JMS | Journées de méthodologie statistique de l'Insee | 2018

Tax avoidance in firms

Arthur BAUER(*), Martin ROTEMBERG(**)
(*)INSEE
(**) New York University
arthur.bauer@insee.fr and mrotemberg@nyu.edu

Mots-clés. Business Taxes, Tax Evasion, Firm Production

Résumé en 350 mots

Abstract

Tax codes can have notches ; regions in which after-tax profits decrease 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 and decrease in sales.

1 Introduction

Firms avoid paying taxes through a variety of strategies, including misreporting, transfer pricing, and changing behavior (Desai, Foley, and Hines, 2006; Liu, Schmidt-Eisenlohr, and Guo, 2017; Zucman, 2014). Even within the category of changed behavior, there are a vareity of options available to firms : they can adjust their input mix or shift sales over time. We leverage a 1997 tax increase in France in order to quantify if and how firms adjusted their behavior to avoid paying taxes. The reform increased profit taxes for firms with over 50 million Francs in turnover by 15%. As a result, firms with fairly similar sales faced different marginal tax rates.

We focus our attention on two questions. First, do firms distort their behavior in order to avoid paying additional taxes ? Second, are there specific types of firms who are more likely to be affected and how do they avoid taxation ? Since there are no (believable) survey questions asking firms if and how they avoid taxes, we use reported behavior on real firm behavior in order to back out manipulation responses.

For the first question, we note that firms have strong incentives to have a size just below the threshold. Endogenously, this leads to excess mass in the firm-size distribution : "too many firms" just below the cutoff, and correspondingly "too few" just above [1]. We use complementary but distinct approaches to estimate the extent of excess mass. First, following a large literature [2] we assume that the firm size distribution is "well behaved" in any given year. This implies that firms far away from the tax threshold do not manipulate their behavior in response to the policy, and so we can use the distribution (of turnover) of those firms in order to estimate a counterfactual manipulation-free firm size distribution. We can then compare the manipulation-free counterfactual distribution to the actual firm-size distribution around the cutoff to back out how firms change their behavior. As a placebo check, we show that this method identifies excess mass neither in the years before the policy was enacted nor in the years after it was phased out.

We also use the time-series dimension of the data in order to estimate excess mass by estimating a counterfactual manipulation-free distribution using years when the policy was not in effect.

Using the cross-sectional information we find that there were around 150 firms who changed their sales in reponse to the new tax regime. We find similar results using the time-series information. This corresponds to a tax elasticity of sales of 0.16.

We then turn to identifying which firms manipulate their sales. We do this by studying the ex-ante characteristics of firms below the tax cutoff (as in Diamond and Person (2016)). If, for instance, firms who normally have high profits manipulate their sales more, then we would observe more of those types of firms just below the cutoff than we would expect (for instance, by using their prevalence when the tax cutoff was not in effect).

In addition to finding that high-profit firms are more likely to manipulate their sales, we also find that firms with larger adjustment costs are less likely to bunch. We start with capital adjustment cost (Asker, Collard-Wexler, and Loecker, 2014). We show that bunchers are, depending on the years, between 3% and 15% more likely to have capital adjustment cost within the lowest tercile. Furthermore, consistent with the logic that material inputs generally more flexible than capital inputs, we find that firms who bunch tend to have larger elasticity of output with respect to material and lower elasticity of output with respect to capital.

We then turn to estimating how firms manipulate their sales. Firms have many potential margins of adjustment, including affecting their production decisions, prices, and inventories. As in (Diamond and Person, 2016; Bachas and Soto, 2015; Thomas Dee and Rockoff, 2016), we note that a similar logic to the bunching estimator can help us back out firm behavior. We compare firms in the manipulation region to firms outside it (either firms who are too small/large, or firms of similar sizes in different years). If firms change some characteristics in order to shrink, then we will see differences in that characteristic in the manipulation. For instance, we find that firms avoid increasing their sales by instead increasing their inventories. In the data, this shows up as firms in the manipulation region overall having more inventories than would be predicted by the out-of-sample counterfactuals. Note that an RD-type of estimate (comparing firms just below to just above the cutoff) is unlikely to generate the causal mechanisms for tax avoidance since the firms who choose to manipulate are ex-ante different along a variety of dimensions. In order to avoid the selection issue, we compare all of the firms who might be affected by manipulation

1. Almunia and Lopez-Rodriguez (2017); Bachas and Soto (2015); Garicano, Lelarge, and Van Reenen (2016); Gourio and Roys (2014); Kleven and Waseem (2013); Liu and Lockwood (2016); Onji (2009)

2. Aghion et al. (2017); Bach (2015); Raj Chetty and Pistaferri (2011); Thomas Dee and Rockoff (2016); Diamond and Person (2016); Kleven (2016); Kleven and Waseem (2013); Lardeux (2018); Barbanchon (2016); Saez (2010)

to those who do not.

We find that while firms do lower their production as they lower sales, the primary driver of the manipulation is an increase in change in inventories and capitalized production. This suggests that adjusting production is relatively more costly than either stocking the production or reinvesting it in the production process. Since the changes "add up," this implies that the values we find are real effects of tax avoidance, not just tax evasion through misreporting or fraud (Best et al., 2015).

2 Institutional setting and Data

As in other countries, when entrepreneurs in France chose between two kinds of tax regimes : taxes on income (IR) or enterprise taxes (IS). Around 2/3 of firms - but 1/5 of value added, is in the former group. Each firms ? sector determines what category of taxes it pays, with firms with benefits mostly from services (BNC) in one category, and firms in trade and manufacturing activity (BIC) or agricultural activity (BA) choosing between a regular (BRN) or simplified (RSI) setup. Around half of firms - but only around 5% of value added, are in the latter group : around 90% of value added is in the the regular tax regime (See Table 1 for the exact values).

For the entire sample of our data (1995-2000), the French corporate income tax was been of 33.3%. In 1997, the rate was increased by 15% [3] (to 38.3% for enterprise firms with sales above fifty million Francs (around seven million Euros). [4] In 1999, the exceptional rate was decreased to 36.6%, and in 2000 it was removed. (JORF nb262, p16387) [5]

Simultaneously, in 1997 the marginal tax rate for profits below 200,000 Francs was lowered to 19% if the firms reinvested their extra net income in capital. This policy remained until 2000.

As of 2001, the condition of reinvestment into capital was removed, the marginal tax rate was increased to 25% and the limit of profit to 250,000 Franks. Empirically, over the 1997-1999 period it is difficult to disentangle the impact of these two policy changes. However in 2000 only one of the two policies remain. As a consequence, it is possible to assess independently its impact and subtract it back from the combined effect of the two policies over the 1997-1999 period.

The incentives to distort one firm ?s behavior to avoid this exceptional contribution were small but significant. Figure 1 shows the differential of tax rates above and below the threshold. [6] This allows us to infer the gains by profit to becoming eligible to an exemption of the contribution and to taking advantage of the reduced marginal tax rate. For a firm with profits of 2 millions of Francs (roughly the sample mean), avoiding the tax would save 100 thousands of Francs.

2.1 Data

We use three datasets collected by the Direction Générale des Impôts and the French National Institute of Statistics (INSEE) : the BRN files, the RSI files as well as *Enquête sur les liaisons financières* (LiFi). In France, each firm has an identifier, SIREN, which facilitates its interaction

3. On 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

4. Furthermore, eligibility was conditioned on 75% of firms' share capital owned by physical people and all of share capital paid-up, we only include eligible firms in our sample.

5. 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`

6. Tax rates are the combination of the marginal tax rates, the additional contribution and the exceptional contribution.

with the different administrations. This identifier is present in these three datasets which allows us to merge them, and to follow firms across the years. The BRN files have often been used in academic research (see for instance (Caliendo, Monte, and Rossi-Hansberg, 2015)). Recent work has integrated smaller firms as those affiliated to the RSI regime in their analysis as well (see for instance (Garicano, Lelarge, and Van Reenen, 2016).

With the BRN-RSI files we build a panel of firms that spans the 1995-2000 period. This period is well-suited for the analysis of the 1997-1999 reform as it contains two years before its implementation that we can use as counterfactual and one year after to analyze the persistence of its consequences. [7]. This dataset is, to the best of our knowledge, the first to allow an analysis of the dynamic dimension of firms' bunching. The BRN RSI files contain about 1.6 million firms each year. They cover the universe of firms within the BIC category affiliated to the normal and simplified regimes. They don ?t cover the BNC category, as firms within this category only pay taxes on revenues and are therefore not concerned by the reform we are studying. The BA categories has always been put aside in economic analysis as it covers very specific types of firms mostly in the agricultural sector.

The BRN-RSI files provide all the relevant firms characteristics, namely inputs of productions, value-added, turnover, profits, and inventories. We use nonimal values as the eligibility threshold of 50 millions of Francs was in not indexed to inflation. An advantage of the BRN-RSI files over FICUS is that it provides information on the type of tax regime firms have opted for (either IR or IS). Nevertheless a drawback is that it provides rawer information than FICUS.

FICUS is also available for all the 1995-2000 period. It is a production of the French census bureau (INSEE). INSEE confronts the balance sheet information obtained from tax forms collected by the tax office (DGFIP) and gathered in the BRN-RSI files with one of its internal sources of information, the EAE survey. As a consequence we make two restrictions to our main sample. First, we clean the dataset and remove extreme values for input shares. We exclude observations that have capital shares above 10. Second, we exclude firms that report negative values of inputs.

We add information on conglomerate membership from the *Enquête sur les liaisons financières* (LiFi), collected by INSEE since 1980 and available every years of the 1995-2000 period. INSEE surveys every year all firms with sales above 393 million of Francs, equity portfolio above 7.9 million of Francs 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 further restrict the sample to firms with turnover between 20 000 thousands of Francs and 100 000 thousands of Francs, as well as to firms eligible to the tax cut. The BRN-RSI dataset allows us to keep only firms affiliated to the IS regime. 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.

Table 2 presents the descriptive statistics. The average output level is 32941 thousands of Francs, the average level of turnover is 38637 thousands of Francs, material capital and labor's average values are respectively 11042 thousands of Francs, 10206 thousands of Francs and 6966 thousands of Francs. The average number of employees is 50. The average profit of firms in our sample is 1436 thousands of Francs. bFirms in our sample are on average twice as large as than the average French firm.

7. In 2001 there were further changes to the tax code preventing us from doing longer follow-up analyses.

3 Theoretical framework

Firms maximize revenues minus costs, where costs come from production, adjustment, taxes, and potentially inventories. The force encouraging firms to hold inventories is the adjustment costs : if a firm's frictionless optimal sales level is too far from baseline (in the case of a tax increase, firms would want to produce less) then firms might want to increase their sales and put some goods in inventory. [8] Pushing against this force is that inventory is also costly. Given convex inventory and adjustment costs, under some regularity assumptions the optimal choice of the firm is to produce less than the bliss point, but still put some goods in inventory.

The set-up is straightforward. Firms who sell x units of a good earn profits of

$$R(x_i) - c(q_i) - \gamma_{is} A(q_i^\star - q_i) - \tau(R(x_i) - c(x_i)) - I(q_i - x_i) \tag{1}$$

where $R(\cdot)$ is the revenue function, $c(\cdot)$ is the constant-returns-to-scale production cost, $A(\cdot)$ is the twice-differentiable convex adjustment cost coming from producing other that $q^\star$, γ_s is how expensive the adjustment cost is at the sector level, q is the quantity produced $\tau(\cdot)$ is the (profits) tax, and $I(\cdot)$ is the twice-differentiable convex inventory cost. We solve for optimal firm behavior in two steps : first by showing that for a given x, there is an optimal production/inventory decision for the firm, and then solving for x in the profit function.

Given $x < q^\star$, a firm seeks to minimize

$$c(q_i) + \gamma_{is} A(q_i^\star - q_i) + I(q_i - x_i).$$

Defining $c' = \phi_i$,

$$\gamma_s A'(q_i^\star - q_i) = \phi_i + I'(q_i - x_i) \tag{2}$$

If the firm chooses to overproduce to the bliss point ($q_i = q_i^\star$), then the left hand side is zero, and the right hand side will be positive, which will not minimize costs. If x is "far enough" away from $q_i^\star$ then a similar logic holds for if a firm chose to produce $q_i = x$: the left-hand-side would be larger than the right. Given the intermediate value theorem and the contuinuity and convexity of the adjustment and inventory costs, this implies that there is a unique production decision q_i for each optimal sales quantity x_i, and correspondingly a convex and twice differentiable function $C(\cdot)$ which captures the production, adjustment, and inventory costs.

We can therefore rewrite Equation 1 as

$$R(x_i) - C(q_i) - \tau(R(x_i) - c(x_i)) \tag{3}$$

which if $\tau(\cdot)$ is differentiable has the traditional solution of

$$R'(x_i) = C'(q_i) - \tau'(R(x_i) - c(x_i))\left[R'(x_i) - C'(q_i)\right].$$

However, if there is a jump in the tax schedule :

$$\begin{cases} \tau = 0 & x_i \leq \theta \\ \tau = .15 & x_i > \theta \end{cases}$$

then some firms will have to make a discrete choice : comparing profits at $x_i = \theta$ to the best choice at $\tau = .15$. As in (Kleven and Waseem, 2013) there will be a cutoff $\tilde{\theta}_i$ for which a firm would weakly prefer to sell θ then anything in $\left(\theta, \tilde{\theta}_i\right)$. Furthermore, if $q_i^\star > \theta$, then it is

8. For simplicity of exposition, we model the choice of the firm statically : it does not take into account its current production choices on either the bliss point or the ability to draw down future inventories.

straightfoward to show that (a) decreasing γ_{is} increases $\tilde{\theta}_i$, as does decreasing ϕ_i. One way that firms might have lower adjustment costs might be due to idiosyncratic firm-specific features, such as better access to capital markets.

Following Asker, Collard-Wexler, and Loecker (2014), we interpret the dispersion in the marginal product of the inputs of production as a proxy for idiosyncratic adjustment costs (since with no adjustment costs the dispersion would be zero). We follow the production function literature (Ackerberg, Caves, and Frazer, 2015; Asker, Collard-Wexler, and Loecker, 2014; Levinsohn and Petrin, 2003; Olley and Pakes, 1996) 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 since it is cheaper for them to adjust their sales.

4 Empirical Approach

4.1 Discontinuity Estimates

If firms endogenously lower their size in order to avoid the extra taxes, there will be excess mass in the distribution below the threshold and correspondingly too little mass above. We follow the standar approaches in the literature to measure excess mass, McCrary (2008). As a robustness test, we also follow Cattaneo and Ma (2016).

4.2 Bunching Estimators

Using Predictions from Distribution away from the Threshold as Counterfactual An alternative approach to measuring bunching is to use firms far away from the threshold. We estimate a fifth degree polynomial of the density far away, then follow it to the hold-out manipulation region. The difference between the estimated density and the actual density captures the extent of manipulation (and is a placebo check in years with no discontinuous tax threshold).

To be precise we estimate the following estimation :

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

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.[9] We determine z_L by eyeballing the distribution (Figure 4) 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

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

9. When we can't eyeball any manipulation, we estimate bunching to be zero as we can't identify any manipulation region that is necessary to estimate bunching with this methodology.

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

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

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 :

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

Using Past Years as Conterfactual Another way to investigate the apparition of distortions in the firm-size distribution over time is to compare distributions of firms across times around the threshold. 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 :

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

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 years 1997-1998 when the 1995-98 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.

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 in figure 3. 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. 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 200 thousands of Francs 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.

To quantify the size of the distortion we rely on an estimation procedure that looks similar to the one presented in the previous paragraph. 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.

The coefficients of the interacted terms (γ_i) in equation 8 allow us to estimate missing mass and excess bunching. They indicate how many additional firms there is per bin below the cutoff. 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 ($\hat{M}$) :

$$\begin{aligned}\hat{M} &= \sum_{i=z_T}^{z_U} \hat{c}_{iControl} - \hat{c}_{iTreat} \\ \hat{B} &= \sum_{i=z_L}^{z_T-1} \hat{c}_{iTreat} - \hat{c}_{iControl}\end{aligned} \tag{9}$$

(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.

z_L is the turnover level of the group of bins that precedes the bunch of ones with significant coefficients below threshold.[10] In our context, $47,600,000$ Francs is the lower end of the valley. This is an advantage of our 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 4 columns (1) and (2). We estimate $Z_U = 51,800,000$ Francs. The manipulation region is fairly symetric around the threshold. We find that in the two years there are 225 firms that bunch.

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 0.415 that is slightly lower than the mean of average bunching we found with the cross section estimations for years 1997 and 1998.

4.3 Identifying Bunchers' Characteristics

Conditional on identifying bunching in the firm size distribution, we are also interested in understanding what are the ex-ante characteristics of the firms that avoid taxes. Below, we descibe our approach, which builds on Diamond and Person (2016).

Using Predictions from Distribution away from the Threshold as Counterfactual In the language of potential outcomes, we can consider the firms below the cutoff in 1995 (before the policy reform) to be "always takers," and firms above to be both "never takers" and

10. We 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.

"compliers." The difference between the observed characteristics of firms above the cutoff and the no-manipulation counterfactual value is due to the compliers leaving (and similarly below the threshold). As a result, we use the differene between observed and counterfactual characteristics of firms to estimate the types of firms who change their size in response to the policy change. In particular, we estimate the characteristics of the bunchers ($Y^{compliers}$) as the average of the two methods :

$$\bar{Y}^{compliers} = 0.5 * \Big(\frac{N^{tot}_{down}}{N^{tot}_{down} - N_{down}} * \bar{Y}^{down\_all} - \frac{N_{down}}{N^{tot}_{down} - N_{down}} * \bar{Y}^{down}\Big) + 0.5 * \Big(\frac{N_{up}}{N_{up} - N^{tot}_{up}} * \bar{Y}^{up} - \frac{N^{tot}_{up}}{N_{up} - N^{tot}_{up}} * \bar{Y}^{up\_all}\Big) \quad (10)$$

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^{tot}_{up} (resp. N^{tot}_{down}) 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 addition to estimating the counterfactual characteristics of firms in the manipulation region using firms in the same year who are larger, we also can do so using the observed data in the pre-policy years. After the policy change, the compliers move from above to below the cutoff, in order to avoid excess taxation. The difference between the firms above the cutoff before and after the policy change is due to the compliers leaving, and similarly the difference below the cutoff is due to the compliers joining. The expression of bunchers' characteristics become :

$$\bar{Y}^{compliers} = 0.5 * \Big(\frac{N^{tot}_{down}}{N^{tot}_{down} - N_{down}} * \bar{Y}^{down}_{Treatment} - \frac{N_{down}}{N^{tot}_{down} - N_{down}} * \bar{Y}^{down}_{Control}\Big) + 0.5 * \Big(\frac{N_{up}}{N_{up} - N^{tot}_{up}} * \bar{Y}^{up}_{Control} - \frac{N^{tot}_{up}}{N_{up} - N^{tot}_{up}} * \bar{Y}^{up}_{Treatment}\Big) \quad (11)$$

Where $\bar{Y}^{down}_{Treatment}$ 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). $\bar{Y}^{down}_{control}$ 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 $\bar{Y}^{up}_{Treatment}$ 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. $\bar{Y}^{up}_{Control}$ 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^{tot}_{up} (resp. N^{tot}_{down}) 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).

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 ∈ [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 Diamond and Person (2016) notation this means we are interested in estimating :

$$\Delta\tilde{Y} = \tilde{Y}^{compliers} - \tilde{Y}^{up}_{Control} \tag{12}$$

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 :

$$\tilde{Y}_i = \beta \cdot compliers_i + \nu_r + \mu_s + \epsilon_i \tag{13}$$

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.

4.4 Measuring Adjustment Cost and Output Elasticities

One characteristic of firms that we are interested in the cost of adjustment. First, we can measure adjustment costs using the dispersion of the revenue share of capital within each sector. Optimally, marginal products (the production function elasticity) should equal marginal costs (the revenue share, as in Asker, Collard-Wexler, and Loecker (2014); Hall (1988). Adjustment costs can prevent equalization, and so the dispersion of the revenue share is a proxy for industry-level adjustment costs. Formally we define adjustment cost as :

$$Adjustment\ cost_{it} = SD_{it}\big(\beta_i + y_{jt} - k_{jt}\big), \tag{14}$$

where β_i refers to sector i output elasticity with respect to capital, y_{jt} refers to firm j log-level of production on year t and k_{jt} refers to firm j log-level of capital on year t.

Second, following a long tradition in production function estimation, e.g. Levinsohn and Petrin (2003) and Olley and Pakes (1996), we use production function elasticities themselves in order to measure the costs of adjustment. If materials are the most flexible input, the firms for whom the elasticity of output with respect to materials are the highest should have the lowest-cost in quickly adjusting their outputs (and firms with high capital elasticities the highest costs).

In order to estimate production function elasticities, we follow the Wooldridge (2009) adaptation of Levinsohn and Petrin (2003).

4.5 Identifying Bunchers' Choices

In addition to measuring the ex-ante characteristics of the bunchers, we are also interested in what they do in order to lower their sales below the threshold. We use both cross-sectional and time-series information in order to estimate this. However, we cannot compare only firms below the cutoff, or firms only above. This is because conditional on size firms may behave similarly. For instance, if sales were equal to materials inputs, then firms manipulating their sales would correspondingly lower their materials. As a result, conditional on size, compliers would have the same input choices as never-takers, and so the previous estimation strategy would estimate no effect. Note that this isn't an issue for ex-ante charactersics, which are not directly affected by firm choices. To estimate firm choices, we compare all of the firms in the manipulation region to their counterfactual counterparts.

Using Past Years as Counterfactual 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 substracting from it the mean of the outcome of interest in the region neighboring the manipulation region (i.e. the regions just below and just above). Our intent to treat estimate simplifies to :

$$ITT = E(\tilde{Y}|firms\ manipulate) - E(\tilde{Y}|firms\ don't\ manipulate) = E^{Treat}_{Manip}(\tilde{Y}) - E^{Control}_{Manip}(\tilde{Y}) \tag{15}$$

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 from Distribution away from the Threshold as Counterfactual We also follow Diamond and Person (2016) 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 :

$$ITT = E(\tilde{Y}|firms\ manipulate) - E(\tilde{Y}|firms\ don't\ manipulate) = E^{Observed}_{Manip}(\tilde{Y}) - E^{Predicted}_{Manip}(\tilde{Y}) \tag{16}$$

5 Empirical Analysis

In this section we describe the results of our data analysis. First we show the existence of bunching in the firm size distribution consistent with the theory. We then describe the characteristics of firms who bunch and then the way that they do so.

5.1 Excess Mass

Figure 2 shows the raw firm size distribution around the tax cutoff. Before the tax reform (in 1995 and 1996) and after (in 2000) there is no visual break in the firm size distribution, but it is clearly visible in 1997-1999. We calculate a counterfactual distribution far away from the cutoff as a solid line. The vertical lines show the excess mass (and under mass) for those years, where the extra mass on the left is equal to the undermass on the right. Figure 3 shows, just for the years 1997-1998. The observed difference in densities around the cutoff relative to the pre-reform years of 1995-1996.

Table 3 presents discontinuity estimates from McCrary (2008) estimation technique which show the bunching only appear in the year with the discrete jumps in the marginal tax rates. Table 4 shows the size of the manipulation region, where both for the cross sectional and panel estimates we find excess mass of 150 firms just below the increase tax rate.

5.2 Characteristics of Excess Mass

In this subsection we describe they types of firms who shrink their size in order to avoid paying taxes. First we show differential bunching by profit level. Since the tax rate is on profits, firms who, e.g., have no profts should not be affected by the policy. Consistent with this, we see in the right panel of Figure 5 that the lowest profitability firms do not demonstrate excess bunching in any of the years (while in 1997 there is a spike to the left of the cutoff, there is no corresponding valley to the right). For the most profitable firms, however, we do see manipulation, consistent with the theory. In Figure 6 we use firm profitability in 1995 (instead of in the current year) and find a similar result.

In Figure 7 we run the same exercise, but using adjustment costs (as measured using (Asker, Collard-Wexler, and Loecker, 2014)). The results are less clean than for profits, but again consistent with theory that firms with the lowest adjustment costs bunch the most.

Table 5 runs the estimation equation 4, and finds consistently that the higher-profit and lower capital adjustment cost firms show more bunching

5.3 Characteristics of Compliers

An alternative approach to measuring who bunches is to instead look within the manipulation region (Diamond and Person, 2016). In Table 76 we report the corresponding estimates using respectively predictions from firms away from the threshold and previous years as counterfactual. In the first rows, we extend the results from the previous subsection : compliers are more likely to have low adjustment costs and high profits. Results are robust to the inclusion of region and sector fixed effects as shown in the columns headed with "FE". We can also examine the production function characteristics of the compliers : they are more likely to have a lower capital elasticity and a higher materials elasticity This is consistent with the oft-stated argument that materials inputs are more flexible (Ackerberg, Caves, and Frazer, 2015; Asker, Collard-Wexler, and Loecker, 2014; Levinsohn and Petrin, 2003; Olley and Pakes, 1996) : the types of firms who find it easier to manipulate their sales are those whose output is more responsive to materials. As a robustness check we run the same analysis on the sub sample of firms that have input shares below 1 and for which the sum of the input shares is lower than 2. Tables C.2 and C.1 show that our results are stable to this restriction.

5.4 Behavior of Compliers

We can undertake a similar exercise to show the behavioral changes of the compliers (Tables 10 and 9). Here we consider the entire manipulation region relative to it's counterfactual prediction either in the cross section or the panel. Not surprisingly, we see that turnover is lower : this is the direct effect of manipulation. Sold production falls by more than output does : inventories and capitalized production are higher for the firms adjusting their sales. As a robustness check we run the same analysis on the sub sample of firms that have input shares below 1 and for which the sum of the input shares is lower than 2. Tables C.4 and C.3 show that our results are stable to this restriction.

6 Discussion

In this paper we describe how firms respond to increases in the profit tax. The nature of our setting, the introduction of a new tax above a specific sales threshold, allows us to estimate many important parameters determining firm responses. First, we use the firms who manipulate their sales below the threshold to estimate the tax elasticity of sales. We find a value of 0.16. Second, we describe the characteristics of those firms : those with high profits (and therefore high incentives to lower their tax rate) and those with low adjustment costs, measured in several ways. Third, we show that firms adjust their turnover by lowering production and increasing inventories.

Références

Ackerberg, D., K. Caves, and G. Frazer. 2015. "Identification Properties of Recent Production Function Estimators." Econometrica 83 :2411–2451.

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. .

Bachas, P., and M. Soto. 2015. "Not(ch) Your Average Tax System : Corporate Taxation Under Weak Enforcement." Working Paper, pp. .

Barbanchon, T.L. 2016. "Optimal Partial Unemployment Insurance : Evidence from Bunching in the U.S." Working Paper, 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. .

Caliendo, L., F. Monte, and E. Rossi-Hansberg. 2015. "The Anatomy of French Production Hierarchies." Journal of Political Economy 123.

Cattaneo, J., and Ma. 2016. "Simple Local Regression Distribution Estimators with an Application to Manipulation Testing." Working Paper, pp. .

Desai, M.A., C.F. Foley, and J. Hines. 2006. "The demand for tax haven operations." Journal of Public Economics 90 :513–531.

Diamond, and Person. 2016. "The Long-Term Consequences of Teacher Discretion in Grading of High Stakes Test." 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.

Gourio, F., and N. Roys. 2014. "Size-dependent regulations, firm size distribution, and reallocation." Quantitative Economics 5 :377–416.

Hall, R. 1988. "The Relation between Price and Marginal Cost in U.S. Industry." Journal of Political Economy 96 :921–47.

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. .

Lardeux, R. 2018. "Who Understands The French Income Tax ? Bunching Where Tax Liabilities Start." Working Paper, pp. .

Levinsohn, J., and A. Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." Review of economic studies, pp. .

Liu, L., and B. Lockwood. 2016. "VAT notches, voluntary registration and bunching : Theory and UK evidence." Working Paper, pp. .

Liu, L., T. Schmidt-Eisenlohr, and D. Guo. 2017. "International Transfer Pricing and Tax Avoidance : Evidence from Linked Trade-Tax Statistics in the UK." CESifo Working Paper Series No. 6594, CESifo Group Munich.

McCrary, J. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design : A Density Test." Journal of Econometrics 142.

Olley, G.S., and A. Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." Econometrica 64 :1263–97.

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

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.

Saez. 2010. "Do Taxpayers Bunch at Kink Points ?" American Economic Journal : Economic Policy, pp. .

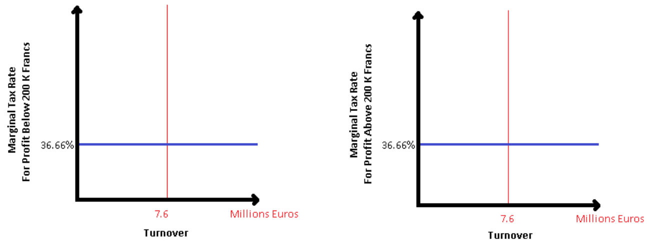
Thomas Dee, B.J., Will Dobbie, and J. Rockoff. 2016. "The Causes and Consequences of Test score Manipulation Evidence from the New York Regents Examinations." Working Paper, pp. .

Wooldridge, J. 2009. “On estimating firm-level production functions using proxy variables to control for unobservables.” Economic letters, pp. .

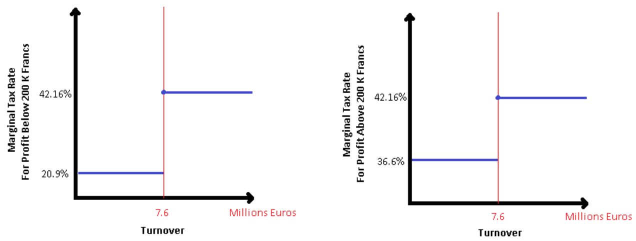
Zucman, G. 2014. “Taxing across Borders : Tracking Personal Wealth and Corporate Profits.” Journal of Economic Perspectives 28.

A Figures

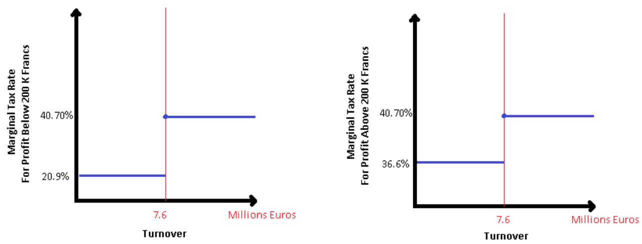
Figure 1 – Evolution of the Tax Schedule



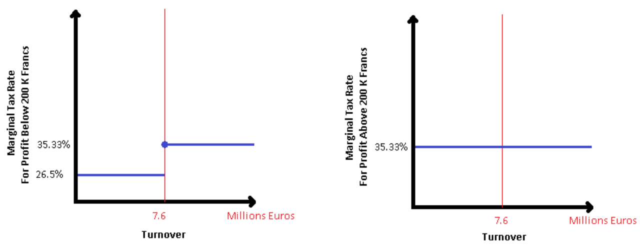
(a) 1995-1996



(b) 1997-1998



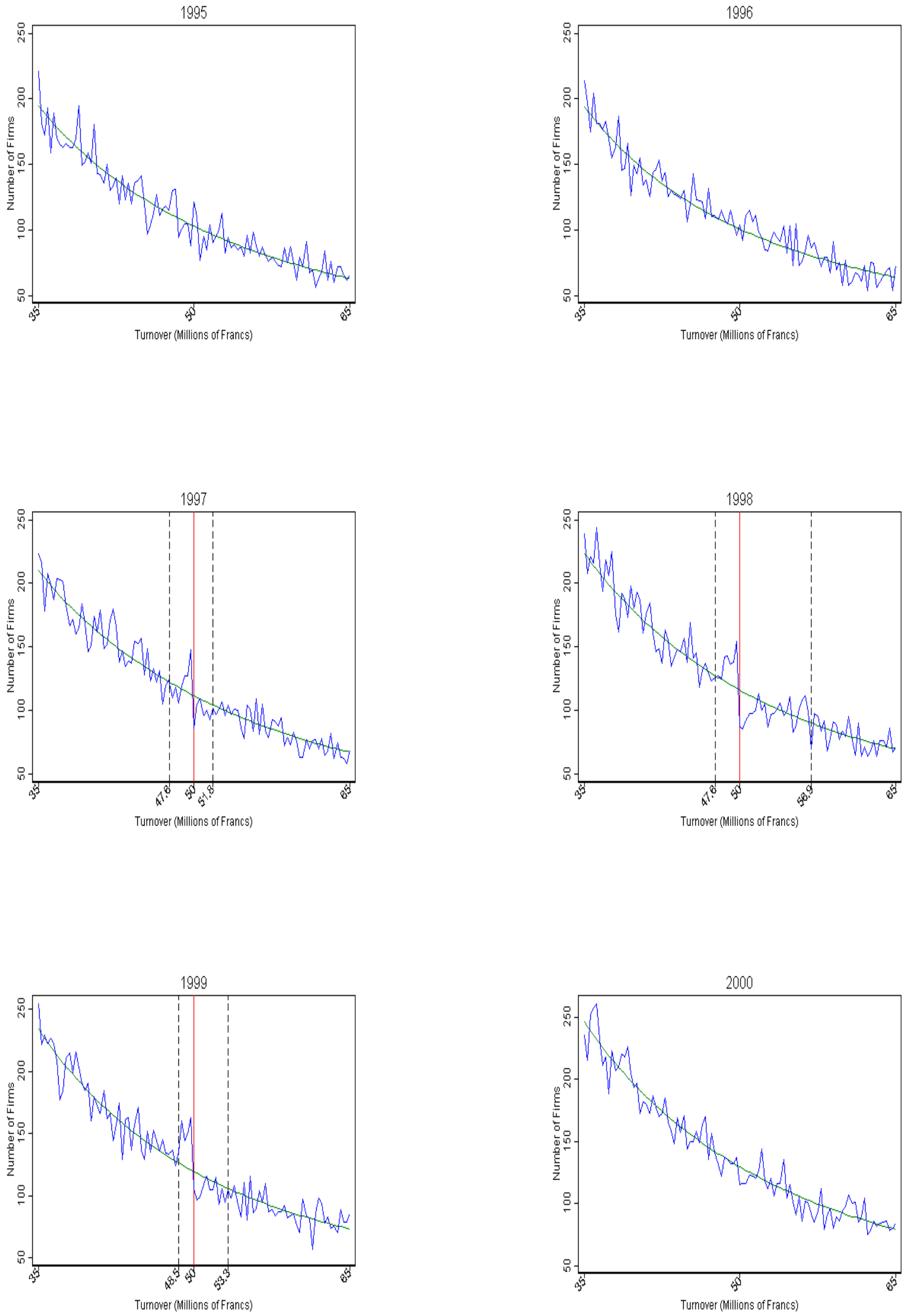
(c) 1999



(d) 2000

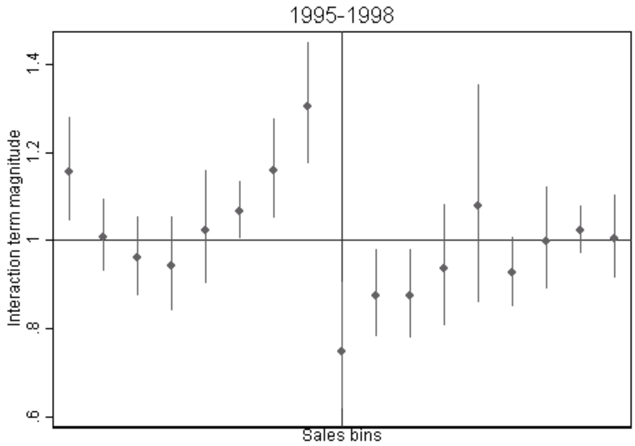
This figure presents the marginal rates of firms eligible to either the reduced corporate income tax or not. The marginal tax rates are the sum of three terms : the *contribution exceptionnelle*, the *contribution additionnelle* and the corporate income tax rate.

FIGURE 2 – A transient discontinuity in firms' sales distribution



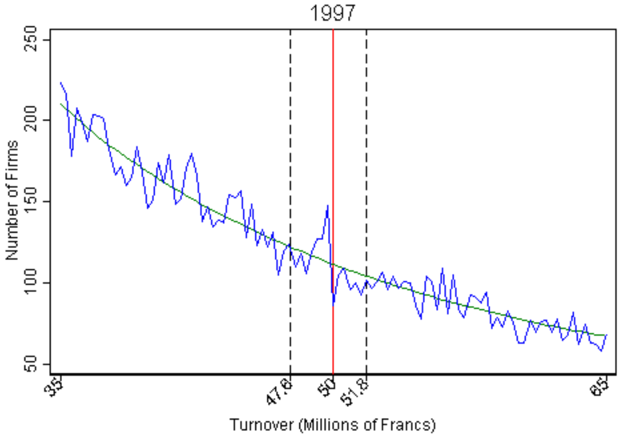
The distribution of firms with sales between kFrancs 20 and kFrancs 100, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate.

Figure 3 – Representing bunching compared to counterfactual obtained from years during which there was no incentive to bunch

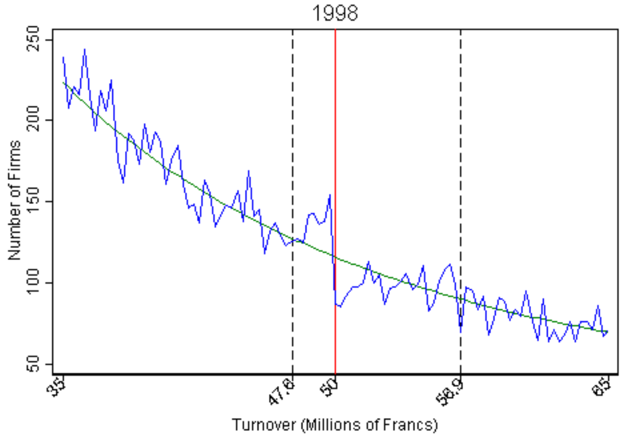


The figure plots the differential number of firms in bins around the threshold in years during which there was incentives to bunch compared to years during which there was no incentives. Each point comes from an interaction term between the bin indicator and the indicator of incentives. 95 % confidence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

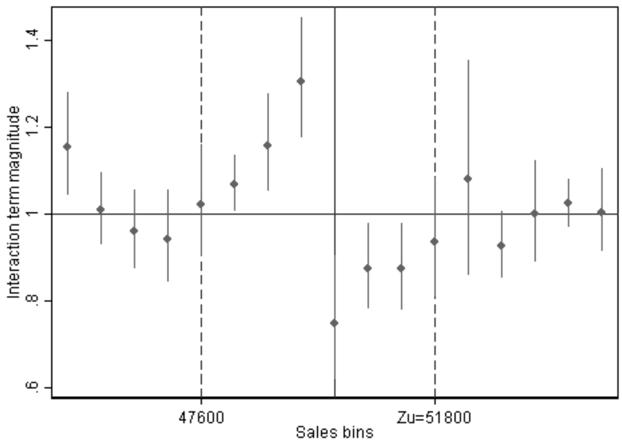
Figure 4 – Estimating the width of the valley



(a) Cross section 1997



(b) Cross section 1998

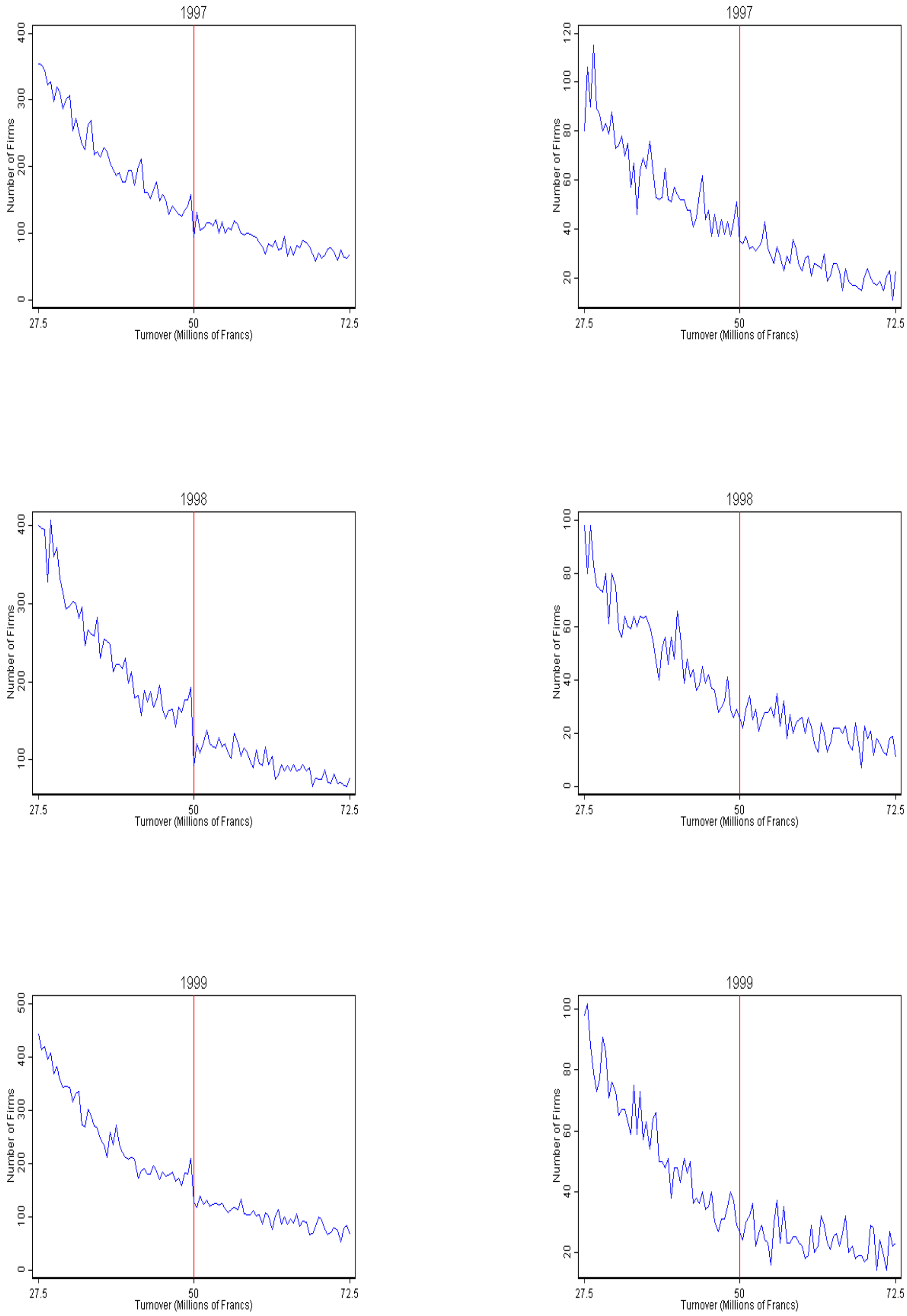


(c) Time series 1997-1998

The figure reports the manipulation region. It illustrates the lower end z_L that is determined either eyeballing where the distribution starts being different from the counterfactual (a) and (b) or as the lower end of the group of bins that precede the bunch of coefficients that are significant below threshold (c). z_U is determined from the equality of missing mass and excess bunching as reported in Table4

FIGURE 5 – Differential bunching by profit level

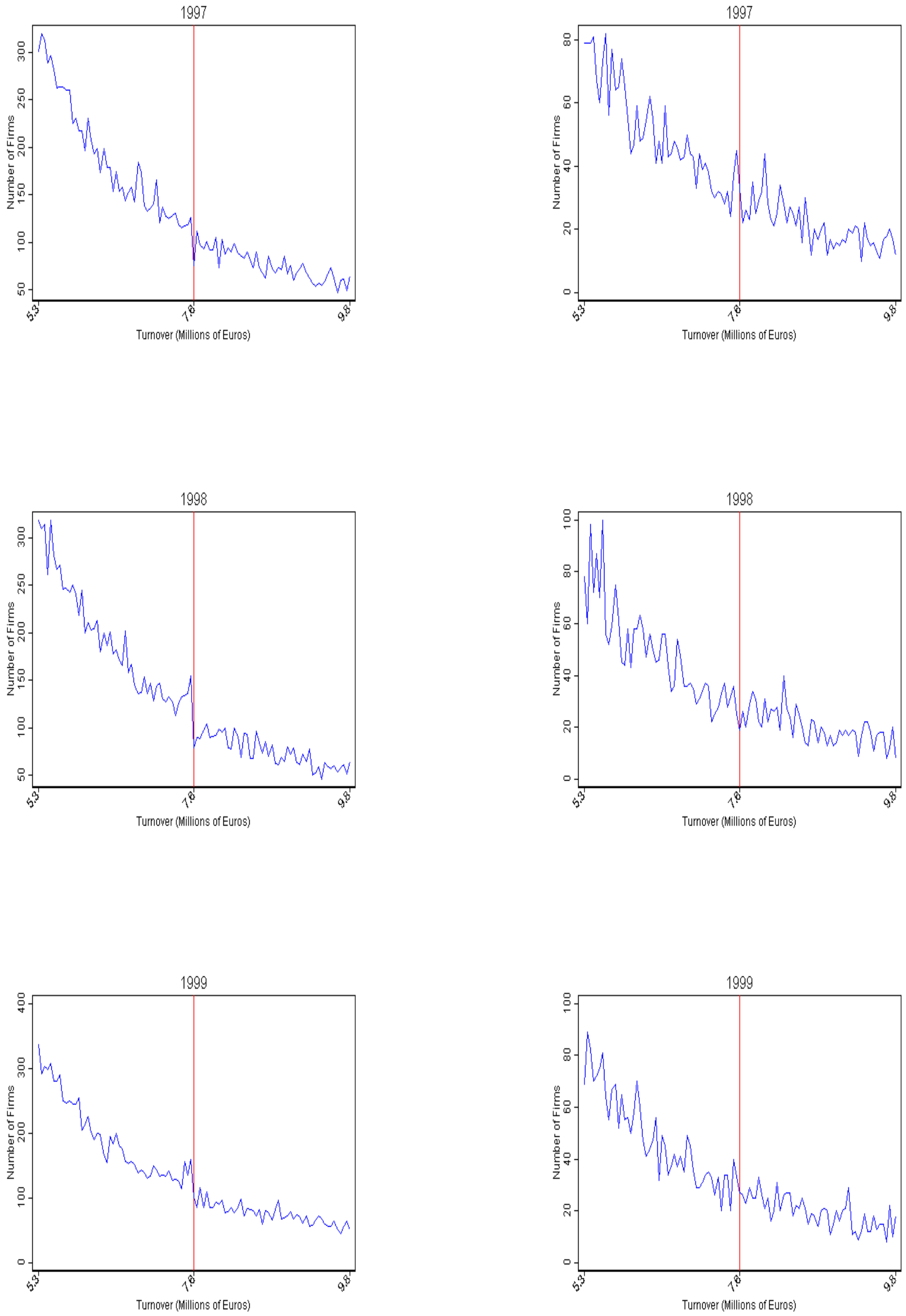
Top tercile | Bottom tercile



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.

Figure 6 – Differential bunching by profit level in 1995

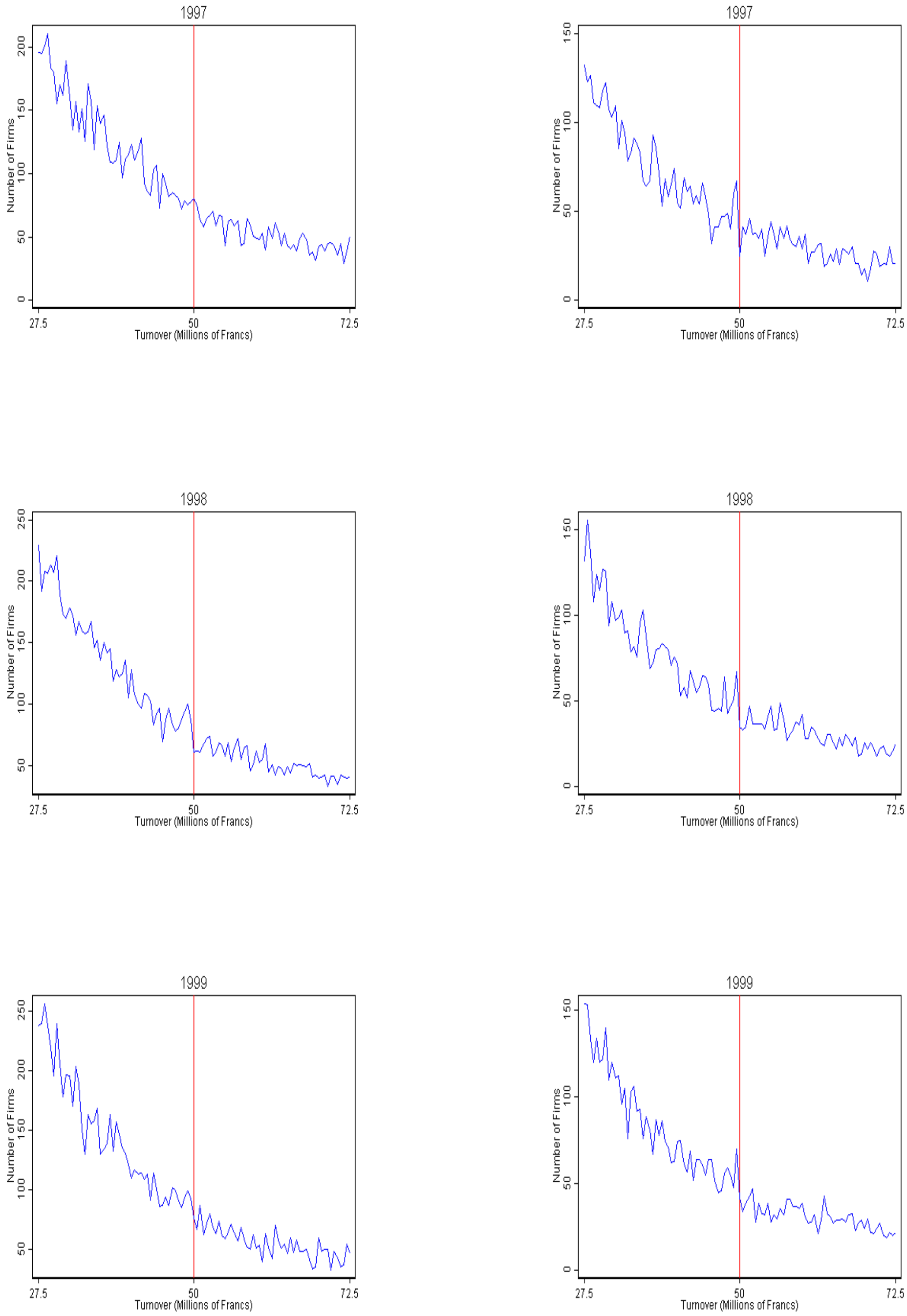
Top tercile Bottom tercile



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, by level of profit.

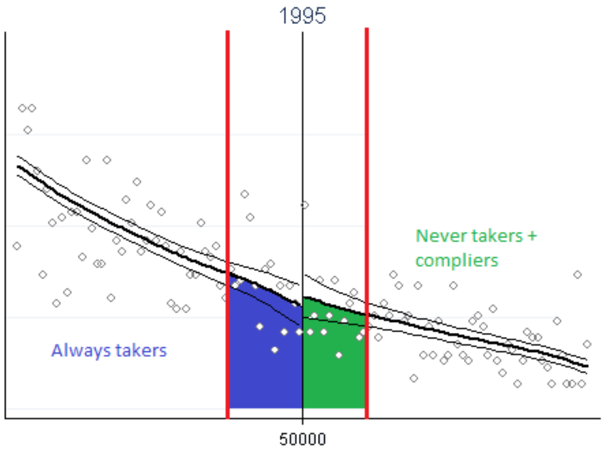
FIGURE 7 – Differential bunching by Adjustment cost

Top tercile | Bottom tercile

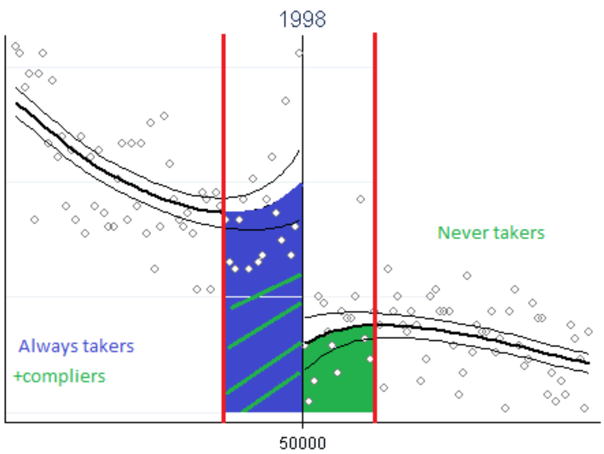


The distribution of firms with sales between kFrancs 20 and kFrancs 100, restricting to firms that are paying corporate income tax and excluding firms that belong to a conglomerate, by level of adjustment cost measured as in (Asker, Collard-Wexler, and Loecker, 2014).

FIGURE 8 – Assessing the Characteristics of the Bunchers



(a) Control



(b) Treatment

These graphs borrow on Diamond and Person (2016). The additional tax can be interpreted as a treatment to which bunching firms are compliers.

B Tables

Table 1 – Share of Firms Affiliated to the Different Tax Regimes

	Obs.	Share					VA value	VA share				
	Nb	BNC	BRN	RSI	IS	IR	B Eur	BNC	BRN	RSI	IS	IR
1995	2,034,117	0.19	0.32	0.49	0.33	0.67	611.33	0.05	0.89	0.06	0.78	0.22
1996	2,226,769	0.18	0.31	0.51	0.33	0.67	627.47	0.05	0.88	0.07	0.80	0.20
1997	2,262,301	0.19	0.30	0.51	0.34	0.66	645.21	0.05	0.88	0.07	0.82	0.18
1998	2,297,619	0.19	0.30	0.51	0.35	0.65	719.77	0.05	0.89	0.07	0.84	0.16
1999	2,323,909	0.20	0.30	0.51	0.35	0.65	792.03	0.04	0.89	0.06	0.80	0.20
2000	2,325,726	0.19	0.30	0.50	0.36	0.64	821.33	0.04	0.89	0.06	0.81	0.19

Table 2 – Descriptive statistics

	mean	count	sd
Sample			
Output	20160.22	191,511	21679.97
Turnover	39546.35	203,609	18869.56
Profit	1304.92	203,609	9377.56
	mean	count	sd
All firms			
Output	8463.42	8,510,565	1022583.07
Turnover	10060.38	10,682,393	925309.87
Profit	1028.11	10,682,393	277244.24

Note: The sample is restricted to eligible firms paying a corporate income tax and with turnover between 20,000 Francs and 100,000 Francs. We also drop observations with capital share larger than 10.

Table 3 – Disontinuity 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 4 – Bunching estimators

	$\hat{B}$	$\hat{M}$	$\hat{b}_{av}$
Panel A : Cross-Section Estimates			
1997	49.724	60.947	0.958*
			(0.633)
1998	114.278	134.700	2.198***
			(0.587)
1999	86.100	95.980	1.511***
			(0.370)
Panel B : Time-series Estimates			
1997-1998	204.264	225.945	0.415

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

Table 5 – Bunching Estimation by Subgroups

	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)

Note: This table reports the bunching estimators for different subgroups of firms.

TABLE 6 – Characteristics of the compliers : Panel

Adjustment cost, Incentives, Ability to bunch						
	Low adjustment cost of capital		Large profit		Large profit in 1995	
	(1)	(2) FE	(3)	(4) FE	(5)	(6) FE
Compliers	0.0364**	0.0341**	0.00519	0.00144	-0.00906	-0.0123
	(0.0152)	(0.0133)	(0.0133)	(0.0133)	(0.0133)	(0.0103)
Observations	6344	6277	6405	6338	5424	5378
Production function characteristics						
	Large elasticity wrt K		Large elasticity wrt L		Large elasticity wrt M	
[1em] Compliers	-0.0404***	-0.0387***	0.00849	0.00890	0.0363**	0.0354**
	(0.0144)	(0.0135)	(0.0131)	(0.0139)	(0.0152)	(0.0142)
Observations	6344	6277	6344	6277	6344	6277

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 Diamond and Person (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 : Cross-Section

Adjustment cost, Incentives, Ability to bunch						
	Low adjustment cost of capital		Large profit		Large profit in 1995	
	(1)	(2) FE	(3)	(4) FE	(5)	(6) FE
Compliers	0.0364**	0.0341**	0.00519	0.00144	-0.00906	-0.0123
	(0.0152)	(0.0133)	(0.0133)	(0.0133)	(0.0133)	(0.0103)
Observations	6344	6277	6405	6338	5424	5378
Production function characteristics						
	Large elasticity wrt K		Large elasticity wrt L		Large elasticity wrt M	
[1em] Compliers	-0.0404***	-0.0387***	0.00849	0.00890	0.0363**	0.0354**
	(0.0144)	(0.0135)	(0.0131)	(0.0139)	(0.0152)	(0.0142)
Observations	6344	6277	6344	6277	6344	6277

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 Diamond and Person (2016) technique. 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 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) FE	(3)	(4) FE	(5)	(6) FE	(7)	(8) FE
Compliers	0.205**	0.154*	0.270***	0.222***	0.181***	0.147***	0.358***	0.324**
	(0.0797)	(0.0834)	(0.0550)	(0.0588)	(0.0477)	(0.0553)	(0.0671)	(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	-0.0171	-0.0118	-0.0417	0.191**	0.155*
	(0.0695)	(0.0831)	(0.0789)	(0.0907)	(0.0752)	(0.0831)
Observations	906	906	906	906	906	906

Table 9 – Consequences of manipulation : panel

	(1) Turnover	(2) Y	(3) Sold production	(4) Change in inventories	(5) Capitalized production
_bs_1	-73.65**	-443.4	-617.5*	85.83*	88.32***
	(33.98)	(327.9)	(325.6)	(50.59)	(21.99)
Observations	4809	4809	4809	4809	4809

Note: This table estimates the consequences of bunching

Table 10 – Consequences of manipulation : cross section

	(1)	(2)	(3)	(4)	(5)
	Turnover	Y	Sold production	Change in inventories	Capitalized production
_bs_1	-7.59496e+22***	-4591.6***	-4750.3***	150.0**	8.700
	(1.89253e+21)	(636.6)	(625.3)	(64.21)	(41.28)
Observations	5685	2274	2274	2274	2274

Note: This table estimates the consequences of bunching

C Appendix

TABLE C.1 – Characteristics of the Compliers : Panel

Adjustment cost, Incentives, Ability to bunch					
	Low adjustment cost of capital		Large profit		Large p in 199
	(1)	(2) FE	(3)	(4) FE	(5)
Compliers	0.0495***	0.0464***	0.0277*	0.0237*	0.00351
	(0.0174)	(0.0160)	(0.0146)	(0.0143)	(0.0165)
Observations	4439	4386	4494	4441	3881
Production function characteristics					
	Large elasticity wrt K		Large elasticity wrt L		Large ela wrt M
[1em] Compliers	-0.0526***	-0.0492***	-0.00200	-0.00163	0.0598***
	(0.0184)	(0.0187)	(0.0188)	(0.0173)	(0.0185)
Observations	4439	4386	4439	4386	4439

Note: This table estimates the different characteristics of the bunchers compared to other firms that were eligible to bunching. The sample is restricted to firms without extreme values of input shares. 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 adapting Diamond and Person (2016) technique to the time series setting. 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 C.2 – Characteristics of the Compliers : Cross-Section

Adjustment cost and Incentives to bunch

	Low adjustment cost of capital		Large profit		Large profit in 1995		Large profit in 1996	
	(1)	(2) FE	(3)	(4) FE	(5)	(6) FE	(7)	(8) FE
Compliers	0.205**	0.154*	0.270***	0.222***	0.181***	0.147***	0.358***	0.324*
	(0.0797)	(0.0834)	(0.0550)	(0.0588)	(0.0477)	(0.0553)	(0.0671)	(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	-0.0171	-0.0118	-0.0417	0.191**	0.155*
	(0.0695)	(0.0831)	(0.0789)	(0.0907)	(0.0752)	(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 Diamond and Person (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 C.3 – Consequences of manipulation : panel

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

Note: This table estimates bunchers' choices. The sample is restricted to firms without extreme values of input shares.

Table C.4 – Consequences of manipulation : cross section

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

Note: This table estimates bunchers' choices. The sample is restricted to firms without extreme values of input shares.