

Tax Avoidance in Firms

Arthur Bauer*
Martin Rotemberg †

January 30, 2018

Abstract

Tax codes can have *notches*; regions in which after-tax profits are decreasing in before-tax sales. Firms endogenously respond to the notches, leading to bunches in the firm-size distribution. We describe a 1997 policy reform in which the French government implemented a transient tax reform that increased profit taxes by 15% by firms with over 50 million Francs in turnover. We use two complementary approaches to estimate the extent of tax avoidance: from a counterfactual distribution generated from firms far away from the tax notch in the same year, and using the entire pre-tax reform distribution. Both results generate similar results for the extent of tax avoidance. We show that the firms who avoid the tax are the ones with the lowest calibrated adjustment costs and those with the largest incentives (the ones with larger profits). The tax avoidance behavior comes mostly from an increase in inventories.

Keywords: Business Taxes, Tax Evasion, Firm Production
JEL: D24, H32

Some of the research in this paper was conducted while the first author was an employee of the National Institute of Statistics and Economic Studies. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the National Institute of Statistics and Economic Studies. No confidential information is disclosed in our results. We are grateful to etienne lehmann and Claire Lelarge for her helpful comments. Contact information: arthur.bauer@insee.fr, 15 Bvd Gabriel Peri 92240 Malakoff. 0141175994 and mrotemberg@nyu.edu

*INSEE

†NYU

1 Introduction

Understanding which firms avoid taxes and their avoidance strategies is a question at the core of the public economics. In particular in France where tax avoidance is estimated to cost about 2 % of GDP, failures of the state to implement manipulation proof tax policy may lower revenue and raise welfare questions.

The 1997 exceptional contribution is an interesting case in point as it applied only for firms with turnover above MFrancs 50. It created a notch: increasing the average tax rate for firms with turnover above the threshold. Besides, it only lasted three years, exhibiting a counterfactual of the firms distribution in years preceding the policy. Finally it created benefits from tax avoidance that were not equally distributed over the distribution of firms. Firms with larger profits would save more money by avoiding this contribution. This created variations in the incentives, useful to assess the impact of the notch.

This paper asks first whether this contribution had distributional effects that suggest that firms tried and avoided to pay this additional tax. Then, it analyses what firms tried and avoided the most the contribution. Third, it describes how firms proceeded to avoid the contribution.

The first step is classical in the analysis of bunching in the context of a notch. It consists in determining whether the number of firms just below the threshold of MFrancs 50 is in excess compared to what it would have been absent the policy, and whether there is an equivalent number of firms that is missing just above the threshold. We therefore implement the classical method of bunching estimation that consists in predicting the counterfactual distribution of turnover, absent the policy, from the distribution of turnover away from the threshold. We then leverage the transient characteristic of the policy and use the distributions during years before its implementation as counterfactual. This allows us to quantify the amount of firms that are bunching, without relying on prediction.

The second step consists in determining the characteristics of firms that are bunching. This leverages the technique developed by Person and Diamond (2016). It consists in considering bunchers as firms that comply to the "treatment" created by the notch. Their average characteristics are obtained from the weighted difference of the average characteristics of firms just below the threshold and the average characteristics of firms that would have been just below the threshold had there been no incentives to bunch. This weighted difference captures the characteristics of the firms that are below the threshold because of the notch.

To identify firms that would have been just below the threshold had there been no incentives to bunch, we consider firms that are just below the threshold during years without any incentives to bunch. The first characteristics we test to determine whether bunchers are different from the rest of the firms that are eligible to bunching are characteristics related to adjustment costs. Our goal is to test whether firms that have lower adjustment cost indeed have a higher tendency to bunch. Then, we test whether firms that have the larger incentives to bunch, i.e. firms that have the larger profits are also over represented among bunchers.

The third step consists in analyzing how firms proceed to bunch. We first ask whether their production level was affected. We then ask whether firms compensated their decrease in sold production by an increase in the change of inventories in order to smooth production that might be costly to adjust. We again rely on the technique developed by (Person and

Diamond, 2016). This technique consists in comparing the average observed levels of production or of its components of production across all firms in the manipulation region to the average predicted levels of components of production in manipulation region had there been no manipulation. Again to predict the outcome's levels in the manipulation region had there been no manipulation we consider outcome's levels for firms that are in the manipulation region in years during which there was no incentives to bunch.

Using standard techniques of bunching, that consist in predicting the counterfactual distribution with a polynomial of turnover, we show that there were 93, 191 and 83 firms bunching each of the three years during which the policy was implemented. The difference in bunching reflects first the learning process that makes bunching slow to appear and second the difference of strength of the incentives to bunch, that were decreased in 1999, namely the third year the policy was implemented. When we combine the two years with the same incentives to bunch 1997 and 1998 and analyze bunching using the two previous years 1995 and 1996 as counterfactual we show that there were 226 firms bunching over the two years, i.e. an average of 113 firms by years to compare to the average 142 firms determined with the standard technique. However since the standard errors obtained by bootstrapping the procedure are larger than 30, we can confirm that the two techniques provide comparable results.

We then show that firms that have the larger adjustment costs are also the firms with the lower likelihood to bunch. We start with capital adjustment cost, as measured by Asker, Collard-Wexler, and Loecker (2014), who show that they are the most important adjustment costs. We show that bunchers are, depending on the years, between 3% and 15 % more likely to have capital adjustment cost within the top tercile. Then we consider elasticity of output with respect to production inputs. In particular we show that bunchers tend to have larger elasticity of output with respect to material and lower elasticity of output with respect to capital, consistent with the idea that capital is fixed and materials are flexible and that it is therefore easier to adjust production when elasticity of output with respect to material is high.

We also show that firms with the larger incentives to bunch are over represented among bunchers. We find that bunchers are, depending on the years, between 2 and 22% more likely to be within the top tercile of profit.

Finally we show that while firms tend to decrease turnover by kFrancs 88 because they bunch, they also tend to decrease production, even though the standard error on this estimate is large. We show that this slight decrease in production is driven by a large decrease in sold production that is consistent with the decrease in turnover. However the latter is compensated by an increase in change in inventories and capitalized production, that respectively increase by kFrancs 153 and 88. As a result we conclude that this strategy to reduce turnover reflects the idea that it is costly to adjust production and that firms prefer to smooth the amount by which they decrease production, by stocking part of their production or investing it in the production process itself. This also suggests that what we observe is a real effect of tax avoidance rather than misreporting. We confirm the latter by computing the output elasticity with respect to the effective tax rate. We show that it is consistent with other estimates found in the literature, and that contrary to Best et al. (2015) we do not need fraud to rationalize it.

Our contribution to the literature is threefold. First we provide evidence of transient

bunching as a consequences of this temporary tax contribution for firms with turnover above MFrancs 50. There is an extensive literature on bunching that started with (Saez, 2010) seminal contribution. (Kleven and Waseem, 2013; Raj Chetty and Pistaferri, 2011) are two famous examples of the analysis of kink and bunch. For a comprehensive review see (Kleven, 2016). In France, several papers have documented bunching in the distribution of firms, three cases in point are (Bach, 2015; Garicano, Lelarge, and Van Reenen, 2016; Aghion et al., 2017).

Second, we suggest an innovative technique to analyze bunching in a dynamic way. While most of the literature has suggested predicting counterfactual distribution using a polynomial regression with data outside the manipulation region that we extrapolate to the manipulation region, we suggest to use past distribution as the counterfactual. The paper the closest to ours is (Brown, 2013). In contrast to ours it relies on past distribution where the threshold was at a different location, while we rely on past distribution in years during which there is no incentive to bunch, alleviating concern of potential overlap of the manipulation regions.

Third we carry an analysis of the characteristics of the bunchers and their bunching strategy. Our methodology is largely inspired from (Person and Diamond, 2016). Other papers that look at the characteristics of the bunchers are (Thomas Dee and Rockoff, 2016) and (Bach, 2015). However they mostly do it with a naive comparison of bunching levels by subgroup. A technique that does not allow a quantitative description of the characteristics of the bunchers. (Person and Diamond, 2016) develops a technique to analyze the consequence of bunching, we use it to determine the strategy of bunching, considering that in the short run, deviations from the counterfactual levels of production components, reflect the strategy firms adopt in order to bunch.

Section 2 presents a theoretical framework that predicts the characteristics of the bunchers. Section 3 describes the institutional setting that resulted in the notch. Section 4 describes the dataset and the sample. Section 5 presents the result of bunching estimations, both in cross-section and using previous years as counterfactual as well as bunching estimation by subgroups. Section 6 determines bunchers characteristics. Section 7 presents bunching mechanisms. Section 8 concludes.

2 Theoretical framework

Let's simplify the situation by supposing that firms only pay taxes if they are above the threshold. This is tantamount to offsetting the level of tax paid so that it is zero below threshold. Firms make before tax profit π . Then their after tax profit is $\pi^{no\ manipulation} = \pi(1 - \tau)$. Let's consider the case where a firm decides to manipulate its turnover level to avoid taxes. Then it incurs an additional cost $c^{adjustment}$ that decreases its profit level to $\pi^{manipulation} = \pi - c^{adjustment}$

As a result, firms will decide to manipulate their profit as long as

$$\pi^{manipulation} - \pi^{no\ manipulation} > 0 \tag{1}$$

i.e. as long as :

$$\pi \cdot \tau - c^{adjustment} > 0 \tag{2}$$

First, we see here that the incentive to bunch is proportional to one firm’s profit. The amount of money saved by bunching is equal to 15 % times the default tax rate of 33% times the amount of profit (which is the tax base). As a result more profitable firms face higher incentives to bunch. We therefore expect bunchers to be the most profitable firms among all firms eligible to bunching.

Second, we see that the lower the adjustment cost, the higher the incentives to bunch. Adjustment costs depend on the sector the firm belongs to. There are sectors with higher adjustment costs than other and there are sectors where production levels are more flexibly set, i.e. sectors where production depends mostly on flexible inputs.

2.1 Production function

The larger the adjustment costs, the more difficult it will be for firms to adjust their turnover to a level that would allow them to avoid taxes. It is well known in the literature (Asker, Collard-Wexler, and Loecker, 2014) that capital adjustment costs are the largest ones, followed by labor adjustment costs and then material adjustment cost.

As a result, firms that rely on material more than other inputs to adjust their production, will face in average lower adjustment costs and be able to adjust to the level of turnover that allows them to avoid taxes. This reasoning leads us to predict the characteristics of firms that are more likely to bunch. That are the firms for which M intervenes the most in the production function, i.e. firms with the largest elasticity of output with respect to material and firms for which K intervenes the less in the production function, i.e firms with the lowest elasticity of output with respect to capital. The higher the elasticity of output to material, the larger the firm’s ability to adjust its production level by adjusting material, and, as a result, the lower the firms’ adjustment costs.

2.2 The role of the heterogeneity of capital adjustment costs

Asker, Collard-Wexler, and Loecker (2014) show that, among the three inputs of production, capital is the one for which adjustment costs are the larger. As a result we focus on these adjustment costs to determine heterogeneity in bunching. In particular equation 2 predicts that firms in sector with the lower adjustment costs of capital will be more likely to bunch.

Asker, Collard-Wexler, and Loecker (2014) suggest that misallocation is driven by adjustment cost. As a result, they interpret the dispersion in the marginal product of the inputs of production, as a proxy for the adjustment cost. We follow them and construct, sectoral level measures of adjustment costs of capital by measuring, within each industry the dispersion of the marginal product of capital:

$$Adjustment\ cost\ capital_{it} = SD_{it}(\beta_i + s_{ft} - k_{ft}) \quad (3)$$

where s_{ft} is the logarithm of the turnover level of firm f, on year t and k is the logarithm of the capital level of firm f on year t. β_i is the logarithm of the elasticity of output with respect to capital. (We estimate it following Levinsohn and Petrin (2003), at the 4 digit level.)

Our model predicts that the lower the adjustment cost of capital, the higher the tendency to bunch.

3 Institutional setting

When entrepreneurs start a business in France they can chose between two kinds of tax regime: either paying taxes on income (IR) or paying taxes for their enterprise (IS). Besides, firms' sectors determine what category of taxes firms are paying (either BIC, BNC or BA: which respectively correspond to profits form trade and manufacturing activity (BIC) , benefits mostly from services (BNC) and benefits from agricultural activity (BA)). Within each category there are two types of tax regime: the normal one (RN) and the simplified one(RSI) From the dataset that covers the universe of firms subject to BIC we can observe that 40.95% of the firms are actually affiliated to the IS regime.

France voted in 1997 an exceptional temporary contribution for firms paying IS taxes with sales above a threshold of MEuros 7.63 (MFrancs 50), which increased by 15% taxes paid by firms on all their profits. This temporal contribution lasted 3 years. In 1999 the contribution was decreased and firms only paid 10% more taxes. Beyond criteria on sales, eligibility was conditioned on 75% of firms' share capital owned by physical people and all of share capital paid-up ¹. (Raspiller, 2007)

The incentives to distort one firm's behavior to avoid this exceptional contribution were small but significant. Figure 1 shows the gains in euros by profit to becoming eligible to an exemption of the contribution. They lie between 0 euros and plus 15000 euros for profits between 0 and kEuros 300. In our sample the mean of profit is equal to kEuros 218. These incentives are therefore of absolute small magnitude.

However, compared to the incentives created by another simultaneous policy that gave incentives to reduce turnover they are relatively large. They are about three times larger. Almost simultaneously there was indeed an important change in the fiscal law. The government changed the marginal tax rate on profits three times between 1996 and 2000. The new lower rates only applied to profits below kEuros 38, limiting maximal potential gains from these tax cuts to Euros 4000. Besides savings from this tax cut were to be reinvested in firm's capital. This meant additional costs (including financial ones and estimated to around Euros 600) to the firms as signaled by the "Rapporteur general" of 1997 reform. Also, despite the fact that both tax cuts and the temporary contribution favored more small enterprises (i.e. with sales below MEuros 7.63) than the larger ones, the incentives to adjust turnover to become eligible to the reduced tax rate were significantly smaller than those to become eligible to the exemption of the exceptional contribution.

4 Data and sample

The dataset we are using is the BRN-RSI file described in (Bertrand, Schoar, and Thesmar, 2015). It is built from tax forms collected by the Direction Générale des Impôts and contains the universe of French firms.

We further restrict the sample to firms with turnover between kFrancs 20 000 and kFrancs 100 000, as well as to firms eligible to the tax cut. The BRN-RSI dataset allows us to keep

¹cf LOI no 97-1026 du 10 novembre 1997 portant mesures urgentes à caractère fiscal et financier <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000185577&categorieLien=id>

only firms affiliated to the IS regim. Due to data limitation on share capital's ownership -on which the second eligibility criteria applied- we make the conservative choice to exclude all firms that belong to a conglomerate.

Information about conglomerate membership is collected in another dataset called *Enquête sur les liaisons financières* (LiFi). This dataset is available for all the 1995-2000 period. It is a production of the French census bureau (INSEE). The bureau surveys every year all firms with sales above MEuros 60, equity portfolio above MEuros 1.2 or with more than 500 employees. Moreover the institute includes in its sample firms that were in the dataset the preceding year or firms that belong to foreign firms.

We make two other restrictions to our main sample. First, we clean the dataset and remove extreme values for input shares. We exclude observations that have input shares above 1 (resp. 10 for capital share). We also exclude observations for which the sum $.1 \cdot K/Y + L/Y + M/Y$ is larger than 2. Second, we exclude firms that report negative values of input.

Table 1 presents the descriptive statistics. The average output level is kFrancs 32941, the average level of turnover is kFrancs 38637, material capital and labor's average values are respectively kFrancs 11042, kFrancs 10206 and kFrancs 6966. The average number of employees is 50. The average profit of firms in our sample is kFrancs 1436.

Firms in our sample are on average larger than the average French firm. Output, turnover, labor, capital and material levels are on average more than two times larger than corresponding average levels in the universe of firms, after exclusion of firms with negative values of inputs.

5 Distortions in the firm size distribution over time

5.1 Cross section

In this section we analyze the distortion in firms' sales distribution from a static point of view. We test whether at the threshold on turnover there is a discontinuity in the probability distribution function of turnover. We run (McCrary, 2008) test for manipulation for each year between 1995 and 2000. Cattaneo and Ma (2016) allows to report p-values for the estimate of the discontinuity. Given the incentives for firms to be below the M Francs 50 threshold when the reform was implemented, we expect additional mass to appear below the threshold between 1997 and 1999.

Figure 2 reports the density plots of turnover density for each year between 1995 and 2000. We clearly see a discontinuity in the density for years 1997, 1998 and 1999. In 1996, there does not seem to be any anticipation effect. The 1996 discontinuity, if any, does not seem significant as the confidence interval around the density plots on the right and on the left of the threshold overlap.

McCrary (2008) allows to quantify the discontinuity and provides the log difference in height of the density on the right and the density on the left. Estimates are reported in the first column of table 2 and the corresponding standard errors in the second. Before the reform, in 1995 and 1996 the discontinuity is estimated at .058 (s.e. .092) and .001 (s.e. .076) respectively. When the policy change was implemented, from 1997 to 1999 the discontinuity

is much larger with point estimates respectively equal to -.331 (s.e. .083), -.598 (s.e. .089) and -.714 (s.e. .114). These estimates seem significant. In 2000 the discontinuity vanishes but does not disappear completely. The point estimate is -.143 (s.e. .110).

In the last column of table 2 we report the p-values associated with the Cattaneo and Ma (2016) test. Those p-values confirm that the discontinuity before the reform is not significant (the p-values are respectively equal to .96 and .69). On the contrary on the three years where the policy change was implemented the p-values are all less than .001, thus the results are significant at the 1% level. The discontinuity we identified in 2000 does not seem to be significant at the 10% level, the corresponding p-value being .128.

Bunching estimation

To further assess the amount of bunching, we conduct the usual estimation of bunching that consists in estimating the number of firms that are in excess just below threshold. To estimate the number of firms that are in excess, we predict the number of firms that should be there had there been no manipulation with a fifth degree polynomial.

To be precise we estimate the following estimation:

$$c_j = \alpha + \sum_{i=1}^5 \beta_i \cdot (z_j)^i + \sum_{i=z_L}^{z_U} \gamma_i \cdot \mathbb{1}[z_j = i] + \epsilon_i \quad (4)$$

where c_j counts the number of firms in bin j . z_j is turnover level in bin j . Given that the variable of interest counts the number of firms per year, the natural choice for the estimation is to rely on a Poisson regression. β_i is the coefficient of order i of the fifth degree polynomial in turnover. γ_i identifies the excess or lack of firms in bin i compared to the counterfactual estimated with the polynomial. z_L is the beginning of the manipulation region and z_U its end. We determine z_L by eyeballing the distribution and z_U is determined such that excess bunching, i.e. the sum of firms in excess below threshold in the manipulation region equals missing mass, i.e. the sum of firms that are missing compared to the counterfactual above threshold in the manipulation region.

Formally we determine z_U as the smallest turnover level such that

$$\hat{M} = \sum_{i=z_T}^{z_U} \hat{c}_j^{cf} - c_j = \sum_{i=z_L}^{z_T-1} c_j - \hat{c}_j^{cf} = \hat{B} \quad (5)$$

Where z_T is turnover level at the threshold, i.e. MFrancs 50. The number of firms per bin in the counterfactual distribution is determined from

$$\hat{c}_j^{cf} = \alpha + \sum_{i=1}^5 \beta_i \cdot (z_j)^i. \quad (6)$$

To normalize the amount of bunching we estimate the average bunching b_{av} that is defined as the ratio of excess bunching over mean density in the manipulation region below threshold. Empirically we define it as:

$$\hat{b}_{av} = \frac{\hat{B}}{\frac{1}{2} \sum_{i=z_L}^{z_T-1} \hat{c}_j^{cf}} \quad (7)$$

Table 3 presents the results of these estimation strategies. We see that bunching level varies from year to year. There are 93 firms bunching in 1997, 191 in 1998 and 83 in 1999. The increase in 1998 might be due to an adaptation effect, firms might not know about the measure in 1997 and learn and adapt in 1998. There may also be some frictions that prevent firms to adapt their turnover on a year-to-year basis. The decrease in bunching in 1999 is probably due to the lower incentives in 1999. The contribution for firms with turnover above MFrancs 50 was decreased to 10 % of the taxes they pay on profit. This represents a one third decrease in the strength of the incentives.

Average bunching follows a similar pattern. It is equal to 1.534 (s.e. 0.614) in 1997, 3.033 (s.e. 0.539) in 1998 and 2.038 (s.e. 0.454) in 1999. It is remarkable that average bunching decreases by one third in 1999 when the strength of the incentive itself decreases by one third. We estimate standard errors bootstrapping the procedure that gives excess bunching and missing mass. We estimate significance by computing the corresponding t-stat. We conclude that all point estimates are significant at the 1 % level. The magnitude of the point estimates we find is within the range of other estimates in the literature. For instance Kleven and Waseem (2013) find average bunching between 5.52 and 0.3, Best et al. (2015) find average bunching between 2 and 4 while Almunia and Lopez-Rodriguez (2017) find average bunching around 0.09.

5.2 Disentangling turnover manipulation and firm creation

The accumulation of density mass below the threshold simultaneous to policy change that we have observed in the previous subsection can be either due to firms manipulating their level of turnover or to the apparition of new firms right below the threshold. Firms that decide to manipulate might indeed split into two in order to avoid the tax contribution. The use of the latter was actually limited by the institutional safeguards. Firms that were eligible to avoid the tax reduction had to have 75 % of their share capital owned by physical people and all of their share capital paid up, which limited firms' ability to split and avoid the tax contribution.

In order to make sure the safeguards were efficient, we reproduce the previous analysis on firms that remain in the sample every years from 1995 to 2000. If bunching was due solely to splits of existing firms into firms of similar size, the density plots would not exhibit any discontinuity once we restrict the analysis to firms that already existed before the introduction of the policy.

To test this, we restrict the sample to a balanced panel of firms that appear every year between 1995 and 2000. Density plots are reported in figure 3. The corresponding discontinuity estimates from (McCrary, 2008) manipulation test are reported in column 1 of Table 2. They are respectively equal for years 1995 to 2000 to .274 (s.e. .210), .110 (s.e. .184), -.647 (s.e. .193), -.919 (s.e. .201), -.617 (s.e. .155) and -.056 (s.e. .150). We find the same pattern as before: all discontinuities are negative during the implementation of the policy and positive or close to zero before and after the policy. They are of comparable magnitude to or larger than previous discontinuities. The corresponding standard errors suggest that the discontinuities we observe from 1997 to 1999 are significant and that those we observe in 1995, 1996 and 2000 are not significant.

(Cattaneo and Ma, 2016)'s test provides p-values for the discontinuity estimates. They

are reported in column 3 of Table 2. We find that the p-values are all larger than .2 in 1995, 1996 and 2000. On the contrary the p-values are smaller than .05 in 1997, 1998 and 1999, i.e. the observed discontinuities are significant at the 5% level between 1997 and 1999.

Discontinuities under the policy, when we restrict the analysis to firms that already existed before the introduction of the policy are not significantly smaller than the ones on all the firms. We therefore conclude that the independence criteria was efficient and firms did not split up in order to avoid from the exceptional contribution.

Bunching estimation

We reproduce the quantification of excess bunching mentioned in the previous section. Results are reported in Table 3. There are respectively 53, 74 and 45 firms that bunch in 1997, 1998 and 1999. This is less than in the dataset with all firms, but there is also less firms in this new dataset where we restrict to firms that were present every year between 1995 and 2000. To present estimates that can be compared, we estimate the average bunching estimates. We find average bunching of 2.346 in 1997, average bunching of 3.018 in 1998 and average bunching of 2.946 in 1999.

The average bunching estimates are, if anything, larger than the ones we estimated on the comprehensive dataset. This confirms that the bunching we observe is not driven by a split of existing firms and that safeguards that prevented such strategy were efficient.

5.3 Using past years as counterfactual

Another way to investigate the existence of distortions in the firm-size distribution over time is to compare distributions of firms across times around the threshold. The method described in the previous section consists in starting from the usual bunching estimation procedure (Kleven and Waseem, 2013; Raj Chetty and Pistaferri, 2011). This technique compares a counterfactual distribution built with information outside the manipulation region to the actual distribution. Here we use distributions in years during which there is no incentive to bunch as counterfactual distributions and compare them to the distribution under the policy.

To be precise we estimate the following equation:

$$c_{jt} = \alpha \cdot Post_t + \sum_{i=z_{-m+1}}^{z_n} \beta_i \cdot \mathbb{1}[z_j \in [i, i + 2]] \mathbb{1}[i \equiv 0[3]] + \sum_{i=z_{-m+1}}^{z_n} \gamma_i \cdot \mathbb{1}[z_j \in [i, i + 2]] \mathbb{1}[i \equiv 0[3]] * Post_t + \epsilon_{it} \quad (8)$$

where m is the number of bins below cutoff and n the number above cutoff. c_{jt} counts the number of firms in bin j in year t . $Post_t$ refers to the second period of analysis, i.e. years 1997, 1998 when the 1995-98 period is under study and year 2000 when the 1998-2000 period is under study. z_j is turnover level in bin j . Given that the variable of interest counts a number of firms per bin, the natural choice for the estimation is to rely on poisson regression.

Even though we pull multiple years with treatment and one or two years as counterfactual, estimations would be subject to too much noise if we put a dummy ($\mathbb{1}[z_j = i]$) per bin.

There would indeed then be at most only two observations per variable. Putting a dummy for three bins ($\mathbb{1}[z_j \in [i, i + 2]]\mathbb{1}[i \equiv 0[3]]$) on the contrary multiplies by three the number of observations per variables. This makes the estimation much more precise. Our choice of three bins is motivated by a search for granularity. Adding more observation per dummy variable would certainly increase the precision but we would lose in granularity, which is essential for instance to assess the width of the manipulation region.

We estimate the regression using either 1995 and 1996 as counterfactual and including years 1997 and 1998 in the sample, or year 2000 as counterfactual and including years 1998-1999 as counterfactual. For the sake of clarity we only report the coefficient of the interaction terms γ_i around the threshold. Given the Poisson estimation we report exponentiated coefficients.

The top panel of figure 4 compares years 1997 and 1998 to years 1995 and 1996. It shows that compared to the 1995 and 1996 distributions, there is excess bunching in the 1997-1998 distribution. We include 1996 as counterfactual year as there is no evidence of anticipation effect in the cross-section results and because we were unable to find any evidence of announcement of the reform prior to year 1997.² Besides the concurrent but smaller measure that reduced marginal rate of imposition for firms with turnover below the same threshold in 1996 was a much smaller incentives as only profits below kEuros 38 were subject to the tax cut. There are three coefficients positive and significant below threshold and three coefficients negative and significant above threshold. This suggests that above threshold, because of the reform there was significantly less firms than there would have been had there been no reform. On the contrary there is significantly more firms just below threshold under the policy than there were before.

The bottom part of figure 4 compares the distributions of turnover in 1998 and 1999 to the 2000 distribution. We see that bunching disappeared in 2000. There are two negative and significant coefficients below threshold. This means that there was significantly less firms below threshold in 2000 than there was under the policy. Similarly, there is one coefficient positive and significant above threshold, suggesting that there were significantly more firms above threshold in 2000 than there were under the policy.

Bunching estimation

To quantify the size of the distortion we rely on an estimation procedure that looks similar to the one presented in section 5.3. We pool 4 consecutive years to increase statistical precision and because we can gather years by the actual incentive level firms face. In particular we gather years 1995 and 1996 where firms face no incentives for manipulation and years 1997 and 1998 where firms above threshold paid 15% more taxes on profit.

²On November 27 1995, France’s Prime Minister Alain Juppé announced a “plan PME pour la France”, i.e. a package of reforms that aimed at alleviating credit constraints for SME, fostering their ability to accumulate capital and to settle in urban areas and finally reducing taxes they pay. There is however no evidence that the exceptional temporary contribution was part of this announcement

We therefore estimate the following equation:

$$c_{jt} = \alpha \cdot Post_t + \sum_{i=z_{-m+1}}^{z_n} \beta_i \cdot \mathbb{1}[z_j \in [i, i + 2]] \mathbb{1}[i \equiv 0[3]] + \sum_{i=z_{-m+1}}^{z_n} \gamma_i \cdot \mathbb{1}[z_j \in [i, i + 2]] \mathbb{1}[i \equiv 0[3]] * Post_t + \epsilon_{it} \quad (9)$$

where m is the number of bins below cutoff and n the number above cutoff. c_{jt} counts the number of firms in bin j in year t . $Post_t$ indicates that year t is one of the two first years where firms faced incentives for manipulation, namely 1997 or 1998. z_j is the turnover level in bin j . Given that the variable of interest counts the number of firms per year, the natural choice for the estimation is to rely again on poisson regression. We run the regression for a bin size of kEuros 150 and report the results in figure 4.

We report only the coefficients of the interacted terms (γ_i). The coefficients of the interacted terms allow us to estimate missing mass and excess bunching. They indicate how many additional firms there is per bin below the cutoff. To be precise the product of the exponential of the coefficients of the interacted term γ_i and of the dummy variable for bin i tells us by how much we must multiply the number of firms in the excluded bin to estimate the number of firms that are in bin i . Subtracting the number of firms that were in this bin in years during which there was no incentives to bunch gives the number of firms that bunch in this bin. The sum of firms that bunch in each bin below the threshold where the interacted coefficients are significant gives us the amount of excess bunching. We similarly estimate the number of missing firms in bins above the threshold to obtain the missing mass.

We can describe the size of the distortion with formal expressions for excess bunching (\hat{B}) and missing mass (\hat{M}):

$$\begin{aligned} \hat{M} &= \sum_{i=z_T}^{z_U} \hat{c}_{iControl} - \hat{c}_{iTreat} \\ \hat{B} &= \sum_{i=z_L}^{z_T-1} \hat{c}_{iTreat} - \hat{c}_{iControl} \end{aligned} \quad (10)$$

(2) where: \hat{c}_{iTreat} refers to the average predicted number of firms per year within bin i during the period of treatment (1997-99) and $\hat{c}_{iControl}$ to the average predicted number of firms per year within this bin during control years.

We develop a statistical test to estimate the value of z_L . It is the turnover level of the group of bins that precedes the bunch of ones with significant coefficients below threshold.

³ In our context, 47600 is therefore the lower end of the valley. This is an advantage of our

³We do not take the smaller level of turnover of this bunch of coefficients because there might be bunchers in the group of bins that precedes it and missing them may misinform us about the characteristics of the bunchers if those that we miss have particular characteristics that would drive a change of the results. On the contrary including non bunchers in the manipulation region does not affect the estimated numbers of bunchers since the coefficient of the interaction term on this group of bins is close to zero. It does not either affect the determination of bunchers characteristics that compare the characteristics of firms below threshold in the manipulating region within years with or without incentives to bunch. Firms below threshold that are not bunching should indeed have the same characteristics in years with and in years without incentives to bunch.

estimation strategy compared to usual techniques of bunching that eyeball the lower end of the manipulation region.

To back out the upper end of the valley we use prediction from the estimations. We follow Kleven and Waseem (2013) and pin down Z_U by the equality $\hat{M} = \hat{B}$. Results of this procedure are presented in Table 3 columns (1) and (2). We estimate $Z_U = kFrancs$ 51800. The manipulation region has the surprising feature to be symmetric around the threshold. The literature has often described a manipulation region that is skewed towards the region where there is no bunching. In our case this might be due to the fact that the manipulation region is overall quite small.

There is however a neat equality in the number of firms that bunch and the number of firms that are missing above the threshold. We find that in the two years there are 225 firms that bunch and 227 firms that are missing above the threshold.

Similarly as we did for the cross-section estimation, we are able to estimate average bunching, dividing by the average density of turnover in the manipulation region below threshold. We find an average bunching of value 3.058 that is slightly larger than the mean of average bunching we found with the cross section estimations for years 1997 and 1998. However it falls within the confidence interval.

5.4 Adjustment costs and share of non-optimizers

In this section we ask whether there is a significant share of non-optimizers, i.e. firms that do not decide to decrease their turnover to be below threshold even though this could help them increase their profits. We first present the distribution of firms for those that have really no interest not to bunch, i.e. those with profit within the top tercile. The distribution is represented in figure 6. We see that even for these firms the distribution exhibits a significant number of firms that are not bunching just above the threshold. In other terms, the empirical distribution is not equal to 0 just above the threshold.

In the literature the presence of firms, just above the threshold of a notch has been interpreted as the evidence for adjustment cost (Almunia and Lopez-Rodriguez, 2017). It is possible to quantify the number of firms that do not bunch because of these adjustment costs. In our setting, firms that have 0 or negative profit have no incentives to bunch either. As a result our estimate captures both the share of firms that have too high adjustment costs and the share of firms that have no incentives to bunch.

Following Almunia and Lopez-Rodriguez (2017) we estimate the share α of non bunchers by $\alpha = \frac{\int_{z_T}^{z_D} g(x) dx}{\int_{z_T}^{z_D} g_0(x) dx}$. Where z_D is the turnover level above which it might not be profitable for firms to bunch, g is the density of turnover and g_0 is the counterfactual density. In our cases, as z_D depends on the profit level we cannot determine it computationally. We rather compute the values of α for different plausible values of z_D and show that they do not differ significantly. Given the estimation setting, where there are only three groups of bins in the manipulation region (including the group that contains the threshold) we can compute alpha for at most three values of z_D . We decide not to compute α for $z_D = 50600$ as this estimation would rely on only one coefficient. We however find comparable values for α for both estimations with either $z_D = 51200$ or $z_D = 51800$. As presented in table 4, we find respectively $\alpha = 0.849$ and $\alpha = 0.872$ for those two values of z_D . These values

are also within the range of the values found in the literature. For instance (Kleven and Waseem, 2013) find α between 0.50 and 0.91 depending on the population and the notch they consider, similarly (Almunia and Lopez-Rodriguez, 2017) find α between 0.47 and 0.98 depending on the sector and the year they consider.

In Table 4 column (2) we report adjusted bunching. This is average bunching for the share of firms that optimize. It is expressed as :

$$b_{adj} = b_{av}/(1 - \alpha) \tag{11}$$

We compute it empirically for each of the two values we have estimated for α . We find a relatively large amount of adjusted bunching, of average 22.071.

In the next section we propose to characterize the firms that bunch and to compare them with the rest of firms that were eligible to bunching.

6 Bunchers' characteristics

6.1 Simple inference from bunching estimates

In this section we compare turnover distributions of firms in different subgroups. In particular we first focus on firms that have high versus low profits, then on firms with high versus low capital adjustment cost. Finally we compare firms with different production function's characteristics, more precisely firms with different level of output elasticity with respect to the different inputs of production.

Figure 8 presents turnover distribution for firms in the top (resp. bottom) tercile of profit in 1997, 1998 and 1999 the three years for which we identified bunching. We see that bunching is only present among high profit firms, i.e. those that face the highest incentives. Actually the mean profit level of firms within the bottom tercile of profit is negative. As a result the average firm in this tercile has no gain from manipulation. The maximum level of profit for firms within this tercile is kFrancs 37. The maximum gain they would therefore expect from manipulation is Francs 1,831 i.e. around Euros 280. We further estimate bunching for firms within the top tercile. We find average bunching of 0.490, 0.879 and 0.653 for years 1997, 1998 and 1999. For firms within the bottom tercile it is impossible to determine z_L and as a consequence to estimate excess bunching. As a consequence we posit bunching is 0. We conclude from this difference of bunching patterns that bunchers are more likely to belong to the top tercile of profits than to the bottom tercile of profit.

Figure 9 presents turnover distribution for firms in the top and the bottom terciles of capital adjustment cost. We see that bunching appears as of 1997 for firms with low adjustment cost while it only appears in 1998 for firms with high adjustment costs. In 1999 bunching seems also larger for firms with low adjustment costs than for firms with high adjustment costs. We find similar average bunching for firms with high adjustment costs than for firms with low adjustment costs in 1998 (of respectively 0.722 and 0.616). In 1999 we find average bunching of 0.719 for firms with low adjustment cost and of 0.376 for firms with high adjustment costs. Here again we posit that bunching is 0 for firms in the high tercile of capital adjustment cost in 1997 where it is impossible to determine z_L . This suggests that in 1997 the bunchers were more likely to belong to the bottom tercile of capital adjustment

cost than to the top tercile. In 1998 bunchers seem to be as likely to belong to the top or the bottom tercile of capital adjustment cost. Similarly in 1999 bunchers were also more likely to belong to the bottom tercile of capital adjustment cost than to the top tercile.

Figure 10 presents turnover distribution for firms in the top and bottom terciles of output elasticity with respect to materials. We see that in 1997 there seems to be some bunching only within firms of the top tercile and not within firms of the bottom tercile. Here again we posit bunching is 0 for the bottom tercile in 1997. In the top tercile we find bunching estimate of 1.056, statistically different from 0. In 1998 it is difficult to conclude from any difference in the two distributions of the top and bottom terciles based only on this graphical evidence. Bunching estimates are respectively of 0.499 for the top tercile and 0.792 for the bottom tercile, though not statistically different from each other. In 1999 bunching estimates are very similar for firms of the top and firms of the bottom terciles. They are respectively equal to 0.717 and 0.710 and not statistically different from each others. We conclude from there that in 1997 bunchers are more likely to belong to the top tercile of output elasticity with respect to material. This is not the case in 1998 and 1999.

Figure 11 presents turnover distribution for firms in the top and bottom terciles of output elasticity with respect to labor. We see that in 1997, firms in the top tercile do not seem to bunch much while those in the bottom tercile do seem to bunch slightly more.⁴ In 1998 it is difficult to determine if there is any difference in the two distributions of turnover for the top and bottom terciles. In 1999, similarly to 1997, there seems to be slightly more bunching among firms in the bottom tercile than among firms in the top tercile.⁵ Bunching estimates are very similar for top and bottom terciles. In 1997 we find average bunching of 0.692 and 0.517 for the top and bottom terciles respectively. In 1998 we find average bunching of 0.494 and 0.620 for the top and bottom terciles respectively. In 1999 we find average bunching of 0.532 and 0.767 for top and bottom terciles respectively. We conclude that bunchers were always as likely to belong to the top and bottom tercile of output elasticity with respect to labor.

Figure 12 presents turnover distribution for firms in the top and bottom terciles of output elasticity with respect to capital. It shows patterns that are similar to the ones of figure 11. We see that firms within the top tercile do not seem to bunch in contrary to firms in the bottom tercile. In 1998 the discontinuity seems higher among firms in the top tercile than among firms in the bottom tercile. However it is difficult to infer from this graphical evidence whether this translates in difference in bunching amount. In 1999 again the discontinuity seems higher among firms in the top tercile than among firms in the bottom tercile. However standard error in the bottom tercile are larger and this impression might be driven by an outlier that is significantly lower than the other points below threshold. Bunching estimates confirm this visual analysis. In 1997, average bunching is equal to 0.364 for firms in the top tercile of output elasticity with respect to capital and to 0.895 for firms in the bottom tercile. In 1998 and 1999 there is less difference between the two subgroups. In 1998 average bunching is equal to 0.787 for firms in the top tercile and to 0.498 for firms in the bottom tercile. In 1999 average bunching is equal 0.596 and to 0.607 for firms in the top and bottom

⁴Even though it is difficult to say whether this visual impression is not driven by the larger size of the standard errors for firms in the bottom tercile than for firms in the top tercile.

⁵Ibid.

tercile respectively. We can conclude from there that bunchers are more likely to have small output elasticity with respect to capital in 1997. In the two other years it is difficult to say anything from this graphical evidence and the average bunching are closer to each other in the two subgroups, suggesting that bunchers are as likely to come from any of the two subgroups.

6.2 Methodology for a quantification

Using predictions as counterfactual

We then analyze more precisely and quantify the characteristics of the bunchers. In that purpose we follow Person and Diamond (2016). Their framework considers bunchers as compliers that do respond to the incentives and adjust their turnover to level just below the threshold, which they would not have done had they faced no incentives. They show that the characteristics of the bunchers ($Y^{compliers}$) can be expressed as:

$$\begin{aligned} \bar{Y}^{compliers} = & 0.5 * \left(\frac{N_{down}^{tot}}{N_{down}^{tot} - N_{down}} * \bar{Y}^{down_all} - \frac{N_{down}}{N_{down}^{tot} - N_{down}} * \bar{Y}^{down} \right) + \\ & 0.5 * \left(\frac{N_{up}}{N_{up} - N_{up}^{tot}} * \bar{Y}^{up} - \frac{N_{up}^{tot}}{N_{up} - N_{up}^{tot}} * \bar{Y}^{up_all} \right) \end{aligned} \quad (12)$$

Where $\bar{Y}^{compliers}$ is defined as the average of the mean values of Y for firms that are "missing" above the threshold and for firms that are bunching. \bar{Y}^{down_all} (resp. \bar{Y}^{up_all}) is the average of characteristic Y for firms that are in the manipulation region below (resp. above) threshold. and \bar{Y}^{down} (resp. \bar{Y}^{up}) is the average of Y for firms that would have been in the manipulation region below threshold (resp. above threshold) had there been no manipulation.

\bar{Y}^{down} and \bar{Y}^{up} are obtained by regressing Y on a polynomial of turnover of order 5 for firms outside the manipulation region and predicting levels of Y within the manipulation region by extrapolating this relationship.

N_{up}^{tot} (resp. N_{down}^{tot}) is the number of firms that fall into the manipulating region above (resp. below) the threshold. They are the number of never takers and the sum of the number of always takers and of compliers. N_{up} (resp. N_{down}) is the number of firms that would have fallen into the manipulating region above (resp. below) the threshold had there been no manipulation. It is the sum of the number of never takers and the number of compliers (resp. the number of always takers).

Using past years as counterfactual

In this paragraph, we extend Person and Diamond (2016) technique and adapt it to this setting where counterfactual distributions are obtained from previous years. The expression of bunchers' characteristics become:

$$\begin{aligned} \bar{Y}^{compliers} = & 0.5 * \left(\frac{N_{down}^{tot}}{N_{down}^{tot} - N_{down}} * \bar{Y}_{Treatment}^{down} - \frac{N_{down}}{N_{down}^{tot} - N_{down}} * \bar{Y}_{Control}^{down} \right) + \\ & 0.5 * \left(\frac{N_{up}}{N_{up} - N_{up}^{tot}} * \bar{Y}_{Control}^{up} - \frac{N_{up}^{tot}}{N_{up} - N_{up}^{tot}} * \bar{Y}_{Treatment}^{up} \right) \end{aligned} \quad (13)$$

Where $\bar{Y}_{Treatment}^{down}$ is the average, when the policy is in place, of the mean values of Y for bunching firms ("compliers") and firms that are naturally present below the threshold absent the policy ("always-takers"). It is obtained by estimating the mean of the characteristic Y of interest in the manipulation region below threshold over years with incentives to bunch (treatment years). $\bar{Y}_{control}^{down}$ is the mean of the characteristic Y of interest in the manipulating region below the threshold over years with no incentives to bunch (control years). It is therefore the mean of the characteristics of interest for the "always takers".

Similarly $\bar{Y}_{Treatment}^{up}$ is the mean of the characteristic of interest Y in the manipulating region above threshold over years with incentives to bunch. It is the average level of Y for never takers. $\bar{Y}_{Control}^{up}$ is the mean of the characteristic of interest Y in the manipulating region above the threshold in control years. It is therefore the average of the means of Y for never takers and compliers.

N_{up}^{tot} (resp. N_{down}^{tot}) is the number of firms that fall into the manipulating region above (resp. below) the threshold in treatment years. They are the number of never takers and the sum of the number of always takers and of compliers. N_{up} (resp. N_{down}) is the number of firms that fall into the manipulating region above (resp. below) the threshold in control years. It is the sum of the number of never takers and the number of compliers (resp. the number of always takers).

In contrast to Person and Diamond (2016) who use a predicted counterfactual instead of the variables we observe during years with no incentives to bunch our analysis might be subject to a change in the variables due to an underlying trend in their evolution. As a result we are interested in the de-trended variables defined as $\tilde{Y} = Y - \bar{Y}_{below}$ where \bar{Y}_{below} is the average of Y in the region neighboring the manipulation region below the threshold (Turnover $\in [45000-47600]$) and during years with no incentives to bunch. We use the region below the part of the manipulation region below the threshold because we are sure there is no compliers in this region in years with no incentives to bunch.

We are interested to compare the characteristics of the compliers to the characteristics of the firms that are eligible to bunching, i.e. those that were in the manipulation region, above the threshold, in years with no policy. With Person and Diamond (2016) notation this means we are interested in estimating:

$$\Delta \tilde{Y} = \tilde{Y}^{compliers} - \tilde{Y}_{Control}^{up} \quad (14)$$

We estimate this raw difference of means as well as a difference of means net of sector and region fixed effects. Sector fixed effects are indicators of the 16-sector French classification of industries. In practice, we estimate the following equation, with and without fixed effects:

$$\tilde{Y}_i = \beta \cdot compliers_i + \nu_r + \mu_s + \epsilon_i \quad (15)$$

The equation is estimated on the set of firms eligible to bunching. The coefficient of interest is the β that directly gives us the difference of means between the two populations of interest. We estimate standard errors by bootstrapping the regression.

6.3 Results

Using past years as counterfactual

In this section we describe the results obtained by estimating equation 15. In accordance with our theoretical framework, we are interested in testing whether bunchers are the firms that face the highest incentives, i.e. the most profitable ones and whether they belong to sectors where adjustment costs are the lowest.

To test the former, we consider before tax profit. After tax profit is indeed affected by the policy, and because of the additional tax burden for firms above the threshold, it would not be surprising that bunchers have larger after tax profit. Table 5 column (9) reveals that bunchers are 2.8% more likely to belong to the top tercile of profit. This difference represents 8.5 % of the share of firms in the top tercile of profit (33%). It is significant at the 5 % level. Column (10) confirms the robustness of the result. The inclusion of fixed effects does not change significantly the point estimates, which stays significant at the 5% level.

As proxy of adjustment costs, we consider output elasticity with respect to the input of production obtained following (Levinsohn and Petrin, 2003) estimation and capital adjustment cost as measured by (Asker, Collard-Wexler, and Loecker, 2014) technique. These different proxies of adjustment costs give all similar results: the higher the adjustment costs the less likely the bunching. Table 5 column (1) (resp. 3) reveals that bunchers are 4.5 % (resp. 5.9 %) less (resp. more) likely to belong to a sector with a large output elasticity with respect to capital (resp. materials). This represents 13.6 % (resp. 17.9%) of the share of firms in the top tercile of output elasticity with respect to capital (resp. materials) (33%). These results are inline with the common assumption that capital is more fixed and that materials are more flexible. Both of these differences are significant at the 1% level. Columns (2) and (6) confirm that these results are robust to the inclusion of region and industry fixed effects⁶. We do not observe any difference between bunchers and eligible firms with respect to their probability to belong to the highest tercile of output elasticity with respect to labor, whether we include fixed effects or not. Labor is indeed often considered as an intermediate input, neither completely fixed nor completely flexible. It is therefore consistent to have it in between capital and material. Even though the point estimate is not significant it is slightly negative. This would be consistent with the idea that labor is rather fixed than flexible. This is in line with the highly regulated labor market in France that induces hiring and separation costs (Abowd and Kramarz, 2003; Kramarz and Michaud, 2010).

Table 5 column (7) confirms that in average bunchers have lower capital adjustment costs. They are 3.17% more likely to be part of the lowest tercile of the adjustment cost of capital, which corresponds to about 10% of the mean value of the dummy that indicates that a firm belong to this lowest tercile. This estimate is significant at the 5 % level. As neither

⁶The granularity of the sectors in the fixed effects is larger than the granularity at which we compute the elasticities of output, which is also the granularity at which we compute the adjustment cost or the one of the datasets with external finance dependence and import demand elasticity

this estimate nor the other ones change significantly when we include region and sector fixed effects(8), we stop reporting estimates without fixed effects. As capital adjustment costs are the larger input adjustment cost, we conclude that the lower the adjustment costs, the easier it is for firms to bunch.

Using predictions as counterfactual

In this section we present the results obtained with (Person and Diamond, 2016) technique. We compare in 1997 the characteristics of the compliers to the characteristics of firms that are eligible to bunching i.e. those that would have been in the manipulation region above threshold had there been no manipulation. To do this we estimate equation 15 using Y as dependent variable (not \tilde{Y}).

Results are presented in Table 7. We see first that manipulating firms are 15 % more likely to have low capital adjustment cost. The coefficient on *compliers* is significant at the 10 % level. They are also 22 % more likely to have large profits. The coefficient on *compliers* is significant at the 1 % level. They also seem less likely to have large output elasticity with respect to capital, though the coefficient is not significant. They are however 15 % more likely to have large output elasticity with respect to material. The coefficient on compliers is positive and significant at the 10 % level.

Overall these results confirm that bunchers are the firms that face the highest incentives to bunch and those that face the lowest adjustment costs. They are in line with the other estimation strategy that uses past years as counterfactual. The exact magnitude of the coefficients however differ for the estimation for profits and for adjustment costs. The coefficients tend to be larger when we use prediction as counterfactual than when we use past years. Using predictions as counterfactual allows us to focus on a specific year (1997) on the contrary using past years as counterfactual provides estimation for years 1997 and 1998. The analysis of the heterogeneity of bunching level revealed that the highest differences were in 1997. This might explain why we observe such differences with the two estimation strategy.

To confirm this we reproduce the analysis using predictions as counterfactual for year 1998 (Table 8). We show that on this year differences between bunchers and non bunchers are less important, suggesting that the difference between the two estimation strategies might be driven by the different time frame they are considering.

7 Bunching mechanisms

7.1 Is it fraud?

Best et al. (2015) develop a technique to back out the existence of tax evasion. They analyze bunching in the Pakistani context, where firms are taxed either on profit or on turnover depending on their level of profits. When firms' profits are over the tax ratio $\frac{\tau_y}{\tau_\pi}$ they pay taxes on profit with a rate τ_π , while when their profits are below this threshold, they pay taxes on turnover (output) with a tax rate of τ_y . At the frontier between those two regimes, the real production incentives is rather continuous. However evasion opportunities are discontinuous: it is much easier to evade while reporting profit than while reporting turnover. As a result

they argue that bunching levels are linked to the existence of evasion. They develop a formal model to link output and evasion elasticities with bunching level. They show that :

$$\Delta\pi \simeq \frac{\tau_y^2}{\tau_\pi} \epsilon_y - \frac{d(\hat{c} - c)}{y} \quad (16)$$

Where $\Delta\pi$ "denotes the profit rate response by the marginal bunching firm", τ_y denotes the tax rate on output, τ_π the tax rate on profit, ϵ_y the output elasticity with respect to the effective tax rate $\hat{c} - c$ is the evasion response. The tax rate ratio, only comes from the feature of the threshold that it is determined by this ratio.

In our setting, there is solely a tax on profit, though the eligibility is determined based on turnover level. (Best et al., 2015) analysis is however informative to the extent that it shows how, when small real incentives translate into large bunching response, they can be attributed to fraud. Indeed, the amount of bunching they observe corresponds to so large output elasticities (between 15 and 135) that only non zero tax evasion can rationalize it.

In our setting, the incentives are also quite low (cf section 2). We are therefore interested to see whether bunching response correspond to plausible elasticity or not. We start by assuming zero tax evasion. We have recourse to the R package `bunchr` that estimates output elasticities in the context of bunching. We estimate output elasticity to be 0.16, which lies within the range of plausible values. Indeed (Best et al., 2015) take values between 0 and 1.5 to calibrate the evasion response. They also mention that Gruber and Rauh (2007) find an elasticity of corporate taxable income of 0.2. They also recall the link between output elasticity and the elasticity of corporate taxable income (ϵ_{CTI}):

$$\epsilon_{CTI} = \frac{\partial CTI / \partial y}{CTI / y} \epsilon_y \quad (17)$$

Our estimates would be consistent with Gruber and Rauh (2007) estimates as long as marginal taxable profits and average taxable profits are of similar magnitude, assuming that the US and France's elasticities of corporate taxable income are similar. As a result we do not need fraud to rationalize the amount of bunching we find.

7.2 Empirical strategy

In this section we describe a strategy to assess what firms do in order to bunch. More specifically we determine what they do because they are bunching and would not do if they were not bunching. We rely on (Person and Diamond, 2016) strategy that assesses the causal impact of manipulation. They develop an intent to treat estimator that compares the average value of the outcome of interest in the manipulation region, had there been no manipulation and the average value of the outcome of interest in the manipulation region. As they say, the difference between the two is due to the fact that compliers within the manipulation region manipulated.

In our setting, in order to compute the average value of the outcome of interest, had there been no manipulation, there is no need to predict the average value of the outcome of interest based on its relationship with the running variable outside the manipulation region. We can simply use the average value of the outcome of interest in years during which there was no

incentives to bunch as counterfactual. To make sure our estimate is not driven by temporal changes in the outcome of interest, we de trend the outcome of interest by subtracting from it the mean of the outcome of interest in the region neighboring the manipulation region (i.e. the region with turnover between 45000 and 47600 kFrancs). Our intent to treat estimate simplifies to :

$$ITT = E(\tilde{Y}|firms\ manipulate) - E(\tilde{Y}|firms\ don't\ manipulate) = E_{Manip}^{Treat}(\tilde{Y}) - E_{Manip}^{Control}(\tilde{Y}) \quad (18)$$

Where *Treat* in exponent indicates that the expectation is taken over observation during the treatment years, while *Control* indicates that the expectation is taken over observation in years during which there was no incentives to bunch. \tilde{Y} is the de-trended outcome of interest.

Using predictions as counterfactual

We also implement Person and Diamond (2016) strategy as a robustness test. Their strategy consists in predicting the value of the outcome of interest for firms in the manipulation region, had there been no manipulation. In that order, we regress the outcome of interest on a polynomial of turnover of order 5 for firms outside the manipulation region and predict levels of Y within the manipulation region by extrapolating this relationship. As a result the Intent To Treat estimator is estimated as :

$$ITT = E(\tilde{Y}|firms\ manipulate) - E(\tilde{Y}|firms\ don't\ manipulate) = E_{Manip}^{Observed}(\tilde{Y}) - E_{Manip}^{Predicted}(\tilde{Y}) \quad (19)$$

7.3 Results

Using past years as counterfactual

We present the results of the first estimation strategy in Table 9. We test several outcomes of interest.

First as a sanity check, we test for turnover level. The coefficient in column (1) is negative and significant at the 5 % level. This means that since they bunch, firms have on average lower turnover levels than they would have, if they did not bunch. To be precise they sell about kFrancs 88 less because they bunch.

Second we analyze whether this has consequences on their production level. As we can see in column (2), bunching decreases bunchers' production level, though the estimate is not significant. It seems to decrease it by about kFrancs 265 in average. This decrease in production is driven by a decrease in sold production (which is lowered by kFrancs 505). The point estimate for sold production is significant at the 10 % level. However the decrease in sold production is larger than the decrease in production because firms stock part of their production in inventories or capitalized production. As we can see the coefficient for change in inventories, is positive and significant at the 5 % level. On average, firms increase their inventories by kFrancs 152 to bunch. Similarly, firms invest their production

to further improve their production process. We can see that the coefficient for capitalized production is also positive and significant at the 1% level. It is however lower in magnitude than the coefficient for change in inventories, since firms in average increase their capitalized production by kFrancs 88.

We can interpret these result as evidence that it is costly for firms to adjust their production. They prefer to decrease their production by a lower level, even if the amount by which they decrease turnover suggest they should decrease more production, and they decide to stock some part of their production. Firms may also anticipate that the reform is only temporary and that they will soon be able to sell the additional production they have been stocking.

Using predictions as counterfactual

As a robustness test we present the results of the second estimation strategy that follows Person and Diamond (2016) in Table 10. We test the same outcome of interest.

It is here not possible to test for the consequences on turnover. Indeed the predicted outcomes are the result of a regression on a polynomial of turnover, that is therefore not identified when the outcome is turnover itself.

The estimate on production is again negative. It is in this context significant and takes value of -2870.6 kFrancs. We confirm that this decrease in production is driven by a decrease in sold production. The coefficient on column (2) is also negative and significant, in this context at the 1% level. Bunchers seem to decrease the value of their sold production by 3163kFrancs. Here also the decrease in sold production is larger than the decrease in production, because firms that manipulate increase their change in inventories. The coefficient for change in inventories is positive and significant at the 5% level. It is equal to 258 and not statistically significant from the coefficient we observe with the previous strategy. Finally the coefficient for capitalized production is also positive thought not significant here.

Overall the two estimation strategies give consistent results. The difference in their magnitude might be driven by the fact they cover different years. The panel strategy that uses past years as counterfactual covers the whole 1995-2000 period, the cross section that uses prediction as counterfactual is restricted to year 1997.

8 Conclusion

This paper presents the consequences of a transient fiscal policy on the distribution of firms with respect to their turnover. We identify subsequent bunching with average estimator falling within the range of the literature's estimates.

We further develop a technique to estimate bunching level, using past years as counterfactual distribution. This technique differs from usual techniques that rely on predictions of the counterfactual based on the link between turnover and density levels outside the manipulation region. Our estimates are in line with the latter technique. This technique gives a more precise estimation of the lower end of the manipulation region.

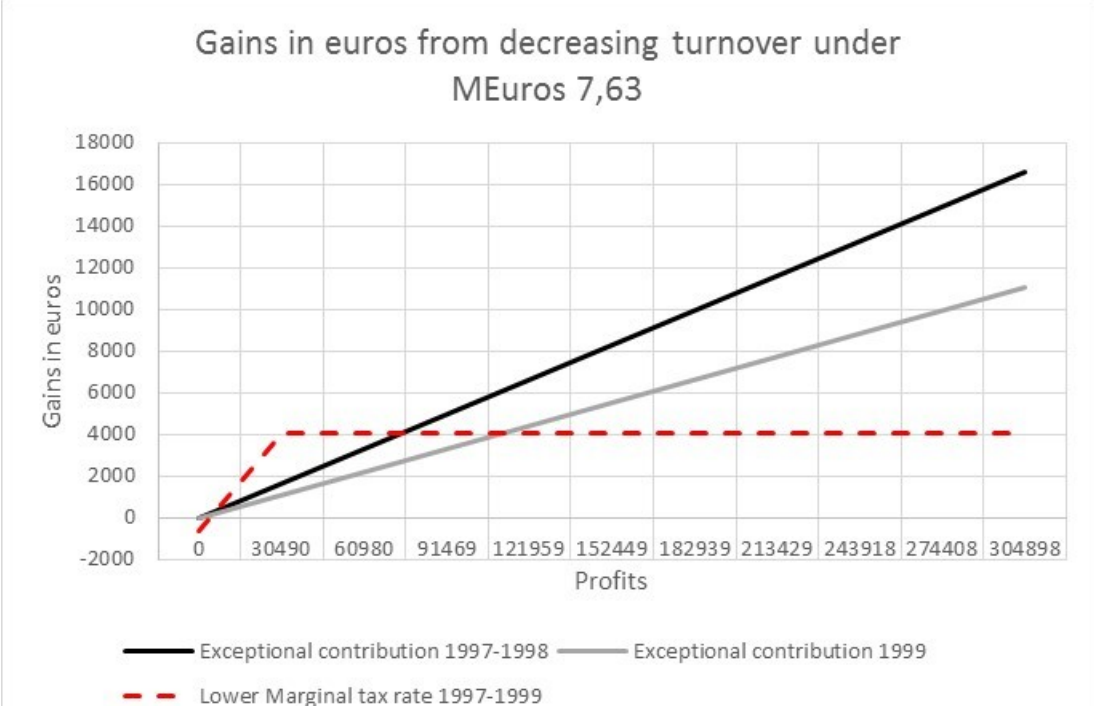
Following (Person and Diamond, 2016), we characterize the firms that bunch. They overall have lower adjustment costs, face higher incentives to bunch.

References

- Abowd, J., and F. Kramarz. 2003. "The Costs of Hiring and Separations." *Labour Economics* 10:499–530.
- Aghion, P., U. Akcigit, M. Lequien, and S. Stantcheva. 2017. "Do Entrepreneurship and Self-Employment Respond to Simpler Fiscal Incentives? Evidence from France." *Working Paper*, pp. .
- Almunia, M., and D. Lopez-Rodriguez. 2017. "Under the Radar: The Effects of Monitoring Firms on Tax Compliance." *American Economic Journal: Economic Policy*, pp. .
- Asker, J., A. Collard-Wexler, and J.D. Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy*, pp. .
- Bach, L. 2015. "Do Sophisticated Entrepreneurs Avoid Taxes More?" *Working paper*, pp. .
- Bertrand, M., A. Schoar, and D. Thesmar. 2015. "Banking Deregulation and Industry Structure: Evidence from the French Banking Reforms of 1985." *Journal of Political Economy*, pp. .
- Best, M., A. Brockmeyer, H. Kleven, J. Spinnewijn, and M. Waseem. 2015. "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy*, pp. .
- Brown, K.M. 2013. "The link between pensions and retirement timing: Lessons from California teachers." *Journal of Public Economics* 98:1–14.
- Cattaneo, J., and Ma. 2016. "Simple Local Regression Distribution Estimators with an Application to Manipulation Testing." *Working Paper*, pp. .
- Garicano, L., C. Lelarge, and J. Van Reenen. 2016. "Firm Size Distortions and the Productivity Distribution: Evidence from France." *American Economic Review* 106:3439–3479.
- Gruber, J., and J. Rauh. 2007. "How Elastic Is the Corporate Income Tax Base?" In J. R. Hines, A. J. Auerbach, and J. Slemrod, eds. *Taxing Corporate Income in the 21st Century*. Cambridge: Cambridge University Press.
- Kleven, H. 2016. "Bunching." *Annual Review of Economics*, pp. .
- Kleven, H., and M. Waseem. 2013. "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan." *Quarterly Journal of Economics*, pp. .
- Kramarz, F., and M.L. Michaud. 2010. "The shape of hiring and separation costs in France." *Labour economics* 17:27–37.
- Levinsohn, J., and A. Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of economic studies*, pp. .

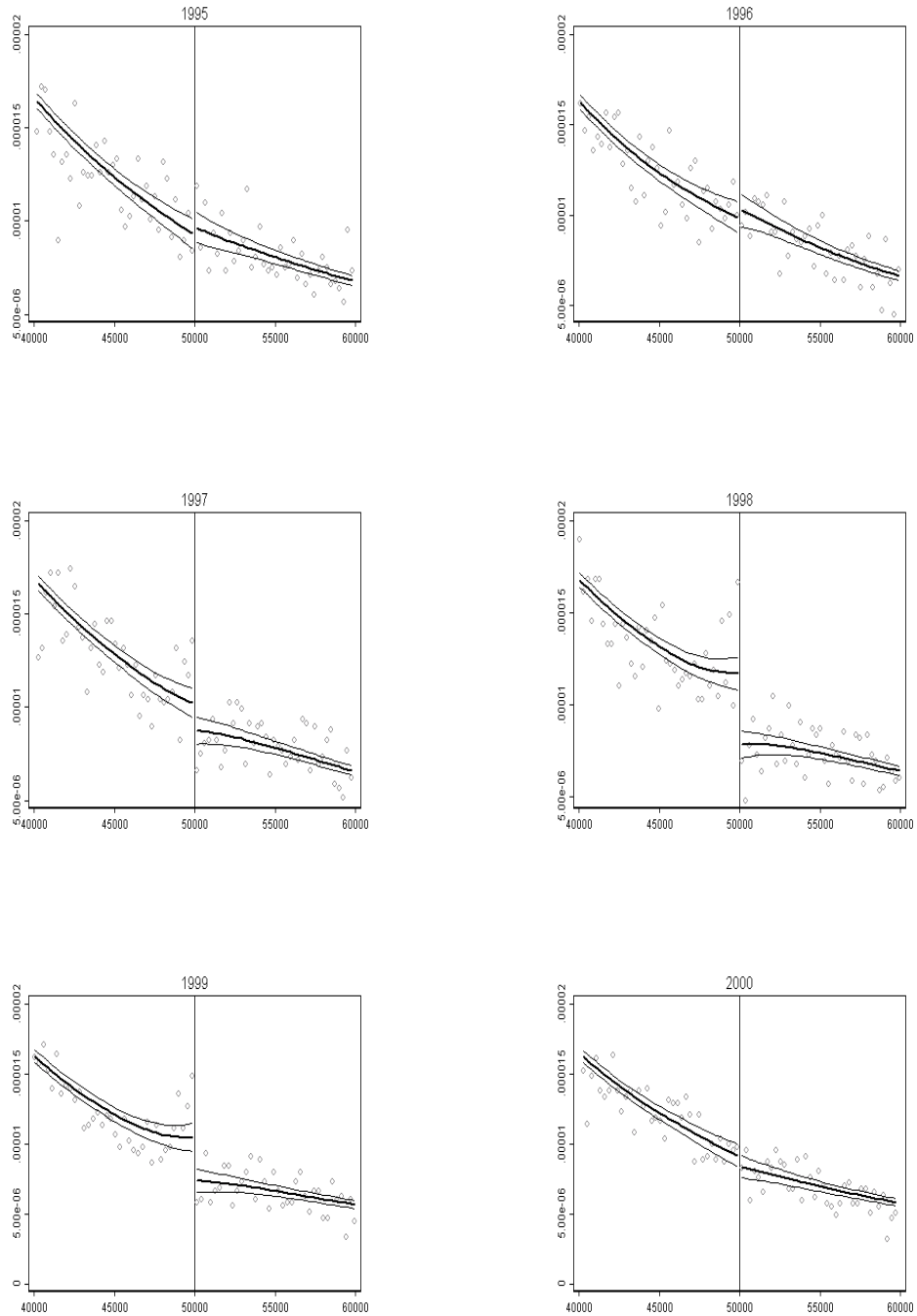
- McCrary, J. 2008. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics* 142.
- Person, and Diamond. 2016. “The Long-Term Consequences of Teacher Discretion in Grading of High Stakes Test.” *Working Paper*, pp. .
- Raj Chetty, O.T., John Friedman, and L. Pistaferri. 2011. “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities.” *Quarterly Journal of Economics* 126:749–804.
- Raspiller, S. 2007. “Le taux réduit d’impôt sur les société pour les PME.” *Trésor éco*, pp. .
- Saez. 2010. “Do Taxpayers Bunch at Kink Points?” *American Economic Journal: Economic Policy*, pp. .
- Thomas Dee, B.J., Will Dobbie, and J. Rockoff. 2016. “The Causes and Consequences of Test score Manipulation Evidence from the New York Regents Examinations.” *Working Paper*, pp. .

Figure 1: Gains from gaining eligibility to the decreased tax cut or to the exemption of the exception contribution



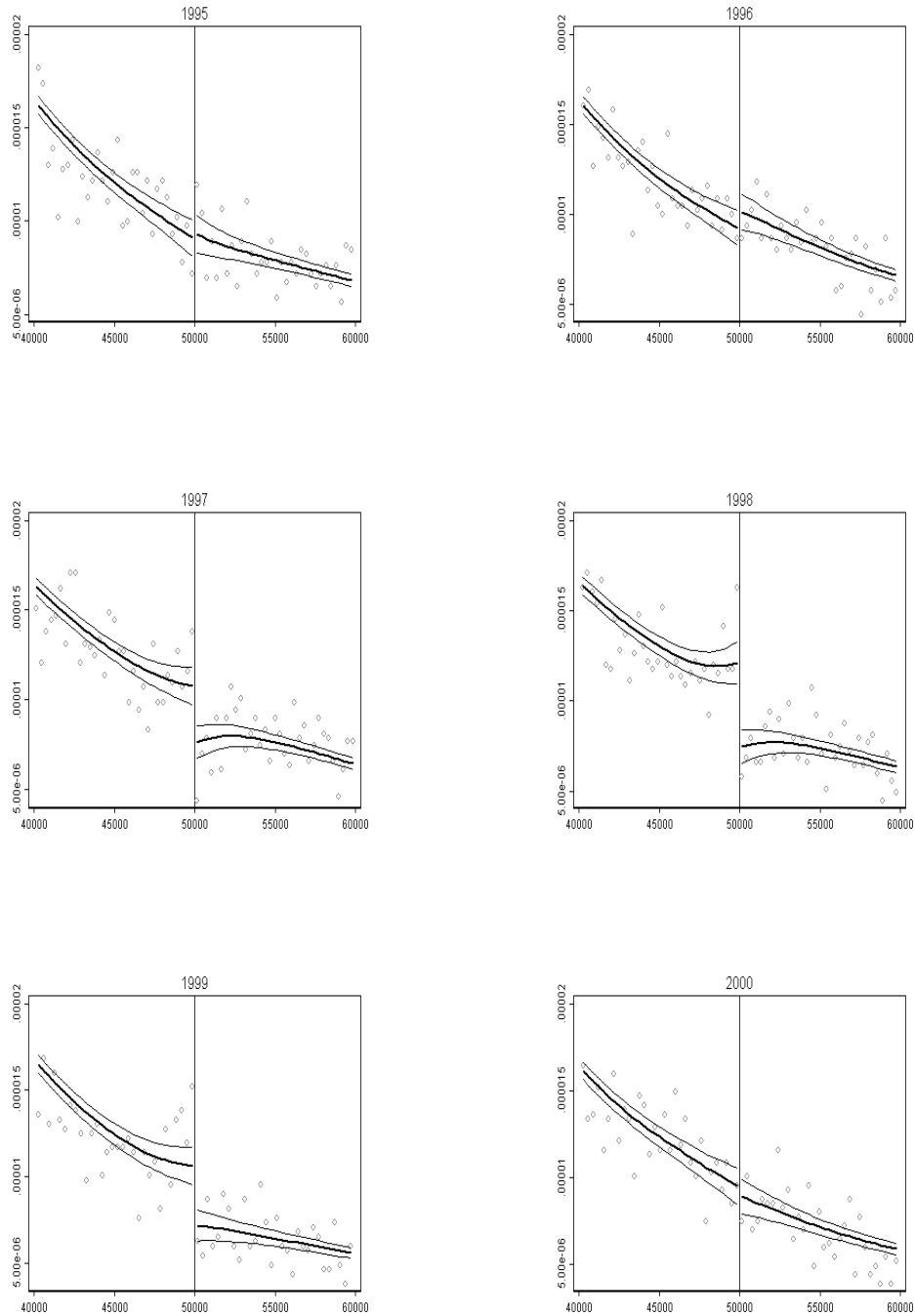
The figure plots the gains from avoiding the exceptional contribution to corporate income tax holding profit constant. We also report in dashed line the benefit from becoming eligible to the reduced marginal tax cut.

Figure 2: A transient discontinuity in firms' sales distribution



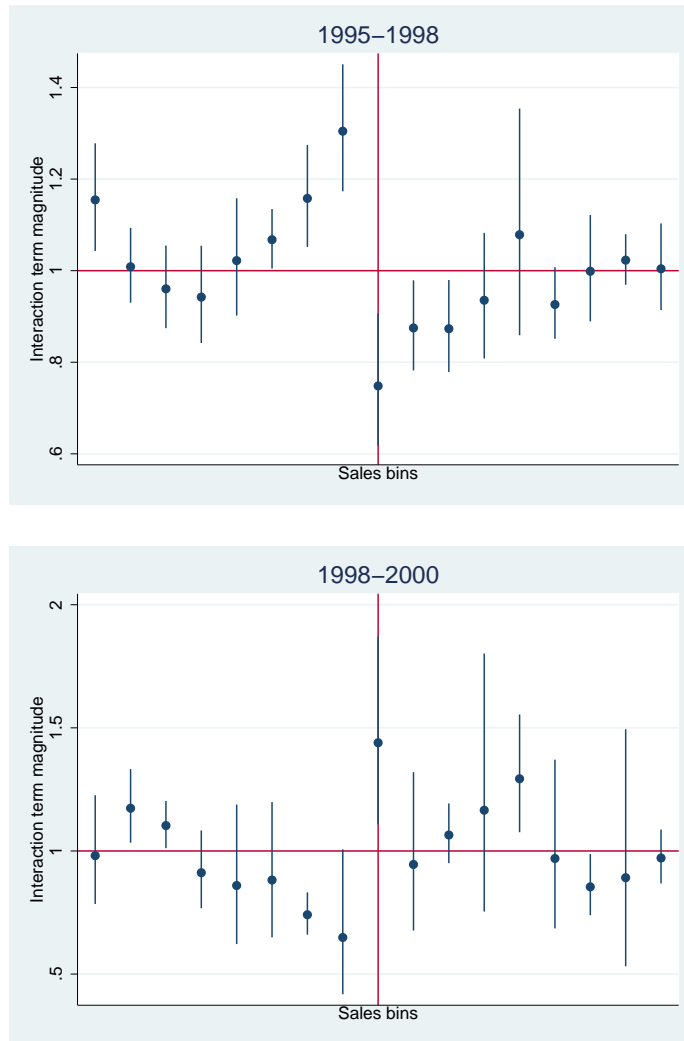
The distribution of firms with sales between kFrancs 20 and kFrancs 100, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 3: A transient discontinuity in firms' sales distribution: balanced dataset



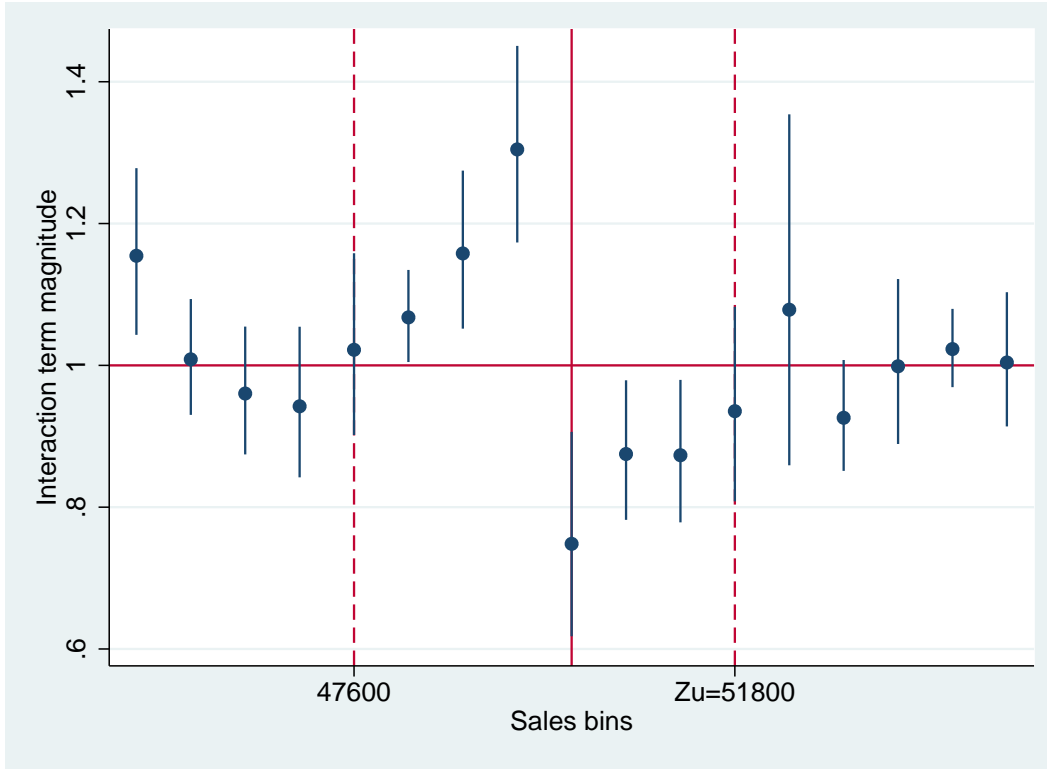
The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 4: Representing bunching compared to counterfactual obtained from years during which there was no incentive to bunch



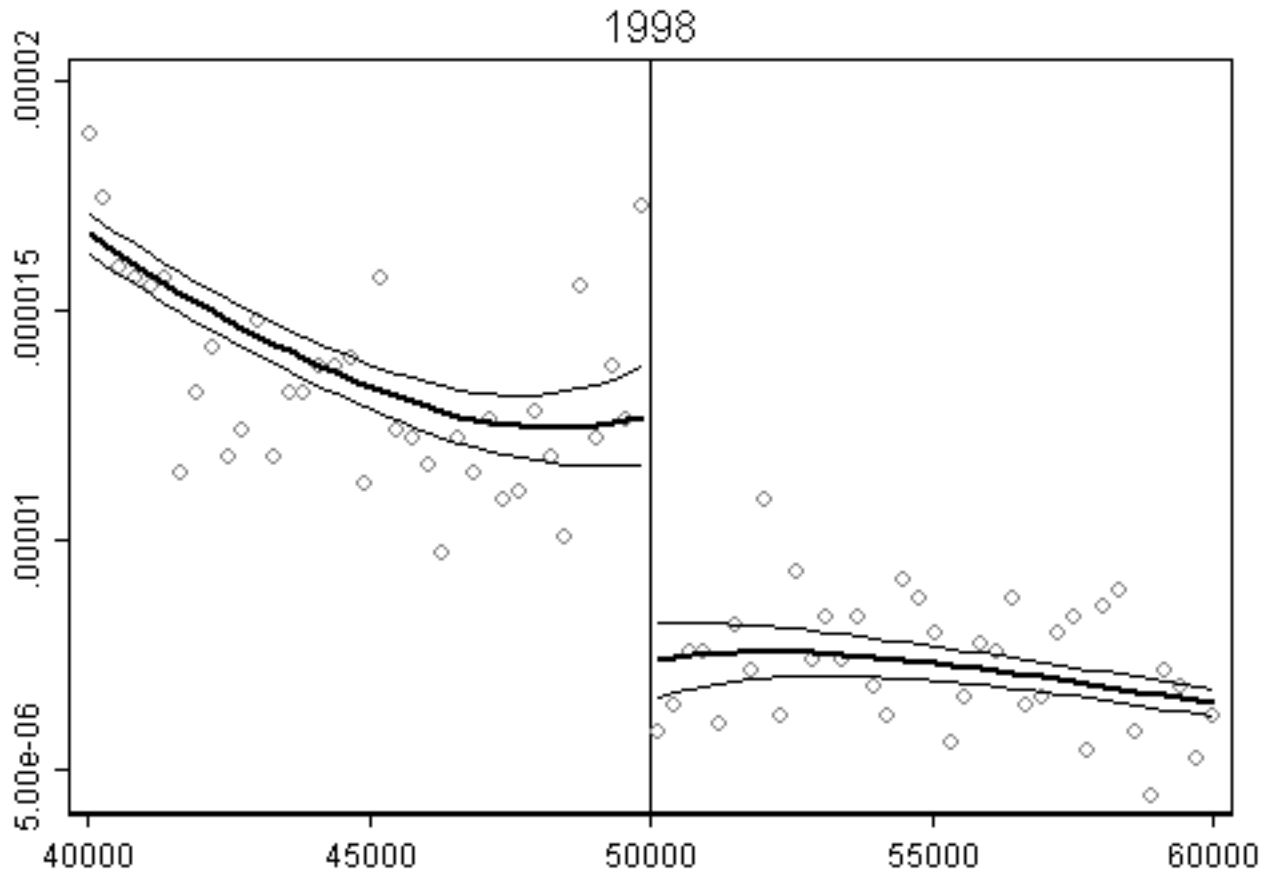
The figure plots the differential number of firms in bins around the threshold in years during which there was incentives to bunch compared to years during which there was no incentives. Each point comes from an interaction term between the bin indicator and the indicator of incentives. 95 % confidence intervals are constructed using robust standard errors clustered at the bin level.

Figure 5: Estimating the width of the valley



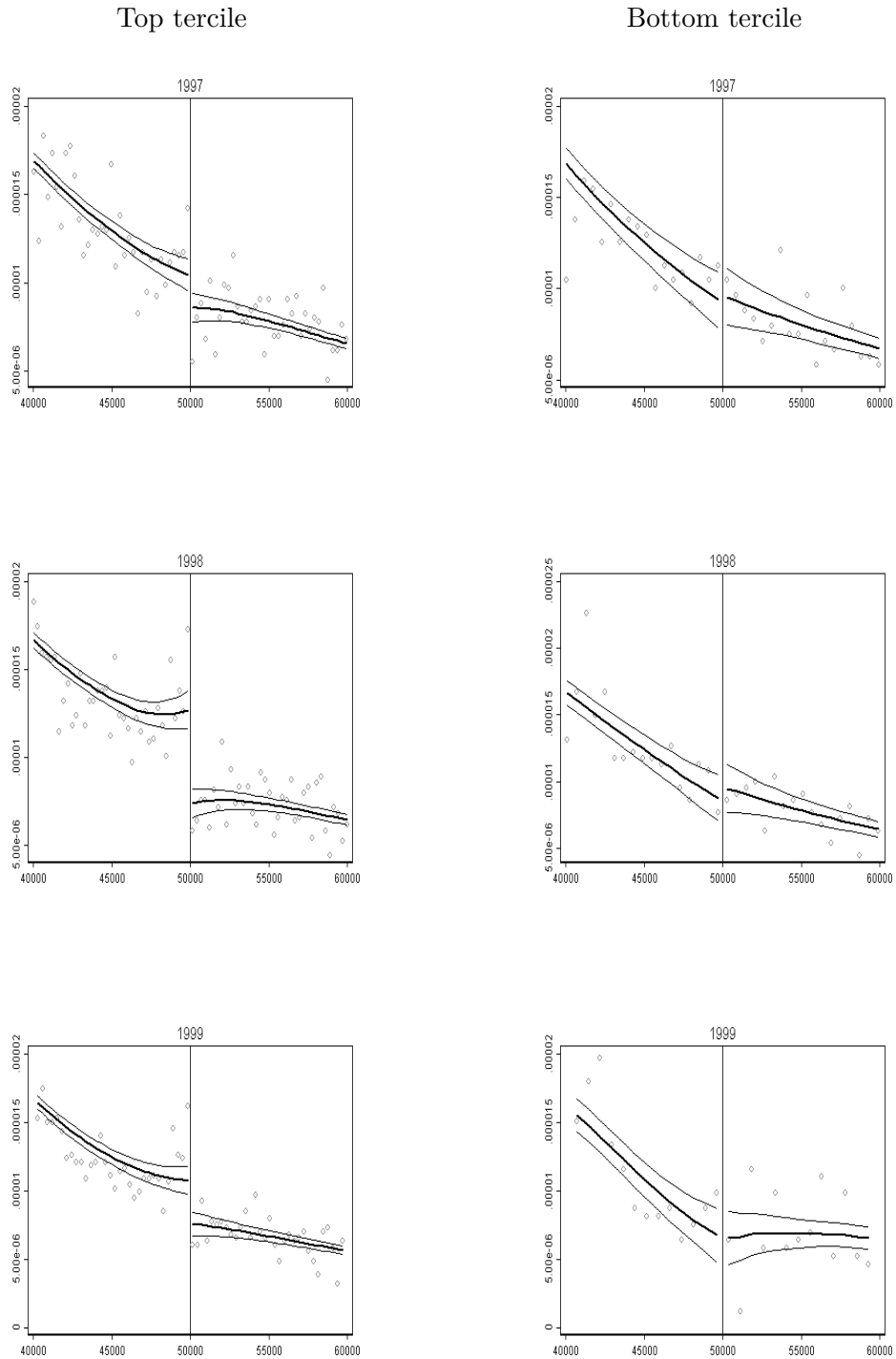
The figure reports the manipulation region. It illustrates the lower end z_L that is determined as the lower end of the group of bins that precede the bunch of coefficients that are significant below threshold. z_U is determined from the equality of missing mass and excess bunching as reported in Table 3

Figure 6: Non optimizers in 1998 among firms with large profit



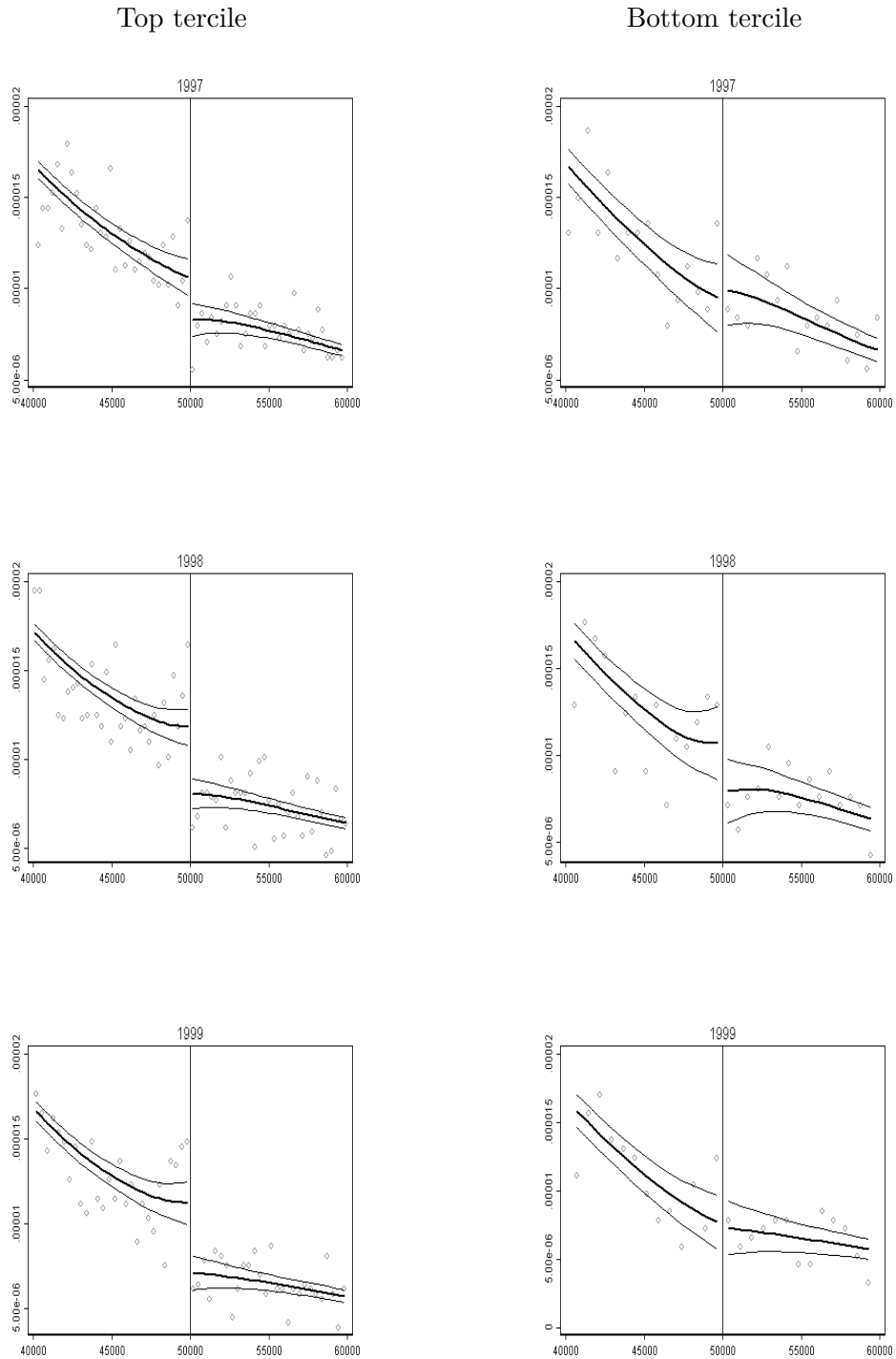
This figure plots the distribution of firms with profits within the top tercile in 1998 and turnover between kFrancs 20 and kFrancs 100. It excludes firms that belong to a conglomerate and restrict the sample to firms that pay the corporate income tax. We zoom on turnover between kFrancs 40 and kFrancs 60, with the threshold in the middle of the graph.

Figure 7: Differential bunching by profit level



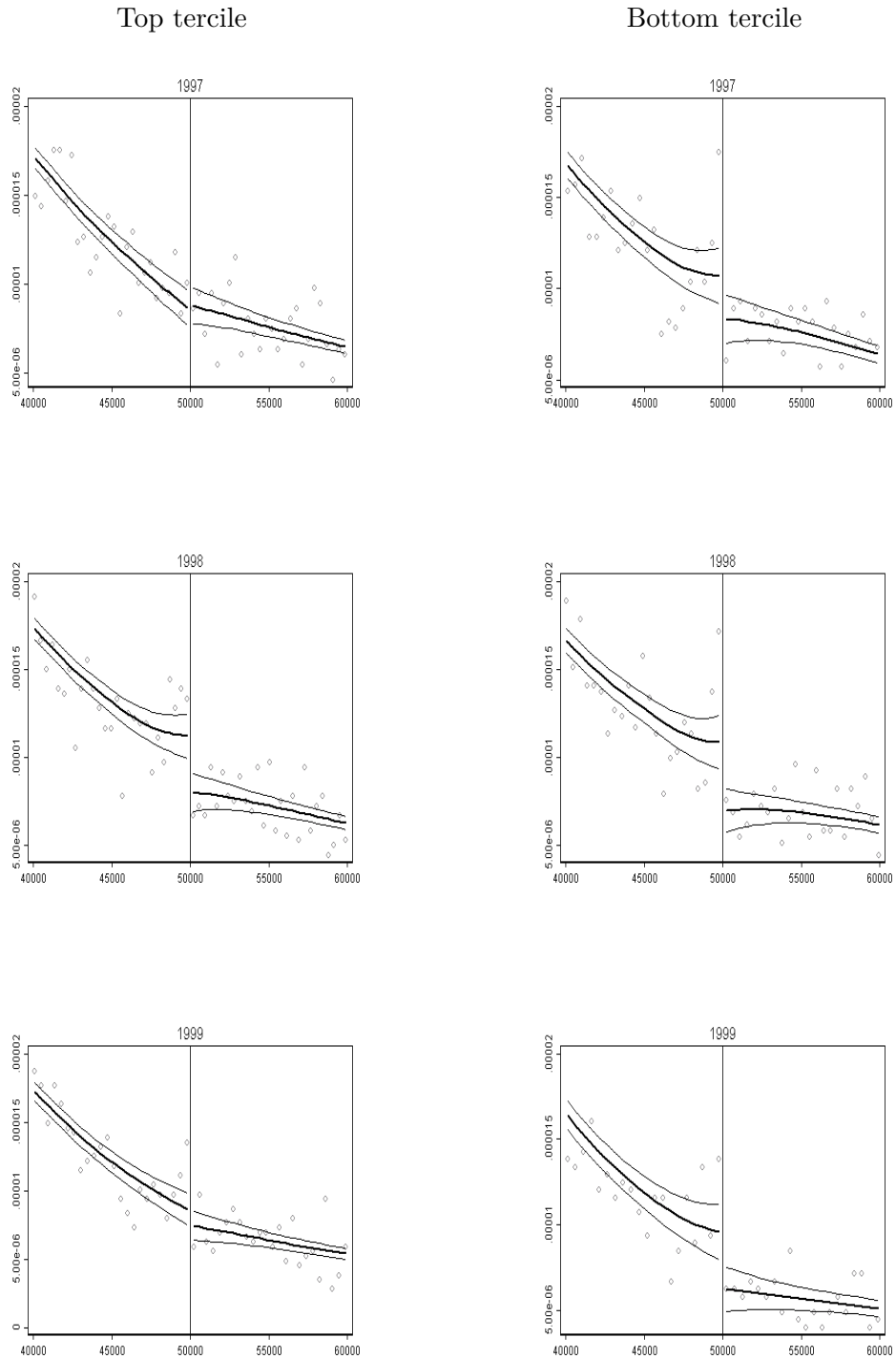
The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 8: Differential bunching by profit level in 1995



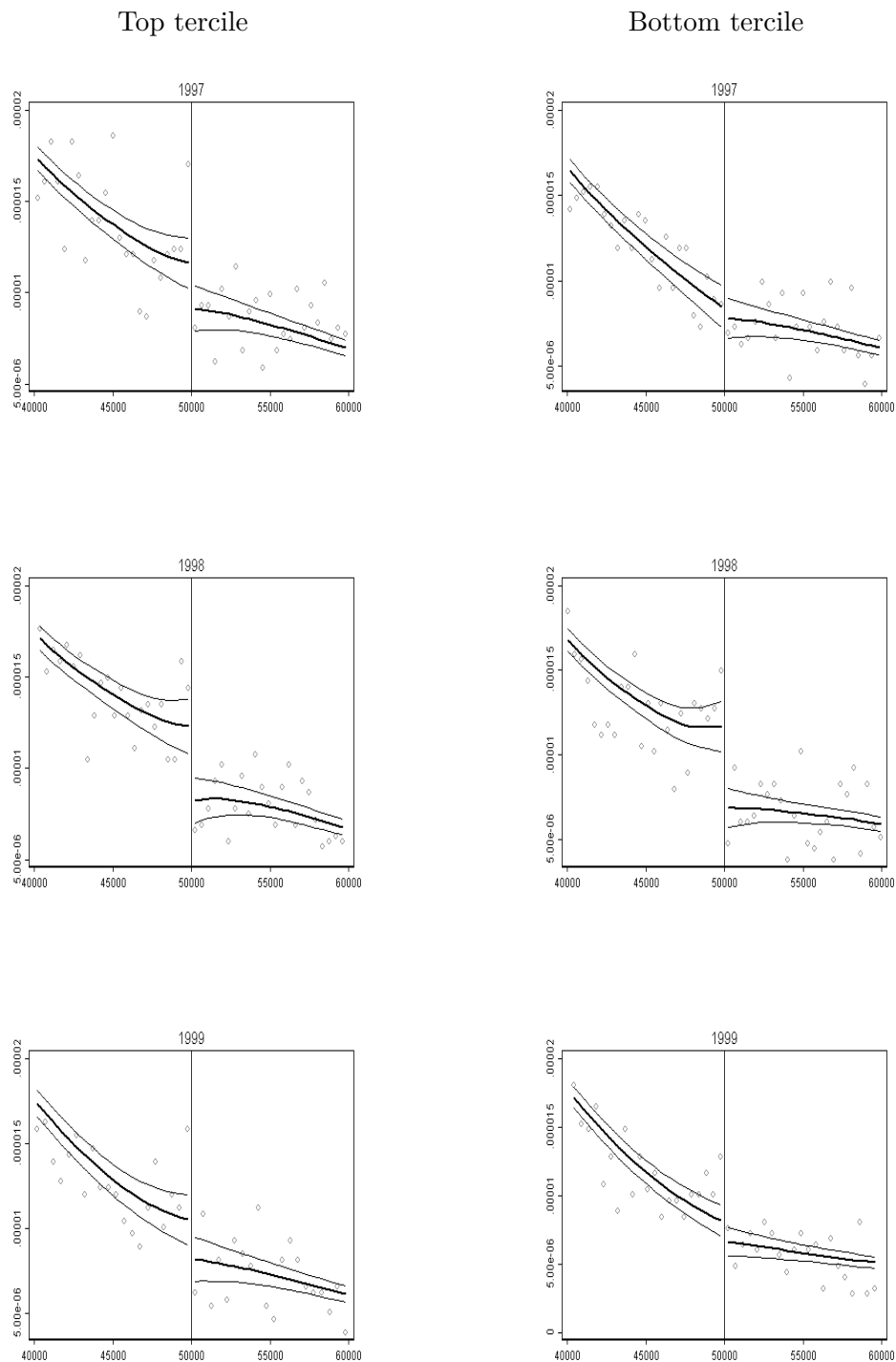
The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 9: Differential bunching by adjustment cost level



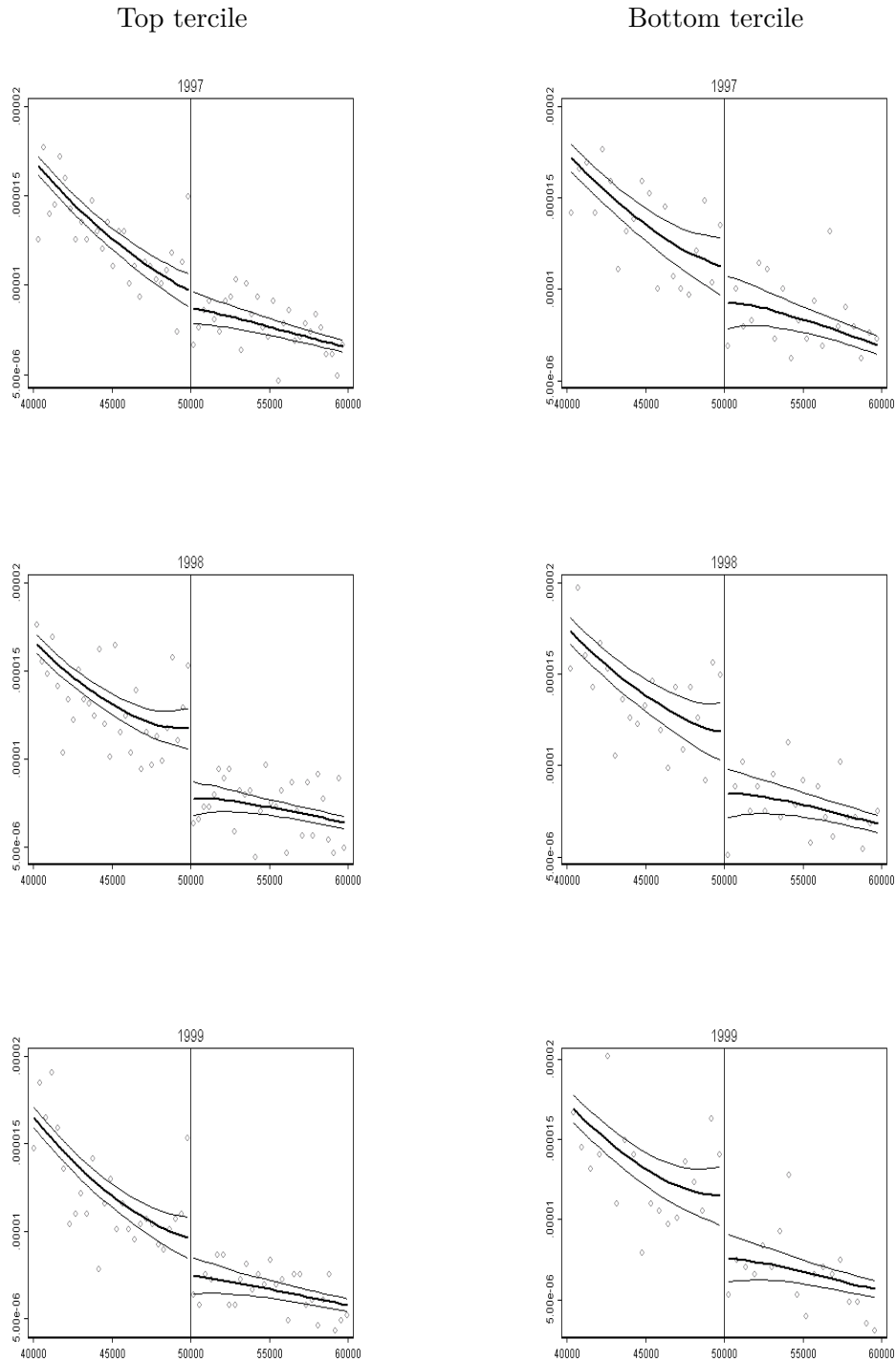
The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 10: Differential bunching by output elasticity with respect to M



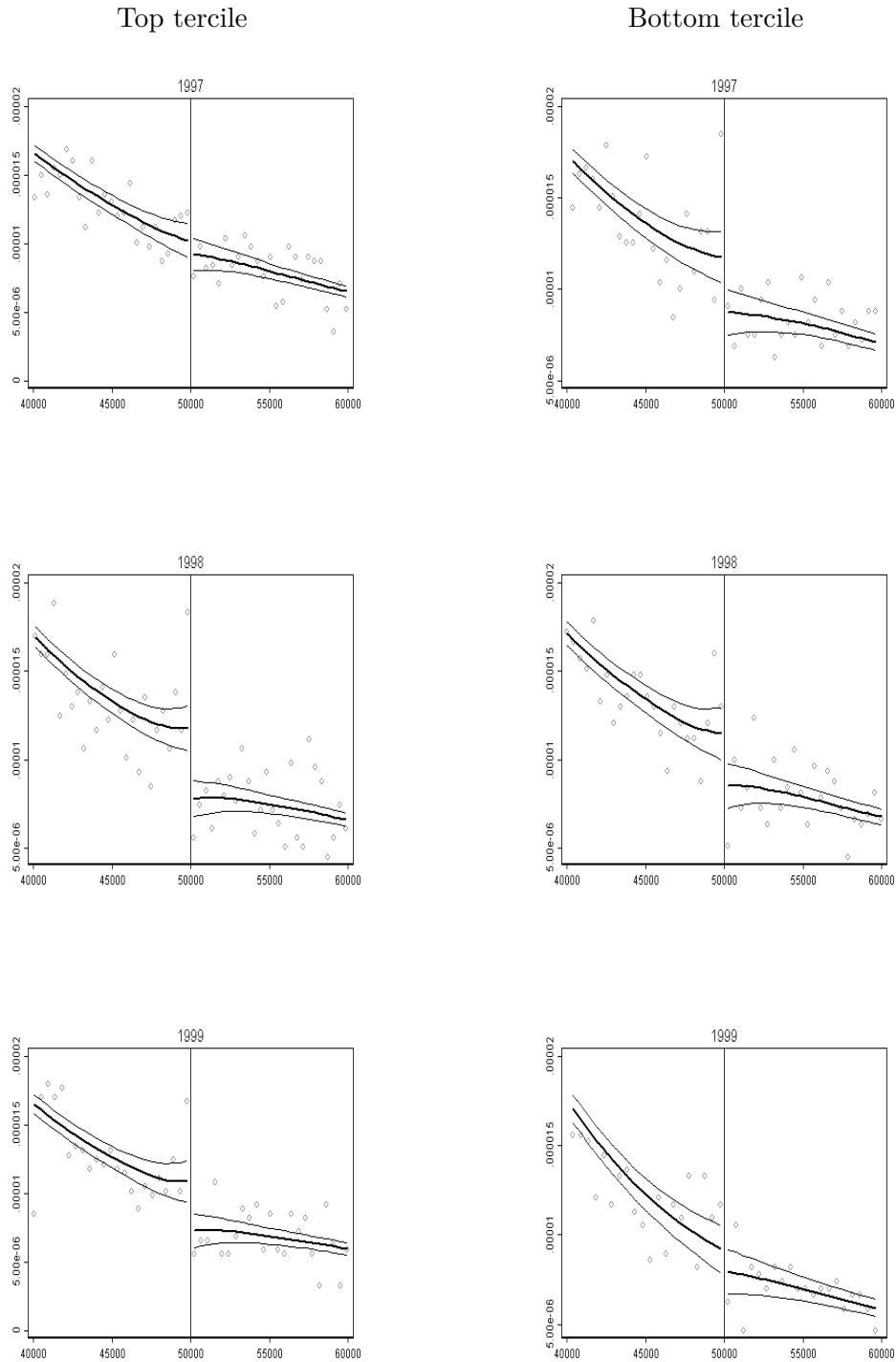
The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 11: Differential bunching by output elasticity with respect to L



The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Figure 12: Differential bunching by output elasticity with respect to K



The distribution of firms with sales between kFrancs 20 and kFrancs 100 that were filled tax forms each year between 1995 and 2000, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate. We zoom on the distribution that is between kFrancs 40 and kFrancs 60 to have the threshold in the middle.

Table 1: Descriptive statistics

	mean	count	sd
Sample			
Output	32941.314	123092	19561.129
Turnover	38636.709	123092	18484.968
Profit	1435.576	123092	8028.399
Adjustment cost of K	1.384	121766	0.415
Import demand elasticity	-1.051	60956	0.162
	mean	count	sd
All firms			
Output	12473.616	4435662	514131.391
Turnover	17763.855	4435662	553525.626
Profit	952.783	4435662	104418.088
Adjustment cost of K	1.420	4416852	0.472
Import demand elasticity	-1.029	1413664	0.202

Table 2: Discontinuity estimates

Raw dataset			
	(McCrary, 2008) estimates	standard errors	(Cattaneo and Ma, 2016) p-values
1995	.058	(.092)	0.9661
1996	.001	(.076)	0.6937
1997	-.331	(.083)	0.0002
1998	-.598	(.089)	0.0000
1999	-.714	(.114)	0.0002
2000	-.143	(.110)	0.1280

Balanced dataset			
	(McCrary, 2008) point estimates	standard errors	(Cattaneo and Ma, 2016) p-value
1995	.274	(.210)	0.6291
1996	.110	(.184)	0.3414
1997	-.647	(.193)	0.0384
1998	-.919	(.201)	0.0153
1999	-.617	(.155)	0.0052
2000	-.056	(.150)	0.2822

Note: This table reports discontinuity estimates for the two samples. The balanced dataset is the dataset restricted to the set of firms that have filled tax forms each years of the 1995-2000 period. Column 1 and 2 report the point estimates and standard errors obtained from (McCrary, 2008) estimation procedure, column (3) reports the p-value.

Table 3: Bunching estimators

Panel A: Raw dataset			
	\hat{B}	\hat{M}	\hat{b}_{av}
1997	91.422	94.019	1.534 (0.614) ^{***}
1998	182.550	162.396	3.033 (0.539) ^{***}
1999	82.609	76.192	2.038 (0.454) ^{***}
Panel B: Balanced dataset			
1997	53.225	52.230	2.346 (0.824) ^{***}
1998	71.652	70.134	3.018 (0.822) ^{***}
1999	44.949	43.375	2.946 (0.744) ^{***}
Panel C: Past years as counterfactual			
1997-1998	225.251	226.814	3.058

Note: This table reports the bunching estimators for different sample (Panel A et B) and the bunching estimators obtained from the technique that uses past years as counterfactual (Panel C). \hat{M} is missing mass and \hat{B} excess bunching \hat{b}_{av} refers to average bunching and \hat{b}_{adj} to adjusted bunching.

Table 4: Proportion of non-optimizers and adjusted bunching

	α	\hat{b}_{adj}
$z_D = 51200$	0.849	23.891
$z_D = 51800$	0.872	20.252

Note: This table reports the sensitivity analysis for the share of optimizers i.e. the values of α for two potential values of z_D . Column (2) reports the corresponding values of adjusted bunching measured as $\hat{b}_{adj} = \frac{\hat{b}_{av}}{(1-\alpha)}$

Table 5: Bunching estimation by subgroup

	Capital adjustment cost		Profits		Profits in 1995	
	Top	Bottom	Top	Bottom	Top	Bottom
1997	0	1.017	0.490	0	0.329	0.651
	-	(0.304)***	(0.332)	-	(0.321)	(0.312)**
1998	0.722	0.616	0.879	0	0.832	0.170
	(0.271)***	(0.270)**	(0.210)***	-	(0.248)***	(0.374)
1999	0.376	0.719	0.653	0	0.657	0.288
	(0.228)*	(0.310)**	(0.194)***	-	(0.198)***	(0.392)
	Elasticity wrt K		Elasticity wrt L		Elasticity wrt M	
	Top	Bottom	Top	Bottom	Top	Bottom
1997	0.364	0.895	0.692	0.517	1.056	0
	(0.244)	(0.351)**	(0.364)*	(0.275)*	(0.388)***	-
1998	0.787	0.498	0.494	0.620	0.499	0.792
	(0.268)***	(0.273)*	(0.219)**	(0.256)**	(0.251)**	(0.386)**
1999	0.596	0.607	0.532	0.767	0.717	0.710
	(0.256)**	(0.289)**	(0.207)**	(0.305)**	(0.286)***	(0.331)**

Note:

Table 6: Characteristics of the compliers.

Adjustment cost, Incentives, Ability to bunch								
	Low adjustment cost of capital		Large profit		Large profit in 1995		Large profit in 1996	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		FE		FE		FE		FE
Compliers	0.0495*** (0.0174)	0.0464*** (0.0160)	0.0277* (0.0146)	0.0237* (0.0143)	0.00351 (0.0165)	0.000422 (0.0163)	0.00194 (0.0160)	0.000422 (0.0160)
Observations	4439	4386	4494	4441	3881	3841	4008	3941
Production function characteristics								
	Large elasticity wrt K		Large elasticity wrt L		Large elasticity wrt M			
[1em] Compliers	-0.0526*** (0.0184)	-0.0492*** (0.0187)	-0.00200 (0.0188)	-0.00163 (0.0173)	0.0598*** (0.0185)	0.0570*** (0.0178)		
Observations	4439	4386	4439	4386	4439	4386		

Note: This table estimates the different characteristics of the bunchers compared to other firms that were eligible to bunching. Firms eligible to bunching are the firms that were above threshold in the manipulation region in years during which there was no incentives to bunch. Characteristics of the bunchers are identified following Person and Diamond (2016) technique. Variables are centered with the mean of the variable in the region just below the manipulating region: Turnover in 45000-47600. Standard errors reported in parentheses are obtained by bootstrapping 500 times the test for the difference of characteristics. Columns (2) (4) and (6) report estimation with region and 16 industry fixed effects. Output elasticities are computed following (Levinsohn and Petrin, 2003) estimation procedure. Adjustment cost of capital is determined from (Asker, Collard-Wexler, and Loecker, 2014) estimation procedure.

Table 7: Characteristics of the compliers.

Adjustment cost and Incentives to bunch								
	Low adjustment cost of capital		Large profit		Large profit in 1995		Large profit in 1996	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		FE		FE		FE		FE
Compliers	0.205** (0.0797)	0.154* (0.0834)	0.270*** (0.0550)	0.222*** (0.0588)	0.181*** (0.0477)	0.147*** (0.0553)	0.358*** (0.0671)	0.324*** (0.0639)
Observations	906	906	912	912	801	801	834	834
Production function characteristics								
	Large elasticity wrt K		Large elasticity wrt L		Large elasticity wrt M			
[1em] Compliers	-0.0512 (0.0695)	-0.0171 (0.0831)	-0.0118 (0.0789)	-0.0417 (0.0907)	0.191** (0.0752)	0.155* (0.0831)		
Observations	906	906	906	906	906	906		

Note: This table estimates the different characteristics of the bunchers in 1997 compared to other firms that were eligible to bunching. Firms eligible to bunching are the firms that were above threshold in the manipulation region in years during which there was no incentives to bunch. Characteristics of the bunchers are identified following Person and Diamond (2016) technique. Standard errors reported in parentheses are obtained by bootstrapping 500 times the test for the difference of characteristics. Column (2) (4) and (6) report estimation with region and 16 industry fixed effects. Output elasticities are computed following (Levinsohn and Petrin, 2003) estimation procedure. Adjustment cost of capital is determined from (Asker, Collard-Wexler, and Loecker, 2014) estimation procedure.

Table 8: Characteristics of the compliers.

Adjustment cost and Incentives to bunch								
	Low adjustment cost of capital		Large profit		Large profit in 1995		Large profit in 1996	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		FE		FE		FE		FE
Compliers	0.0115 (0.0887)	-0.000891 (0.0926)	0.415*** (0.0668)	0.407*** (0.0640)	0.262*** (0.0597)	0.258*** (0.0557)	0.368*** (0.0595)	0.351*** (0.0600)
Observations	1047	1044	1053	1050	885	883	915	913
Production function characteristics								
	Large elasticity wrt K		Large elasticity wrt L		Large elasticity wrt M			
[1em] Compliers	0.160* (0.0847)	0.201** (0.0784)	0.144+ (0.0889)	0.109 (0.0846)	0.0855 (0.0916)	0.113 (0.119)		
Observations	1047	1044	1047	1044	1047	1044		

Note: This table estimates the different characteristics of the bunchers in 1998 compared to other firms that were eligible to bunching. Firms eligible to bunching are the firms that were above threshold in the manipulation region in years during which there was no incentives to bunch. Characteristics of the bunchers are identified following Person and Diamond (2016) technique. Standard errors reported in parentheses are obtained by bootstrapping 500 times the test for the difference of characteristics. Column (2) (4) and (6) report estimation with region and 16 industry fixed effects. Output elasticities are computed following (Levinsohn and Petrin, 2003) estimation procedure. Adjustment cost of capital is determined from (Asker, Collard-Wexler, and Loecker, 2014) estimation procedure.

Table 9: Consequences of manipulation: panel

	(1)	(2)	(3)	(4)	(5)
	Turnover	Y	Sold production	Change in inventories	Capitalized production
_bs_1	-88.63** (39.88)	-265.1 (312.8)	-505.3* (304.9)	152.5** (69.70)	87.79*** (31.13)
Observations	3692	3692	3692	3692	3692

Note: This table estimates the consequences of bunching

Table 10: Consequences of manipulation: cross section

	(1)	(2)	(3)	(4)
	Y	Sold production	Change in inventories	Capitalized production
_bs_1	-2870.6*** (604.4)	-3163.3*** (593.7)	257.8** (116.0)	34.91 (69.76)
Observations	1352	1352	1352	1352

Note: This table estimates the consequences of bunching