definition of the 16 sectors, see the online appendix). Robust standard errors reported in parentheses. Significance levels: *** = 1%, ** = 5%, and * = 10%.

(b) Weighted Least Squares Regressions

	Dependent variable: Average Bunching (b_{av})			
	(1)	(2)	(3)	(4)
Share of Final Consumer Sales	-0.037	-0.039	-0.038	-0.039
	(0.010)***	(0.011)***	(0.010)**	(0.011)***
Median Number of Employees		-0.000		-0.000
		(0.000)		(0.000)
Average Tangible Assets/Revenue			-0.061	-0.054
			(0.036)	(0.040)
Constant	0.097	0.104	0.106	0.108
	(0.004)***	(0.008)***	(0.007)***	(0.008)***
Observations	16	16	16	16
R-squared	0.318	0.334	0.340	0.373

Notes: All regressions are estimated by Weighted Least Squares. The weights are the inverse of the variance of the bunching estimates. The unit of observation is the sector of activity (for details on the definition of the 16 sectors, see the online appendix). Standard errors reported in parentheses. Significance levels: *** = 1%, ** = 5%, and * = 10%.

Online Appendix
Not Intended for Publication

"Under the Radar: The Effects of Monitoring Firms on Tax Compliance"

Miguel Almunia (University of Warwick)
David Lopez-Rodriguez (Banco de España)

A Corporate Evasion with Multiple Taxes

Consider a firm that produces good y combining material acquisitions m and labor n, which are tax-deductible expenditures[53] in the corporate income tax, and non-deductible inputs z according to the production function $y = \psi f(m, n, z)$, where ψ is a productivity parameter and $f(\cdot, \cdot, \cdot)$ is strictly continuous, increasing and concave in inputs use. Firm hires in competitive markets materials at unit cost c, labor at wage rate w, and non-deductible expenditures at unit cost q, and sell their output at the market price p, which is normalized to unity.

Suppose the existence of value added tax (VAT) by the credit method in which firms charge a flat tax rate t^{vat} on their sales y and receive a credit for the monetary value of their material expenditures $e \equiv cm$. Firms must transfer to the tax authority the difference between charged and deductible VAT, that is $t^{vat} \cdot P^{vat}$ with $P^{vat} = y - e$. Government also levies linear payroll taxes on the wage bill $P^{ss} = l \equiv wn$, charging t^{ss_1} on account of employers, that are tax deductible in the corporate income tax, and t^{ss_2} on account of employers. We assume that both payroll taxes are fully born by firms. Firms also consider the regulatory costs associated to hiring labor captured by a convex cost function in the reported wage bill, $\gamma(l)$.[54] Finally, the income generated by the firm is taxed with a proportional rate t^{cit} on taxable profits $P^{cit} = y - e - l \cdot (1 + t^{ss_1})$, so firm's net-of-tax income with truthful reporting is given by $\Pi = (1 - t^{cit})P^{cit} - P^{ss}t^{ss_2} - qz - \gamma(l)$.

Suppose that the tax authority is not able to monitor all transactions in the economy creating incentives for firms to evade taxes by misreporting their tax bases. Consider that an evader firm could underreport the monetary value of their revenue by an amount $u^y \equiv y - \overline{y} \geq 0$, where $\overline{y}$ denotes reported revenue, to reduce taxable corporate income and to appropriate tax revenue from the VAT. Firm may also attempt to inflate the value of their material acquisitions, given by $u^e \equiv \overline{e} - e \geq 0$, where $\overline{e}$ denotes reported expenditures, to claim larger tax credits in both corporate income tax and the VAT. Firms may have incentives to hide a share of their wage bill by an amount $u^l \equiv l - \overline{l} \geq 0$, where $\overline{l}$ denotes reported labor expenditures, to evade payroll taxes and save regulatory costs of hiring labor. Given these potential evasion channels, firm's reported tax bases in the corporate income tax, payroll taxes and the VAT are given, respectively, by

$$\overline{P^{cit}} = [(y - u^y) - (e + u^e) - (l - u^l) \cdot (1 + t^{ss_1})], \tag{18}$$

[53]We make the distinction between this two tax-deductible inputs because the dataset in our empirical application includes accurate measures of firms' total expenditures on material acquisitions and labor wage bill.

[54]The assumptions on the incidence of payroll taxes on account of employers and employees, and the existence of regulatory costs associated to hiring workers seems particularly appropriate for the Spanish case. As an example, Bentolila, Dolado, and Jimeno (2012) discuss the costs and rigidities imposed on Spanish firms by multiple regulations in labor markets.

$$\overline{P^{ss}} = (l - u^l), \tag{19}$$

$$\text{and} \quad \overline{P^{vat}} = [(y - u^y) - (e + u^e)]. \tag{20}$$

Evasion behavior is costly because it requires, for instance, collusion between the firm and its trading partners and employees; the creation of parallel accounting books and payment systems in cash; or it can imply forego business opportunities. We introduce these resource costs of evasion by a reduced form $\kappa(u^y, u^e, u^l)$ that is an increasing, convex and separable function in each of its arguments. The tax authority detects evasion with probability $\delta = \phi h(u^y, u^e, u^l)$, where ϕ is the monitoring effort parameter and the enforcement technology $h(\cdot)$ is a continuos, convex and separable function in each evasion channel. Whenever misreporting is detected, the firm is compelled to pay back the evaded tax plus a proportional penalty rate θ that, for simplicity, is assumed homogeneous for all channels of evasion.

The expected profit of the firm net of corporate and payroll taxes, and augmented by the expected appropriation of VAT revenue, is given by

$$\begin{aligned} E\Pi &= (1 - t^{cit})[\psi f(m, n, z) - e - l(1 + t^{ss_1})] - qz + tr \cdot [u^y + u^e - u^l(1 + t^{ss_1})] \\ &\quad + (t^{ss_1} + t^{ss_2}) \cdot ru^l + t^{vat} \cdot r[u^y + u^e] - \kappa(u^y, u^e, u^l) - \gamma(\bar{l}), \end{aligned} \tag{21}$$

where $r \equiv [1 - \phi h(u^y, u^e, u^l)(1 + \theta)]$ is the expected rate of return of 1 euro evaded. Firms make production and reporting decisions in order to maximize their expected profit such that an interior optimum for firms real and evasion decisions satisfies the system of first-order conditions given by

$$\psi f_m(m^*, n^*, z^*) = c \tag{22}$$

$$\psi f_n(m^*, n^*, z^*) = w \left[1 + t^{ss_1} + \frac{t^{ss_2} + \gamma_{\bar{l}}(\bar{l}^*)}{(1 - t^{cit})} \right] \tag{23}$$

$$\psi f_z(m^*, n^*, z^*) = q/(1 - t) \tag{24}$$

$$\left[t^{cit} + t^{vat}\right] \cdot r = \kappa_{u^y}(u^{y*}) + (1 + \theta)\phi h_{u^y}(u^{y*}) \cdot \widehat{T} \tag{25}$$

$$\left[t^{cit} + t^{vat}\right] \cdot r = \kappa_{u^e}(u^{e*}) + (1 + \theta)\phi h_{u^e}(u^{e*}) \cdot \widehat{T} \tag{26}$$

$$\left[(t^{ss_1} + t^{ss_2}) - t^{cit}(1 + t^{ss_1})\right] \cdot r + \gamma_{\bar{l}}(\bar{l}^*) = \kappa_{u^l}(u^{l*}) + (1 + \theta)\phi h_{u^l}(u^{l*}) \cdot \widehat{T} \tag{27}$$

where $\widehat{T} \equiv [t^{vat} \cdot (u^{y*} + u^{e*}) + (t^{ss_1} + t^{ss_2}) \cdot u^{l*} + t^{cit} \cdot (u^{y*} + u^{e*} - u^{l*}(1 + t^{ss_1}))]$ is the total evaded taxes by the multiple misreporting channels. The system of optimal conditions shows that positive tax rates on corporate income distort inputs demand decisions reducing revenue from potential production at zero tax rates. These conditions also indicate that the existence of both payroll taxes and labor regulatory costs create distortions increasing the marginal cost of hiring employees and thus reducing labor demand.

The optimal evasion conditions for each misreporting channel predict that firm evades taxes to the point where the marginal expected return of misreporting transactions is equal to the expected costs associated to tax evasion. The latter is the result of the marginal resource costs born in each misreporting channel plus the deterrence effect created by tax enforcement that results from the interaction between monitoring effort and the enforcement technology. The systematic matching of tax returns from multiple taxpayers implies that a marginal unit of misreporting in one channel increases the chances of being detected, and thus paying back the total amount evaded, in multiple channels.

The expected returns of misreporting revenue and expenditures are positively related with the tax rates. The larger the tax rates on both the VAT and the corporate income tax are, the higher the incentives to hide revenue and inflate material acquisitions to reduce those tax bases. Notice that when firms have scope to misreport their transactions they do not act as fiscal intermediaries, that just transfer collected VAT to the tax agency, but instead firms have incentives to appropriate a share of VAT revenue. Finally, the optimal condition for hidden labor bill indicates that firms could have incentives to misreport it when the marginal savings in payroll taxes and regulatory costs were larger than the foregoing tax credits in corporate income tax due misreporting of labor costs. Overall, the model identifies two channels that create positive returns for labor misreporting: i) the existence of a significant gap between payroll taxes and corporate tax rates; and ii) the presence of large regulatory costs associated to hiring workers.

B Anatomy of the LTU Response: Input Ratios and Tax Bases

Consider the model with heterogeneous monitoring effort and resource costs presented in subsection 2.2. Before the introduction of a LTU, the system of optimal conditions indicates that the demand of tax-deductible inputs (e.g. materials and labor) is smoothly increasing in productivity, $dm/d\psi > 0$ and $dn/d\psi > 0$. Hence, the reported ratios of input expenditures over revenue, $\overline{M} \equiv cm/\overline{y}$ and $\overline{L} \equiv wn/\overline{y}$, are continuous in ψ over the range $[\underline{\psi}, \overline{\psi}]$. This implies that in the neighborhood of y^L defined by the small interval $(y', y^L + d\overline{y}_\phi^M)$ the average reported ratios of inputs expenditures over revenue are *almost* *equal*, that is,

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_0(\overline{y}) d\overline{y}} \cong \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{M} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_0(\overline{y}) d\overline{y}} \quad \text{and} \quad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_0(\overline{y}) d\overline{y}} \cong \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{L} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_0(\overline{y}) d\overline{y}}. \tag{28}$$

Suppose the LTU is introduced. In the presence of heterogeneity in both monitoring effort effectiveness and resource costs, there is a subset of firms in the pre-LTU density interval $(y^L, y^L + d\overline{y}_\phi^M)$, the *nonbuchers*, with prohibitive resource costs to respond reducing reported revenue. In contrast, the complementary subset of firms also located in that interval in the pre-LTU situation, the *bunchers*, reduce their reported revenue to stay below the threshold because that results in larger expected profits, i.e. $E\Pi_0(m, n, z, u \mid \psi^M, \phi_0) > E\Pi_1(m, n, z, u \mid \psi^M, \phi_1)$. Considering that **due to frictions** the bunchers locate **within/along** the interval(y', y^L), the model provides different predictions on the expected average reported ratios of input expenditures over revenue around the LTU threshold. These predictions depend on whether bunchers' reaction is due to real (i.e. reduction of production) or evasion (i.e. increase of concealed revenue) responses to the enforcement threshold.[55]

Real Response. Bunchers can avoid the threshold lowering their production, and thus their inputs demand, without bearing additional resource costs of evasion. This reaction implies that **within/in** the interval (y', y^L) below the threshold there are firms with $\psi \in [\psi', \psi^L]$ that hire more inputs than bunchers with $\psi \in [\psi^L, \psi^M]$. This causes that both average reported ratios of expenditures over revenue are not continuous at the threshold y^L. Indeed, the real reaction of the bunchers to the LTU results in i) a downward trend of both ratios in the interval (y', y^L); and ii) a discrete upward jump of these ratios at the threshold such that

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})\, d\overline{y}} < \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}} \quad \text{and} \quad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y}) d\overline{y}} < \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}}. \tag{29}$$

Evasion Response. Bunchers can avoid the threshold increasing their concealed revenue, and thus without modifying their inputs demand, paying resource costs of additional evasion. This response implies that in the interval (y', y^L) below the threshold firms with $\psi \in [\psi', \psi^L]$ hire lower inputs than bunchers with $\psi \in [\psi^L, \psi^M]$. The evasion response of the bunchers thus creates i) an upward trend of the average reported ratios

[55]The bunchers' reaction to the LTU threshold could be a combination of both potential responses, real and evasion, but we discuss the two polar responses for analytical simplicity. This simplification provides predictions on the expected average patterns of input ratios around the threshold when the reaction is dominated by either the real or the evasion channel.

of expenditures in the interval (y', y^L); and ii) a discontinuous downward jump of these ratios at the threshold such that

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y}) d\overline{y}} > \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}} \quad \text{and} \quad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y}) d\overline{y}} > \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}} . \tag{30}$$

Evasion Response with Inputs Misreporting. Considering the extended model with inputs misreporting and multiple taxes presented above (Appendix B), evader firms have incentives to i) inflate their material acquisitions in an amount u^e to claim larger tax credits in both the VAT and the CIT; and ii) hide part of their wage bill, u^l, to save both payroll taxes and the regulatory costs of labor. The optimal amount of evasion in each expenditure channel is heterogeneous among firms because it depends negatively on the effectiveness of monitoring effort and the resource cost of evasion. Firms thus have larger incentives to avoid stricter tax enforcement when their expenditures misreporting is higher in the pre-LTU situation (i.e. larger expected profits of bunching). Bunchers that react to avoid the LTU increasing their concealed revenue therefore also report a higher (lower) proportion of materials (labor) to evade taxes than firms with $\psi \in [\psi', \psi^L]$ also located in the interval (y', y^L). Define the ratios of reported inputs expenditures over renenue with inputs misreporting by $\overline{M} \equiv (cm + u^e)/\overline{y}$ and $\overline{L} \equiv (wn - u^l)/\overline{y}$, respectively. The evasion response of bunchers that also misreport expenditures in a larger proportion creates i) an upward (downward) trend of the average ratio of materials (labor) in the interval (y', y^L); and ii) a downward (upward) jump of the materials (labor) ratio at the threshold such that

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y}) d\overline{y}} >> \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}} \quad \text{and} \quad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y}) d\overline{y}} < \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}} . \tag{31}$$

LTU Effectiveness: Tax Bases. High productivity firms that are in the interior of the LTU have lower scope to evade taxes when the LTU is effective ($d\phi > 0$). The extended model predicts that these firms reduce concealed outcome, u^y, inflated materials, u^e, and hidden wage bill, u^l, in a magnitude that depends on the effectiveness of the LTU. An effective LTU thus raises the reported tax bases of the corporate income tax, $\overline{P^{cit}}$, the payroll tax, $\overline{P^{ss}}$, and the value-added tax, $\overline{P^{vat}}$, with a break at the LTU threshold creating i) a downard trend of the average tax bases in the interval (y', y^L) due to bunchers' misreporting; and ii) an upward parallel shift of tax bases for high productivity firms such that

$$\frac{\int_{y'}^{y^L} \overline{P^{cit}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} << \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{P^{cit}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \quad , \tag{32}$$

$$\frac{\int_{y'}^{y^L} \overline{P^{ss}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} << \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{P^{ss}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \quad , \tag{33}$$

$$\frac{\int_{y'}^{y^L} \overline{P^{vat}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} << \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{P^{vat}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \quad . \tag{34}$$

C Additional Institutional Background

Exceptions to LTU Eligibility Rule Exporting firms that claim a VAT refund are automatically included in the LTU census, regardless of their operating revenue. We do not have data on VAT claims related to exports that allows us to identify these firms accurately, so we cannot exclude these firms from the analysis nor can we use this set of firms as a comparison group.

Two regions in Spain, Navarra and País Vasco, have their own independent tax authorities. Firms with headquarters located in each of these regions are monitored by those independent tax authorities, unless they obtain more than 75% of their operating revenue from transactions in other Spanish regions, in which case they are monitored by the national LTU. Based on our conversations with tax auditors, most of the administrative structures of the independent tax authorities are very similar to the national one, but we could not confirm whether they also consider the €6 million threshold to classify “large” firms. The distribution of reported revenue features modest, but statistically significant, bunching in the two regions. We choose to exclude them from the main analysis because of the uncertainty about how many firms are subject to the LTU and also because they represent a small proportion(7.2%) of firms with revenue between €3 and €9 million.

Corporate Income Tax Threshold. The standard rate in the corporate income tax was 35% of taxable profits in the period 1995-2007. A lower rate of 30% was applied to firms under a revenue threshold that was modified over time: from €1.5 million in 1999 up to €10 million in 2010 (full details provided in Table A.1). The cutoff for this tax break overlapped with the LTU threshold in 2004, but was different in the rest of the years. The lower rate was applied only to the first €90,121 of taxable profits (€120,202 since 2005) creating a notch for eligible firms with low taxable profits, and a kink for those with high profits.

External Audit and Abbreviated Returns Threshold. Firms are required by law to have their annual accounts audited by an external private firm if they fulfill two of the following criteria for two consecutive years: (i) annual revenue above €4.75 million; (ii) total assets above €2.4 million;[56] and (iii) more than 50 employees on average during the year. These criteria also determine whether a firm can use the abbreviated form of the corporate income tax return, rather than the standard (long) version. These requirements create compliance costs,[57] and the private audit information could complement tax enforcement because auditors face legal responsibility if any misreporting is found.

[56] The revenue limit was originally 790 million pesetas (€4.748 million), and the assets limit was 395 million pesetas (€2.374 million).

[57] The yearly fee charged by private audit firms is in the range €10,000 - €30,000 for firms with revenue close to €4.75 million, a small but non-negligible expenditure (0.2 to 0.6% of total revenue, but 4 to 12% of reported *profits* on average).

D Data: Further Details

Definition of Sectors of Activity

The table below provides the sector definitions that we use in section 5.2 in terms of the 2009 version of the National Classification of Economic Activities (in Spanish, CNAE), which follows the Eurostat standard NACE Rev. 2. We use 2-digit CNAE codes to define sectors. The third column shows the number of firms in each sector for the 1995-2007 pooled CBB dataset, and the last column shows the percentage they represent overall.

Sector	**CNAE-2009 Codes**	**# of Firms**	
PrimarySector	01-09,19 - Agriculture, forestry, fishing, and mining	3,738	1.63%
Manuf_FoodBev	10,11,12 - Manufacture of food, beverages and tobacco	9,257	4.03%
Manuf_NonMetals	22,23 - Manufacture of plastics and non-metallic minerals	8,583	3,74%
Manuf_Metals	24,25 - Metal products, machinery	7,358	3.20%
Manuf_Equipment	26-28,33 - Manufacture of computers, electronics, equipment	5,506	2.40%
Manuf_Others	13-17,20,21,29-32 - Textiles, clothing, wood, paper, chemicals	17,367	7.56%
Const_Buildings	41 - Construction of buildings	25,888	11.27%
Const_SpecializedAct	43 - Specialized construction activities	10,327	4.50%
MotorVehicles	45 - Wholesale trade and repair of motor vehicles	12,134	5.28%
WholesaleTrade	46 - Wholesale trade (except motor vehicles)	66,406	28.92%
RetailTrade	47 - Retail trade	11,715	5.10%
RestHotels	55,56,79 - Hotels, restaurants and travel agencies	19,977	8.70%
Transportation	49-52 - Transportation by land, water, air, support activities	7,101	3.09%
CulturalActiv	18,58-60,90,93 - Publishing, movies, radio & TV, sports	4,934	2.15%
RealEstate	68,77- Real estate, rental and leasing	4,334	1.89%
OtherServices	53,61-64,69-75,78,80-82,85-88,92,95-96 - Other services	15,004	6.53%

Source: http://www.ine.es/daco/daco42/clasificaciones/cnae09/estructuraen.pdf

Original CBB data and final dataset

We start from the original CBB data as provided by the *Banco de España* in September 2014. We include data for the years 1995 through 2007, both included. In order to construct the final dataset for our analysis, we take several steps. First, we drop observations from two regions where tax collection is independent of the federal tax authority and hence the LTU threshold does not apply (Pais Vasco and Navarra). Second, we choose a bin width of €60,101, which is one-hundredth of the revenue level of the LTU threshold. For symmetry, we keep 50 bins below and 50 bins above the threshold, so in total there are 100 bins. Hence, our final dataset has firms with reported revenue between €3.005

million and €9.015 million. Within this range, we define some of the ratios that we use in the section on input expenditures: materials and labor expenditures as % of revenue, average gross wages (defined as the total wage bill divided by the number of employees), and fixed assets as % of revenue. Finally, we drop the top and bottom 1% of observations from each of these variables, in order to avoid the presence of outliers in the data. There is some overlap in the extreme values, such that a firm with abnormally high materials is likely to have abnormally low labor expenditures. The final dataset contains 285,570 observations, and summary statistics are reported on Table 1. The Stata do-files used to process the original data to arrive at the final dataset are available upon request.

Input-Output Tables

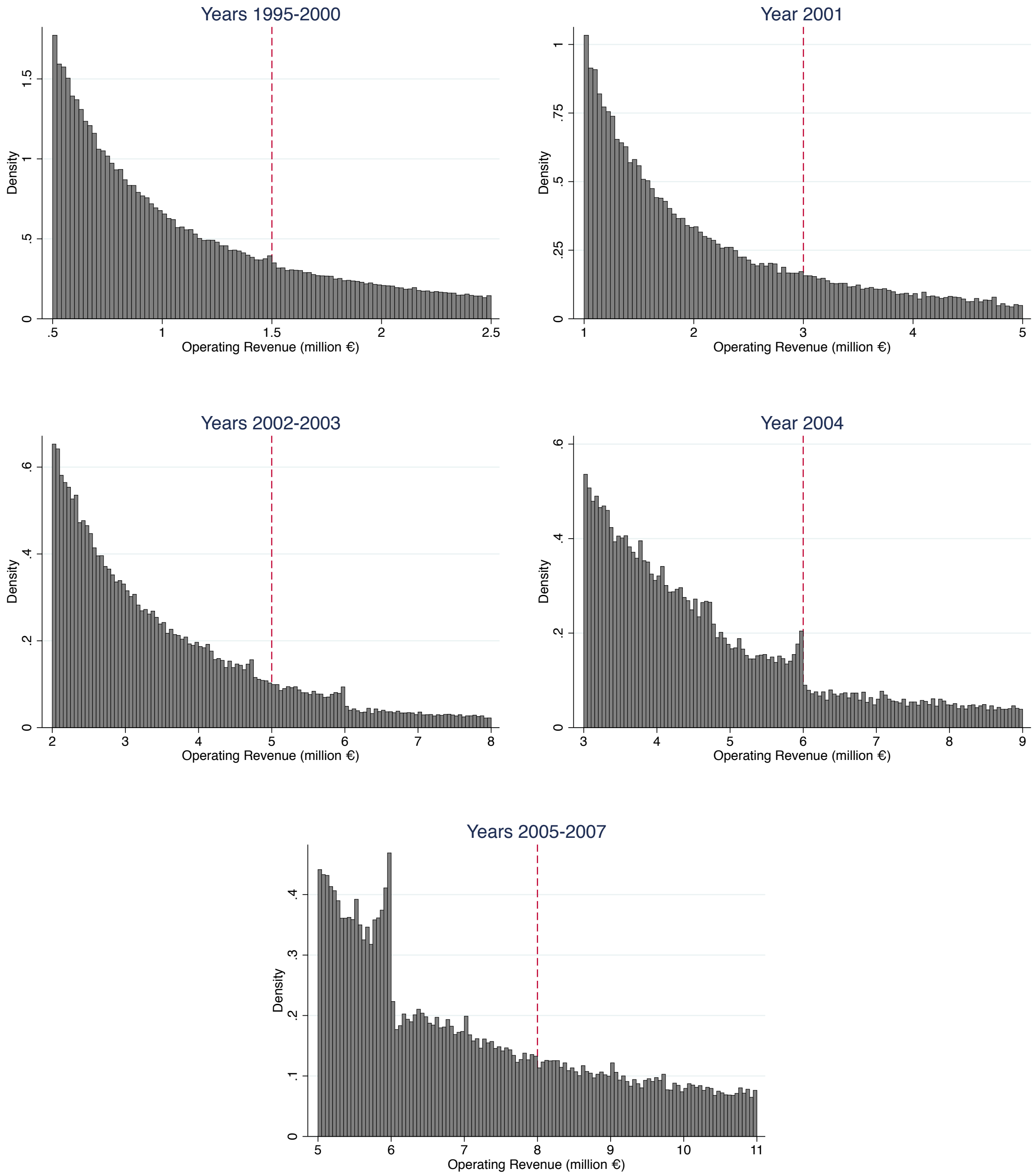
We use the input-output tables produced by the National Statistics Institute (*Instituto Nacional de Estadística, INE*) for the year 2000. Sectors of activity are defined according to Spanish industry classification (TSIO), which does not match CNAE 2009 codes exactly but has substantial overlap. To calculate the share of sales made to final consumers by sector, we divide the column labelled *"Consumo final de los hogares, interno"* ("Households' final consumption, domestic") by the column *"Total empleos"* ("Total uses"). The original table used for the calculations can be downloaded from:

www.ine.es/daco/daco42/cne00/simetrica2000.xls

The table we provide together with our main dataset contains, additionally, the correspondence between our sector definitions (based on CNAE 2009 codes) and the sectors defined in the input-output tables.

Appendix Figures

Figure A.1: Behavioral (Non)response at the Corporate Income Tax Threshold



Notes: these graphs show the operating revenue distribution for different periods, around the threshold for the corporate income tax cut for small firms. There is no bunching at this threshold in any year except for 2004, the year in which this cutoff overlapped with the LTU threshold discussed in the main text.

Figure A.2: Distribution of Reported Revenue, by Year

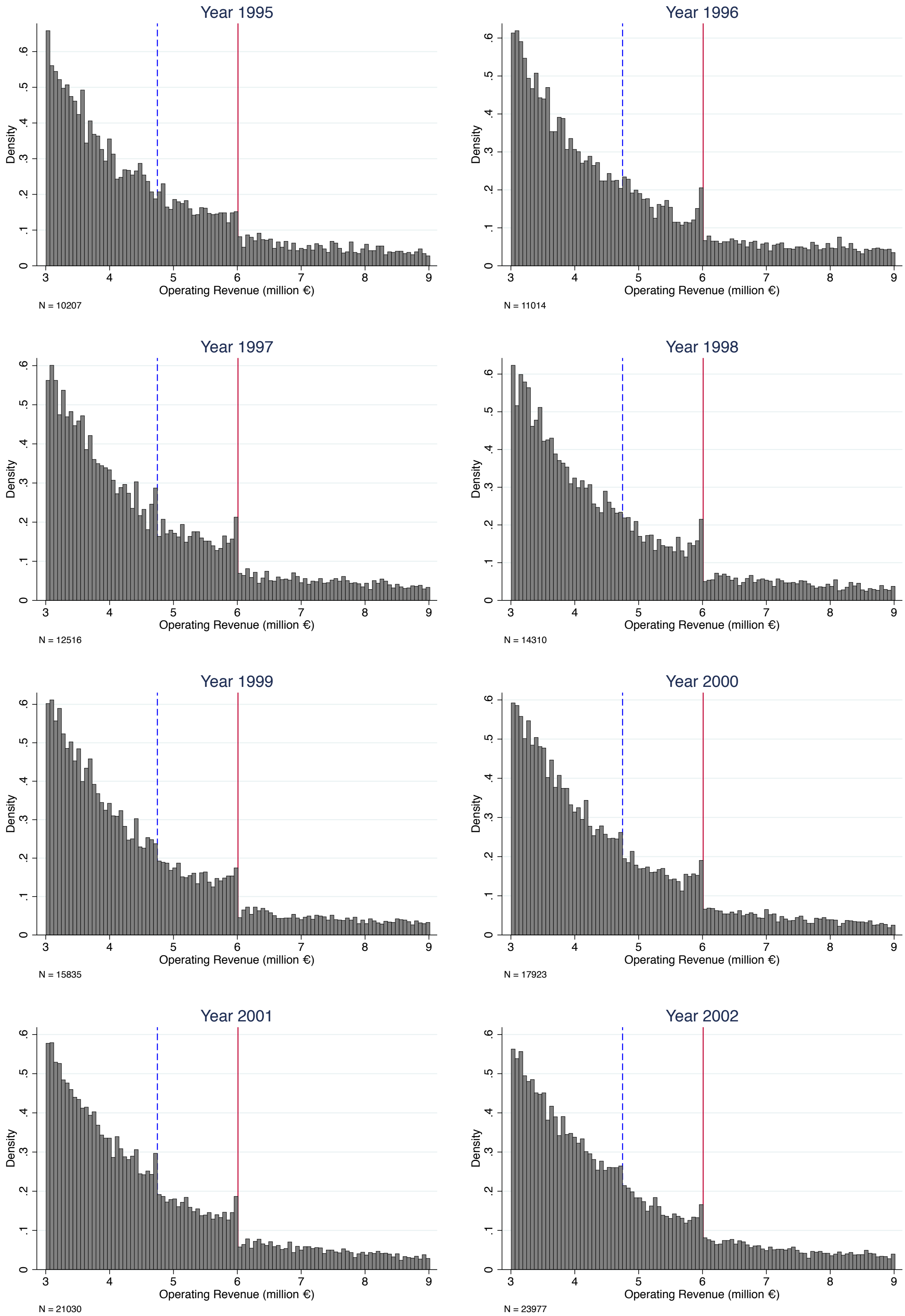
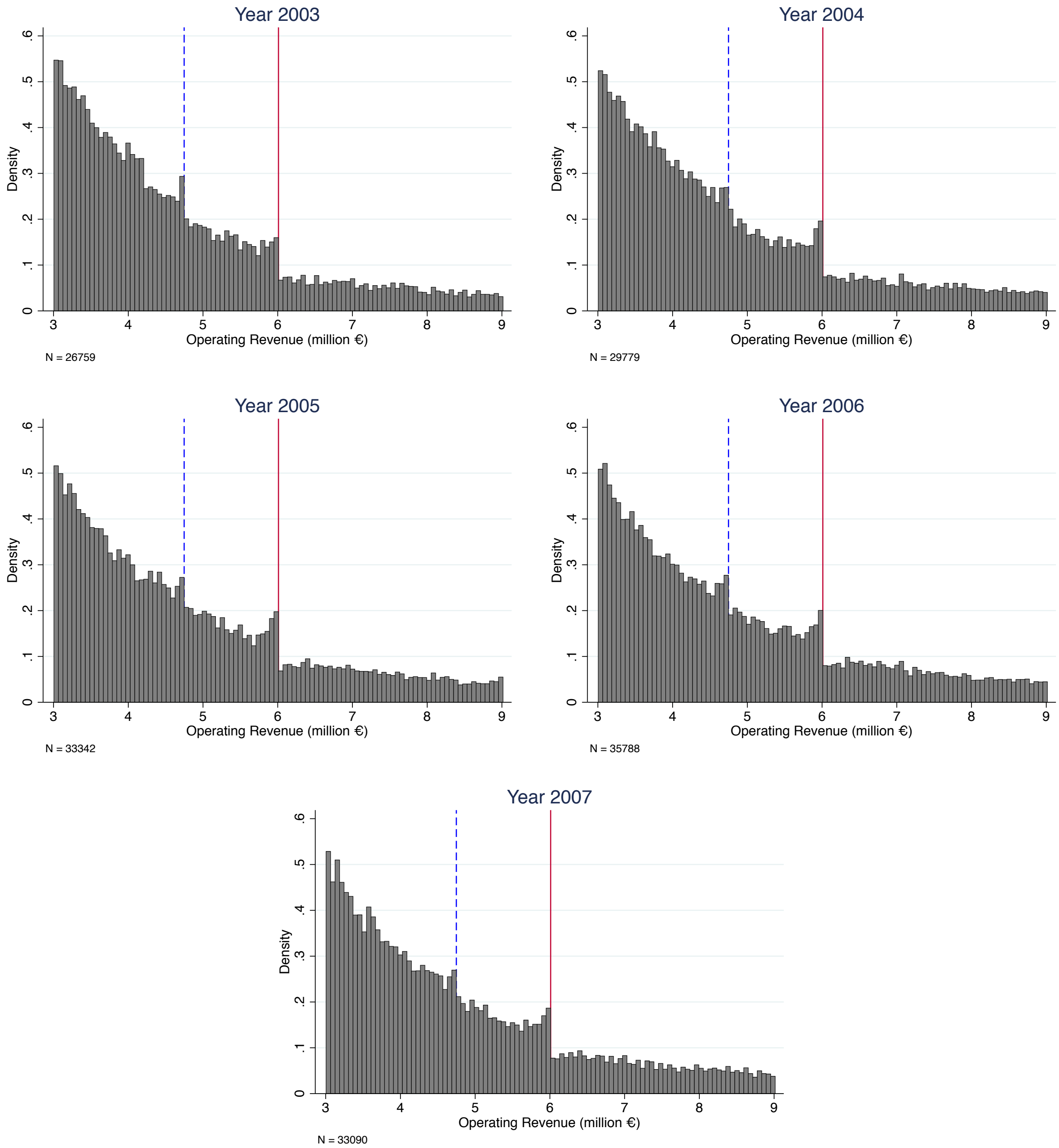


Figure A.2: Distribution of Reported Revenue, by Year (continued)



Notes: these graphs show the operating revenue distribution for each year in the period under study (1995-2007).

Figure A.3: Bunching Response by Sector

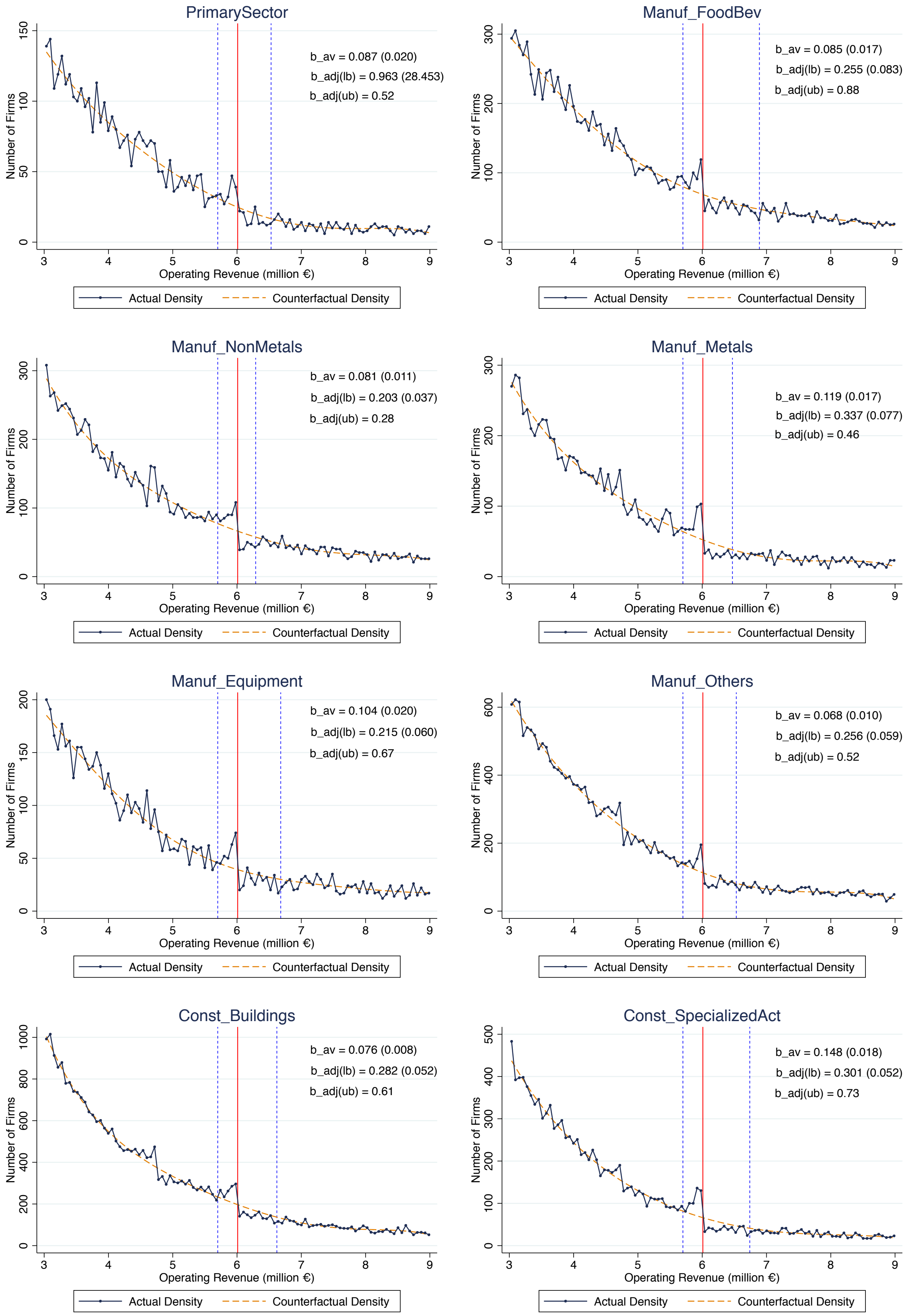
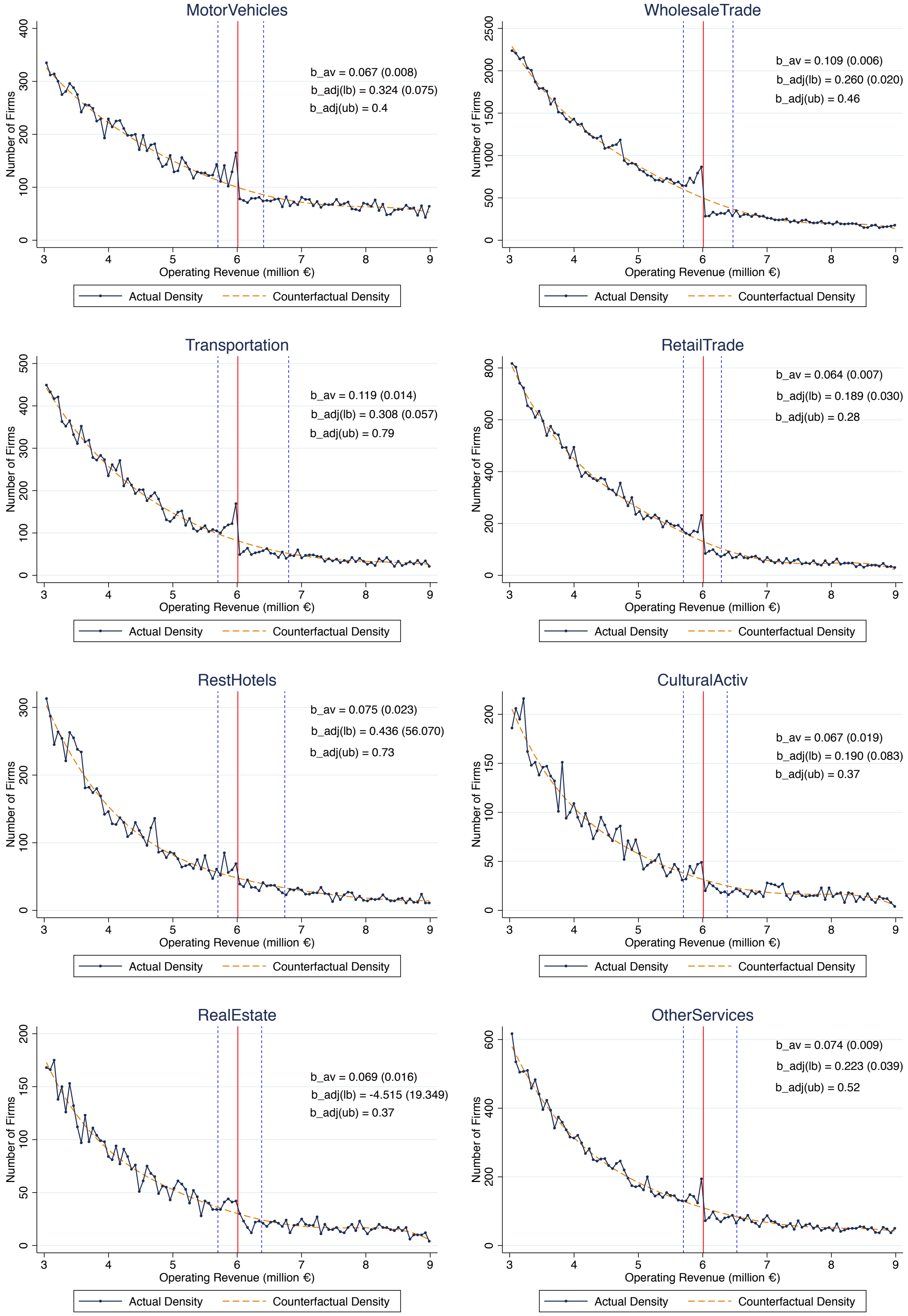
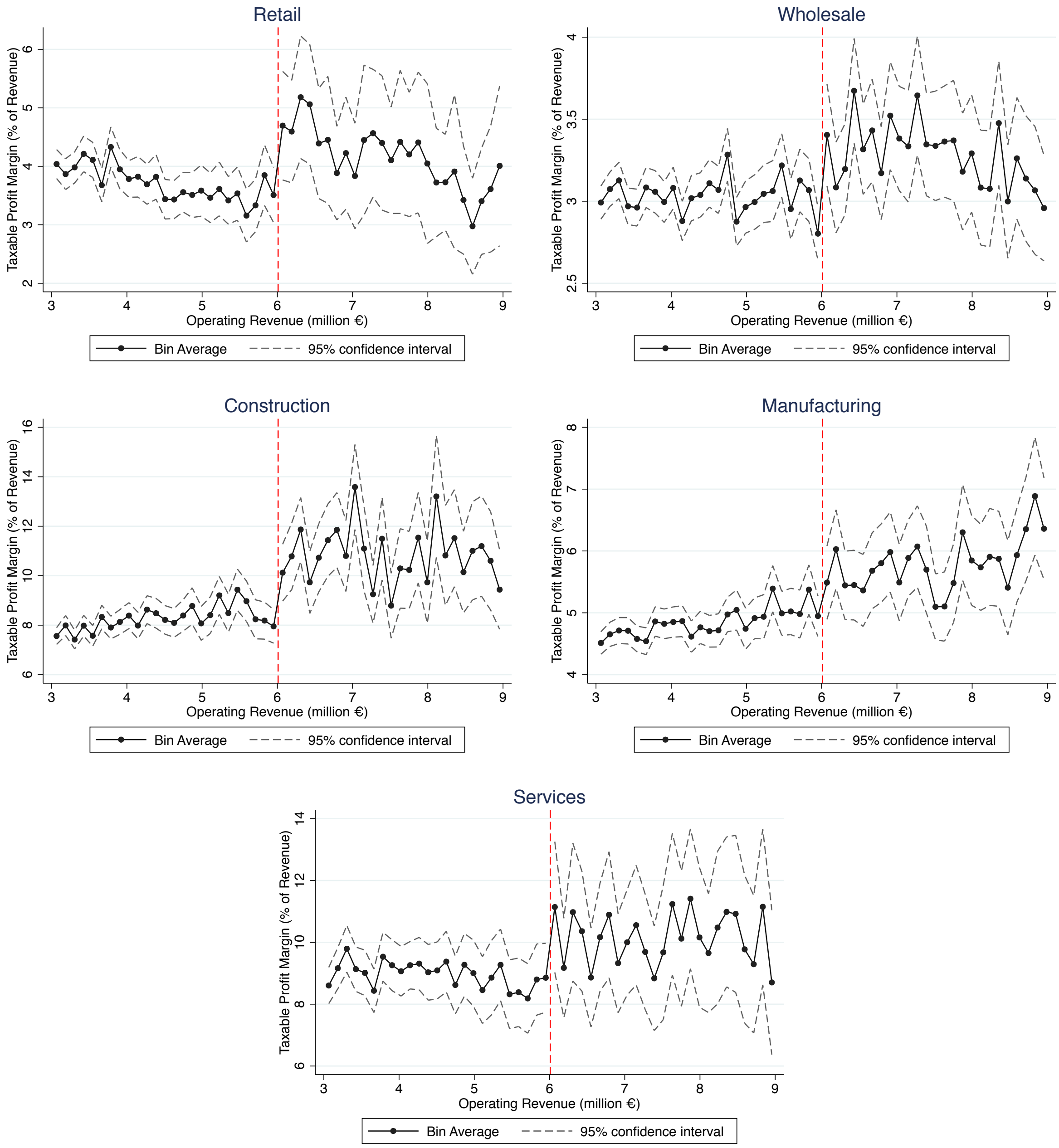


Figure A.3: Bunching Response by Sector (continued)



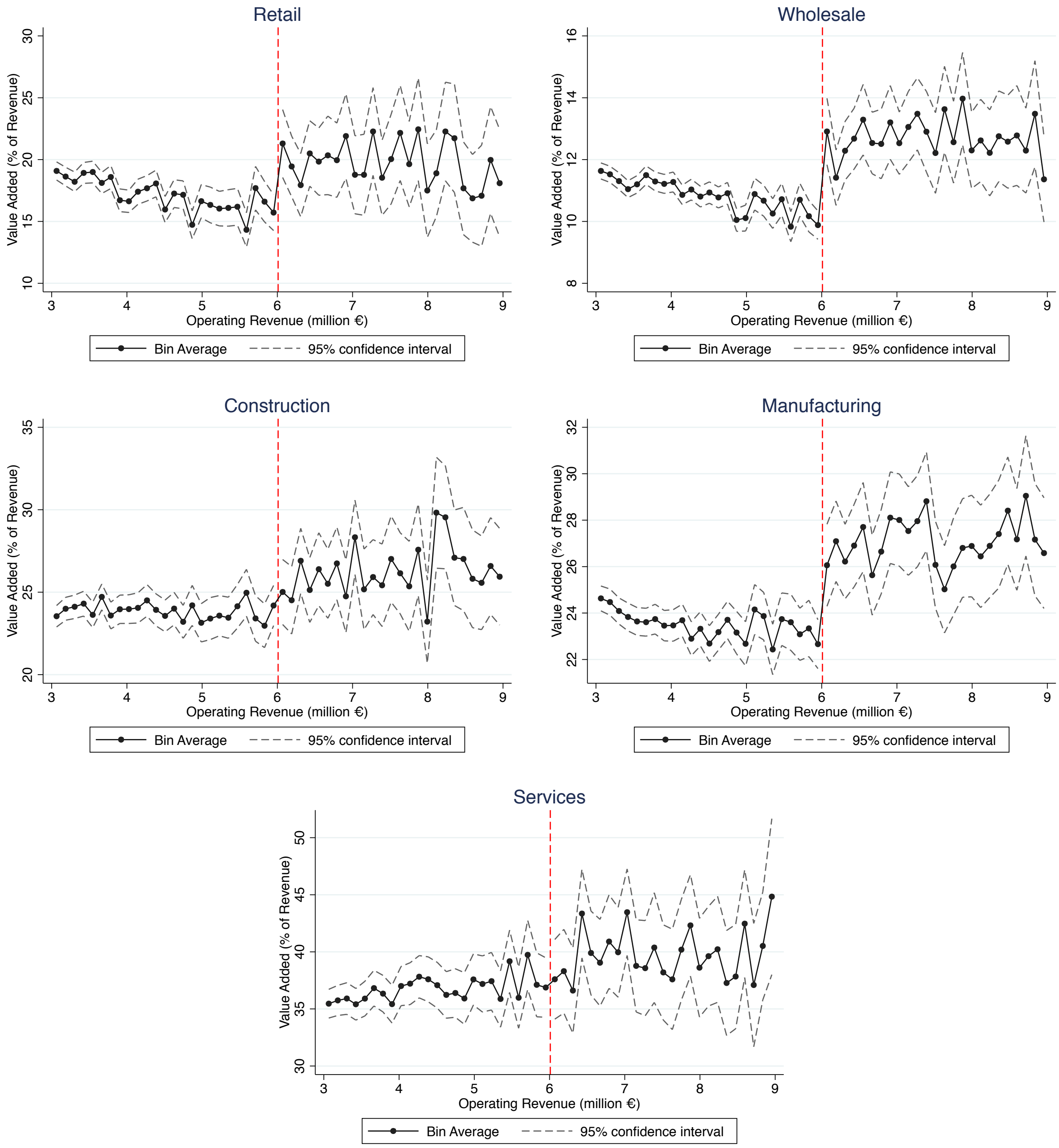
Notes: these graphs show the average observed and counterfactual operating revenue distribution for each sectors in the period under study (1995-2007).

Figure A.4: Reported Corporate Income Tax Base by Sector



Notes: these graphs show the average CIT tax base (taxable profit) as a percentage of total revenue by sector of activity. We distinguish 5 broad sectors of activity to ensure that there is statistical power to compare the behavior of firms below and above the LTU threshold, indicated by the the dashed (red) vertical line. The black dotted lines denote bin averages and the grey dashed lines show 95% confidence intervals for each bin average. We trim outliers in the data by dropping the observations in the top and bottom 1% of the outcome variable in each €1-million interval in the range $\overline{y} \in (3, 9)$ million. The "CIT tax base" is estimated as explained in the note to Figure 7. (Bin width= €120,202).

Figure A.5: Reported Value Added Tax Base by Sector



Notes: these graphs show the average value-added tax (VAT) base as a percentage of total revenue by sector of activity. We distinguish 5 broad sectors of activity to ensure that there is statistical power to compare the behavior of firms below and above the LTU threshold, indicated by the the dashed (red) vertical line. The black dotted lines denote bin averages and the grey dashed lines show 95% confidence intervals for each bin average. We trim outliers in the data by dropping the observations in the top and bottom 1% of the outcome variable in each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

Appendix Tables

Table A.1: Revenue Threshold: Corporate Income Tax Benefit for Small Firms

<table>
<tr><th>Year</th><th>Threshold</th><th>Standard tax rate</th><th>Special tax rate</th><th>Applicable range</th></tr>
<tr><td>1999</td><td rowspan="2">€1.5 million</td><td rowspan="8">35%</td><td rowspan="8">30%</td><td rowspan="6">Up to €90,151 of taxable profits</td></tr>
<tr><td>2000</td></tr>
<tr><td>2001</td><td>€3 million</td></tr>
<tr><td>2002</td><td rowspan="2">€5 million</td></tr>
<tr><td>2003</td></tr>
<tr><td>2004</td><td>€6 million</td></tr>
<tr><td>2005</td><td rowspan="3">€8 million</td><td rowspan="3">Up to €120,202 of taxable profits</td></tr>
<tr><td>2006</td></tr>
<tr><td>2007</td><td>32.5%</td><td>27.5%</td></tr>
</table>

Source: the applicable laws are: Law 43/1995 (Article 122), Law 6/2000 (Article 122), Law 24/2001 (Article 122), Law 4/2004 (Article 108), Law 2/2004 of the *Presupuestos Generales del Estado* (Annual Government Budget Law, Article 108).

Table A.2: Overview of the Spanish Tax System

	Top tax rate	*Share of tax revenue*
Social Security Contributions (PRT)	38%	33%
Personal Income Tax (PIT)	48% (46%)	22%
Value-Added Tax (VAT)	16%	19%
Corporate Income Tax (CIT)	35% (30%)	13%
Other indirect taxes and fees	-	13%
Federal Tax Revenue / GDP	30-37%	

Sources: Instituto de Estudios Fiscales (2011). The top marginal rate of the individual income tax was reduced to 46% 2005. The top marginal rate of the corporate income tax was reduced to 32.5% in 2006 and 30% in 2007. The data on tax revenues reflects averages for the period 1999-2007 and includes regional-level revenues in all calculations.

Table A.3: CBB Dataset Compared to Official CIT Statistics

		All Firms		€3-€10 million	
1995	Official Statistics	564,146		20,686	
	CBB Database	435,482	77.9%	12,592	60.9%
1996	Official Statistics	607,186		22,216	
	CBB Database	483,028	80.4%	13,924	62.7%
1997	Official Statistics	651,510		23,892	
	CBB Database	530,590	82.2%	16,216	67.9%
1998	Official Statistics	700,169		25,659	
	CBB Database	591,974	85.3%	18,453	71.9%
1999	Official Statistics	743,660		26,199	
	CBB Database	604,744	81.3%	20,083	76.7%
2000	Official Statistics	823,659		31,294	
	CBB Database	635,627	77.2%	22,468	71.8%
2001	Official Statistics	872,713		34,391	
	CBB Database	726,119	83.2%	25,561	74.6%
2002	Official Statistics	942,148		37,157	
	CBB Database	813,516	86.3%	29,003	78.1%
2003	Official Statistics	971,756		39,786	
	CBB Database	879,042	90.5%	32,191	80.9%
2004	Official Statistics	1,042,725		43,062	
	CBB Database	953,153	91.4%	35,846	83.2%
2005	Official Statistics	1,121,879		46,977	
	CBB Database	1,024,183	91.3%	40,422	86.0%
2006	Official Statistics	1,267,542		52,396	
	CBB Database	1,054,238	83.2%	43,325	82.7%
2007	Official Statistics	1,330,911		55,843	
	CBB Database	1,068,001	80.2%	39,728	71.1%

Notes: The percentages indicate the proportion of firms with a legal status of *Sociedad Anónima* (SA, equivalent to Corporation) or *Sociedad Limitad* (SL, equivalent to Limited Liability Company) in the CBB dataset compared to the number of firms with the same legal status that submitted a corporate income tax return that year. Official statistics have been compiled by the fiscal division of *Banco de España* based on several issues of *"Memoria de Administración Tributaria"*, an annual report published by the Spanish tax agency (AEAT, 1995-2007). The CBB dataset is described in detail in section 4.

Table A.4: Sensitivity Analysis, Pooled 1995-2007 data

Polynomial degree q	Excluded Interval y_{lb}	y_{ub}	Bunching Estimators b_{av}	b_{adj}	Obs.
4	5.30	6.68	0.106	0.306	285,570
			(0.015)***	(0.058)***	
5	5.30	6.68	0.094	0.277	285,570
			(0.013)***	(0.051)***	
4	5.40	6.68	0.108	0.313	285,570
			(0.012)***	(0.049)***	
5	5.40	6.71	0.101	0.297	285,570
			(0.011)***	(0.045)***	
4	5.50	6.59	0.106	0.308	285,570
			(0.009)***	(0.037)***	
5	5.50	6.62	0.099	0.289	285,570
			(0.008)***	(0.031)***	
4	5.60	6.53	0.102	0.296	285,570
			(0.008)***	(0.032)***	
5	5.60	6.59	0.096	0.279	285,570
			(0.007)***	(0.028)***	
4	5.70	6.47	0.098	0.285	285,570
			(0.006)***	(0.026)***	
5	**5.70**	**6.53**	**0.095**	**0.276**	285,570
			(0.005)***	**(0.023)*****	
4	5.80	6.38	0.090	0.257	285,570
			(0.004)***	(0.017)***	
5	5.80	6.41	0.089	0.256	285,570
			(0.004)***	(0.017)***	

Notes: this table shows the sensitivity of the bunching estimators to different assumptions on the excluded region used to estimate the counterfactual and the order of the polynomial. In all rows, we use the pooled 1995-2007 sample including all firms with reported revenue $y \in$(€3.01,€9.01). We pick different values of q, as shown in the first column, and y_{lb}, as shown in the second column. We obtain the corresponding values for y_{ub} and the point estimates for the bunching estimators b_{av} and b_{adj} using the methods described in the main text. The results are very similar for all the reasonable choices of the lower bound (y_{lb}), and for polynomials of degree 4 and 5. We highlight the results for $y_{lb} = 5.70$ and $q = 5$, which are the values chosen to produce the main estimation results. Significance levels: *** = 1%, ** = 5%, and * = 10%.