CESifo
CONFERENCES 2021

12th Norwegian German Seminar on Public Economics

Munich, 5 – 6 November 2021

Bunching with multiple response variables: Estimating firm response to thin capitalization rules

Martin Eckhoff Andresen and Lars Thorvaldsen

CESifo NHH OFS Oslo Fiscal Studies

Bunching with multiple response variables: Estimating firm response to thin capitalization rules

Martin Eckhoff Andresen* Lars Thorvaldsen†

Preliminary version: November 3, 2021
(Do not circulate)

Abstract

We expand the bunching framework and develop simple reduced form methods for estimating how bunchers respond along a range of endogenous variables. We show that it is possible to back out means among bunchers and their counterfactual allocation for any outcome of interest by extrapolating from the relationship between the outcome and the running variable into the bunching region. We use this framework to examine how bunching firms adjust to the thin capitalization rules introduced in Norway in 2014. We find that firms bunch at all thresholds generated by the rule and that the bunching is driven by a restructuring of debt and equity and leads to an average reduction in the debt-to-asset ratio of 20 percent among the bunching firms.

Keywords: Bunching, capital structure, debt bias

JEL Codes: H00, C01, H25, G30

*Statistics Norway, martin.andresen@ssb.no
†University of Oslo, lars.thorvaldsen@econ.uio.no

1 Introduction

At odds with one of the most fundamental principles of tax neutrality, most countries grant deductions of interest cost while the user cost of equity is a non-deductable expense. This debt bias is increasing in the corporate tax rate and lowers the after-tax price of debt relative to equity and thus incentivize firms to choose a higher leverage ratio than they otherwise would. Overleveraged firms has been identified by policymakers as a key macroeconomic stability concern because they are believed to amplify economic distress during financial crisis and economic downturns (IMF, 2016).

Although it is well established in the literature that corporate taxes affects firms' capital structure, the order of magnitude is still in dispute. The key issue with identification of the effect has been a lack of exogenous variation in tax rates, with most empirical research relying on variation over time or across countries to identify the effects of corporate taxes on firm capital structure (de Mooij, 2011). This type of variation may suffer from correlation with unobserved determinants of firm capital structure and may even be endogenous to firm capital structure itself, causing trouble for causal interpretation of resulting estimates.

In this paper, we improve the current knowledge of how firms respond to the debt bias in the tax scheme by exploiting the incentives to bunch at the threshold for the thin capitalization rule introduced in Norway in 2014. This setting is particularly useful for the purpose because it involves a much larger shift in the debt bias incentives compared to a small reduction in the corporate income tax rate, and because we have access to firms tax returns which have been shown to differ quite substantially from public financial statements that is most frequently used in the literature (Devereux et al., 2018).

We first document that firms respond to the new regulation by reducing net interest cost to bunch at the thresholds where the rule applies. Depending on the sample and threshold, we find that firms on average reduce interest costs or the ratio of interest costs to EBITDA by between 5-9 percent. Next, we try to understand the anatomy of the bunching response by expanding the bunching framework to analyze adjustments in other outcomes than the bunching variable. Building on ideas in Diamond and Persson (2016), we formalize how we can back means among bunchers and their counterfactual for any outcome of interest by extrapolating from the observed relationship between the outcome and the bunching variable. In cases where the outcome measure is predetermined or unmanipulable by the agent, this can be used to characterize bunchers. In cases where the outcome variable is endogenous and adjusted in response to the incentive change, the difference between the mean of bunchers and their counterfactual may be interpreted as a measure of a local treatment effect for bunchers. The intuition for the approach can be explained as follows. The bunching region is populated by bunchers and non-bunchers, that is, responding agents and agents that would allocate in the region independent of the incentive change. Extrapolation from unaffected firms into the bunching region allow us to back out the mean for non-bunchers in the bunching region and the counterfactual mean for bunchers. Because the mean in the bunching region is a weighted average of bunchers and non-bunchers outcomes, and the number of bunchers are already estimated by standard bunching estimators, it is straightforward to back out the mean for bunchers.

Using this methodology we find that the bunching response in interest costs is accompanied by a conversion of debt to equity and an offloading of internal corporate group debt to the bunching entity which allows bunching firms to avoid the thin capitalization rules. The changes leads to an average reduction in the debt-to-asset ratio of 20 percent among bunching firms which is substantially higher than previous estimates in the literature (Buslei and Simmler, 2012; Alberternst and Sureth, 2015).

Because reductions in debt to other members of the corporate group is a crucial response of the bunching firms, we next analyze total capital structure measures at the level of the corporate group. We aggregate capital structure outcomes on the corporate level and find that the observed effects among bunching firms are large enough to also impact the group level. In particular, we estimate that the corporate groups in the sample significantly reduce the debt-to-asset ratio with 15 percent. This finding implies that the Norwegian earnings stripping rule could mitigate the debt bias inherent in the corporate income tax scheme, contrary to the belief in the literature.

This paper contributes to several strands of literature. First, we contribute to the literature on the tax sensitivity of capital structure and corporate debt. Since Modigliani and Miller formulated the capital structure irrelevance theorem in 1958 (Modigliani and Miller, 1958), there has been a debate on which principles that guide the financing decision. The two most influential theories have been the trade-off theory (starting with Kraus and Litzenberger (1973)), in which firms balance costs and gains by the use of debt, and the pecking order theory (starting with Myers (1984)), in which firms refrain from issuing equity to avoid signaling that the company shares are over-valued. While both theories have been used with some success to explain observed variation in the firms debt-to-assets ratio, none of them are fully consistent in explaining variation across different types of firms and across industries (Graham and Leary, 2011). Nonetheless, there seems to be a consensus in the literature that the corporate income tax does matter, but its importance is disputed. While earlier studies found that a one percentage point increase in the marginal tax lead to a 0.28 percentage point increase in the debt-to-assets-ratio (de Mooij, 2011; Feld et al., 2013) later studies have found much larger effects ranging from 0.41 (Faccio and Xu, 2015) to 0.53 (Heider and Ljungqvist, 2015).

Second, we connect to the relatively new literature on thin capitalization rules, introduced in many European countries during the last decade. For long, the most popular rule has been the so-called safe harbor rule, in which deductions are removed above a cutoff in the debt-to-equity-ratio. In recent years, however, many countries have switched to limiting deductions relative to a percentage share of the firms' EBITDA, frequently referred to as the earnings stripping rule (examples include Germany, Finland and Norway) (Ruf and Schindler, 2015). A large number of studies have looked at the effects of the safe harbor rule, and found that firms respond by reducing the debt-to-assets ratio (see for example (Buettner et al., 2012, 2016; Overesch and Wamser, 2010)). At the same time, the response is found to mostly concern internal debt. Buettner et al. (2018) finds that there is an association between the introduction of the safe harbor rule and a reduction in FDI if introduced in countries with above average CIT rate. Only a handful of studies have analyzed the earnings stripping rule, and results have been mixed. Harju et al. (2017) study the Finnish case and find that firms respond by decreasing financial costs, but no effect on the debt level. This is in contrast to studies of the German case, in which it is found that firms respond by reducing the debt-to-assets ratio (Buslei and Simmler, 2012; Alberternst and Sureth, 2015). A recent study of the Norwegian earnings stripping rule finds that firms respond by reducing internal debt without substitution towards external debt, leading to an increase in the CIT tax revenue (Andresen and Kvamme, 2019). In contrast to these studies we rely on bunching methods rather than differences-in-differences between exposed and unexposed firms.

Third, we add to the vast literature using the bunching methodology, pioneered by contributions from Saez (2010), Chetty et al. (2011) and Kleven and Waseem (2013). While many researchers have used the bunching framework in later years, there have been few attempts to expand the methodology. A notable exception is Diamond and Persson (2016) who uses observed bunching in scores after national test in Sweden to analyze the effect of benevolent teacher grading on earnings. To obtain causal estimates, they use teachers manipulation of grades as an instrument together with estimates of counterfactual distributions to obtain averages among bunchers and counterfactual means in the bunching region. Homonoff et al. (2020) later adopted this framework in a bunching setting more similar to the current study. In this paper, we build on and formalize these contributions to a standard bunching setting. We also refine the estimation procedure of the counterfactual distribution above the threshold, first developed by Chetty et al. (2011), relevant for settings in which the incentive change applies for all agents crossing the threshold (e.g. kinks and proportional notches). In addition, we connect with a growing subgroup of papers using the bunching approach in a firm setting. At this point, this paper is most closely related to

Bachas and Soto (2018) who develop a bunching framework in a setting with Costa Rician firms facing a shift in average corporate taxes at different revenue levels, to estimate the elasticity of taxable income. Other notable studies that have applied a bunching framework in a firm setting includes Coles and Smith (2019); Devereux et al. (2014); Boonzaaier et al. (2019).

The paper proceeds as follows. In section 2 we outline the reduced form estimation procedure for both the bunching variable and other outcomes of interest and detail our refinement of the adjustment to the bunching variable above the cutoff. Section 3 describes the institutional setting for the application of the estimation method to the introduction of thin-cap rules. In section 4 we show the data and sample selection, and section 5 shows the estimation results. Section 6 concludes.

2 Reduced form bunching estimation

In reduced form bunching applications, the goal is commonly to estimate the excess number of agents allocating at a particular point or in a region of the distribution of the running variable z, to avoid the consequence of some incentive change at the threshold z^*. In the canonical labor supply example, the marginal tax rate on earnings changes discontinuously, generating a kink (Saez, 2010; Chetty et al., 2011) or the average tax rate changes, generating a notch (Kleven and Waseem, 2013). In order to estimate the excess number of agents, referred to as bunchers, we need to provide an estimate of the counterfactual distribution of the running variable in the absence of the change in incentives that happen at the cutoff z^*.

The estimation procedure is guided by some theoretical assumptions, routinely applied in the literature. First, it is assumed that bunching agents only come from one side of the threshold. In the standard labor case, this implies that no agent below the threshold respond to the change by increasing earnings to get closer to the cutoff. Second, it is assumed that the population size stays constant before and after the change, i.e., no extensive margin effects. This does not rule out entry or exit of agents, but rules out entry or exit caused by the cutoff. Third, it is crucially assumed that the distribution of the running variable is smooth in the absence of the incentive change at z^*.

In this setting, estimation of the counterfactual is complicated by two problems: First, the distribution might be distorted by the presence of bunchers around the cutoff, and secondly the observations above the cutoff might be distorted because agents in this region are subject to a different set of incentives than under the counterfactual. To address the first issue, we follow the literature (Saez (2010); Chetty et al. (2011); Kleven and Waseem (2013), see also Bertanha et al. (2021)) and estimate the density of the running variable z outside of the bunching region as a polynomial, excluding the region that is contaminated by bunching agents. Start by choosing a bandwidth ω and separate observations into bins centered at the cutoff, so that bin j contains observations with $[z^* + j\omega < z_i \leq z^* + (j+1)\omega]$, where negative numbers j index bins to the left of the cutoff and positive numbers index bins to the right of the cutoff. We use m_j to refer to the midpoint of bin j. Before proceeding, we visually inspect the distribution of z to determine the number of excluded bins L and H below and above the cutoff so that $z_L = z^* - \omega L$ and $z_H = z^* + \omega H$, respectively.[1] Then estimate a polynomial density, excluding the bunching segment where $z_L < z \leq z_H$:

$$c_j = \sum_i \mathbb{1}[z^* + j\omega < z_i \leq z^* + (j+1)\omega] = \sum_{p=0}^{P} \beta_p m_j^p + \epsilon_j$$

This regression estimates the counterfactual density in the case of no change at the cutoff as a polynomial function, excluding observations in the bunching window where the density may be contaminated by the presence of bunchers.

The second problem occurs because agents have incentives to reduce the use of z above z^*, even when they don't find it optimal to bunch at z^*, simply because they are exposed to the change in incentives. As an example, this incentive is present in the standard labor

[1] Data-driven procedures for this exist, see e.g. (Bosch et al., 2020). We do not pursue this in this paper.

supply case with a kink studied by Chetty et al. (2011), as well as in the thin capitalization rules studied in this paper. In these cases, the observed value of z for agents who allocate above the cutoff does not represent the choice in the counterfactual scenario, but rather something smaller. In the standard labor supply model used in bunching applications as well as in the thin capitalization rules application in this paper, z will be depressed by a constant factor for all agents above the cutoff.

To address this distortion, Chetty et al. (2011) iteratively adjust the frequency in all bins above the cutoff until the bunching mass equals the missing mass above the cutoff. Intuitively, the density above the cutoff must be adjusted upwards until the excess mass in the bunching region matches the missing mass above, because this is where the bunching agents came from. To implement this, Chetty et al. (2011) therefore estimate

$$(1 + \mathbb{1}(m_j > z^*)\kappa)c_j = \sum_{p=0}^{P} \beta_p m_j^p + \epsilon_j$$

, excluding the bunching region, and iterate by slightly increasing κ until the following integration constraint holds

$$\int_{z_L}^{z_{\max}} h_0(z)dz = \omega \sum_i \mathbb{1}[z_i > z_L]$$
$$\implies \sum_{p=0}^{P} \frac{\beta_p}{p+1}(z_{\max}^{p+1} - z_L^{p+1}) = \omega \sum_i \mathbb{1}[z_i > z_L] \quad (1)$$

where i indexes agents and the right hand side is simply the observed number of agents with running variable above the lower limit z_L. Intuitively, the area under the estimated counterfactual density from z_L to $z_{\max}$ must be equal to the total number of agents observed in this region, because no agent has incentives to respond by moving out of this region.

This procedure is inconsistent with most theoretical models that are used to explain the underlying incentives to reduce z when allocating above the threshold. The simple labor supply model in Saez (2010) and Chetty et al. (2011), for example, predicts that the observed z is depressed relative to the counterfactual above the cutoff.

Thus, rather than adjusting bin counts up to account for this distortion, the consistent way to adjust is to shift observed values of z for agents above the cutoff until the integration constraints hold. This amounts to shifting and stretching the distribution to the right until

$$\sum_{p=0}^{P} \frac{\beta_p}{p+1}(((1+\kappa)z_{\max})^{p+1} - z_L^{p+1}) = \omega \sum_i \mathbb{1}[z_i > z_L]$$

Where the only change from eq. (1) is that the integral needs to be evaluated up until the new, shifted maximum value of z. We thus estimate

$$\tilde{c_j} = \sum_i \mathbb{1}[z^* + j\omega < z_i(1 + \mathbb{1}[z_i > z^*]\kappa) \leq z^* + (j+1)\omega] = \sum_{p=0}^{P} \beta_p m_j^p + \epsilon_j$$

which is simply the bin counts from the shifted values of z, where observations above the cutoff have been shifted up by $(1+\kappa)$. An improvement of this procedure compared to the procedure in Chetty et al. (2011) is that the resulting shift parameter has a clear interpretation: The running variable is increased by $100\kappa\%$ in the tax regime below the threshold relative to the regime above the threshold. No such interpretation may be attached to the shift parameter estimated in Chetty et al. (2011).

While this adjustment is consistent with the simple labor supply model in Chetty et al. (2011), we note that this is unlikely to have affected the main takeaway from this and other bunching papers because the observed amount of bunching, and therefore the required shift κ, is very small. Nonetheless, we show in a Monte Carlo simulation in appendix B, that a bunching estimator based on this adjustment of labor supply above the cutoff performs better than the adjustment in Chetty et al. (2011) when the true labor supply elasticity is substantial, and that the adjustment from Chetty et al. (2011) may perform worse than no adjustment at all.

Having accounted for the distortion in z above the cutoff, we arrive at estimates of the counterfactual number of agents $h_0(z) = \frac{1}{\omega}\sum_{p=0}^{P}\beta_p z^p$. With these in hand, it is straightforward to follow the existing bunching literature. Specifically, we construct $B = \sum_i \mathbb{1}[z_L < z_i \leq z_H] - \frac{1}{\omega}\sum_{p=0}^{P}\frac{\beta_p}{p+1}({z_H}^{p+1} - z_L^{p+1})$ as an estimate of the number of bunching agents. We can also relate the number of bunching agents to the estimated counterfactual at the threshold, $b = \frac{\omega B}{\sum_{p=0}^{P}\beta_p z^{*p}}$, a measure of excess mass. Finally, we follow Kleven and Waseem (2013) to estimate Δz^*, the reduction in z from the marginal buncher who is indifferent between allocating at the optimal point above the cutoff, and allocating at the cutoff. Intuitively, the bunching agents would have allocated somewhere above the cutoff in the absence of the incentive change, and when we distribute the bunching agents below the counterfactual density starting at z^*, we can exactly fit the bunchers up until the point where the marginal buncher would have allocated under the counterfactual. Thus,

$$
\begin{aligned}
B = & \quad \frac{1}{\omega}\int_{z^*}^{z^*+\Delta z^*} h_0(z)dz \\
B = & \quad \frac{1}{\omega}\sum_{p=0}^{P}\frac{\beta_p}{p+1}((z^* + \Delta z^*)^{p+1} - z^{*p+1})
\end{aligned}
$$

This is a polynomial in the unknown Δz^*, using the estimates for B and $h_0(z)$, and we solve this by finding the first real root above z^*. Kleven and Waseem (2013) uses a constant approximation to the true density at this stage, and simplify to $\Delta z^* = \frac{\omega B}{h_0(z^*)}$. This approximation does not matter as long as the bunching segment is small or the true density is constant in the bunching segment, as noted by Kleven (2016), but using the same polynomial throughout is, in our view, more consistent. It is also better in cases where the true density is nonconstant and the bunching segment is sizable, as we illustrate in a Monte Carlo simulation in the appendix B. In the empirical application in this paper the estimated density is strongly downward sloping in the bunching segment, leading to appreciable differences in Δz^* between the constant approximation and the full polynomial density.

The reduced form estimation procedure is illustrated in figure 1, where the density of the running variable is plotted using blue bars. Vertical lines at z^* and z_L indicate the cutoff and visually determined lower limit of the excluded range, respectively, while for illustration purposes the bunching above is assumed to be sharp so that there is no need to exclude any part of the observations above the cutoff ($z_H = z^*$). As described above, a polynomial is fit to the density, excluding observations in (z_L, z^*). To account for the distortion above the threshold, all observations above z^* are then shifted slightly, bins are reconstructed and the polynomial re-estimated until the integration constraint in (1) holds. The final data above the cutoff for which the polynomial is fit is illustrated in red, and the resulting estimated counterfactual is plotted as a solid red line. Based on this, the number of bunchers is simply the excess mass in the bunching region, colored in light blue. To find the response of the marginal buncher Δz^* from this estimate, we fit the light blue mass of agents under the solid red line, starting at z^*.

So far, our approach has closely followed the literature with a few changes to account for distortions above the cutoff and nonconstant density over the bunching segment. In many applications, however, there are other ways for the agent to adapt to the cutoff than reducing the running variable z. As an example, in the classical labor supply application of bunching, bunching agents might plausibly increase capital income to compensate for the lost labor income from the bunching response, or might optimally change effort at the job. In our application to the thin capitalization rules, firms may adjust by increasing the use of inputs that are substitutes to debt, such as equity, or reduce the use of inputs that are complements to debt. Understanding these substitution patterns are crucial when investigating bunching agents' response to sharp discontinuities in incentives.

To estimate response along other endogenous variables x using reduced form methods, we note that under mild smoothness assumptions on the heterogeneity in agent's preferences and constraints, the relationship between a particular variable x and the running variable z

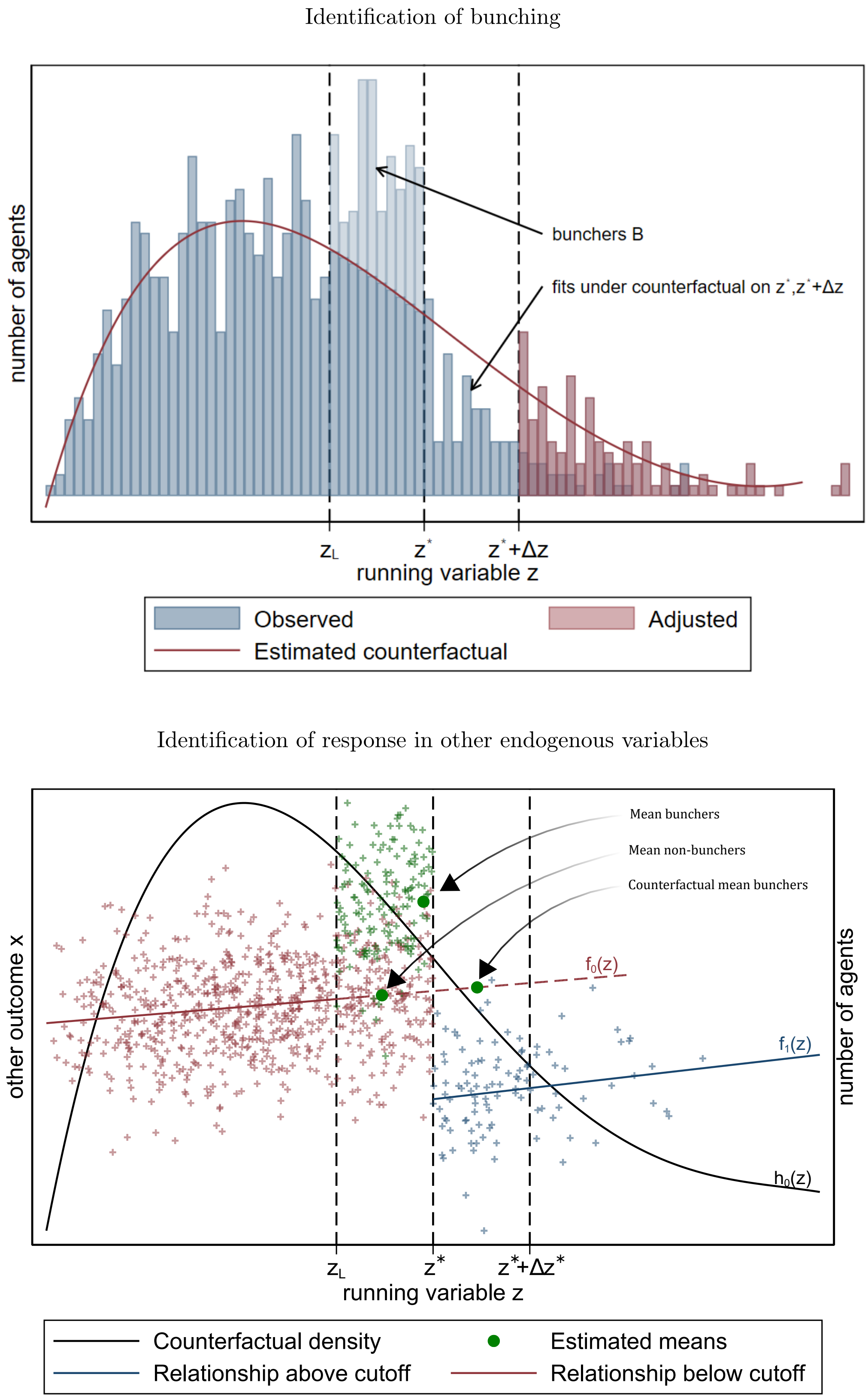


Figure 1: Identification of counterfactual densities in the bunching approach

will be smooth under the counterfactual. Thus, sharp discontinuities in this relationship in the bunching region is evidence that bunching agents behave systematically different than agents who would have allocated in the bunching region even in the absence of a cutoff. We use this intuition to estimate simple polynomial regressions of the relationship between each potential endogenous variable x and the running variable z on the form

$$x_i = f_g(z) = \sum_{h=0}^{H} \gamma_{gh} z_i^h \quad + \mu_{gi}$$

, separately in samples below ($g = 0$) and above ($g = 1$) the cutoff, excluding the bunching region. Notice that there is no need to bin observations when estimating this regression, unlike the baseline bunching estimates. Next, we estimate the excess value of x in the bunching segment by extrapolating from the relationship below the excluded region into the bunching segment,

$$\begin{aligned} C &= \textstyle\sum_{z_L < z_i \leq z_H} (x_i - f_0(z_i)) \\ C &= \textstyle\sum_{z_L < z_i \leq z_H} (x_i - \sum_{h=0}^{H} \gamma_{0h} z_i^h) \end{aligned}$$

Intuitively, sharp deviations from the relationship below the cutoff when we enter the bunching region is evidence that bunchers behave differently than agents who would have allocated in the bunching region independently of the cutoff, because the agents in this region is a mix of bunchers and nonbunchers. Next, we know that the overall average of x in the bunching region must be a weighted average of the means of bunchers and nonbunchers in this region. As we have already estimated the number of bunching and nonbunching agents, and it is straightforward to back out the mean of x among bunchers and nonbunchers:

$$
\begin{aligned}
x_{NB} =& \frac{\int_{z_L}^{z_H} h_0(z) f_0(z) dz}{\int_{z_L}^{z_H} h_0(z) dz} \\
x_B =& \frac{C}{B} + x_{NB} \\
x_k^{CF} =& \frac{\int_{z^*}^{z^*+\Delta z^*} h_0(z) f_0(z) dz}{\int_{z^*}^{z^*+\Delta z^*} h_0(z) dz}
\end{aligned}
$$

The final equation above estimates the counterfactual mean of x among the buncher by extrapolating further into the region $(z^*, z + \Delta z^*)$. The differences between these two means among bunchers, $x_B - x_{CF}$ is one measure of average response in x among bunchers to the change in incentives. Unlike the estimates of the bunching mass itself, this approach relies on extrapolation from one side of the bunching region only.

The estimation procedure for other outcomes is illustrated in the lower panel of figure 1. We fit x as a (linear) polynomial in z, using only observations below the excluded range so that $z_i < z_L$, plotting the result as a solid red line. Extrapolating from this into the excluded range, illustrated by the dotted red line, produces estimates of the values of x for nonbunching agents in this region. There are two types of agents in this region: Bunching and nonbunching agents, illustrated by green and red dots in this figure, but unobservable to the econometrician. We have already estimated the number of bunching agents B above. Therefore, the excess value C in the bunching region must reflect bunchers' deviating behavior, and we can estimate the mean value of x among bunching agents simply by dividing this excess value by the number of bunchers and adding the estimated mean among nonbunchers.

These estimated means for the bunchers in the bunching region and their counterfactual mean is plotted using green dots. Despite being unable to identify who the bunchers are, we can identify the means of other endogenous variables for these agents. In the example provided, we see that the mean of x for bunchers are considerably larger than their counterfactual allocation (and that of nonbunchers in the bunching region), indicating that bunching agents compensate for the reduction in z in order to bunch by increasing the use of x.[2]

When x is a predetermined variable or is otherwise unmanipulable by the agent, the approach above can be adjusted to connect the relationship below and above the cutoff. In particular, we can adjust the running variable z by multiplying it with the already estimated shift parameter $1+\kappa$ for observations above the threshold. Because x in this case is unchanged, we can then estimate the polynomial across the entire support of x excluding the bunching region, in essence relying on interpolation into the bunching region from both sides rather than extrapolation from below. In this case, the estimated means among bunchers can be used to characterize which agents are more likely to bunch in response to the change in incentives.

3 Institutional setting

Norway employs a worldwide corporate income tax with a rate of 22 percent, slightly above the OECD average (20.6 % in 2020), but close to neighboring countries. From the years 1992-2013 the rate was held constant at 28 percent, before slowly being lowered down to 22 percent in 2019. As with most corporate income tax schemes, interest costs are deductible, while there is no allowance for the cost of equity financing. This bias is counteracted in the personal income tax scheme with a more lenient taxation of dividends compared to interest income.[3] For a shareholder originating in Norway, these two systems could neutralize the financing decision entirely. However, for a small open country as Norway, it is normally assumed that the rate of return requirement on corporate equity before personal taxes is determined by international investors on the world capital markets (Lindhe and Södersten, 2012).

[2]The reduced form bunching estimator used in this paper is implemented in the Stata package rfbunch, available at https://github.com/martin-andresen/rfbunch.

[3]Individuals are only liable to a tax above a rate-of-return allowance for dividends.

Beginning in 2014, restrictions on interest deductions for firms were imposed. In line with many other European countries, Norway responded to the OECD Base Erosion and Profit Shifting (BEPS) project by initiating regulations to reduce the problem of profit shifting. One of the channels for such practices has long been recognized as debt shifting, where corporate groups inflate interest costs in high tax countries and transfer revenue to low tax countries through internal lending, taking advantage of the tax differential. The rules apply for all firms with both 1) net interest cost above NOK 5 million (approximately 500'000 Euro) and 2) net interest cost above 30 percent of earnings before interest, tax, depreciation and amortization(EBITDA). [4] Once a firm exceeds the two thresholds, its tax liability may increase. The aim of the regulation is to exempt the allowance for internal interest costs above the pre-defined ceiling. Internal interest costs is defined in the tax act as interest payments from a firm to a firm or person which directly or indirectly owns or control at least 50 percent of the firm. The most important categories here are loans between entities in the same corporate group and loans guaranteed for by a controlling entity.

Although there are two conditions that has to be met by a firm to be included in the rule, only one will be binding for each firm. For small firms with sufficiently low EBITDA, 30% (25% from 2016) of EBITDA can never be larger than NOK 5 million and the binding threshold will therefore be when total interest cost exceeds NOK 5 million. For firms with larger EBITDA, it is the other way around. To avoid any extra tax costs, firms with a small EBITDA thus have an incentive to bunch at the threshold of NOK 5 million in total interest costs, while larger firms have an incentive to keep total interest costs below 30% or 25% of its own EBITDA.

The increase in the tax base decided by the rule:

$$max(0, min(z_I, z - \eta \cdot EBITDA))$$

[4] The regulation has been changed several times since its inception. In 2016, the threshold was lowered to 25 percent of EBITDA. In 2019, special rules for corporate groups were added with targeting of external interest costs in addition to internal interest costs. The 2019 version is not studied in this paper.

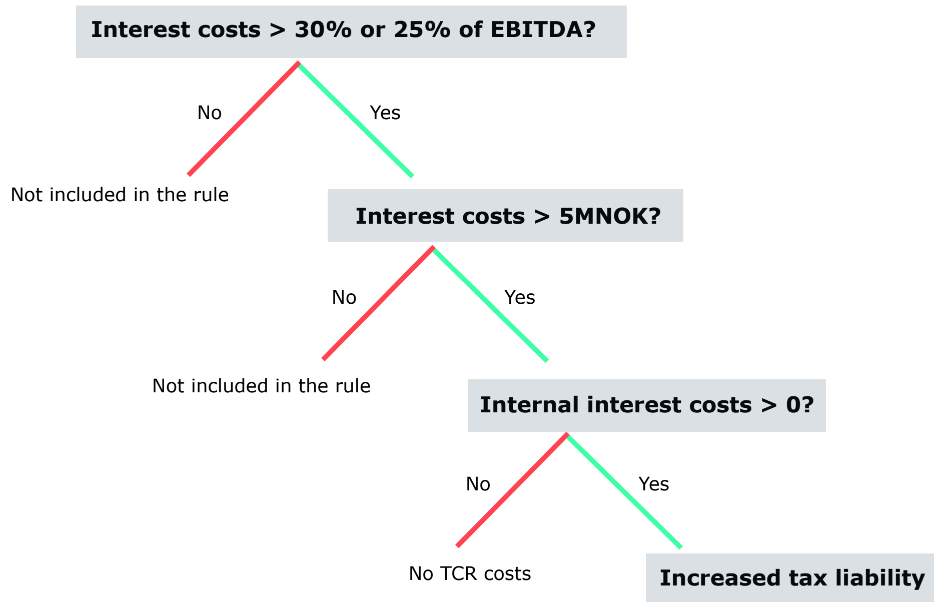


Figure 2: Firms with both interest costs above 30% (25% from 2016) of EBITDA and interest costs$>$ NOK 5 million are included in the TCR. However, if the firm has no internal interest costs, there are no consequences of crossing the thresholds.

where z_I is internal interest cost and z is total interest cost. η is either 30% or 25% depending on the year. In words, the smaller of internal interest costs or the difference between total interest costs and the maximum allowance is added to the tax base as long as it is nonnegative.

Based on the design of this rule, we make several important observations: 1) The increase in tax base at the threshold is in general different for each firm, as it depends on the firm-specific capital structure 2) A necessary condition for increased tax costs is positive internal interest costs. 3) The difference between total interest cost and the maximum allowance can only be smaller if external interest cost are lower than the maximum allowance ($z - z_I < \eta \cdot EBITDA$). To see this, note that internal interest costs enters in both arguments, and the decisive factor is the size of external interest costs relative to the maximum allowance.

Finally, the consequences of crossing the threshold are different at the two thresholds. In particular, smaller firms facing the NOK 5 million threshold can potentially accumulate tax costs below the threshold. To see this, note that to reach NOK 5 million in total interest costs, small firms will first exceed the condition of total interest cost above the allowance ceiling, making the second term argument in the rule positive. As a result, the added tax cost at the threshold can be quite large, and the firms facing this threshold will in general have a notch in the tax function at the threshold due to an increase in the average tax.[5]

At the threshold for the large firms it is different. Conditioning on positive internal interest costs, we can observe that the second argument must be the smallest argument when starting with total interest costs below the allowance ceiling. Consequently, there is no potential for accumulation of costs below the threshold in this case, and the large firms face a tax kink.

4 Data and sample selection

For the empirical analysis we combine rich register data and tax return data for the universe of public and non-public limited liability companies for the years 2014-2018.[6] For placebo tests we also use data for 2012 and 2013. The main variables comes from the firms' income statement and balance sheet as reported to tax authorities as part if tax returns each year. In addition, we link these data with the national register for business enterprises and the shareholder register which gives us useful information such as the sector type, the number of employees and ownership structure.

Firms that exceeds the two thin-cap thresholds are also required to give a detailed report on interest costs and affiliations to related parties to the tax authorities. We have access to these reports, but since we mainly study firms that allocate below the thresholds to avoid the TCR, this information is only relevant for a small part of our estimation samples. Beginning

[5]There are two scenarios in which small firms face a marginal tax increase at the threshold instead of a notch: 1) A small firm facing the NOK 5 million threshold issues only external debt below the threshold, but begins to use internal debt shortly after the crossing. The result is zero extra tax costs at the crossing, but each new unit of internal interest cost is added to the tax base, and the implicit marginal tax rate increase. 2) A firm with EBITDA close to the allowance ceiling (NOK 20 million in the years 2014-2016 and NOK 16.7 million in the years 2017-2019) will be constrained by the NOK 5 million threshold, but without a potential to accumulate much costs below the threshold, since the distance between the first threshold and second threshold is very small. In practice, this firm will thus have approximately zero added tax costs at the threshold. After the crossing, each new unit of either internal or external interest cost leads to an increase in the tax base, implicitly increasing the marginal tax rate. However, these two scenarios exemplify highly unlikely behavior and we believe that small firms in general face a tax notch at the relevant threshold.

[6]Firms in the financial, oil and shipping sectors are taxed under special schemes and not included in the analysis.

in 2008, some of the firms also became required to disclose information on intra-company transfers, in an effort by the Norwegian authorities to monitor transfer pricing. This applies for related establishments if the aggregate value of the transfers exceeds NOK 10 million (approximately 1 million Euro) over the year, which is the case for the majority of firms in our estimation samples.[7] Together with the registers mentioned above, we use these forms to separate foreign controlled multinational corporations (FMNC's) from domestic controlled multinationals (DMNC) and purely domestic corporate groups and firms.

We do not directly observe interest rates, but we observe debt and interest costs as reported annually in the firms' tax return. To separate between internal and external interest costs, we exploit that firms are asked to separate intra-company interest costs from total interest costs in the same report. A caveat at this point is that the definition of intra-company interest cost is wider under the thin-cap rule, as it includes loans guaranteed for by a related firm. Since the bunching firms do not cross the relevant threshold, they do not have to report information on guaranteed loans. However, by examining the detailed reports from the firms that do cross the threshold (approximately 1,000 firms each year), we find that only about one percent have such arrangements, and the discrepancy between two figures should therefore be very small.

EBITDA is not directly reported by firms below the threshold, but it is straightforward to construct this measure from variables in the tax records. To get at this figure, we start with reported income and subtract reported intra-group contribution, amortization and interest costs. We compare this estimate with the reported EBITDA for firms above the threshold, and find a very strong correlation (0.98).

To construct the two estimation samples we exclude firms that are not subject to the relevant threshold due to its EBITDA, and we zoom in on data points around the relevant threshold. In the notch case, this implies that we drop firms with EBITDA above NOK 16.7 million in the two first years and above NOK 20 million in the following years (remember

[7]Formally, the two forms referred to are known as RF-1315 and RF-1123.

that the EBITDA-relative threshold was changed from 30% to 25% in 2016.). Since there is a large mass of firms with zero net interest costs, we limit the notch sample to firms with net interest costs from NOK 2-14 million. In the kink sample we exclude firms with EBITDA under NOK 16.7 or 20 million for the relevant years. In addition, we only keep firms with interest costs relative to EBITDA in the window of 0.05 to 0.65. Table 1 below show descriptive statistics for the samples, along with the full sample of firms for comparison.

Table 1: Descriptive statistics

	Notch sample		Kink sample		Full sample	
	Mean	Se	Mean	Se	Mean	Se
EBITDA	3.15	(22)	88.70	(225)	1.01	(30)
Taxable income	-3.82	(23)	33.54	(150)	0.37	(27)
Total debt	167.75	(281)	1210.67	(12,225)	20.23	(982)
Total internal debt	56.95	(193)	455.87	(4,139)	6.46	(332)
Equity	104.89	(525)	777.31	(8,423)	18.03	(730)
Interest costs	4.18	(2.4)	18.51	(57)	0.23	(17)
Debt-to-assets-ratio	0.81	(.33)	0.67	(.22)	0.70	(.54)
Investment	8.28	(36)	51.67	(167)	1.20	(27)
Employees	35.33	(148)	200.63	(682)	10.09	(91)
Number of shareholders	19.67	(299)	183.67	(3,381)	6.48	(323)
Assets	272.65	(668)	1987.97	(20,487)	38.26	(1,678)
Received intragroup contribution	3.56	(15)	38.67	(214)	0.65	(39)
Given intragroup contribution	2.71	(25)	21.22	(103)	0.62	(31)
Share foreign-controlled multinational	0.12	(.33)	0.24	(.43)	0.03	(.18)
Age	3.38	(1.5)	3.41	(1.6)	4.40	(1.1)
Taxes	0.17	(.51)	9.05	(36)	0.20	(4)
Share with 0 revenue	0.20	(.4)	0.08	(.27)	0.25	(.43)
Share private limited company	0.99	(.075)	0.99	(.11)	1.00	(.035)
Share in corporate group	0.89	(.31)	0.93	(.25)	0.42	(.49)
Share subsidiaries	0.75	(.43)	0.83	(.38)	0.31	(.46)
Share windsorized	0.02	(.15)	0.00	(.033)	0.11	(.31)
Observations	16,502		4,714		1,286,591	

Notes: Data for the years 2014-2018. Non-shares are given in NOK million. Foreign multinational is defined by foreign shareholders owning at least 50 percent of total shares. Age is the number of years a firm is present in the data, i.e., with a maximum value of 5. Debt-to-assets ratio is windsorized at the 95/5-th percentile and one outlier based on equity is removed.

The table shows that the two subsamples consists of mainly private limited companies and 9 out of 10 firms are part of a corporate group. Compared to the full population, we also note that the share of foreign owned multinational corporations is substantially larger in both subsamples.

Firms in the notch sample are smaller (by construction) than those in the kink sample, but still larger than the average firm in the population. Firms in the notch sample are also characterized by having on average a negative taxable income because deductions are larger than earnings. We believe that this is mainly due to a reallocation of profits within corporate groups to concentrate the tax position at the top level. Although a firm with negative taxable income will have no de facto additional tax payment when crossing the TCR threshold in the same fiscal year, the firm still has an incentive to avoid being included in the scheme in order to accumulate a larger loss-carry forward for preceding years, as well as possibly using less intra-group contributions which increases tax liabilities at receiving firm.

5 Reduced form results

In this section we show reduced form evidence of response to the thin capitalization rule. We begin with evidence of bunching at both thresholds generated by the rule. In the next sections we analyze important changes in the capital structure for firms in the notch sample, by estimating responses for other endogenous variables at both the firm and corporate group level.

5.1 Bunching estimates

To estimate number of bunching firms at each threshold we follow the procedure explained in detail in section 2. We set bandwidths to balance bias and variance and use bands of NOK 0.15 million in the notch sample and 0.0071 in the kink-sample, which yields about the same number of bins in both samples. The bunching window is identified by visual inspection, excluding five bins below the cutoff in the notch case. In this case, the TCR incentives suggest that bunching will be strictly below the threshold because large tax costs can be realized immediately after crossing. We therefore decide not to exclude data above the threshold in this sample. At the threshold for the larger firms we observe a more diffuse mass of bunching around the threshold and exclude two bins below and above the cutoff. In both applications we use a 7th order polynomial to estimate the density. We show robustness to the choice of polynomial order in the appendix. Standard errors are constructed by bootstrap with 200 replications, clustering at the firm to account for any within-firm correlation of errors across years. Figure 3 depict the main bunching estimates at all thresholds with pooled data, while further bunching estimates are reported in Table 3.

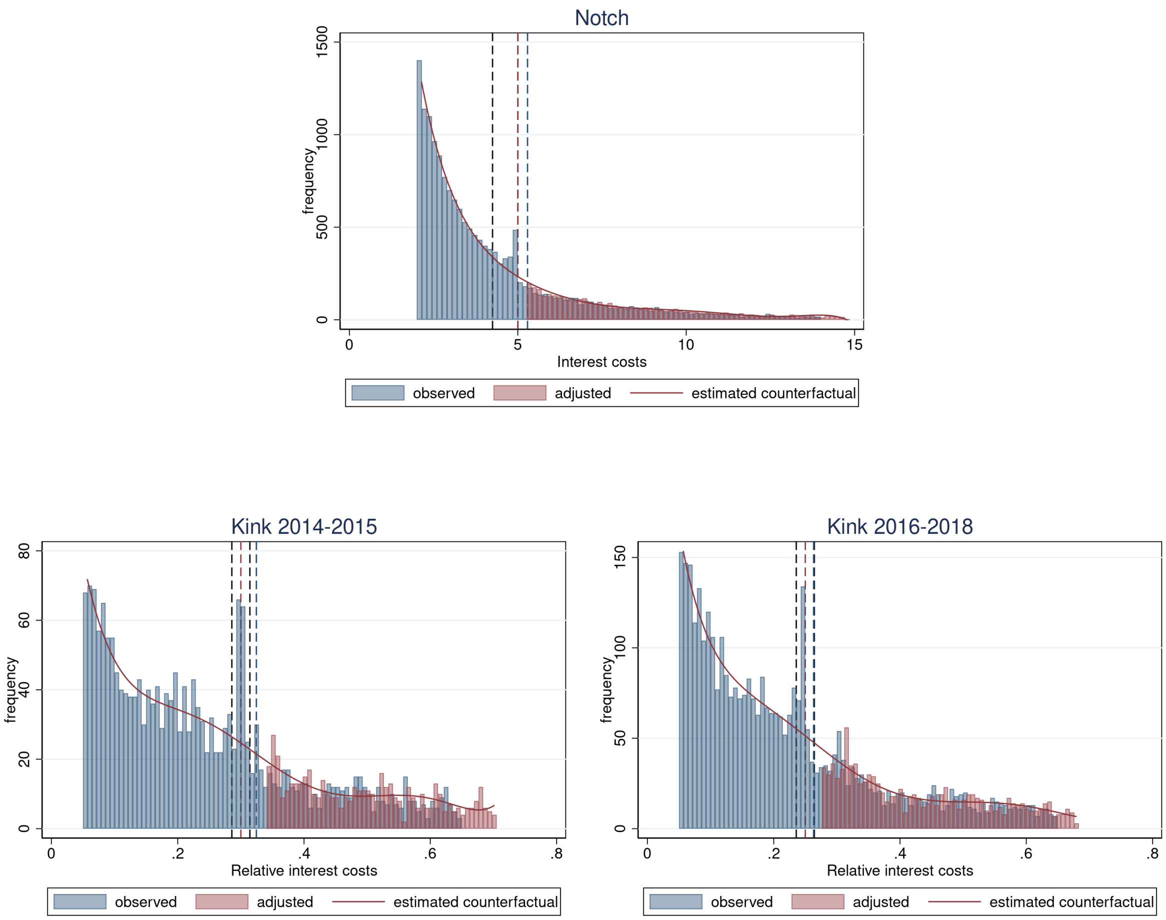


Figure 3: Main bunching estimates

Notes: Observed distribution (blue bars) at the notch and kink thresholds. Black lines mark the bunching region (in the notch case only below the threshold), and red lines mark the threshold. Blue lines mark $z^* + \Delta z^*$, the counterfactual location of the marginal bunching firm. Red bars show the final bin counts after shifting and stretching of the distribution above the threshold to account for a behavioral response for firms impacted by the thin-cap rule.

We observe that the distributions of interest costs / relative interest costs are falling smoothly in both samples before spiking at the relevant threshold. Throughout the period there are 420 bunching firms in the notch case and 79 and 106 firms in the kink case, depending on the period and allowance ceiling. By placing the excess mass under the counterfactual distribution, we infer that the marginal bunching firm in the notch sample reduces interest costs by NOK 290,000 to allocate at the threshold. This amounts to around 5.4 percent of the interest cost of the marginal firm.

In the kink case the estimated marginal response is in interest costs relative to EBITDA, which could be manipulated by both interest costs and the size of earnings. If we assume that EBITDA is fixed and equal to the average in the kink sample, a response of 0.02 implies a NOK 1.7 million reduction in interest costs. This implies that the marginal bunching firm reduces the ratio of interest cost to EBITDA by around 7.6 percent for the first period and 8.4 percent for the second.

The above estimates relates to the response of bunching firms themselves. As argued above, firms who choose not to bunch also have incentives to reduce interest costs, as these are no longer tax deductible. Our estimates of the shift parameter $1 + \kappa$ required to fit the bunchers under the counterfactual is informative of the reduction in interest cost among non-bunching firms affected by the thin capitalization rule. In particular, our estimates imply that interest costs are reduced by $1\frac{1}{1-\kappa}$, or 4.9 percent, when debt is no longer tax deductible.

Who are the bunchers? Using the methodology described in section 2 we can obtain characteristics for bunching firms based on variables that are not likely manipulated in the bunching response. This allows us to characterize who are more likely to be bunchers based on predetermined characteristics. We shift observations above the threshold by $1 + \kappa$ and estimate a regression of a characteristic x on a 7th order polynomial of x, excluding the bunching region, and back out means among bunchers and their estimated counterfactual. The results shown in table 2 show that the bunching firms stand out in some characteristics. Focusing on the notch sample, we find that 99 percent of the bunching firms are subsidiaries and 87 percent belong in the real estate sector, compared to these shares being 73 and 55 percent under the counterfactual. This implies that subsidiaries and firms in the real estate sectors are more likely than average to respond by bunching.

The bunching firms in the notch sample are also on average members of larger corporate groups measured by the number of members of their corporate group, and they are slightly younger firms than their estimated counterfactual, which could indicate a small organizational effect of the introduction of thin capitalization rules. We explore this further in robustness tests in subsection 5.4.

Table 2: Descriptive statistics bunchers

	Notch		Kink 2014-2015		Kink 2016-2018	
	Bunchers	Cf	Bunchers	Cf	Bunchers	Cf
Sector: Manufacturing	0.004	0.006	0.10	0.09	-0.01	0.1
	(0.03)	(0.01)	(0.04)	(0.02)	(0.03)	(0.02)
Sector: Real estate	0.87	0.55	0.36	0.34	0.57	0.30
	(0.08)	(0.02)	(0.06)	(0.04)	(0.08)	(0.03)
Sector: Construction	0.02	0.11	0.07	0.5	0.02	0.05
	(0.04)	(0.05)	(0.03)	(0.02)	(0.03)	(0.02)
Share subsidiaries	0.99	0.73	0.88	0.81	0.90	0.80
	(0.06)	(0.01)	(0.04)	(0.03)	(0.05)	(0.03)
Share FMNC	0.16	0.15	0.28	0.28	0.29	0.23
	(0.6)	(0.01)	(0.06)	(0.03)	(0.07)	(0.03)
Age	4.46	4.76	4.05	4.85	5.20	4.67
	(0.29)	(0.06)	(0.29)	(0.17)	(0.033)	0.15
Number of group members	6.47	2.27	3.06	2.12	3.49	1.71
	(0.84)	(0.09)	(0.40)	(0.17)	(0.45)	(0.14)

Notes: Estimated with data for 2014-2018. Standard error in parenthesis. Foreign multinational is defined by foreign shareholders owning at least 50 percent of total shares. Age is the number of years a firm is present in the data from 2014.

Figure 5 and 6 in the appendix plot the evolution of firms bunching year by year after the introduction of the thin-cap rule. Despite the fact that the tax rate on taxable profit is falling over time, mechanically decreasing the cost of crossing the threshold, we see a small tendency of sharper bunching at all thresholds throughout the period. This can indicate a rigidity of capital caused by for example adjustment costs, but it could also be caused by other mechanisms such as inattention. At the same time, the evolution of the observed bunching at the kink illustrates that the bunchers are able to swiftly readjust when the threshold moves from 0.3 in 2015 to 0.25 in 2016.

Table 3: Bunching estimates

	Notch	Kink 2014-2015	Kink 2016-2018
Required shift κ	1.05***	1.09***	1.08***
	(0.010)	(0.019)	(0.020)
Number of bunchers B	420.03***	79.36***	105.89***
	(55.105)	(14.064)	(19.520)
Bunchers' share of sample B/N	0.03***	0.04***	0.04***
	(0.003)	(0.007)	(0.007)
Normalized bunching $B/\omega \int_{z_L}^{z^*} h_0(z)dz$	0.30***	0.80***	0.75***
	(0.042)	(0.161)	(0.154)
Response of marginal buncher Δz^*	0.29***	0.02***	0.01***
	(0.041)	(0.005)	(0.003)
Response of average buncher $(\int_{z^*}^{z^*+\Delta z} zh_0(z)dz)/\omega B - z^*$	0.14***	0.01***	0.01***
	(0.021)	(0.002)	(0.002)
Observations	16,501	2,027	2,687

Notes: Standard errors in parenthesis. Required shift is the shift of the counterfactual polynomial above the threshold required to fit all agents under the counterfactual (see section 2 for details). The measure of normalized bunching compares the size of the excess mass to the counterfactual mass in the same region. The response of the marginal buncher shows the longest distance traveled by a bunching firm, while the response for the average buncher is the average reduction in interest costs/ relative interest costs for the bunching firms.

5.2 Anatomy of the bunching response

Having documented that firms respond to the incentives in the thin-cap rule, we now consider changes in capital structure among bunching firms. We focus on the firms in the notch sample in this section. To examine possible response channels, we estimate the relationship between the dependent variable and interest costs outside the bunching region, and plot the result together with binned means of the dependent variable (a detailed description of the procedure can be found in section 2). Figure 4 show the results from this exercise for the most important outcome variables.

Assuming that the bunching firms come from above the threshold, they will have to reduce total debt, the price of the debt, or a combination of both to land at the cutoff. For some forms of debt, such as intra-group debt, the firm likely has large discretion over the terms of the loan and the interest rate, even when the arms-length principle mandates that internal interest rates should mimic that between external partners. For others, such as external debt to financial institutions, the only way to adjust is through the total amount of debt.

Accordingly, debt is arguably the most important adjustment channel for firms to bunch at the threshold. In line with this reasoning, the first panel in figure 4, shows that the estimated average debt level among bunchers is 29 percent lower than the counterfactual (also see detailed estimation results in table 4). If we assume an average interest rate of 4 percent and that debt was constant throughout the year, the estimated debt reduction would on average lower interest costs with NOK 2.3 million, which is substantially higher than the estimated average response of NOK 0.14 million reported in table 3. This discrepancy could be caused by short-term debt without interest accrued during the year, which constitutes a large share of the estimated debt reduction.

In addition to the overall debt reduction, we observe important structural changes in debt types held by bunching firms. Specifically, there is a large and significant shift from long-term bank-debt towards long-term intra-company debt. We believe that there are two main explanations for this behavior. First, since the cost of crossing the thin-cap threshold is proportional to internal interest costs, a corporate group can avoid thin-cap costs by channeling internal debt towards a bunching group member that is not affected by the new regulation. This story is supported by the last panel of figure 4, which shows that total internal debt is considerably reduced for firms that choose to allocate above the cutoff, likely in order to reduce the tax liability. Second, when facing a new regulation with thresholds in interest costs, the corporate group obtains more flexibility by replacing external debt with internal debt since bank-debt is routinely subject to higher costs with changes in size and maturity.

If the corporate group aims to avoid a net reduction in assets for the bunching firms, we would expect a debt reduction to be accompanied with an increase in equity. To explore this dynamic, we report effects for equity outcomes in table 4 below and find a large and significant effect on equity mainly driven by an increase in firms share premium.

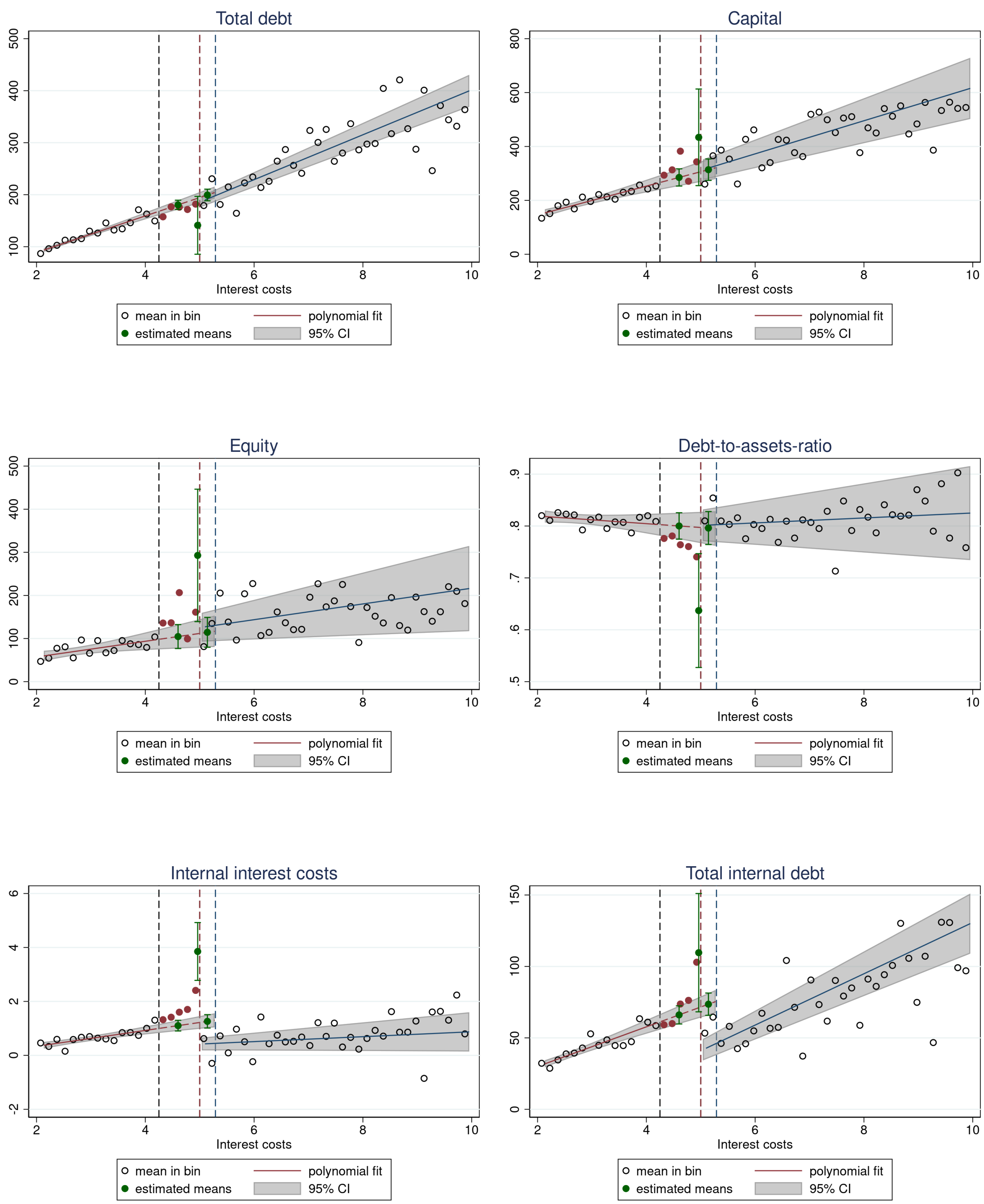


Figure 4: Alternative endogenous variables

Illustrating potential response and substitution margins for firms' capital structure to the TCR. A 1-st order polynomial used in all estimations.

Table 4: Endogenous variables

	Bunchers		Counterfactual		Difference	
	Mean	Se	Mean	Se	Mean	Se
Debt outcomes						
Total debt	141.17	(30.66)	199.46	(6.18)	-58.29*	(33.92)
Total internal debt	109.68	(24.37)	73.56	(4.27)	36.13	(26.38)
Total short-term debt	-0.77	(16.92)	56.54	(4.61)	-57.31***	(20.28)
Total long-term debt	141.94	(25.32)	142.91	(4.30)	-0.98	(27.38)
Long-term bank debt	5.29	(15.25)	74.36	(3.29)	-69.07***	(17.31)
Long-term intra-group debt	106.62	(22.14)	49.89	(3.13)	56.72**	(23.20)
Other long-term debt	6.83	(6.55)	9.75	(1.69)	-2.92	(7.48)
Short-term bank debt	-0.22	(6.74)	10.15	(1.22)	-10.36	(7.32)
Short-term intra-group debt	3.07	(9.58)	23.67	(2.83)	-20.60*	(11.83)
Accrued interest	0.75	(0.32)	0.57	(0.05)	0.18	(0.36)
Other short-term debt	2.22	(3.30)	5.40	(0.75)	-3.18	(3.77)
Debt-to-assets-ratio	0.64	(0.06)	0.80	(0.02)	-0.16**	(0.05)
Equity outcomes						
Equity	293.04	(83.63)	114.34	(16.87)	178.69**	(87.10)
Share premium	207.66	(62.21)	59.36	(10.32)	148.30**	(67.09)
Dividends	-0.91	(2.62)	1.91	(0.92)	-2.82	(3.47)
Other outcomes						
EBITDA	3.27	(2.27)	5.35	(0.52)	-2.08	(2.56)
Capital	434.20	(100.49)	313.80	(19.63)	120.40	(106.38)
Operating income	-72.59	(45.22)	91.46	(14.43)	-164.05***	(57.69)
Operating costs	-75.37	(45.28)	87.01	(14.30)	-162.39***	(57.59)
Internal interest costs	3.85	(0.56)	1.26	(0.12)	2.59***	(0.61)
Taxable income	-3.50	(2.35)	-3.69	(0.50)	0.19	(2.62)
Received intra-group contribution	1.84	(2.02)	4.63	(0.54)	-2.79	(2.41)
Given intra-group contribution	-0.28	(2.14)	4.08	(0.55)	-4.36*	(2.56)
Loss carry-forward	1.18	(1.11)	1.98	(0.22)	-0.80	(1.23)
TCR tax base increase	2.43	(0.28)	1.15	(0.06)	1.28***	(0.31)

Notes: The table shows estimates of averages among bunching firms, the estimated counterfactual average and a test for difference in means. Non-shares are reported in MNOK. Standard errors in parenthesis. Stars only used in the difference test-column: * $p<0.10$, ** $p<0.05$, *** $p<0.01$

As a consequence of the restructuring from debt towards equity, we observe a sharp reduction in the debt-to-asset ratio of 20 percent among bunching firms (16 percentage points). Compared to previous studies of thin capitalization rules of this type, this response in the debt-to-asset ratio is quite large. Buslei and Simmler (2012) and Alberternst and

Sureth (2015) both study responses to a similar thin capitalization rule in Germany and estimate a treatment effect on the debt-to-asset ratio of 10 and 5 percent, respectively. Another body of literature studies the effect of the so-called safe-harbor rules and find an effect of around 6 percent at the highest (see Ruf and Schindler (2015) for an overview). The differences in estimates could be explained by several factors. First, it is likely that the effect is most prominent for firms in the vicinity of the threshold (Buslei and Simmler (2012) obtain smaller estimates when they use firms further away from the threshold). Second, earlier research have shown that financial statement data tend to underestimate the tax effect on capital structure Devereux et al. (2018).

5.3 Corporate group level results

Our results so far has focused solely on the response of the firm that is exposed to the thin capitalization rules. Corporate groups, however, may have strong incentives to distribute debt and contributions differently within the group in order to minimize total tax burden in the group. Furthermore, to assess the impact of the CIT debt bias on macroeconomic stability concerns, it is usually not the leverage ratio of subsidiaries that we are interested in. A daughter is only in financial trouble if the mother company cannot bail her out, so what matters for the bigger picture is the leverage and stability of the corporate group. To understand if the capital movements we observe is a result of a reshuffle within groups with little real consequences or if they are part of a real response in the capital structure, we repeat the analysis from above using aggregates at the corporate group level. With ownership data from several sources we are able to connect entities within groups and separate between domestic and multinational organizations based on the jurisdiction of the corporate head. We only observe data reported for the Norwegian members of the corporate groups, so that our aggregates will measure the Norwegian part of the corporate group only.

Table A1 in the appendix show descriptive statistics for the corporate groups. The sample consists of around 2100 corporate groups each year, many of which are relatively small when compared with larger firms in the Norwegian corporate sector. We measure the average ratio of bunching firm assets to its corporate group aggregate assets and find that the bunching firms holds between 30-60 percent of total corporate group assets. In other words, since both the bunching subsidiary is significant in size compared to the corporate group, and the fact that the average group is not very large, we find that responses among bunching firms are important, also on the group level.

Table 5: Corporate group level results

	Bunchers		Counterfactual		Difference	
	Mean	Se	Mean	Se	Mean	Se
Equity	1474.00	(541.65)	370.91	(219.15)	1103.17	(722.94)
Total debt	1237.27	(572.90)	686.13	(208.81)	551.14	(743.43)
Long-term bank debt	26.17	(86.77)	248.41	(26.69)	-222.24**	(104.95)
Long-term intra-group debt	775.00	(269.25)	173.15	(66.85)	601.85**	(295.03)
External interest costs	0.34	(7.20)	14.87	(2.41)	-14.54	(9.14)
Internal interest costs	15.77	(5.89)	2.27	(1.51)	13.50**	(6.38)
Debt-to-asset-ratio	0.60	(0.04)	0.71	(0.01)	-0.11**	(0.05)
Observations	14,764					

Notes: Corporate group outcomes for groups with bunching firms compared to the estimated counterfactual outcome. Standard error in parenthesis, clustered at the group level. Group outcomes are calculated by summing the relevant outcome for each group in each year, except for the debt to asset ratio which is the ratio of aggregate debt over assets. About 10 percent of the firms in the notch sample are not in a group and not included in these results. Debt-to-asset ratio is windsorized at the 95-th/5th percent level. Significance stars only used in the difference-column.* $p<0.10$, ** $p<0.05$, *** $p<0.01$

The significant effects on corporate group level are similar to those at the firm level; a restructuring of long-term debt and an increase in internal interest costs. We also find that the corporate groups on average reduce the debt-to-asset ratio of 15 percent. Since we have multinational corporate groups in the sample, it could be that some of these corporate groups still do not change capital structure on the group level, if the unobserved member abroad fully eliminate the effect we observe in Norway. The finding that the TCR has an impact on the corporate group level seems to contradict previous findings related to thin-cap rules that limits intra-group borrowing (de Mooij and Hebous, 2017). However, a plausible explanation for the different results could be that the population of firms are somewhat different. While most earlier studies of thin capitalization rules use samples with a high presence of multinational corporations, the sample in this study is more dominated by smaller domestic groups.

5.4 Robustness checks

In this section we perform a number of checks and tests to address potential concerns with estimates of both the response in the bunching variable and in other outcomes.

For the bunching estimates we make the crucial assumption of smoothness in the counterfactual distribution in the absence of the thin-cap rule. To control that the results are not driven by for example round number bunching, we reproduce the bunching plots using data for the two years prior to the change. Figure 9 in the appendix show no indication of an excess mass at the thresholds before the reform is implemented. We also explore the sensitivity of the polynomial order by re-estimating all bunching estimates in both samples with different polynomials. Results reported in table A2-A4 clearly show that the main conclusion does not depend on this decision.

A second important assumption for the counterfactual estimation is that the bunching firms come from one side of the threshold only. In the case studied in this paper, we can come up with examples of how a firm could move to the right in the distribution by increasing interest costs to bunch at the threshold [8] While there is no easy way to test this assumption, we do believe that the changes in capital structure among bunchers as described above lend support to this premise. In particular, we find that the bunching firms convert debt to equity, a behavior consistent with the idea that they come from above the threshold. If instead they came from the other side, we would expect to observe an increase in debt to raise interest costs, and it would be less likely to see movement in equity.

[8] Two firms in the same corporate group have net internal interest costs NOK 7 million and NOK 3 million and relatively low EBITDA such that the NOK 5 million threshold in the thin-cap rule is binding. A tax-minimizing response in this case could be to transfer interest costs internally in the corporate group by allocating both firms at NOK 5 million.

6 Conclusion

In this paper we study firms' behavioral response to the introduction of a thin capitalization rule in Norway in 2014. We find that corporate groups use subsidiaries to bunch at the threshold for the rule to avoid potentially additional tax costs. Both the bunching firms and the corporate groups restructure debt in the face of a new regime, and we estimate a significant decline in leverage ratios both at the entity and group level. The main implications of our findings is that firms seem to react strongly to the tax bias within the corporate income tax scheme and that the capitalization rule could potentially mitigate this effect.

References

Alberternst, S. and Sureth, C. (2015). The effect of taxes on corporate financing decisions: Evidence from the German interest barrier. arqus Discussion Papers in Quantitative Tax Research 182, arqus - Arbeitskreis Quantitative Steuerlehre.

Andresen, M. E. and Kvamme, F. (2019). Rentebegrensningsregelen: En empirisk analyse. *SSB Discussion papers*.

Bachas, P. J. and Soto, M. (2018). Not(ch) your average tax system : corporate taxation under weak enforcement. Policy Research Working Paper Series 8524, The World Bank.

Bertanha, M., McCallum, A. H., and Seegert, N. (2021). Better Bunching, Nicer Notching. Finance and Economics Discussion Series 2021-002, Board of Governors of the Federal Reserve System (U.S.).

Boonzaaier, W., Harju, J., Matikka, T., and Pirttilä, J. (2019). How do small firms respond to tax schedule discontinuities? Evidence from South African tax registers. *International Tax and Public Finance*, 26(5):1104–1136.

Bosch, N., Dekker, V., and Strohmaier, K. (2020). A data-driven procedure to determine the bunching window: an application to the Netherlands. *International Tax and Public Finance*, 27(4):951–979.

Buettner, T., Overesch, M., Schreiber, U., and Wamser, G. (2012). The impact of thin-capitalization rules on the capital structure of multinational firms. *Journal of Public Economics*, 96(11):930–938.

Buettner, T., Overesch, M., and Wamser, G. (2016). Restricted interest deductibility and multinationals' use of internal debt finance. *International Tax and Public Finance*, 23(5):785–797.

Buettner, T., Overesch, M., and Wamser, G. (2018). Anti profit-shifting rules and foreign direct investment. *International Tax and Public Finance*, 25(3):553–580.

Buslei, H. and Simmler, M. (2012). The Impact of Introducing an Interest Barrier: Evidence from the German Corporation Tax Reform 2008. Discussion Papers of DIW Berlin 1215, DIW Berlin, German Institute for Economic Research.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics*, 126(2):749–804.

Coles, Patel, S. and Smith (2019). How do firms respond to corporate taxes.

de Mooij, R. A. (2011). The Tax Elasticity of Corporate Debt; A Synthesis of Size and Variations. IMF Working Papers 11/95, International Monetary Fund.

de Mooij, R. A. and Hebous, S. (2017). Curbing Corporate Debt Bias. IMF Working Papers 17/22, International Monetary Fund.

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

Devereux, M. P., Maffini, G., and Xing, J. (2018). Corporate tax incentives and capital structure: New evidence from UK firm-level tax returns. *Journal of Banking & Finance*, 88(C):250–266.

Diamond, R. and Persson, P. (2016). The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests. NBER Working Papers 22207, National Bureau of Economic Research, Inc.

Faccio, M. and Xu, J. (2015). Taxes and Capital Structure. *Journal of Financial and Quantitative Analysis*, 50(3):277–300.

Feld, L. P., Heckemeyer, J. H., and Overesch, M. (2013). Capital structure choice and company taxation: A meta-study. *Journal of Banking & Finance*, 37(8):2850–2866.

Graham, J. R. and Leary, M. T. (2011). A Review of Empirical Capital Structure Research and Directions for the Future. *Annual Review of Financial Economics*, 3(1):309–345.

Harju, J., Kauppinen, I., and Ropponen, O. (2017). Firm Responses to an Interest Barrier: Empirical Evidence. Working Papers 90, VATT Institute for Economic Research.

Heider, F. and Ljungqvist, A. (2015). As certain as debt and taxes: Estimating the tax sensitivity of leverage from state tax changes. *Journal of Financial Economics*, 118(3):684–712.

Homonoff, T., Spreen, T. L., and St. Clair, T. (2020). Balance sheet insolvency and contribution revenue in public charities. *Journal of Public Economics*, 186(C).

IMF (2016). Tax policy, leverage and macroeconomic stability. *IMF Policy Paper*.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8(1):435–464.

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

Kraus, A. and Litzenberger, R. H. (1973). A state-preference model of optimal financial leverage. *The Journal of Finance*, 28(4):911–922.

Lindhe, T. and Södersten, J. (2012). The Norwegian shareholder tax reconsidered. *International Tax and Public Finance*, 19(3):424–441.

Modigliani, F. and Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American Economic Review*, 48(3):261–297.

Myers, S. C. (1984). The capital structure puzzle. *The Journal of Finance*, 39(3):574–592.

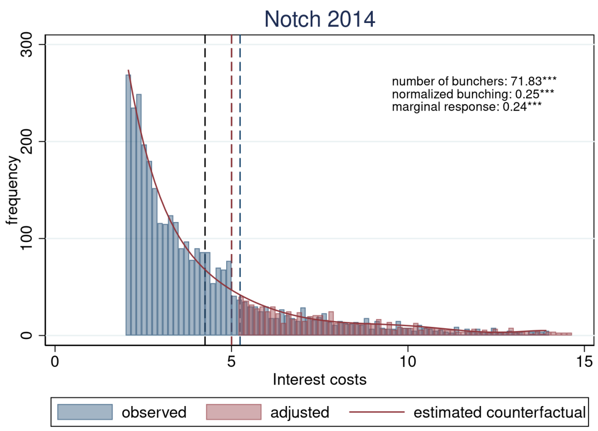
Overesch, M. and Wamser, G. (2010). Corporate tax planning and thin-capitalization rules: evidence from a quasi-experiment. *Applied Economics*, 42(5):563–573.

Ruf, M. and Schindler, D. (2015). Debt Shifting and Thin-Capitalization Rules – German Experience and Alternative Approaches. *Nordic Tax Journal*, 2015(1):17–33.

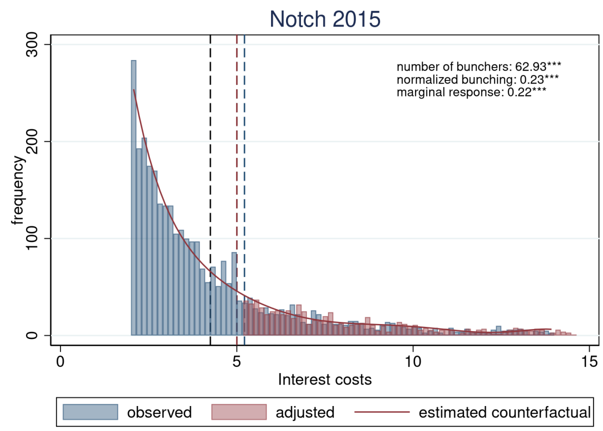
Saez, E. (2010). Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, 2(3):180–212.

Appendix

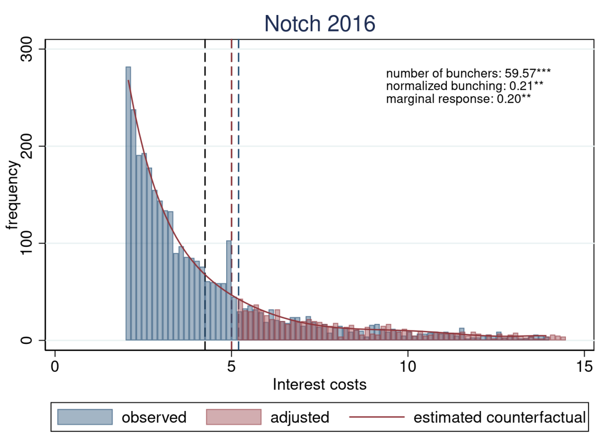
A



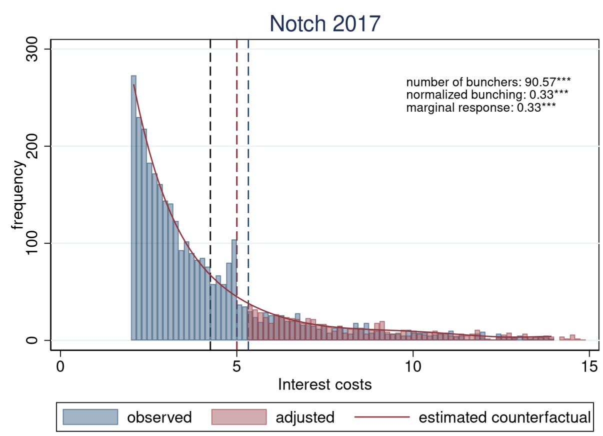
Notch 2014
number of bunchers: 71.83***
normalized bunching: 0.25***
marginal response: 0.24***
frequency
Interest costs
observed
adjusted
estimated counterfactual



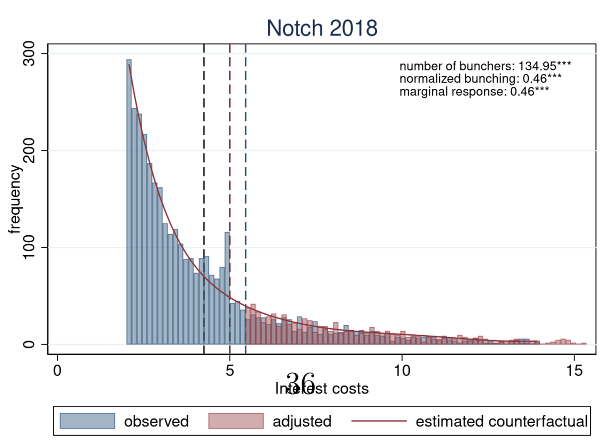
Notch 2015
number of bunchers: 62.93***
normalized bunching: 0.23***
marginal response: 0.22***
frequency
Interest costs
observed
adjusted
estimated counterfactual



Notch 2016
number of bunchers: 59.57***
normalized bunching: 0.21**
marginal response: 0.20**
frequency
Interest costs
observed
adjusted
estimated counterfactual



Notch 2017
number of bunchers: 90.57***
normalized bunching: 0.33***
marginal response: 0.33***
frequency
Interest costs
observed
adjusted
estimated counterfactual



Notch 2018
number of bunchers: 134.95***
normalized bunching: 0.46***
marginal response: 0.46***
frequency
Interest costs
observed
adjusted
estimated counterfactual

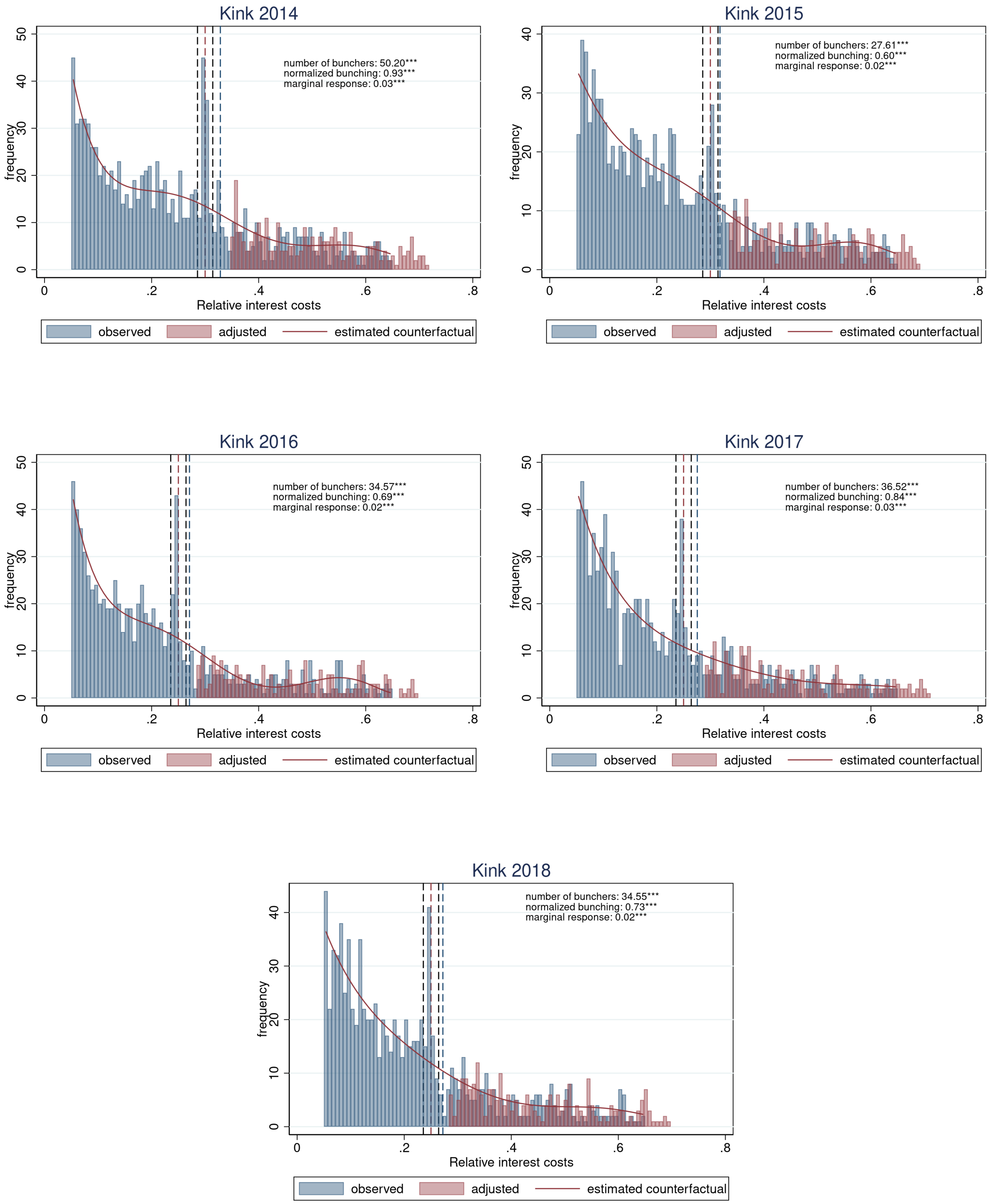


Figure 6: Illustrating the evolution of bunching in the kink sample.

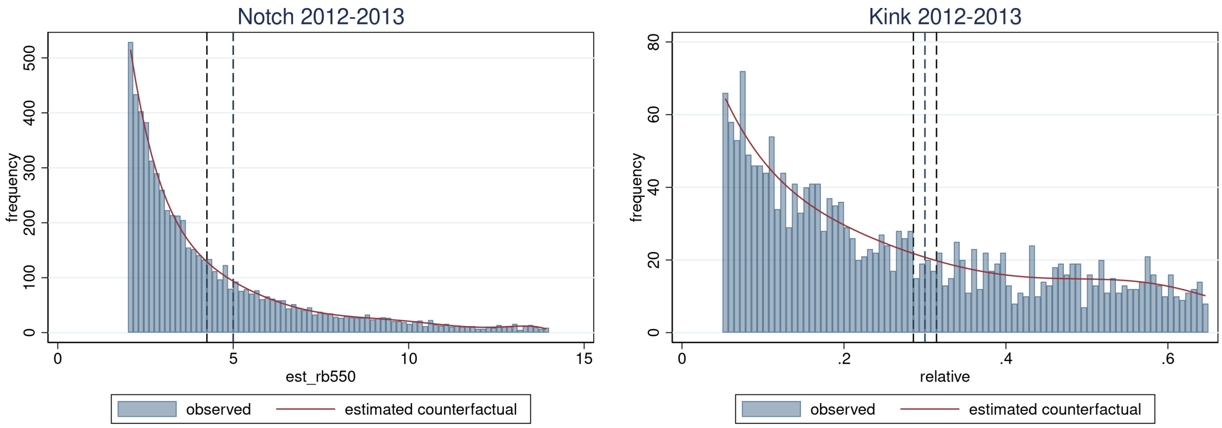


Figure 7: Placebo-test: The distribution of interest costs/relative interest costs in the two years (2012-2013) prior to the implementation of the TCR.

Table A1: Descriptive statistics corporate groups notch sample

	Mean	Se
Equity	799.81	(8378.44)
Total debt	1038.43	(9202.47)
Total interest costs	14.50	(90.24)
Internal interest costs	1.96	(103.95)
External interest costs	12.54	(60.94)
Debt-to-assets ratio	0.71	(.23)
Revenue	490.75	(6674.60)
Operational costs	384.40	(3747.68)
Employees	16.36	(200.00)
Observations	14,764	

Table A2: Polynomial sensitivity notch

Polynomial order	3	4	5	6	7
Required shift κ	1.05***	1.10***	1.08***	1.07***	1.05***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Number of bunchers B	117.42**	434.91***	526.70***	483.01***	419.79***
	(38.70)	(55.26)	(60.32)	(54.11)	(55.93)
Bunchers' share of sample B/N	0.01**	0.03***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Normalized bunching					
$B/\omega \int_{z\_L}^{z^*} h_0(z)dz$	0.07**	0.31***	0.40***	0.36***	0.30***
	(0.02)	(0.04)	(0.05)	(0.04)	(0.04)
Response of marginal buncher Δz	0.07**	0.33***	0.42***	0.35***	0.29***
	(0.02)	(0.05)	(0.06)	(0.05)	(0.04)
Response of average buncher					
$(\int_{z^*}^{z^*+\Delta z} zh_0(z)dz)/\omega B - z^*$	0.03**	0.16***	0.20***	0.17***	0.14***
	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)
Observations	16,502	16,502	16,502	16,502	16,502

Notes: Bunching results with the notch sample and different order of polynomials. We can observe that the estimated number of bunching firms is quite stable across polynomial orders.

Table A3: Polynomial sensitivity kink 2014-2015

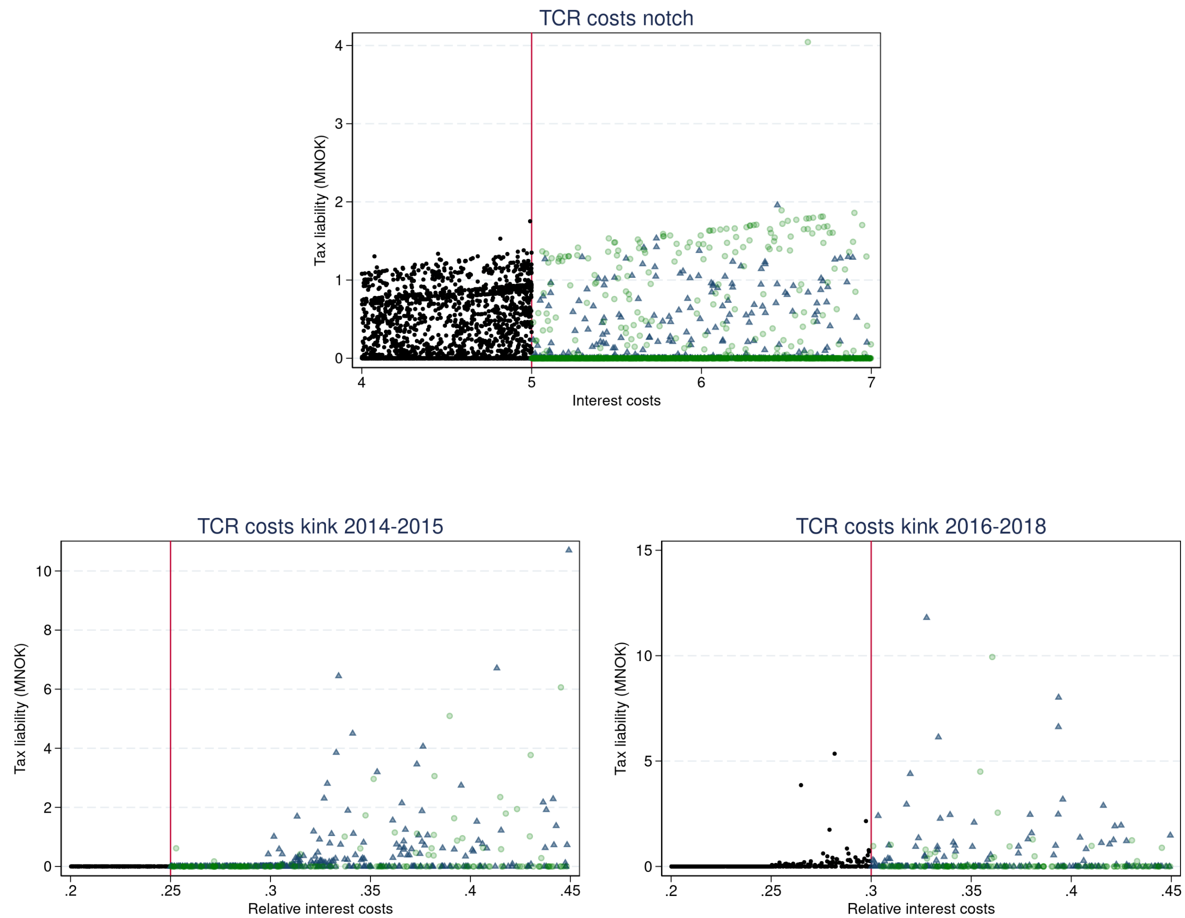
Polynomial order	3	4	5	6	7
Required shift κ	1.02***	1.01***	1.01***	1.03***	1.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Number of bunchers B	80.07***	74.73***	49.94**	54.43***	51.13**
	(18.67)	(18.15)	(18.21)	(15.74)	(17.19)
Bunchers' share of sample B/N	0.02***	0.02***	0.01**	0.01***	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Normalized bunching					
$B/\omega \int_{z\_L}^{z^*} h_0(z)dz$	0.42***	0.38***	0.23**	0.25**	0.23**
	(0.10)	(0.10)	(0.09)	(0.08)	(0.08)
Response of marginal buncher Δz	0.01***	0.01***	0.01*	0.01**	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Response of average buncher					
$(\int_{z^*}^{z^*+\Delta z} zh_0(z)dz)/\omega B - z^*$	0.01***	0.01***	0.00*	0.00**	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4,714	4,714	4,714	4,714	4,714

Notes: Bunching results with the kink sample in the two first reform years.

Table A4: Polynomial sensitivity kink 2016-2018

Polynomial order	3	4	5	6	7
Required shift κ	1.07***	1.05***	1.03***	1.03***	1.03***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Number of bunchers B	109.70***	112.60***	87.79***	74.99***	74.23***
	(23.22)	(24.06)	(22.29)	(21.68)	(21.84)
Bunchers' share of sample B/N	0.02***	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Normalized bunching					
$B/\omega \int_{z\_L}^{z^*} h_0(z)dz$	0.45***	0.47***	0.33***	0.27***	0.27**
	(0.10)	(0.10)	(0.09)	(0.08)	(0.08)
Response of marginal buncher Δz	0.01***	0.01***	0.01***	0.01***	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Response of average buncher					
$(\int_{z^*}^{z^*+\Delta z} zh_0(z)dz)/\omega B - z^*$	0.01***	0.01***	0.00***	0.00***	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4,714	4,714	4,714	4,714	4,714

Notes: Bunching results with the kink sample and different order of polynomials in the years 2016-2018.



Notes: Illustrating potential thin-cap tax liability (black dots below thresholds) and actual increased tax liability for the years 2014-2018. Tax liability is calculated by multiplying the added tax base with the tax rate for every year. Increased tax base is based on the rule $max(0, min(z_I, z - \eta \cdot EBITDA))$ (see section 3 for details). Above the threshold, blue triangles are year-firm observations with positive taxable income and green circles are year-firms with negative taxable income. In all figures we observe that many observations fall on the x-axis, implying that the firm had either zero internal interest costs or large EBITDA relative to net interest costs. In the notch case there is no strong indication of a response to minimize additional tax abpve the threshold by e.g. converting all internal debt with external debt. In the kink case, however, we note that there are very few positive firm-year observations in the vicinity of the threshold, especially in the first two years.

B

Table 1 provides a summary of potential parameters of interest when using reduced form methods to estimate bunching behaviour.

Table 1: Reduced form bunching: Parameters of interest

Parameter	Name	Description	Definition / details
		A: The bunching variable z	
κ	Shift	Shift required to fit all agents under counterfactual	
$h_0(z)$	Counterfactual density	Estimated counterfactual in the absence of tax change	$\sum_{p=0}^{P} \beta_p z^p$
B	Number of bunchers	Excess number of agents in bunching region	$\sum_i \mathbb{1}[z_L < z_i \leq z_H] - \frac{1}{\omega} \int_{z_L}^{z_H} h_0(z)dz$
	Normalized bunching	Bunching relative to total in bunching region	$\frac{\omega B}{\int_{z_L}^{z^H} h_0(z)dz}$
b	Excess mass	Bunching relative to counterfactual at the cutoff	$\frac{\omega B}{h_0(z^*)}$
Δz	Marginal response	Reduction in z from marginal buncher	solution to $B = \frac{1}{\omega} \int_{z}^{z^*+\Delta z} h_0(z)dz$
z_{NB}	Mean nonbunchers	Mean z among nonbunchers in the bunching region	$\frac{\int_{z_L}^{z_H} z h_0(z)dz}{\int_{z_L}^{z_H} h_0(z)dz}$
z_{CF}	Mean counterfactual	Counterfactual mean of z among bunchers	$\frac{\int_{z^*}^{z^*+\Delta z} z h_0(z)dz}{\int_{z^*}^{z^*+\Delta z} h_0(z)dz}$
		B: Other response variables x	
$f_0(z)$	Relationship below	Estimated relationship between x and z below z^*	$\sum_{h=0}^{H} \gamma_{0h} z^h$
$f_1(z)$	Relationship above	Estimated relationship between x and z above z^*	$\sum_{h=0}^{H} \gamma_{1h} z^h$
x_{NB}	Mean, nonbunchers	Mean of x among nonbunchers in the bunching region	$\frac{\int_{z_L}^{z_H} h_0(z) f_0(z)dz}{\int_{z_L}^{z_H} h_0(z)dz}$
x_B	Mean, bunchers	Mean of x among bunchers	$\frac{C}{B} + x_{NB}$
x_{CF}	Mean counterfactual	Counterfactual mean of x among bunchers	$\frac{\int_{z^*}^{z^*+\Delta z} h_0(z) f_0(z)dz}{\int_{z^*}^{z^*+\Delta z} h_0(z)dz}$

In order to evaluate the impact of the two improvements to the bunching estimator, we provide evidence from Monte Carlo simulations based on the traditional labor supply setting where the agent faces a kink (Chetty et al., 2011) to evaluate the importance of adjusting the running variable above the cutoff and a notch (Kleven and Waseem, 2013) to evaluate the impact of the constant approximation to the density in the bunching segment.

In both cases, the true elasticity $\epsilon = 0.4$ and the cutoff is situated at $z^* = 0.35$. Utility is log-linear and disutility of labor is iso-elastic, so that $U(z) = (1-t)z - \frac{\alpha^{-\frac{1}{\epsilon}}}{1+\frac{1}{\epsilon}} z^{1+\frac{1}{\epsilon}}$, where α is individual productivity. The tax rates are $t_0 = 0.15$ below the cutoff and $t_1 = 0.25$ above the cutoff, with the difference that the tax rate applies only to income above z^* in case 1 and to all income in case 2. This implies that optimal income is given by

$$z = \begin{cases} \alpha(1-t_0)^{\epsilon} & \text{if } \alpha < \alpha_L \\ z^* & \text{if } \alpha_L < \alpha \leq \alpha_H \\ \alpha(1-t_1)^{\epsilon} & \text{if } \alpha > \alpha_H \end{cases}$$

Where $\alpha_L = z^*(1-t_0)^{-\epsilon}$ determines the productivity of the marginal unaffected agent and α_H determines the productivity of the marginal buncher. The latter is equal to $z^*(1-t_1)^{-\epsilon} \approx 0.393$ in the kink case, and is found by solving $U(z^*) = U(\alpha(1-t_1)^{\epsilon})$ numerically for the notch case, yielding $\alpha_H = 0.558$. Productivity is drawn from a beta distribution, $\alpha \sim \beta(2, 5)$. All simulations use a bandwidth of 0.01, seventh order polynomials and use a single bin from $(0.225, 0.25]$ as the excluded region.

Figure 8 provides the result for the analysis of adjustment mentods. The top panel illustrates the comparison of the adjustment method in Chetty et al. (2011) to no adjustment and the adjustment method in our paper. In blue, we see that the no adjustment method overestimates the elasticity in this example. The adjustment method in Chetty et al. (2011) illustrated in red, however, makes matters worse: Elasticities are severely underestimated.

Our adjustment method, illustrated in green, centers on the true value.

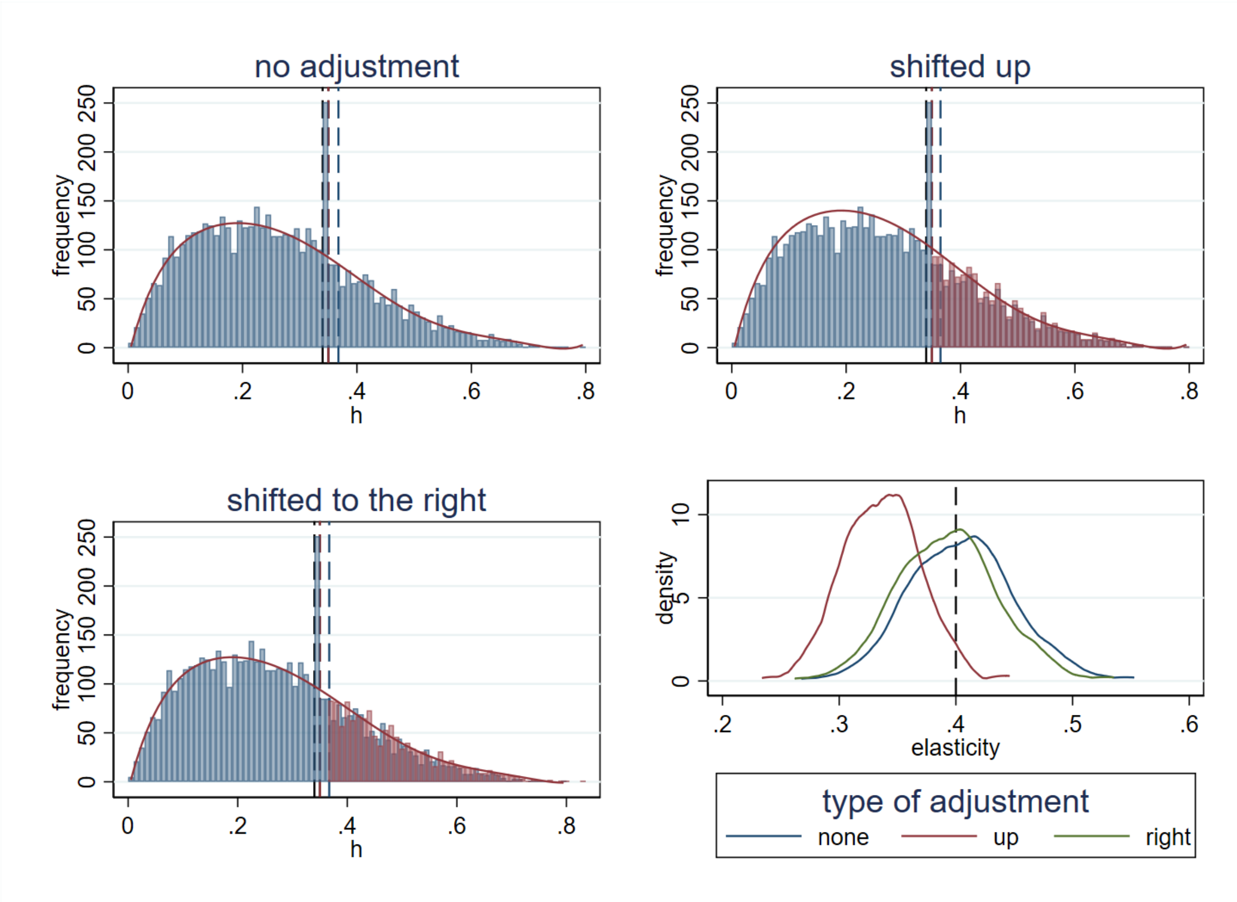


Figure 8: Monte Carlo simulations of alternative adjustment methods, kink case

Figure 9 shows the comparisons of the constant approximation to the bunching segment from Kleven and Waseem (2013) and the full polynomial approximation from this paper. Both use the adjustment method from this paper, adjusting observed values to the right before reconstructing bins. Because the density is non-constant in the bunching region and the bunching segment is substantial when the true elasticity is sizable, it is not surprising to see that the constant approximation illustrated in blue does a poor job. In contrast, our estimates in red center on the true value.

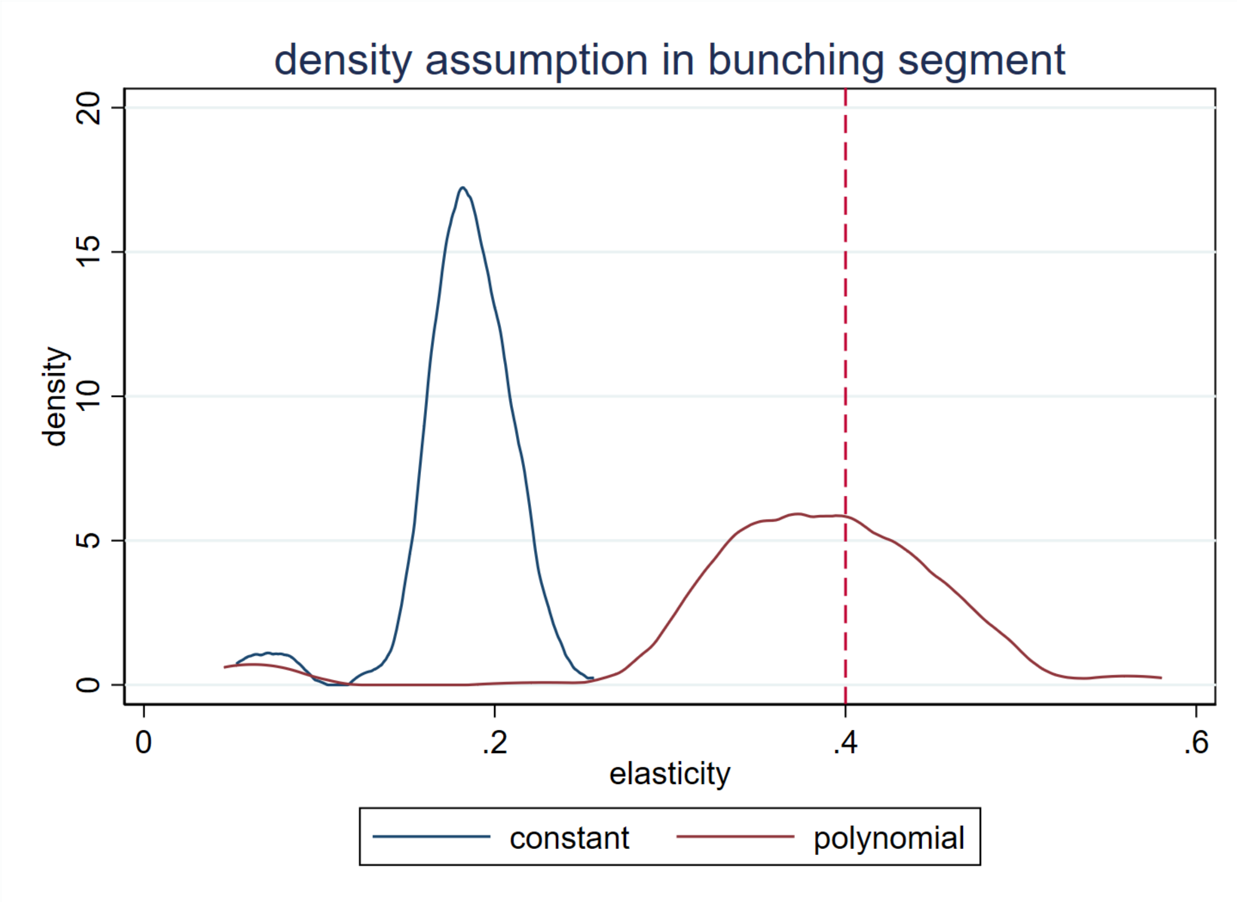


Figure 9: Monte Carlo simulations of naive vs. constant approximation to density in the bunching region, notch case