Under the Radar:

The Effects of Monitoring Firms on Tax Compliance

ONLINE APPENDIX

NOT INTENDED FOR JOURNAL PUBLICATION*

Miguel Almunia
University of Warwick

David Lopez-Rodriguez
Banco de España

March 2017

*The published version of the paper can be found at [URL], with DOI XXX.

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[1] 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, $\gamma(l)$.[2] 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

[1] 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.

[2] 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.

the corporate income tax, payroll taxes and the VAT are given, respectively, by

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

$$\overline{P^{ss}} = (l - u^l), \tag{2}$$

$$\text{and} \quad \overline{P^{vat}} = [(y - u^y) - (e + u^e)]. \tag{3}$$

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 continuous, 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

$$\begin{aligned} E\Pi \;=\; & (1 - t^{cit})[\psi f(m, n, z) - e - l(1 + t^{ss_1})] - qz + tr \cdot [u^y + u^e - u^l(1 + t^{ss_1})] \\ & + (t^{ss_1} + t^{ss_2}) \cdot ru^l + t^{vat} \cdot r[u^y + u^e] - \kappa(u^y, u^e, u^l) - \gamma(\bar{l}), \end{aligned} \tag{4}$$

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

$$\psi f_m(m^*, n^*, z^*) = c \tag{5}$$

$$\psi f_n(m^*, n^*, z^*) = w \left[1 + t^{ss_1} + \frac{t^{ss_2} + \gamma_{\bar{l}}(\bar{l}^*)}{(1 - t^{cit})} \right] \tag{6}$$

$$\psi f_z(m^*, n^*, z^*) = q/(1 - t) \tag{7}$$

$$\left[t^{cit} + t^{vat} \right] \cdot r = \kappa_{u^y}(u^{y*}) + (1 + \theta)\phi h_{u^y}(u^{y*}) \cdot \widehat{T} \tag{8}$$

$$\left[t^{cit} + t^{vat} \right] \cdot r = \kappa_{u^e}(u^{e*}) + (1 + \theta)\phi h_{u^e}(u^{e*}) \cdot \widehat{T} \tag{9}$$

$$\left[(t^{ss_1} + t^{ss_2}) - t^{cit}(1 + t^{ss_1}) \right] \cdot r + \gamma_{\bar{l}}(\bar{l}^*) = \kappa_{u^l}(u^{l*}) + (1 + \theta)\phi h_{u^l}(u^{l*}) \cdot \widehat{T} \tag{10}$$

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 Section 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,

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_0(\overline{y}) d\overline{y}} \cong \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{M} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_0(\overline{y}) d\overline{y}} \quad \text{and} \quad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_0(\overline{y}) d\overline{y}} \cong \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{L} \cdot g_0(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_0(\overline{y}) d\overline{y}}. \tag{11}$$

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 *non-optimizers*, 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 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.[3]

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 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

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})\, d\overline{y}} < \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{M} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}} \quad \text{and} \quad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y}) d\overline{y}} < \frac{\int_{y^L}^{y^L + d\overline{y}_\phi^M} \overline{L} \cdot g_1(\overline{y}) d\overline{y}}{\int_{y^L}^{y^L + d\overline{y}_\phi^M} g_1(\overline{y}) d\overline{y}}. \tag{12}$$

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

[3]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.

response of the bunchers thus creates i) an upward trend of the average reported ratios of expenditures in the interval (y', y^L); and ii) a discontinuous downward jump of these ratios at the threshold such that

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} > \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{M} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \qquad \text{and} \qquad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} > \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{L} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} . \tag{13}$$

Evasion Response with Inputs Misreporting. Considering the extended model with inputs misreporting and multiple taxes presented in Section A above, 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 revenue 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

$$\frac{\int_{y'}^{y^L} \overline{M} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} >> \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{M} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \qquad \text{and} \qquad \frac{\int_{y'}^{y^L} \overline{L} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} < \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{L} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} . \tag{14}$$

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 downward 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

$$\frac{\int_{y'}^{y^L} \overline{P^{cit}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} << \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{P^{cit}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \quad , \tag{15}$$

$$\frac{\int_{y'}^{y^L} \overline{P^{ss}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} << \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{P^{ss}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \quad , \tag{16}$$

$$\frac{\int_{y'}^{y^L} \overline{P^{vat}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y'}^{y^L} g_1(\overline{y})d\overline{y}} << \frac{\int_{y^L}^{y^L+d\overline{y}_\phi^M} \overline{P^{vat}} \cdot g_1(\overline{y})d\overline{y}}{\int_{y^L}^{y^L+d\overline{y}_\phi^M} g_1(\overline{y})d\overline{y}} \quad . \tag{17}$$

C Additional Institutional Background

LTU Eligibility Rule

The Spanish Tax Authority fixes a yearly revenue criteria to allocate firms under LTU monitoring. In particular, firms report to the tax agency their total revenue of year t in January of $t+1$ when their annual VAT summary must be submitted. Firms with revenue in year t that exceeds the €6.01 million threshold are monitored by the LTU since $t+1$, and the LTU is also in charge of reviewing all their tax returns from year t (e.g. CIT, VAT, wages withholding). Similarly, if revenue falls below the threshold in year s, the firm is removed from the LTU census in $s+1$ and tax returns from year s are not monitored by the LTU.

Groups

According to Spanish tax and business laws, there are two types of business groups: corporate groups and consolidated fiscal groups. The latter are more narrowly defined than the former, which require sharing the same activity and that the dominant firm owns at least 75% of the subsidiary's capital (see article 67 of the Royal Decree 4/2004 of the CIT for details). The LTU's revenue criterion refers to individual legal entities or consolidated fiscal groups (article 121 of the VAT Law), not to corporate groups. Therefore, a firm with annual revenue below €6.01 million that belongs to a large corporate group will not be included in the LTU census, but it would be included if it were part of a consolidated fiscal group. Note that consolidated fiscal groups are typically formed by the largest corporations, whereas corporate groups are more common. We do not have access to data on which firms belong to fiscal groups, but according to the Statistics published by the Spanish Tax Authority from 2004 to 2007, less than 1% of firms in the revenue range €1.5-€6 million are considered part of the LTU due to belonging to fiscal groups. Since

this is a very small share, we do not expect that including these firms in the analysis will introduce a noticeable bias in our bunching estimates.

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. Since we do not have information on the geographic destination of sales at firm level, we are unable to identify which large firms in these regions are within the LTU stricter tax monitoring. 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 the total number 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.2). 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;[4] and (iii) more than 50 employees on average during the year. These criteria also determine whether a firm can

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

use the abbreviated form of the corporate income tax return, rather than the standard (long) version. These requirements create compliance costs,[5] and the private audit information could complement tax enforcement. Private auditors have a legal responsibility to communicate tax misreporting to the authorities only in the (extreme) case of detecting systematic fraud and criminal activities. However, they do not have to report neither their statement nor detected accounting inconsistencies to the tax authority.

Third-Party Information Reported in Tax Returns

The tax authority introduced in the 1980's a mandatory information form (*Modelo 347*) in which all firms, both below and above the LTU threshold, 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. Note that we do not have access to any data from these forms.

D Data: Further Details

Main Variables Used in the Analysis

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, i.e. the cost of all raw materials and services purchased by the firm in the production process; (iii) labor expenditures, which accounts for the total wage bill of a firm, excluding social security contributions; (iv) number of employees; (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. Table A.5 reports summary statistics for all these variables.

Definition of Sectors of Activity

Table A.1 provides the sector definitions that we use in Section 5.3 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.

[5]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).

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 (País 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 A.5. The Stata do-files used to process the original data to arrive at the final dataset are available on the *Journal*'s website.

Input-Output Tables

We use the input-output tables produced by the National Statistics Institute (*Instituto Nacional de Estadistica, 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.

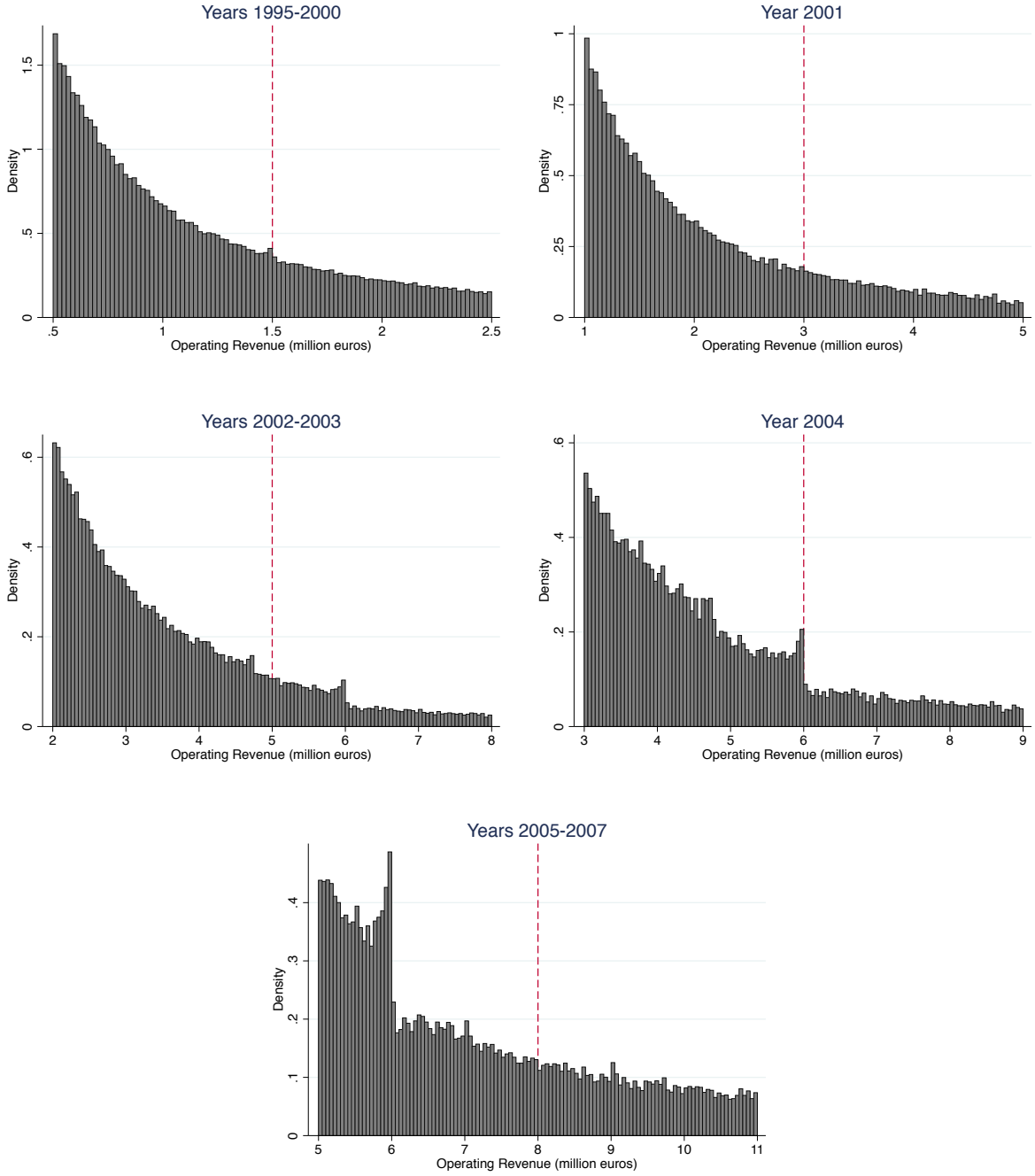
References

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

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).

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. Only firms with positive taxable profits are included (about 80% of the full sample), because the tax rate is irrelevant for them. 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. The results are essentially identical when using the full sample of firms.

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

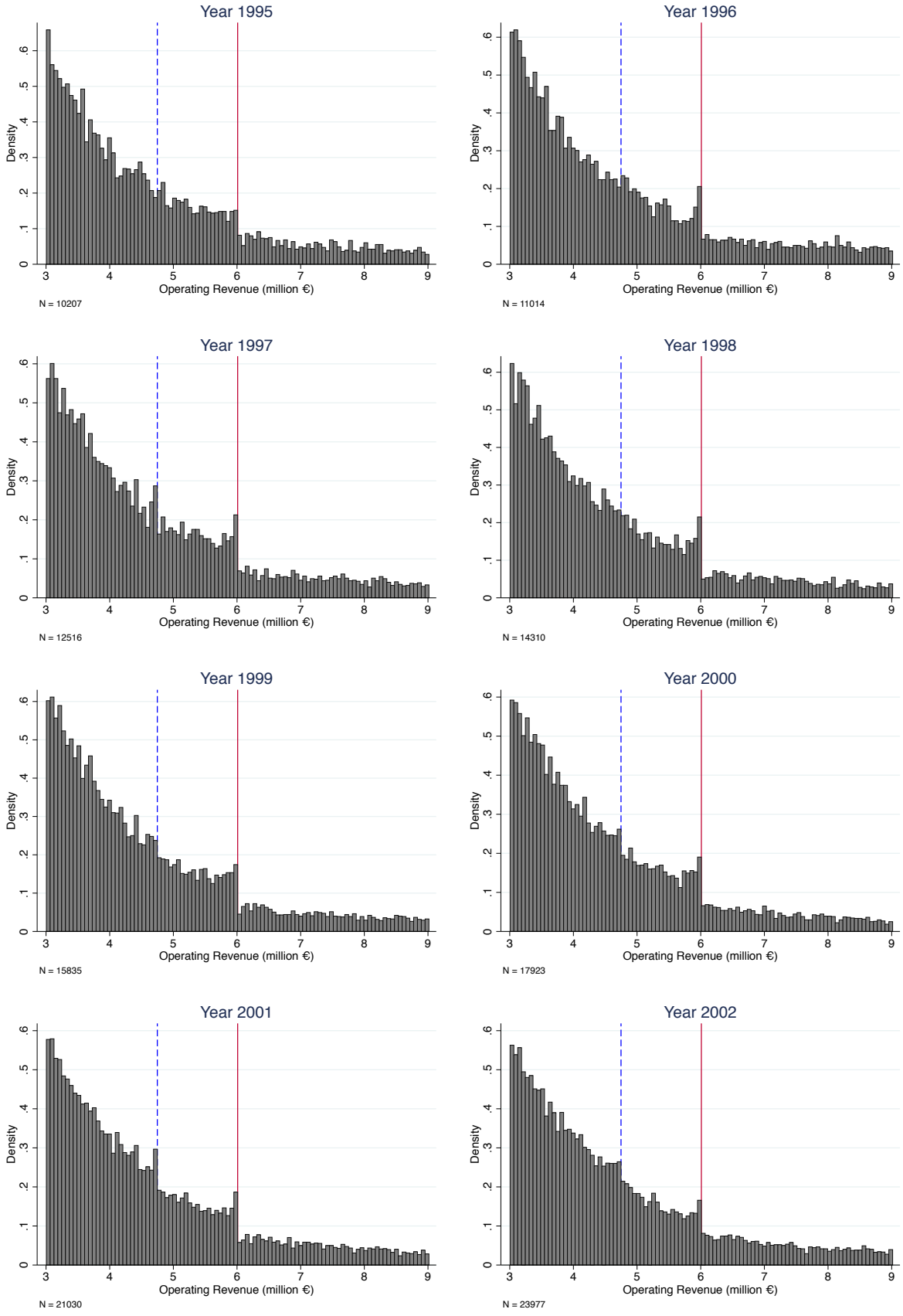
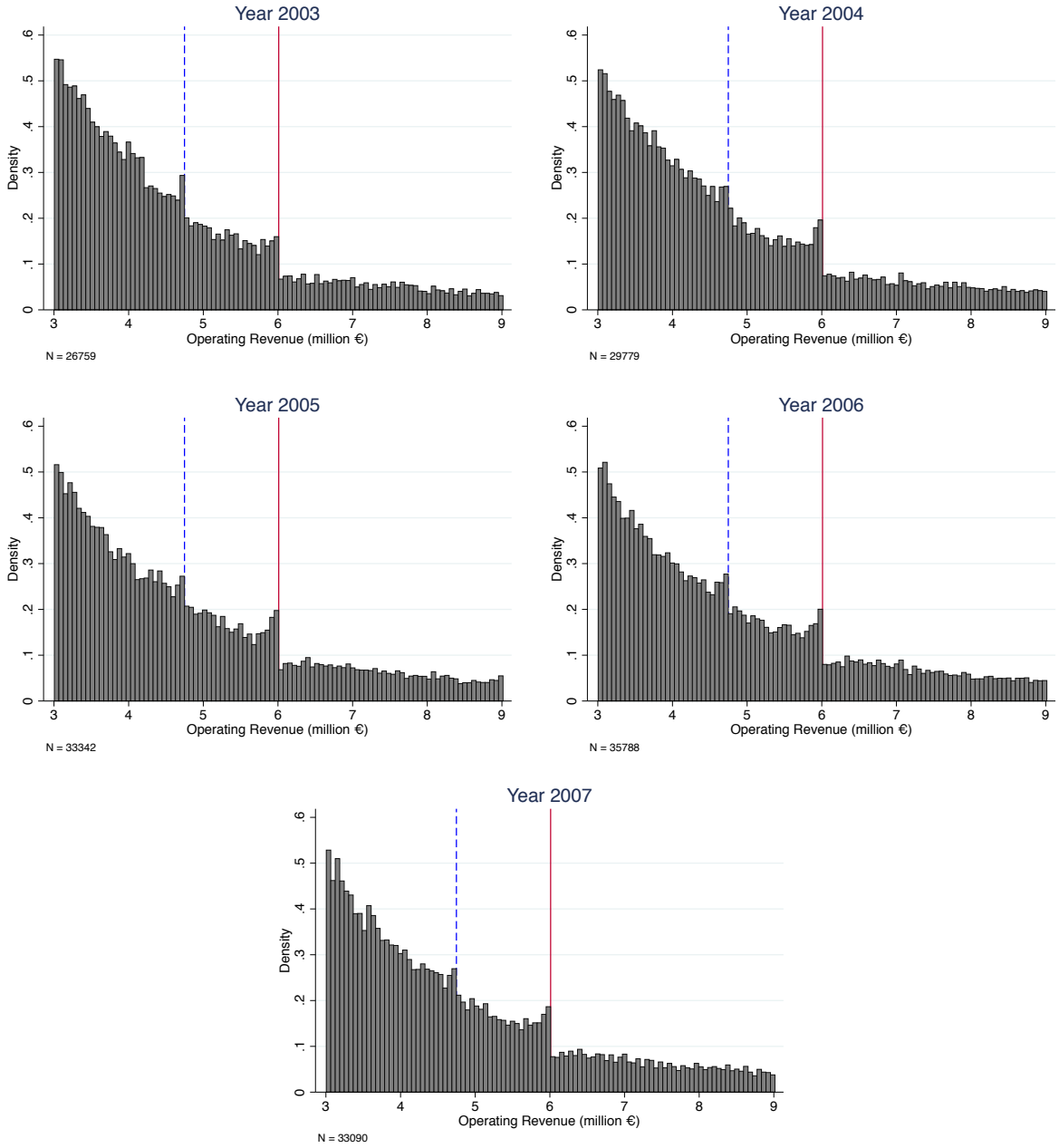


Figure A.2: Distribution of Reported Revenue, by Year (continued)



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

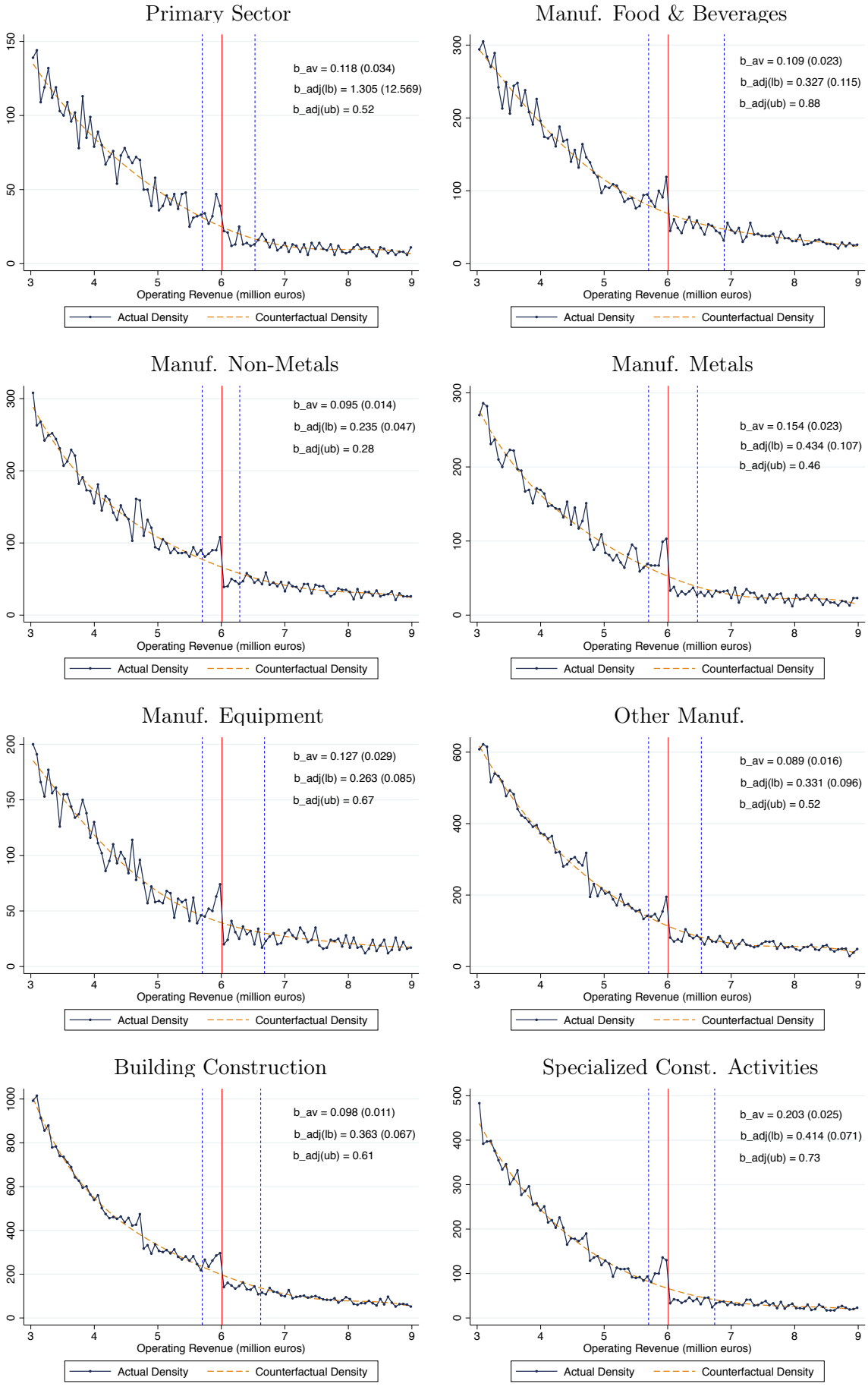
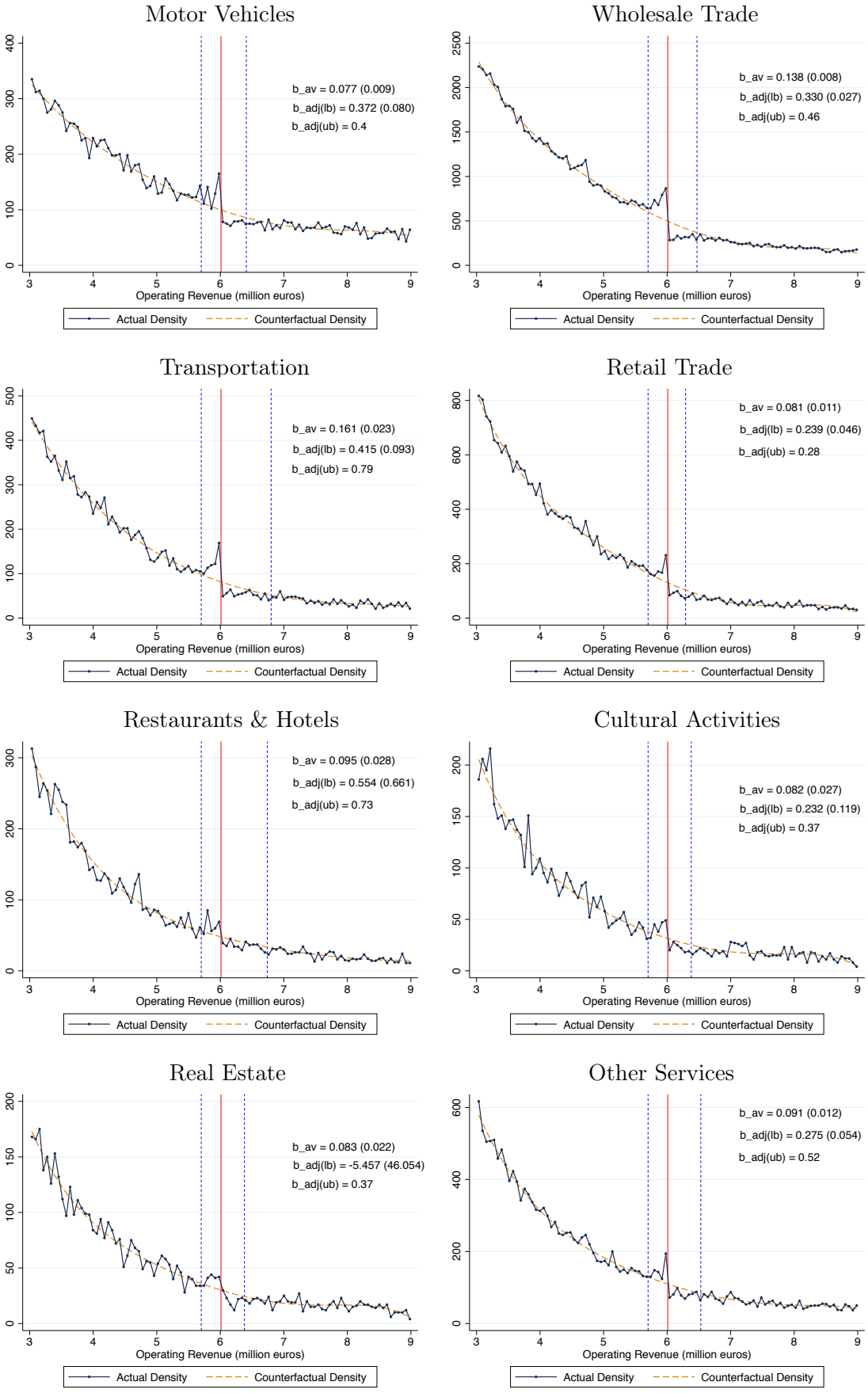


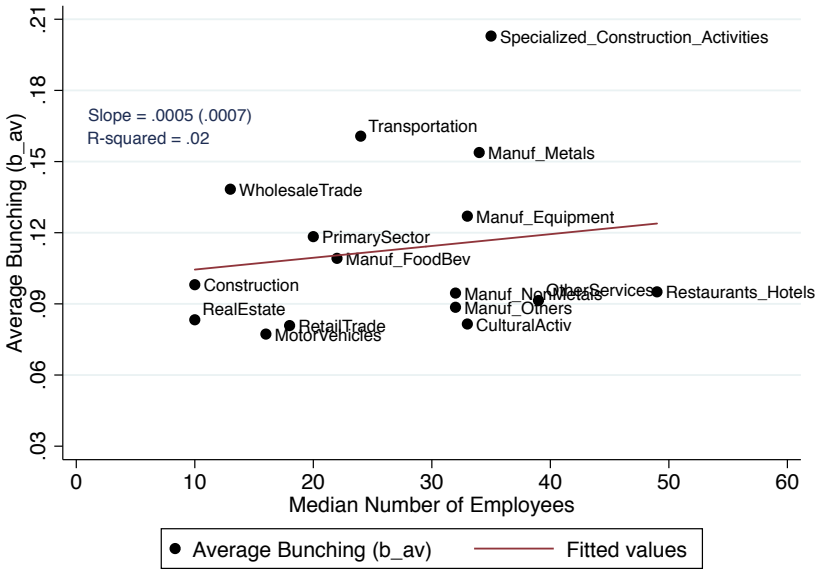
Figure A.3: Bunching Response by Sector (continued)



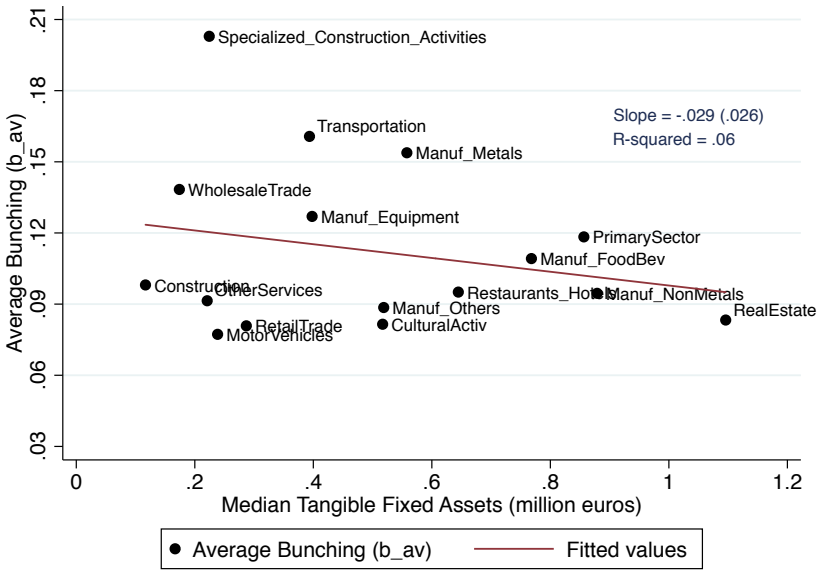
Notes: these graphs show the observed and counterfactual operating revenue distribution for each sector in the period 1995-2007. The dashed (red) vertical line indicates the LTU threshold. The estimation method is identical to that applied in Figure 2 in the paper and explained in the main text.

Figure A.4: Average Bunching by Firm Size Characteristics

(a) Median Number of Employees



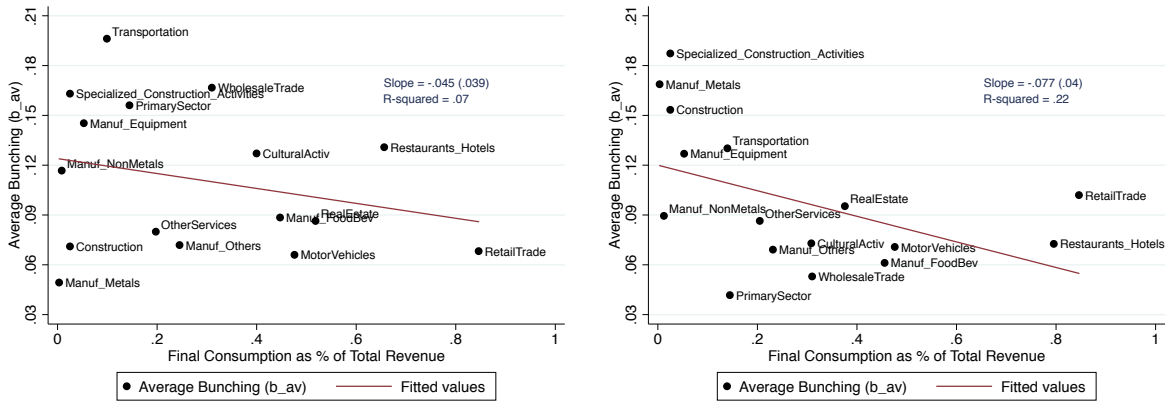
(b) Median Tangible Fixed Assets



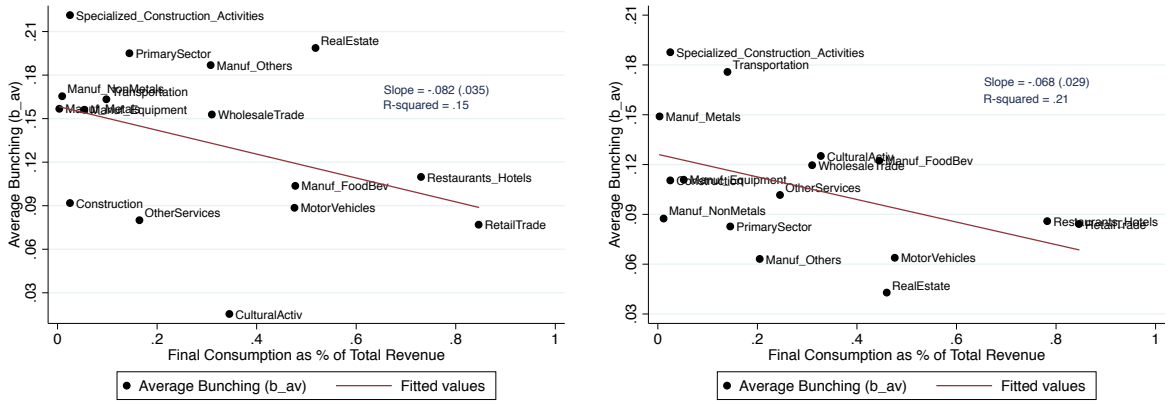
Notes: these graphs are robustness checks for the complementarity result depicted in Figure 3 in the paper. The top panel shows average bunching against the median number of employees by sector. The bottom panel shows average bunching against median tangible fixed assets by sector. The slope of the relationship is close to zero and not statistically significant in either case, suggesting that firm size characteristics such as employment or tangible assets are not strongly related to bunching behavior at the sector level.

Figure A.5: Robustness of Complementarity Result

(a) Below (left) or Above (right) Median Number of Employees



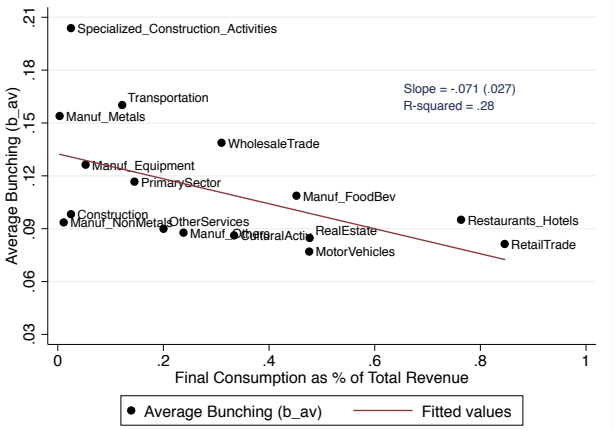
(b) Below (left) or Above (right) Median Tangible Fixed Assets



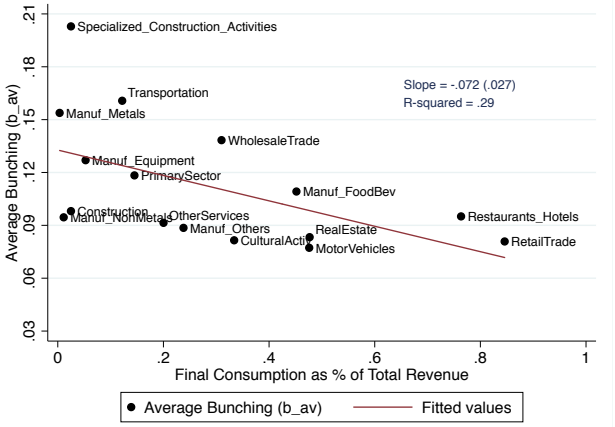
Notes: these graphs are robustness checks for the complementarity result depicted in Figure 3 in the paper. The top panel shows the average bunching estimates by sector, using only firms below (left panel) and above (right panel) the median number of employees in the overall sample. Similarly, the bottom panel shows the average bunching estimates by sector, using only firms below (left panel) and above (right panel) the median of tangible fixed assets in the overall sample. The slope of the relationship is negative in all cases as in Figure 3, and it is statistically different from zero in all cases except for the firms below the median number of employees.

Figure A.6: Robustness of Complementarity Result to Bin Size

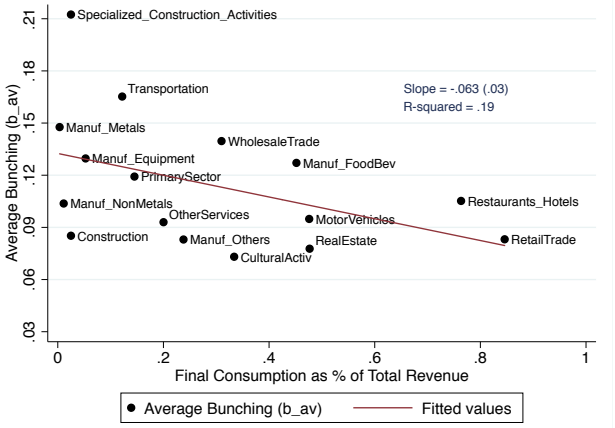
(a) Bin width = €30,050



(b) Bin width = €60,101 (as reported in Figure 3 in the paper)

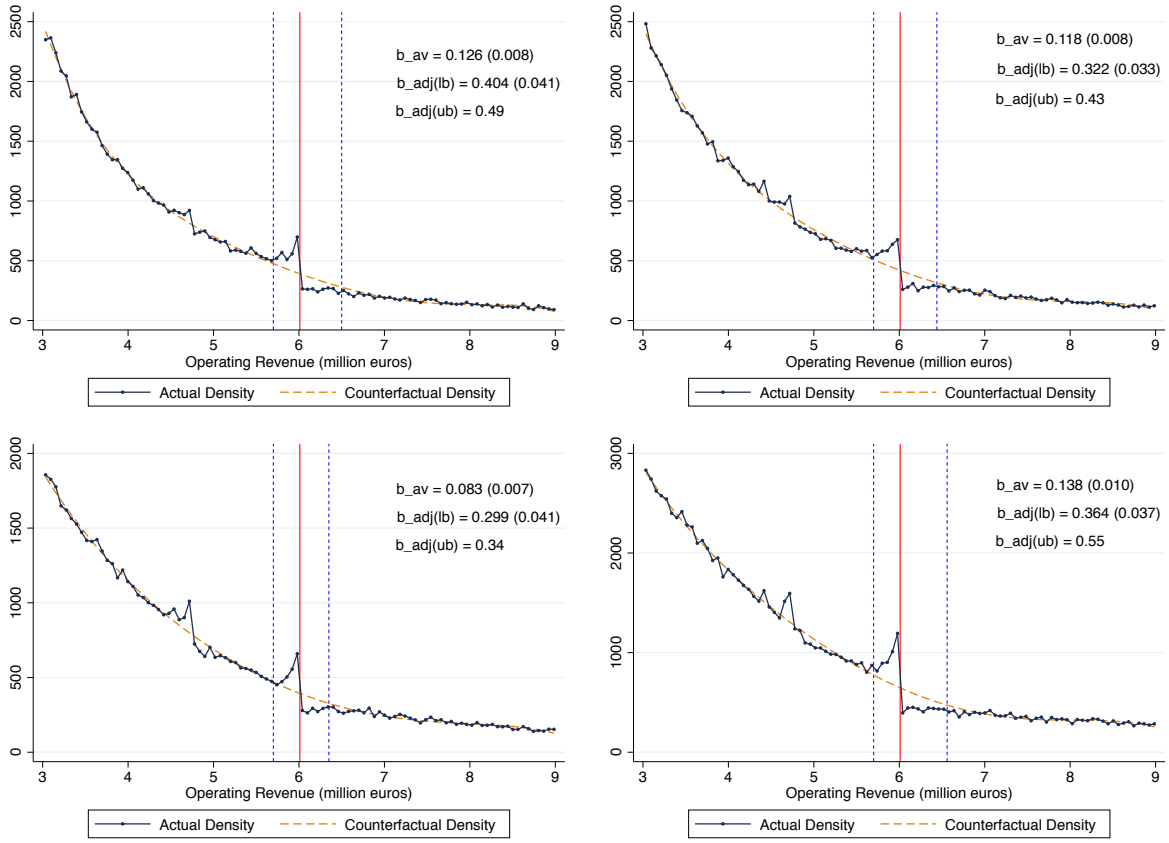


(c) Bin width = €120,202



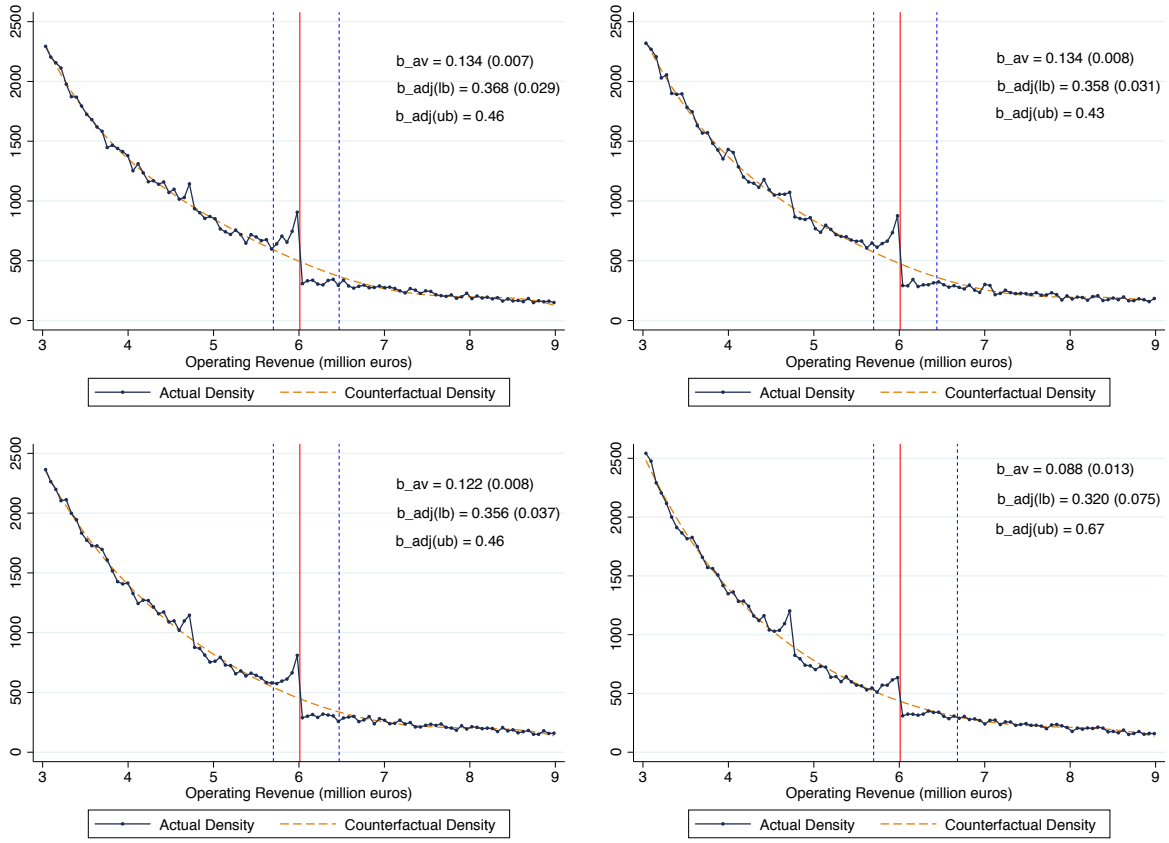
Notes: these graphs are robustness checks for the complementarity result depicted in Figure 3 in the paper, showing how the results change using different bin sizes. The top panel shows the correlation between average bunching by sector and share of final sales using a bin width of €30,050 in the bunching estimation. The middle panel uses a bin width of €60,101 (exactly the same as Figure 3). The bottom panel uses a bin width of €120,202. The slope of the relationship is negative and significant in all cases, and the magnitude is very similar.

Figure A.7: Firm Size Distribution by Quartile of Employment



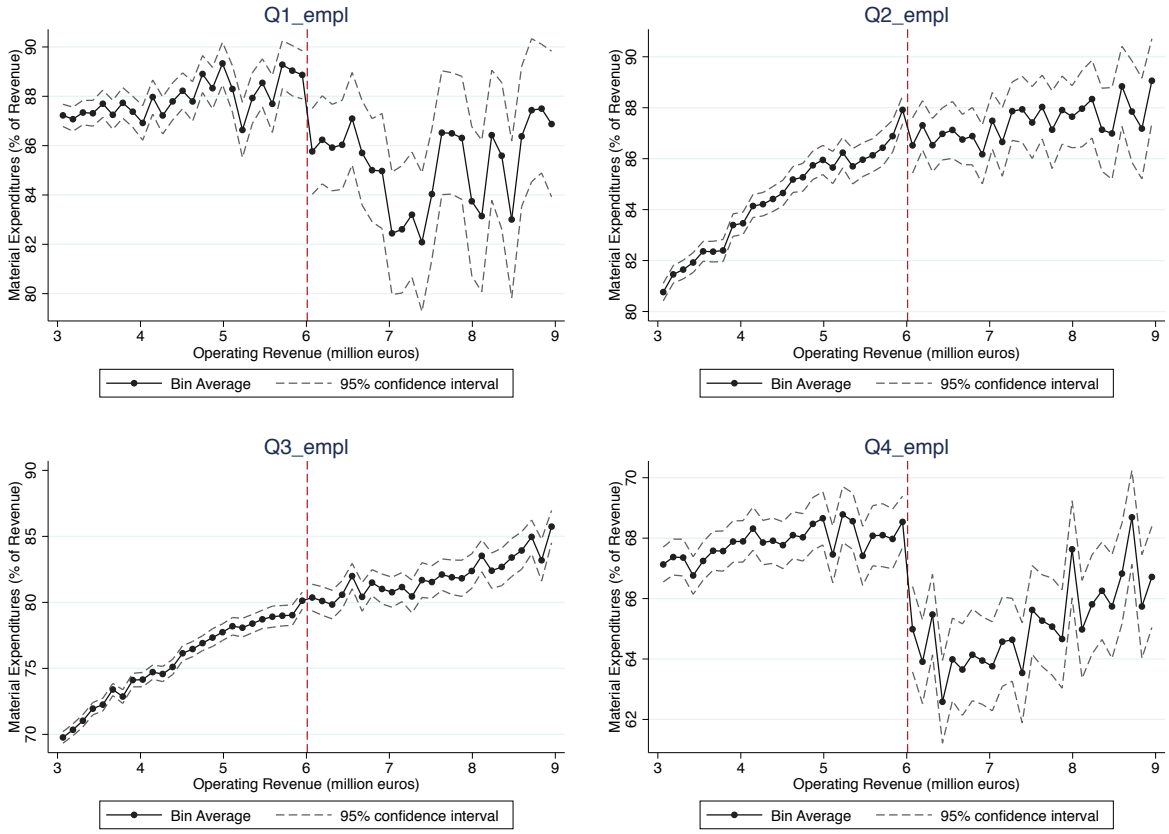
Notes: these graphs show the observed and counterfactual operating revenue distribution by quartiles of the employment distribution in the period 1995-2007. We divide all firms in the sample into four quartiles based on the employment distribution. Q1 = 0-9 employees; Q2 = 10-21 employees; Q3 = 22-38 employees; Q4 = 39 or more employees. The dashed (red) vertical line indicates the LTU threshold. The estimation method is identical to that applied in Figure 2 and explained in the main text. (Bin width= €60,101).

Figure A.8: Firm Size Distribution by Quartile of Tangible Fixed Assets



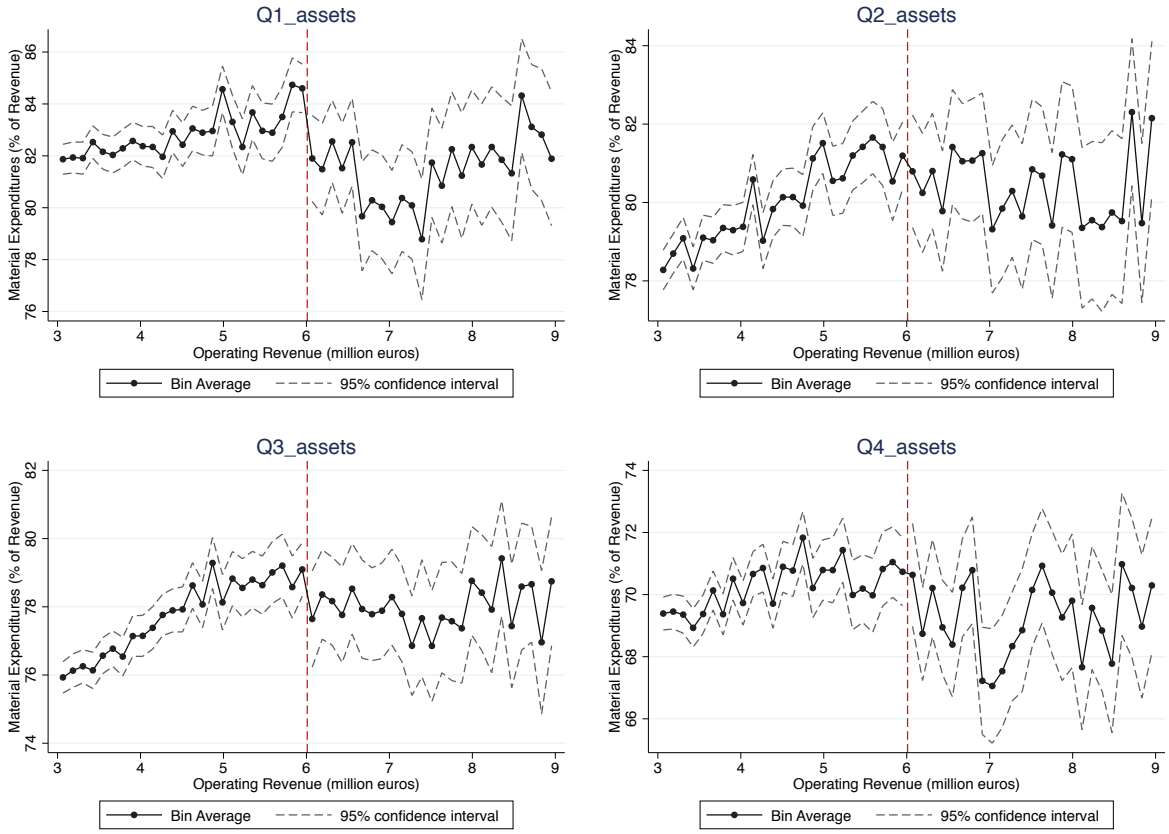
Notes: these graphs show the observed and counterfactual operating revenue distribution by quartiles of tangible fixed assets in the period 1995-2007. We divide all firms in the sample into four quartiles based on the distribution of tangible fixed assets. Q1 = (€0, €0.154) million in fixed assets; Q2 = (€0.145,€0.463); Q3 = (€0.463,€1.163); Q4 = (€1.163,∞). The dashed (red) vertical line indicates the LTU threshold. The estimation method is identical to that applied in Figure 2 and explained in the main text. (Bin width= €60,101).

Figure A.9: Reported Material Expenditures by Quartiles of Employment



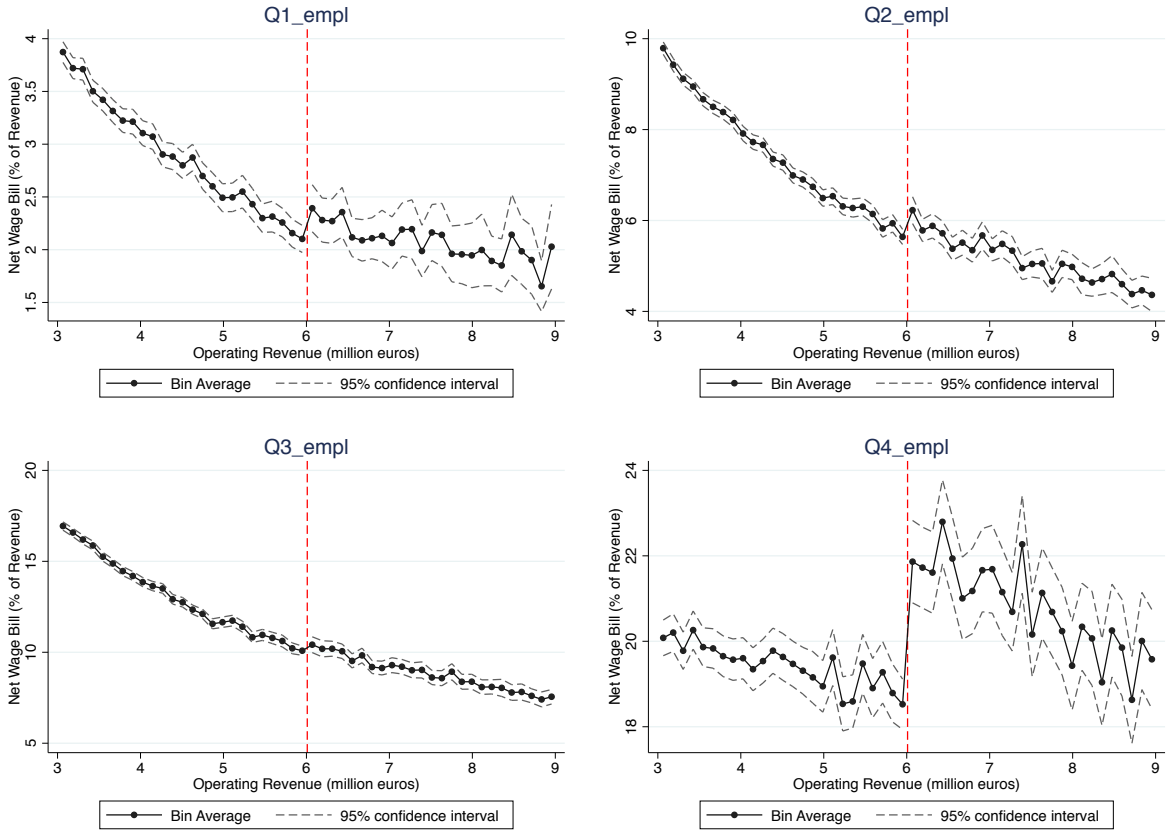
Notes: these graphs show the average ratio of material expenditures as a percentage of total revenue by sector of activity, the same outcome as Figure 4 in the main text. We divide all firms in the sample into four quartiles based on the employment distribution. Q1 = 0-9 employees; Q2 = 10-21 employees; Q3 = 22-38 employees; Q4 = 39 or more employees. 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. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the 1st and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

Figure A.10: Reported Material Expenditures by Quartiles of Tangible Fixed Assets



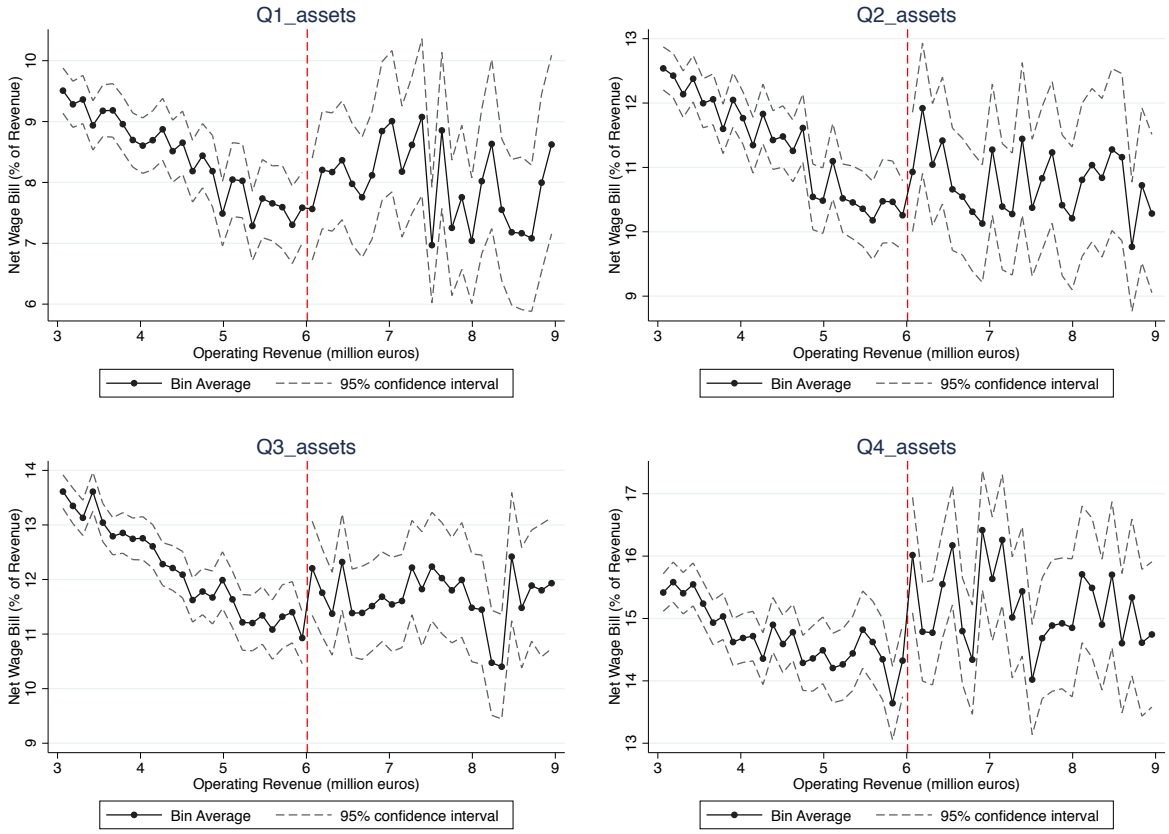
Notes: these graphs show the average ratio of material expenditures as a percentage of total revenue, the same outcome as Figure 4 in the main text. We divide all firms in the sample into four quartiles based on the distribution of total fixed assets. Q1 = (€0, €0.154) million in fixed assets; Q2 = (€0.145,€0.463); Q3 = (€0.463,€1.163); Q4 = (€1.163,∞). 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. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the 1st and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

Figure A.11: Reported Labor Expenditures by Quartiles of Employment



Notes: these graphs show the average ratio of the net wage bill as a percentage of total revenue. The net wage bill is defined as the total wage bill excluding employee-contributed payroll taxes (social security contributions), as in Figure 4 in the main text. We divide all firms in the sample into four quartiles based on the employment distribution. Q1 = 0-9 employees; Q2 = 10-21 employees; Q3 = 22-38 employees; Q4 = 39 or more employees. The dashed (red) vertical line indicates the LTU threshold. The black dotted lines denote bin averages and the gray dashed lines show 95% confidence intervals for each bin average. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the 1st and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

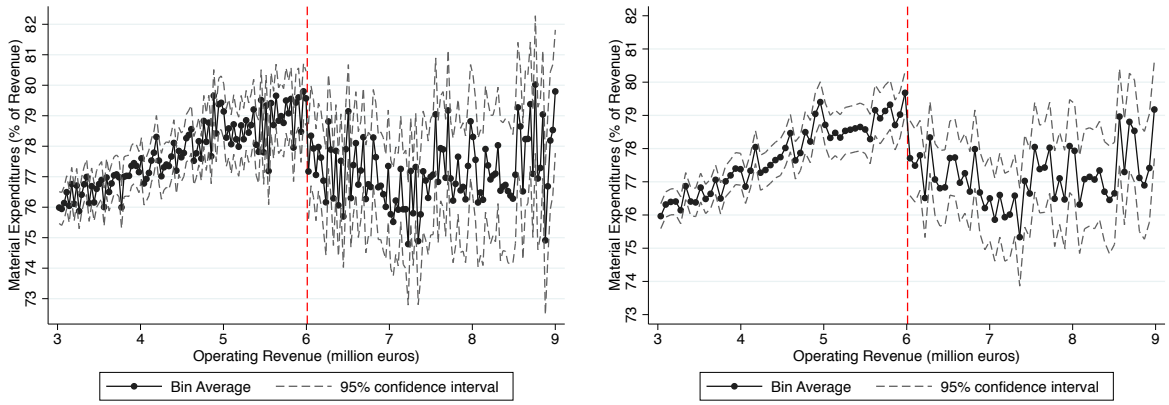
Figure A.12: Reported Labor Expenditures by Quartiles of Tangible Fixed Assets



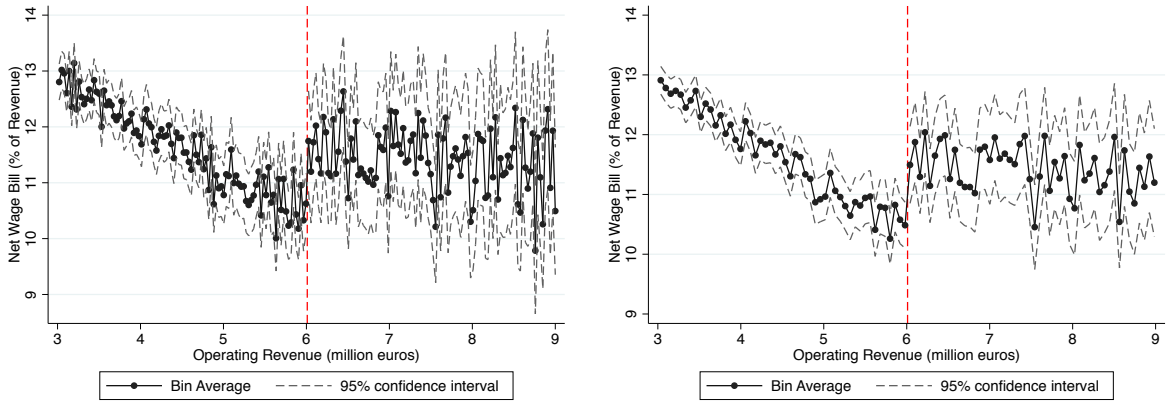
Notes: these graphs show the average ratio of the net wage bill as a percentage of total revenue by sector of activity. The net wage bill is defined as the total wage bill excluding employee-contributed payroll taxes (social security contributions), as in Figure 4. We divide all firms in the sample into four quartiles based on the distribution of total fixed assets. Q1 = (€0, €0.154) million in fixed assets; Q2 = (€0.145,€0.463); Q3 = (€0.463,€1.163); Q4 = (€1.163,∞). The dashed (red) vertical line indicates the LTU threshold. The black dotted lines denote bin averages and the gray dashed lines show 95% confidence intervals for each bin average. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the 1st and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

Figure A.13: Reported Input Expenditures: Different Bin Sizes

(a) Reported Material Expenditures

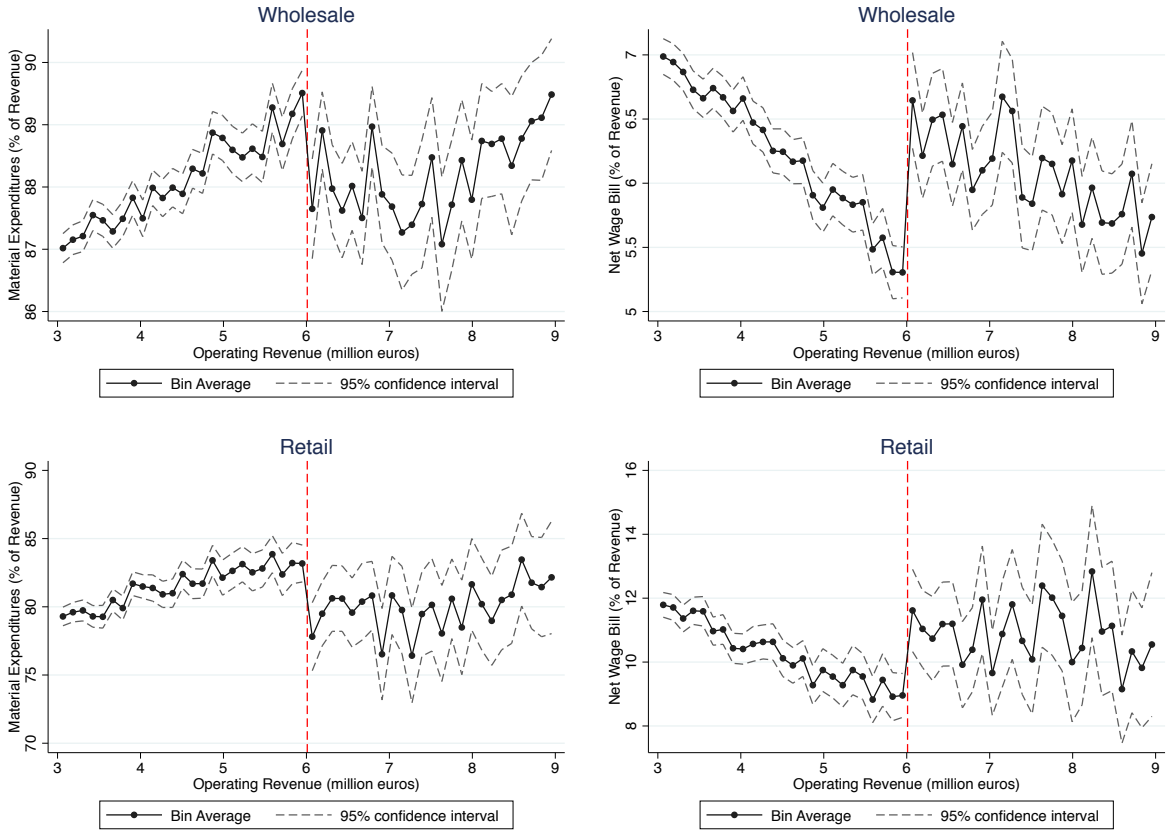


(b) Reported Labor Expenditures



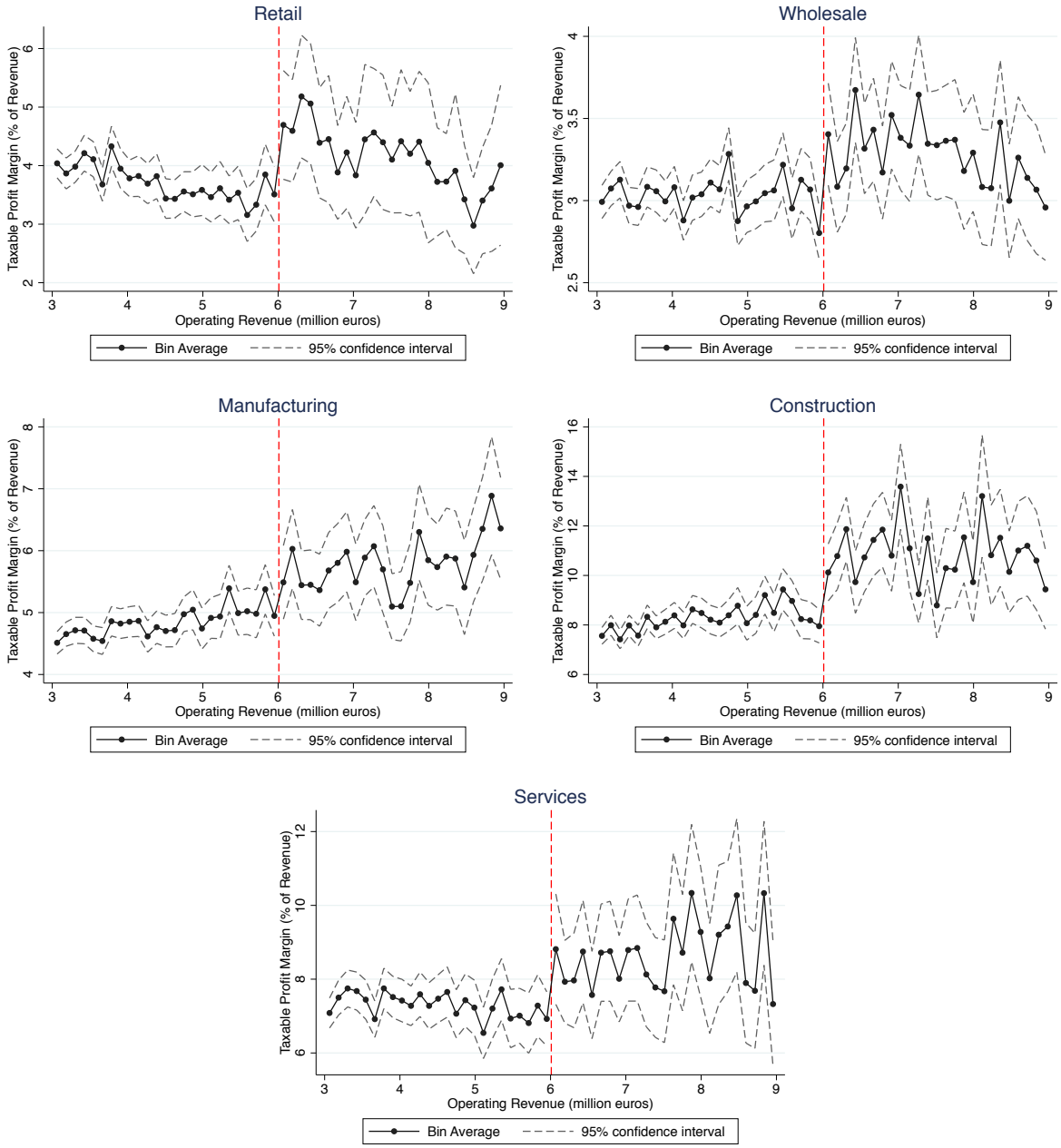
Notes: these graphs show the averages of reported inputs (materials and labor) using smaller bin sizes than in Figure 4. For each outcome, the figure on the left uses a bin width of €30,050, and the figure on the right uses a bin width of €60,101. The patterns are essentially the same as in the figure from the main text, where the bin size is €120,202.

Figure A.14: Reported Input Expenditures: Wholesale vs. Retail



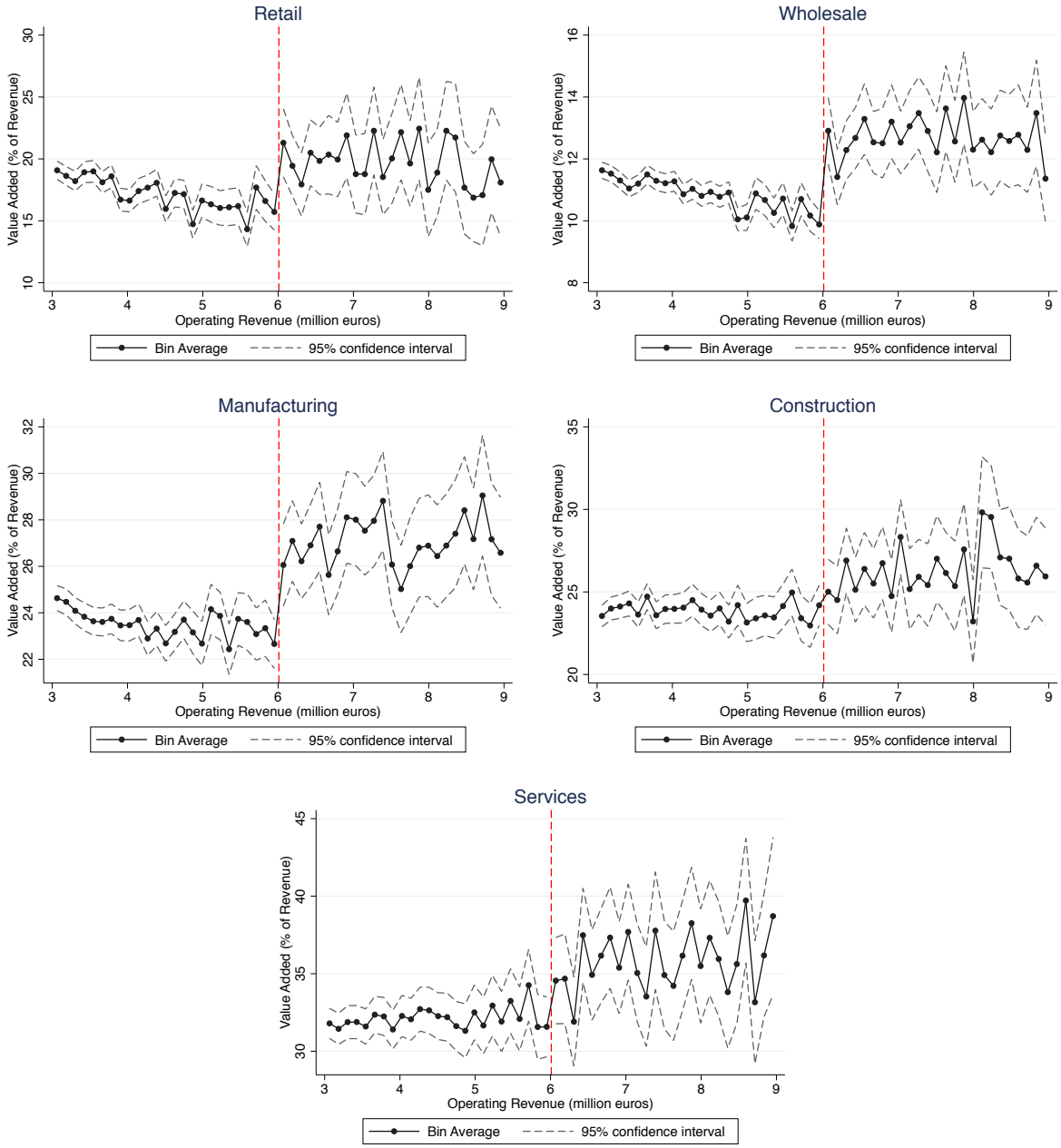
Notes: these graphs show average material and labor expenditures (the same outcomes as in Figure 4 in the main text) for two broad sectors, wholesale and retail. These sectors are chosen because they represent two polar cases in terms of the share of sales made to final consumers. The sector definitions are broader than in some of the previous exercises to ensure enough sample size. Wholesale includes wholesale trade and motor vehicles. Retail includes retail trade and restaurants and hotels. The dashed (red) vertical line indicates the LTU threshold. The black dotted lines denote bin averages and the gray dashed lines show 95% confidence intervals for each bin average. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the 1st and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

Figure A.15: 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 enough 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 gray dashed lines show 95% confidence intervals for each bin average. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the 1st and 99th percentile. We do this for 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 5. (Bin width= €120,202).

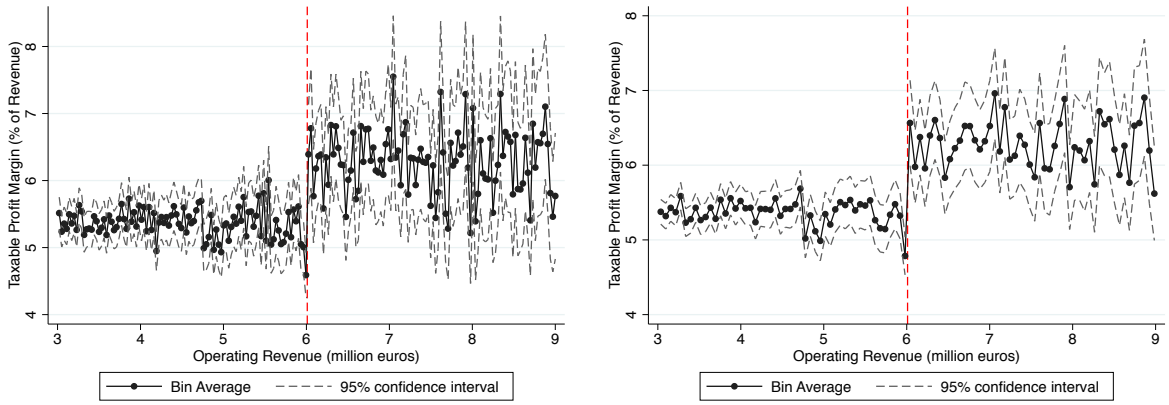
Figure A.16: Reported Value Added Tax Base by Sector



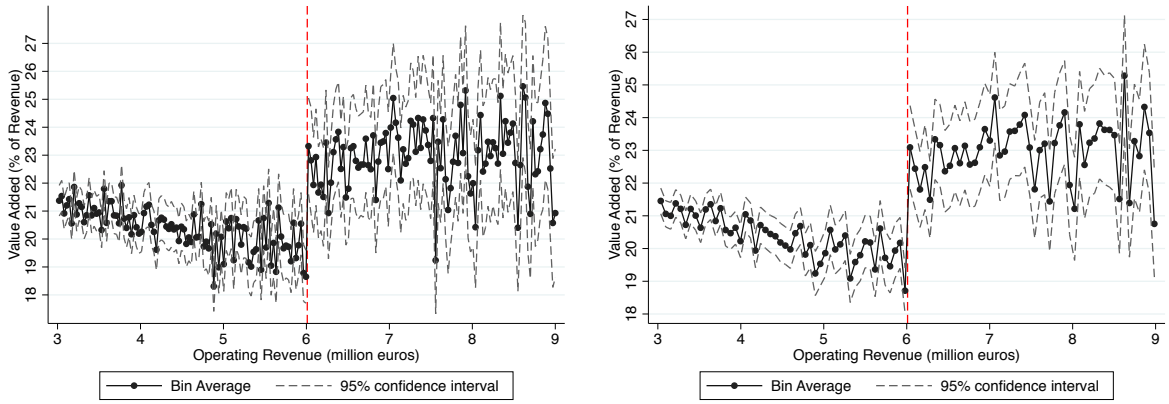
Notes: these graphs show the average value-added tax (VAT) base as a percentage of total revenue by sector of activity. We distinguish 5 broad sectors of activity to ensure that there is enough statistical power to compare the behavior of firms below and above the LTU threshold, indicated by the the dashed (red) vertical line. The black dotted lines denote bin averages and the gray dashed lines show 95% confidence intervals for each bin average. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the first and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million. (Bin width= €120,202).

Figure A.17: Reported Tax Bases: Different Bin Sizes

(a) Reported Taxable Profits (as % of Total Revenue)



(b) Reported Value Added (as % of Total Revenue)



Notes: these graphs show the averages of reported tax bases (taxable profits and value added) using smaller bin sizes than in Figure 5 in the main text. For each outcome, the figure on the left uses a bin width of €30,050, and the figure on the right uses a bin width of €60,101. The patterns are essentially the same as in Figure 5, where the bin size is €120,202. The dashed (red) vertical line indicates the LTU threshold. The black dotted lines denote bin averages and the gray dashed lines show 95% confidence intervals for each bin average. To avoid the spurious effect of extreme values, we winsorize observations in the top and bottom 1% of the outcome variable, meaning that we set those values equal to the first and 99th percentile. We do this for each €1-million interval in the range $\overline{y} \in (3, 9)$ million.

Appendix Tables

Table A.1: Number of Observations by 2-digit CNAE Sector codes

Sector	**CNAE-2009 Sector Codes**	**# of Observations**	
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: CBB dataset described in the main text for the number of observations. For the sector classifications, see http://www.ine.es/daco/daco42/clasificaciones/cnae09/estructuraen.pdf.

Table A.2: Revenue Threshold: Corporate Income Tax Benefit for Small Firms

Year	Threshold	Standard tax rate	Special tax rate	Applicable range
1999	€1.5 million	35%	30%	Up to €90,151 of taxable profits
2000				
2001	€3 million			
2002	€5 million			
2003				
2004	€6 million			
2005	€8 million			Up to €120,202 of taxable profits
2006				
2007		32.5%	27.5%	

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.3: 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.4: 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 Limitada* (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 2.3.

Table A.5: 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's 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 online appendix D.

Table A.6: 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	0.151	0.439	285,570
			(0.015)***	(0.079)***	
5	5.30	6.68	0.140	0.411	285,570
			(0.013)***	(0.072)***	
4	5.40	6.68	0.149	0.433	285,570
			(0.012)***	(0.068)***	
5	5.40	6.71	0.147	0.431	285,570
			(0.011)***	(0.064)***	
4	5.50	6.59	0.140	0.408	285,570
			(0.009)***	(0.050)***	
5	5.50	6.62	0.135	0.394	285,570
			(0.008)***	(0.048)***	
4	5.60	6.53	0.131	0.381	285,570
			(0.008)***	(0.041)***	
5	5.60	6.59	0.129	0.375	285,570
			(0.007)***	(0.040)***	
4	5.70	6.47	0.120	0.350	285,570
			(0.006)***	(0.035)***	
5	**5.70**	**6.53**	**0.121**	**0.382**	285,570
			(0.007)* **	**(0.036)*** **	
4	5.80	6.38	0.106	0.301	285,570
			(0.004)***	(0.023)***	
5	5.80	6.41	0.108	0.312	285,570
			(0.004)***	(0.020)***	

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%.

Table A.7: Sensitivity of Adjusted Bunching Estimator by Year

Year	Upper Bound of Estimation Interval $Y^D = 6.07$	$Y^D = 6.13$	$Y^D = 6.19$
Pooled data			
1995-2007	0.353	0.382	0.424
	(0.032)	(0.036)	(0.043)
Annual data			
1995	0.408	0.257	0.331
	(0.170)	(0.079)	(0.121)
1996	0.303	0.395	0.379
	(0.069)	(0.111)	(0.105)
1997	0.538	0.515	0.634
	(0.125)	(0.118)	(0.173)
1998	0.331	0.351	0.368
	(0.069)	(0.077)	(0.084)
1999	0.326	0.416	0.508
	(0.065)	(0.097)	(0.139)
2000	0.601	0.660	0.698
	(0.138)	(0.166)	(0.188)
2001	0.313	0.350	0.437
	(0.060)	(0.073)	(0.105)
2002	0.429	0.410	0.401
	(0.116)	(0.109)	(0.107)
2003	0.302	0.348	0.381
	(0.064)	(0.080)	(0.095)
2004	0.497	0.550	0.566
	(0.109)	(0.132)	(0.142)
2005	0.268	0.321	0.356
	(0.034)	(0.045)	(0.054)
2006	0.292	0.298	0.317
	(0.042)	(0.044)	(0.048)
2007	0.301	0.305	0.351
	(0.051)	(0.053)	(0.067)

Notes: this table shows the sensitivity of the adjusted bunching estimator b^{lb}_{adj} to different values of the upper bound of the interval on which this parameter is estimated, y^D. The main estimates reported in Table 1 are reported in the central column here. When $y^D = 6.07$, the proportion of non-bunchers α is estimated as the ratio of the counterfactual frequency to the actual frequency in the first bin to the right of the LTU threshold. When $y^D = 6.13$ and $y^D = 6.20$, we use the first two and three bins, respectively.

Table A.8: Sensitivity of Adjusted Bunching Estimator by Sector

Sector	Upper Bound of Estimation Interval $Y^D = 6.07$	$Y^D = 6.13$	$Y^D = 6.19$
Primary Sector	1.305	1.301	0.571
	(12.569)	(21.917)	(0.595)
Manuf. Food and Beverages	0.327	0.533	0.510
	(0.115)	(0.320)	(0.299)
Manuf. Non-Metals	0.235	0.245	0.294
	(0.047)	(0.050)	(0.067)
Manuf. Metals	0.434	0.528	0.450
	(0.107)	(0.152)	(0.117)
Manuf. Equipment	0.263	0.300	0.513
	(0.085)	(0.107)	(0.374)
Manuf. Others	0.331	0.291	0.306
	(0.096)	(0.079)	(0.087)
Construction of Buildings	0.363	0.478	0.499
	(0.067)	(0.105)	(0.114)
Specialized Constr. Activ.	0.414	0.496	0.529
	(0.071)	(0.097)	(0.109)
Motor Vehicles	0.372	0.361	0.345
	(0.080)	(0.077)	(0.072)
Wholesale (exc. Motor V.)	0.330	0.342	0.384
	(0.027)	(0.029)	(0.035)
Transportation	0.415	0.485	0.595
	(0.093)	(0.124)	(0.190)
Retail Trade	0.239	0.283	0.338
	(0.046)	(0.061)	(0.082)
Restaurants and Hotels	0.554	0.470	0.719
	(0.661)	(0.389)	(4.679)
Cultural Activities	0.232	0.393	0.449
	(0.119)	(0.439)	(2.375)
Real Estate	-5.457	0.952	0.456
	(46.054)	(5.939)	(0.729)
Other Services	0.275	0.326	0.455
	(0.054)	(0.071)	(0.125)

Notes: this table shows the sensitivity of the adjusted bunching estimator b^{lb}_{adj} to different values of the upper bound of the interval on which this parameter is estimated, y^D. The main estimates reported in Table 2 are reported in the central column here. When $y^D = 6.07$, the proportion of non-bunchers α is estimated as the ratio of the counterfactual frequency to the actual frequency in the first bin to the right of the LTU threshold. When $y^D = 6.13$ and $y^D = 6.20$, we use the first two and three bins, respectively.

Table A.9: Robustness Checks for Complementarity Result

	Dependent variable: Average Bunching (b_{av})							
	Empl. below median		Empl. above median		Assets below median		Assets above median	
	OLS	WLS	OLS	WLS	OLS	WLS	OLS	WLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of Final Consumer Sales	-0.042	-0.029	-0.062	-0.099	-0.096	-0.034	-0.069	-0.035
	(0.041)	(0.017)*	(0.034)*	(0.025)***	(0.042)**	(0.017)**	(0.029)**	(0.020)*
Median Number of Employees					-0.001	-0.000	-0.001	-0.000
					(0.002)	(0.001)	(0.001)	(0.000)
Median Tangible Fixed Assets	-0.018	-0.008	-0.037	-0.002				
	(0.018)	(0.027)	(0.021)	(0.012)				
Constant	0.129	0.105	0.143	0.100	0.184	0.131	0.104	0.106
	(0.021)***	(0.009)***	(0.021)***	(0.011)***	(0.042)***	(0.013)***	(0.018)***	(0.016)***
Observations	16	16	16	16	16	16	16	16
Total obs. in subsample	124,864	124,864	121,122	121,122	142,737	142,737	142,833	142,833

Notes: this table shows robustness checks for the complementarity result, which corresponds to Figure 3 and Tables 3 and 4 in the main text. Columns 1, 3, 5 and 7 are estimated by OLS, and columns 2, 4, 6 and 8 are estimated by WLS. In the latter, the weights are the inverse of the variance of the bunching estimates. Columns 1-2 use data for firms below the median number of employees (in the overall sample), while columns 3-4 use data for firm above the median number of employees. Similarly, columns 5-6 use data for firms below the median level of tangible fixed assets, while columns 7-8 use data for firms above the median level of tangible fixed assets. The unit of observation is the sector of activity (for details on the definition of the 16 sectors, see section D above). Robust standard errors reported in parentheses. Significance levels: *** = 1%, ** = 5%, and * = 10%.