The Corporate taxation of Small-Sized Enterprises*

Théo Valentin†

February 25, 2025

First draft, please do not distribute

Abstract

This paper studies how firms learn changes in the tax code and its impact on firm behavior. To do so, I use a 30-years panel of tax administrative data on French companies. I leverage a policy reform which reduced the marginal corporate income tax (CIT) rate for firms below a threshold in addition with the firm's reported tax liability allowing to know which firm does or doesn't understand the reform. On average, it takes 4 years for firms to be correct on their tax liability. Firms make binary mistakes, and the learning process is heterogeneous among them. The reform also allows to measure the elasticity of corporate income and thus the contribution of learning to the elasticity. Finally, I use the difference in the learning status to uncover the effects of the policy through a staggered diff-in-diff.

Keywords: Corporate tax, Firms behavior, Bunching, Kinks, Elasticity

JEL codes: D22, D83, H25, H32

*I thank Laurent Bach, Pierre Bachas, Geoffrey Barrows, Suzanne Bellue, Pierre Boyer, Marion Brouard, Bertrand Garbinti, Jonathan Goupille-Lebret, Xavier d'Haultfoeuille, Hilary Hoynes, Claire Leroy, Clément Malgouyres, Andreas Peichl, Stefan Pollinger, Nina Roussille, Emmanuel Saez, David Sraer, Michael Visser, Gabriel Zucman and participants of the CESifo Public Economics and ZEW Public Finance Conferences for valuable comments and fruitful conversations. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 - Centre d'accès sécurisé aux données – CASD)

†Centre de Recherche en Économie et de Statistiques (CREST), CNRS, École polytechnique, GENES, ENSAE Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France. theo.valentin@ensae.fr

1 Introduction

Tax revenues amount for between 25% and 45% of GDP in Western economies[1] and are levied through complex systems [Benzarti and Wallossek, 2024]. Economic agents are usually thought to need a stable and clear environment to make optimal decisions. Frequent or large changes in the tax code can hinder this decision-making or slow it down, for instance through the length of the learning process. This learning process might also be heterogeneous across agents such that changes in the environment may reallocate opportunities. The literature has focused on tax (mis)learning from households and individuals [Chetty and Saez, 2013, Chetty et al., 2013, Feldman et al., 2016, Benzarti, 2020, Rees-Jones and Taubinsky, 2020]. Except the paper of Zwick [2021] on tax complexity, evidence of firms learning process after tax changes remain scarce.

In this paper, I use a policy reform implemented in France from 2001 onward that introduced a bracket in the corporate income tax to provide evidence on how firms learn changes in the tax code and how does that affect their overall reaction to the taxes. The reform introduced three discontinuities: a kink at 38.120 € on the taxable income, that is a change in the marginal tax rate, a notch at 7.63m € on the revenue (turnover or total sales) and another notch at 75% of the ratio of natural persons among all shareholders, that is a change in the average tax rate. Below the kink is introduced a reduced tax rate of 25% in 2001, and 15% from 2002 onward, instead of the standard 33,3% CIT rate that was applied uniformly before. Only firms that have a revenue below the notch and satisfy the notch of shareholders can benefit from this reduced rate. The tax gain for firms above the kink is around 7000 €, for firms below their average tax rate is divided by more than two. Thus, as more than 95% of firms are targeted by the policy and more than 75% are below the kink, this constitutes a sizeable change in the CIT. This paper tackles three main questions: How do firms learn the tax code? How does this process impact behavioral responses? How does this process impact real responses to taxes?

First of all, I rely on accounting data from the French tax administration and a specific feature of it, namely the *expected amount of due taxes* reported by firm, to provide new evidence on the learning process of the tax code of firms. Each year, they have to report in year N+1 their expected tax liability on the realized taxable income of year N on their accounting documents for the tax administration. Though this has no impact on what will be paid by the firm at the end of the day, it allows me to check whether firms correctly understood the modification of rates on the tax code.

I document that after the introduction of the kink in 2001 it took several years for firms to understand correctly the change in the tax schedule. I show that it takes approximately 5 years for the average expected tax liability to converge to the true tax liability. For firms below the kink, only the

[1] Tax revenue by country, OECD

marginal tax rate is modified, not the average one such that by computing a counterfactual amount of due taxes with no reform I am able to observe which firms still used the pre-reform tax schedule to report their taxes. In 2001, the first year of the reform, there is as much firms reporting using the pre-reform code than firms using the correct one. In 2002, there are still a substantial part of firms making mistakes, around 40%, but this time two errors emerge: firms that are still using the pre-reform code and firms using the 2001 code. It is only by 2005 that firms completely stopped using the pre-reform tax code. Using this counterfactual measure, I can study the type of mistakes firms make and answer the question of the type of the misunderstanding. I find that firms make binary mistakes: either they do learn or they don't, they don't apply rates that were never used in the legislation. For firms above the kink however, because there are two changes, namely the reduced rate and the value of the kink, it is not possible to understand the type of mistakes they are making. I then construct a measure of the amount of due taxes firms should have reported if they were using the correct tax code. I apply this tax code to the reported taxable income. In order to check the validity of such measure I check whether it follows what is reported before the reform and whether it follows what is reported by firms unaffected after the reform. The measure closely follows each trend suggesting it is reliable. Then I use this measure to tag firms by learning status. If a firm reports an amount of due taxes above 10% of the correct due taxes, I tag them as being wrong. I use this indicator to understand what are the characteristics of the learners.

Second, I analyze the reaction of firms at the discontinuities. I show that firms do react to the kink on the taxable income, but very gradually. From 2001 onward, firms start to bunch at the threshold of 38.120€. Here, contrary to the literature, firms do not react very quickly but the excess mass builds year after year. The bunching mass increased sharply from 2001 to 2006 and then continue growing but much more slowly until 2012 when the maximum is reached. This first gradual increase is very consistent with learning. Using the tagging of learning status defined above I show that it is indeed learning that drives the behavioral responses at the kink through differential bunching up until 2004. From 2005, *correct* firms don't exhibit any dynamic, they bunch sharply but the extent remains the same. *Wrong* firms however start to bunch, suggesting that they understood where the kink is, but exhxibit substantial frictions. The dynamic is also driven by new business creation. There is a bunching mass at the first-year taxable income distribution, suggesting either that firms are created and directly optimize or that they are created as part of a tax optimization strategy, which would suggest the presence of firm splitting.

To study the channels of behavioral responses, I use different angles. I study how many consecutive years firms stay at the kink and show that this dominate the number for counterfactual bins around it. I also show that the intra-firm volatility decreases at the kink while being very linear across taxable

income. This suggests that firms at the kink are able to adjust their taxable income much more than non bunching firms. Furthermore, using the specific timing of a change in the location of the kink in 2022-2023, I show that half of the bunching mass moved even though the production decisions were already realized which suggests that at least half of the bunching mass is due to tax avoidance and income shifting. Studying the differential bunching by legal status also points toward tax avoidance as the main driver of bunching. The legal status exhibiting the most intense bunching is Professions Limited Liability Company (LLC). In France these are types of activities that are usually levied at the PIT but it is possible to choose to be levied at the CIT through this specific legal status. Such firms are thus sophisticated entrepreneurs that choose their legal status so as to minimized their overall taxes. Finally, matching the tax administrative data with linked employer-employee data, I show that firms reporting no employees bunch through the wage channel: owner-managers pay themselves or their relatives higher wages to remain below the kink. This is also validated by anecdotal evidence on websites dedicated to advice for business-owners. This body of evidence strongly suggests that the main driver of behavioral responses is avoidance.

I then compute the elasticity of corporate taxable income (CETI) relying on the behavioral responses at the kink. Using the Bertanha et al. [2023] method, I compute one elasticity per year. The numbers closely follow the evolution of the bunching mass, increasing since 2012 and goes down afterward. This suggests that the elasticity is sensitive to the different channels behind it and to the tax environment Slemrod and Kopczuk [2002]. This allows me to give bounds on the welfare effects of the reform. Relying on a conceptual framework, the elasticity and the share of it coming from avoidance, I can give local welfare effects in the spirit of Chetty [2009] and Devereux et al. [2014].

Third, I evaluate the effect of the policy on firms' outcomes. In order to do so, I take advantage of the discontinuities and the different learning status among firms. I rely on Regression-Discontinuity Design (RDD) to measure whether there are local effects of the policy on firms when they stop or start to gain access to this reduced tax rate. I also rely on a Regression-Kink Design (RKD) to measure whether the marginal tax rate change at the kink modifies the production decisions of firms. Because the assumption of the absence of manipulation is violated in my case, I use the donut-hole approach. I show that the donut-hole approach is reliable in my case: the linearity assumption of the link function between a given outcome and the running variable (here the taxable income) seems highly plausible, which makes the extrapolation much more plausible too. I show that for the kink, the change in the marginal tax rate did not trigger any effect. More precisely, there is an effect when including the units at the kink but a counter-intuitive one: firms would increase their taxable income when facing a higher marginal tax rate. When excluding the units in the bunching region, none of the estimates are significant.

Then, I use the difference in learning to compare firms in a staggered setting. The identification relies on the fact that though firms are treated the same year, the year they understand the treatment differs so as the year they are able to adjust and react to this reform. I rely on the Callaway and Sant'Anna [2021] method that requires conditional parallel trends but does not need the date to be random, which is not the case here.

This paper contributes to three strands of the literature.

First, I relate to the literature on the understanding of the tax code. The literature has focused on households and individuals [Chetty and Saez, 2013, Chetty et al., 2013, Feldman et al., 2016, Benzarti, 2020, Rees-Jones and Taubinsky, 2020] and showed substantial heterogeneity across individuals and states. Regarding firms, Zwick [2021] studies the claiming of tax refunds and sheds light on the low take-up rate, suggesting high cost of claiming. He also shows that sophisticated tax preparers help firms to claim more refunds, a stable provision of tax codes. I contribute to this rising literature by giving new evidence on how do firms learn the tax code after a reform. Using a specific feature of my data, I do not have to rely on bunching intensity or the claiming of refunds to study this learning process.

Second, I contribute to the literature on the corporate elasticity of taxable income. The literature on firms CETI has been booming [Brockmeyer, 2014, Devereux et al., 2014, Best et al., 2015, Boonzaaier et al., 2019, Chen et al., 2021, Bachas and Soto, 2021, Bukovina et al., Forthcoming, Lobel et al., 2024, Massenz, 2024, Agostini et al., 2024]. The main findings are that elasiticities can be large whenever firms have opportunities to avoid taxes. These papers use bunching methods to quantify both the elasticity and the extent of avoidance. I provide new evidence for the bunching and elasticity of firms in a developed country such as France. Especially, I show how dynamic behavioral responses can be and that structural elasticities might be hard to grasp from them. I also contribute by giving new evidence on avoidance from firms both with other data like Askenazy et al. [2022], Lapeyre [2024] and with the timing of the reform. The extent of such a channel is important for welfare effects but also for its impacts on market competition and firm growth Di Marzio et al. [2023].

Third, I contribute to the literature on the effect of taxes on firms' real responses. Unlike accelerated depreciation, for which the literature finds overall positive and significant effects [House and Shapiro, 2008, Zwick and Mahon, 2017, Guceri and Liu, 2019], and dividend taxation, for which the literature finds overall null effects [Yagan, 2015, Alstadsæter et al., 2017, Bach et al., 2021], results on the effect of the statutory corporate income tax rate are still debated [Cui et al., 2021, Harju et al., 2022, Chodorow-Reich et al., 2024]. I provide new evidence on both the effect of modifying the

marginal tax rate for small companies, but also on reducing the average tax rate.

The paper is structured as follows. Section 2 describes the institutional background and the successive reforms. Section 3 presents the conceptual framework. Section 4 describes the different sources of data and some descriptive statistics. In section 5 I then present the behavioral responses to the reform in section 6 and the real responses in section 7

2 Institutional background and setting

2.1 The corporate income tax schedule in France until 2000

The corporate income tax is levied on the taxable income (*résultat fiscal*), as opposed to the accounting income (*résultat comptable*) from which it is computed. The accounting income is defined as the turnover minus the different type of costs. It can be thought of as the profit a firm is making. To go from this to the taxable income, a firm has to add back some costs that they paid but are not tax deductible and on which they have to pay taxes (luxury expenses, non-deductible personnal uses of the firm's assets etc...)[2]; the firm can deduct some other revenues for example capital gains that are taxed differently.

Firms have to declare their accounting information on a yearly-basis to the tax administration. They have until late April or beginning of May to return the documents for the previous year when they close the fiscal year on December 31^{st}, the exact date depends on the year. If a firm ends its fiscal year before, they have 3 months to send the information. The amount of details that have to be declared depend on the corporation system (*régime fiscal*) known as "simplified" (RSI, *régime simplifié d'imposition*)[3] and "normal" (RN, *régime du réel normal*)[4]. The RSI is standard for firms with turnover between €176.000 and €818.000 operating in business and housing, and between €72.600 and €247.000 for firms oprating in the provision of services. If a firm cross these thresholds, they can still apply the RSI but only for the first year above, then they have to declare at the RN. Below the lower part of the brackets, a firm is under the *Micro-entreprise* system[5] (see Aghion et al. [2024] for a study of these thresholds).

The timing of the remittance of the taxes is as follows. Each 15^{th} of March, June, September and December, firms at the CIT have to pay a quarter of their taxes which are down payments. Because they still don't know the final amount of their current year taxable income, the tax administration

[2] Quelles charges peuvent être déduites du résultat fiscal d'une entreprise ? French Ministry of the Economy website
[3] Régime réel simplifié.
[4] Quelles différences y a-t-il entre le régime simplifié d'imposition et le réel normal ?
[5] Seuils des régimes d'imposition de la micro-entreprise.

bases its computation on the last taxable income available, so the one of the year before[6] or simply on the last CIT due taxes. In May of the coming year, firms have to pay (or being paid back) the net amount of their true tax liability minus what they have already paid.

Firms can, under some conditions, choose whether to be taxed at the CIT or at the personal income tax (PIT) of their owners[7]. For instance, SARL (Limited liability companies), SA (limited companies) and SAS (simplified joint-stock company), can choose PIT instead of CIT if they do not have more than 50 employees, have a turnover below €10millions, are owned at least at 50% by a natural person and at least 34% of the voting rights are owned by one or several people in command. Note that, if a firm decides to move from CIT to PIT, it cannot move back to CIT.

The French corporate income tax was introduced in 1948 at a rate of 50%[8]. The rate remained the same until 1985 and was then decreased to 45%[9] in 1986. Between 1986 and 1993, the rate is gradually decreased to 33,3%[10]. However, all the companies, whatever their status or income, faced the same rate. In 1995, an additional contribution (*contribution additionnelle*, also known as *surtaxe Juppé*) is introduced. It is defined as 10% of the total corporate income taxes due by each firm, such that the final rate is $(1+10\%)\times 33.3\% = 36.63\%$[11]. In 1997, an exceptional contribution (*contribution exceptionnelle*) was introduced, in the same way as the additional one from 1995, at a rate of 15% for the years 1997 and 1998, 10% for 1999[12]. This contribution only concerned firms with turnover above 50 million of Francs (€7.63 millions) and was removed from 2000 onward. Thus, there were two marginal rates at that time, a 36.63% rate below the threshold and a $41.62\%(=(1+10\%+15\%)\times 33.3\%)$ or $39.96\%(=(1+10\%+10\%)\times 33.3\%)$ above, in 1997 or in 1998-1999 respectively.

In 1997 came another reform: a reduced rate of 19% was implemented for the part of profit that was incorporated to the capital of a firm, up to a quarter of the profit or 200.000 of Francs (€30.000). Only firms with a turnover below 50 millions of Francs were concerned, and at most for three consecutive in credit (strictly positive profit) fiscal years[13]. This reduced rate was removed from 2001 onward. For now, we abstract from this reform (that has been studied by Bauer and Rotemberg [2018]).

[6] L'impôt sur les sociétés, comment ça marche ? French Ministry of the Economy
[7] Entreprises soumises à l'impôt sur les sociétés : comment opter pour l'impôt sur le revenu ?
[8] Décret n°48-1986 du 9/12/1948, art. 1
[9] Article 219 du CGI, Version en vigueur du 12 juillet 1986 au 18 juin 1987
[10] Article 219 du CGI, version en vigueur du 01 janvier 1993 au 02 septembre 1994
[11] Loi n° 95-885 du 4 août 1995 de finances rectificative pour 1995 (1)
[12] LOI no 97-1026 du 10 novembre 1997 portant mesures urgentes à caractère fiscal et financier (1)
[13] Loi de finances pour 1997, Art 10-I

2.2 The 2001 reform

In 2001, a kink is introduced in the CIT schedule at €38.120 of taxable income for firms with turnover below €7.63 million and detained at least at 75% by natural person or by firms themselves detained at least at such a threshold. The rate below the threshold of taxable income is 25% instead of 33.3%. The public finance law was voted on December 31^{st}, proposed on September 20^{th}. Thus, theoretically, the information is available to everyone starting from this date. However, the salience of this reduced tax rate is surely not perfect until some years though every firm entitled to it benefited from it. This reduced rate is then decreased to 15% from 2002 onward[14]. This reduction, however, is already announced in the law of 2001, and already in the proposition of September 2000. Though not perfectly salient, firms have the information at least one year and a half before the fiscal year of 2002. Note that in 2001, the additional contribution presented in the previous section is reduced to 6%, in 2002 to 3%, and removed from 2006 onward. Thus, the real CIT rate in 2001 is $25\% \times (1+6\%) = 26.5\%$ for firms targeted by the reduced rate, 35.3% otherwise. In 2002, it is respectively 15.45% and 34.3% up until 2005 included (see table **??** or figure 1).

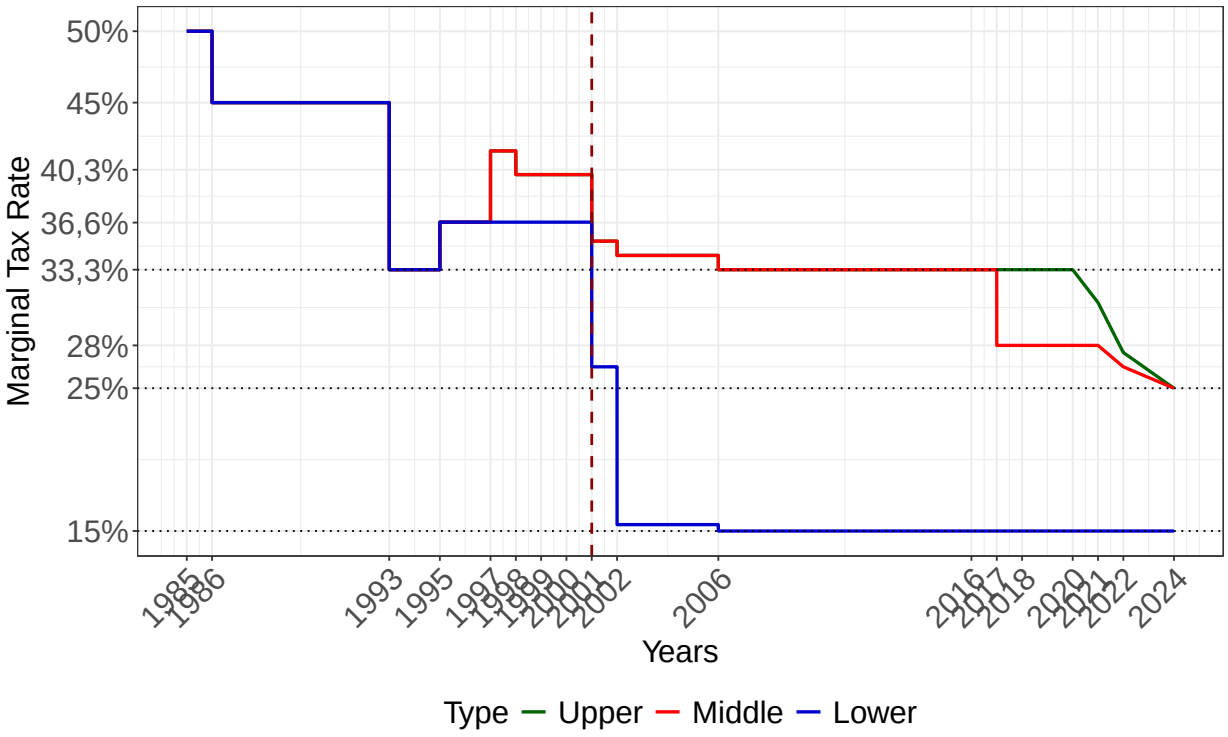

Figure 1: Marginal corporate income tax rates across time

Denoting the kink by k^*, τ_0 the rate below the kink, τ_1 the rate above, x the taxable income, one

[14] Article 219 du CGI I.b.

can write the tax system as:

$$T_{cit}(x) = \tau_0 \times min(k^*, x) + \tau_1 \times max(0, x - k^*)$$

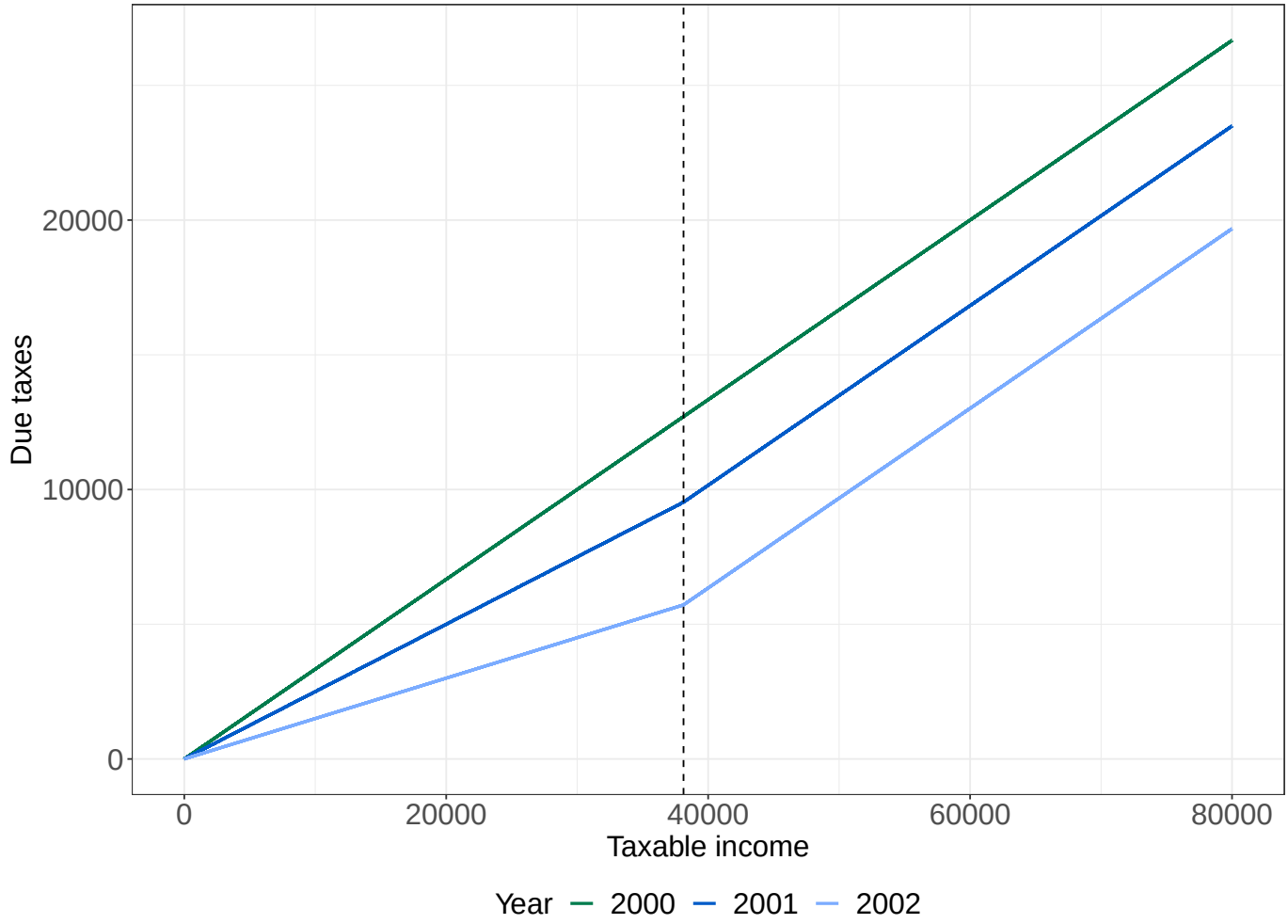

Figure 2: Tax schedule before and after the 2001-2002 tax reform

As written above, this tax schedule only concerned firms with turnover below €7.63m (see figure 10). If we write the the general tax system, with y the turnover, N^* the turnover notch, then:

$$T(x,y) = T_{cit}(x)\mathbb{1}_{\{y \leq N^*\}} + \tau_1 \cdot x \cdot \mathbb{1}_{\{y > N^*\}}$$

Finally, the last condition is the 75% natural person threshold. To benefit from the reduced rate, a firm needs to be detained at least at 75% by natural person or by firms themselves detained at least at such a threshold. For instance if a firm is detained at 34% by a natural person, by 48% by a firm itself detained at 75% by natural person and the rest by other legal persons, then this firm can benefit from the 15% rate. Indeed, adding the two percentages, the firm is detained at 82% by natural persons. If the ratio of detention went from 48% to 40% then the firm could not benefit from it, since it would be

detained only at 74% by natural persons[15]. This creates another notch.

Consider a firm close to the threshold N^*, that satisfies the requirement regarding natural persons and which needs to choose between being above this threshold or not, and suppose that their taxable income is strictly above k^* so that the incentive is not strong enough to make them move on the taxable income dimension. The gain of bunching at the notch N^* from 2002 to 2016 is:

$$\begin{aligned} Gain &= \tau_1 \cdot x - \tau_1 \cdot x - k^*(\tau_0 - \tau_1) = k^* \cdot (\tau_1 - \tau_0) \\ &= 38120 \times (0.33 - 0.15) = 6861.6 \end{aligned}$$

Thus each firm just above the threshold N^* faces an incentive to bunch that is equal to €6861.

In 2013, a corporate income tax credit based on payroll was implemented. This tax credit allowed firm to get a fraction of a part of their payroll back. This policy was targeting wages below 2.5 times the minimum wage. The fraction given back to firms was 4% in 2013, 6% in 2014, 7% in 2017 and back to 6% in 2018. Thus in year 2013, firms with at least one worker paid below 2.5 times the minimum wage could get back 4% of the total amount of wages paid that are below 2.5 times the minimum wage. This policy was put in place to reduce the cost of employing low-skilled workers[16] [Carbonnier et al., 2018, 2022].

In 2017, a phased-in decrease in the CIT rate was implemented. The rate went from 33.3% to 28% for SMBs (Small and Medium size Businesses) only on the first €75.000 of taxable income, which in France are defined by a turnover below €50millions and less than 250 employees. For the year 2018, this rate was levied up until €500.000 of taxable income.

Finally, in 2021 and 2022 the government implemented a double expansion of this reduced rate. The notch was moved from 7.63m€ to 10m€ in 2021. The kink was moved from 38.120€ to 42.500€ for fiscal years that end from December 31st, 2022[17]. This yield an interesting variation since the date from which this new threshold was available was not clear to all companies according to anecdotal evidence[18]. Indeed, some official websites were misleading: as they presented this threshold modification as something that would start by January 1st, 2023[19].

[15] See III.b.2. for a graphical representation. BOFIP 06/2023 IS - Liquidation et taux - Taux réduit applicable au bénéfice des petites et moyennes entreprises - Redevables concernés

[16] What is the CICE? French Economic ministry website

[17] Imposition des résultats - Impôt sur les sociétés

[18] Business owners websites on tax legislation

[19] What change from January 1st, 2023? Official website of the Ministry of the Economy

3 Conceptual Framework

This section is dedicated to shed light on how the tax policies presented in the previous section changed the incentives of firms toward their choice of production.

Suppose a firm i produces an outcome y using non-labor inputs $c(y)$ and labor inputs w. This generates an accounting income of $y-w-c$. For simplicity suppose that w is the wage of the owner-manager and that c includes all other wages. If the firm is levied at the CIT, then they face a tax schedule on their taxable income $y-w-\alpha \times c$ with α a parameter representing the different costs that the firm needs to add back (see institutional background in section 2.1). This parameter is thought of as being below 1 (see also Devereux et al. [2014]). The owner can decide how to receive this outcome: through a wage that they can deduct from the corporate taxable income (w), through dividends (denoted d) or both. On both type of income the owner will pay the PIT. If the owner choses to receive dividends there is an abatement of 40% such that they will pay taxes on $w+0.6\times d$. The firm can also evade taxes through over-reporting of its costs c to pay less corporate taxes. Finally, the firm can also avoid taxes by shifting income through higher wages. Firms can also be levied at the PIT in which case they can pay taxes on the whole profite including the owner-manager's wage (non-deductible) and there is no abatement on dividends.

First we will study the production decision of a firm with and without evasion. Second, we will study the distribution decision with and without avoidance including the choice of the tax schedule.

3.1 The production decision

The profit of the firm can be written as

$$\pi = y - c(y) - w - T_{cit}(y - w - \alpha \times c(y))$$

As of 2001, the CIT schedule is kinked such that there are two different marginal tax rates based on the value of the taxable income (see section 2.2). For a given marginal tax rate τ, the optimal production decision will be made according to the following equation:

$$c'(y) = \frac{1-\tau}{1-\alpha\tau}$$

Here, bunching at the kink could already emerge based on the shape of the cost function. If the cost function is still quite linear around the kink, then the firm is somehow still productive enough to go above the kink. However, if the cost function is very convex at the kink, then a marginal increase of

the output makes the cost very different on each side of the kink such that it is better for the firm to stop at the kink. This mirrors the standard explanation of bunching through productivity.

It is possible to introduce evasion as in Bachas and Soto [2021], Lobel et al. [2024]. For simplicity I will only introduce it for the cost, c still being the real cost and $\tilde{c}$ being the reported cost. The firm can then reduce its taxable income facing a cost of evasion $g(\tilde{c}-c(y))$. Another first order condition is then added which is

$$g'(\tilde{c}-c(y)) = \alpha\tau$$

If we compare before and after the policy reform, firms that are evading through costs over-reporting below the kink will pay less taxes since the marginal (and average) tax rate is lower. Thus they can reduce the extent of cost over-reporting so as to still pay less taxes and face a lower cost of evasion while having revenues unchanged. This is another channel of bunching.

Finally, the reduced marginal tax rate might play a role in the decision to enter the market. As it reduces the average tax rate, another part of the distribution of firms that would not enter under a 33,3% rate could enter under a 15% rate.

3.2 The distribution of income

Once the output has been produced, the owner-manager can decide how to distribute the realized income. Anecdotal evidence on business-owners' website show that at least some firms adjust wages so as to remain below the kink[20]. It is fairly easy to adjust the wage so as to prevent the higher marginal tax rate to apply on the part of profit above the kink as it is tax-deductible, though under some conditions. The wage needs to correspond to a real work done by the owner-manager for the firm and it should not be immoderate with respect to the kind of tasks done or to the total of sales for instance[21]. These conditions are quite vague. Formally, the distribution of income can be written as the program that follows:

$$\max_{w,d} w + d - T_{pit}(w + 0.6 \times d) \text{ where } d = y - w - c - T_{cit}(y - w - \alpha \times \tilde{c})$$

Firms levied at the PIT can decide to turn to the CIT due to the reduced tax. These firms are levied on the entire income, there is no abatement on dividends and they cannot deduct the wage they pay themselves. Depending on the legal status, they can either deduct the other type of costs or apply an abatement on the taxable income fixed by the government. The latter is preferred if a firm has few

[20] Comparison between the CIT and PIT schedule in a business owner website.
[21] Régime fiscal des rémunérations des dirigeants de société (personne physique)

costs. If one write it formally:

$$\max_{w,d} w + d - T_{pit}(w + y - \tilde{c}) \text{ where } d = y - c - w$$

so it can simply be written: $y - c - T_{pit}(w + y - \tilde{c})$. In which case an owner-manager has no interest in paying themselves a wage.

4 Data and descriptive statistics

This paper aims at

4.1 Data

To study the reaction of firms to changes in the tax code, one needs to rely on tax administrative data. In this respect I use French Tax Administrative data

Tax administrative accounting data I rely mostly on reported accounting data provided by the tax administration. As explained in section 2.1, firms have to report their accounting information to the tax administration every year. The documents filled are named *liasses fiscales* and display the total of sales, the payroll, the breakdown of assets (current and fixed), of costs and so on. The information is available for all normal-schemed firms (RN) in the BIC-RN data from 1999 to 2016, and then for all RN and RSI (simplified scheme) from 2016 to 2023 using the BIC-IS. In order to retrieve the RSI before 2016, I use both the BIC-RS data from 2010 to 2016 and the newly available data FICAS and BRN. This allows me to have the same variables from 1993 to 2008 for both firms at the RN and at the RSI. Since firms are indexed by a number (the *SIREN*) I can track them during all the period. In the end I have a 30-year exhaustive panel for RN firms and a 30-year exhaustive panel with a gap of 2 years (2009-2010) for RSI firms.

In addition with the information mentionned above, firms also have to report their accounting income (profit) and their taxable income plus all the deductions and costs added back. Furthermore, as firms have to report their breakdown of costs, they also have to report the amount of due CIT. Since they report in May N+1 (if they ended their fiscal year in December N, 4 months after otherwise) for the realized outcomes of year N, they already know their taxable income. They have to apply the correct tax schedule to their taxable income and report it in the row *hk*. Note that this *hk* variable includes all the possible tax credits. It is supposed to be the amount of due CIT after all has been realized and computed. Note that because it is reported by the firm, all the computations are made

by the firm not by the tax administration. The drawback is that I don't observe the correct amount of due taxes paid effectively by the firm, as it is not provided by the tax administration. However, what I do instead is applying the correct tax schedule to the reported taxable income. This allows me to understand which firms understand the tax code and its changes across time and tag them by their understanding. I need to make the assumption that the way I compute the due amount of CIT is sufficiently accurate for my tagging to be accurate too. My main assumption is that my computation is correct enough so that after a substantial change in the tax code, I can track which firms understand in which year. Fortunately, I can check before the reform whether my method correctly aligns with what is reported by firms. For my computation to be correct even without a reform, I would need to micro-simulate the due taxes and use all the tax credits available. This is possible but is for now left for a further stage of this paper.

Employer-employee data I exploit the French linked employer-employee data (DADS) to compare the reported payroll in the accounting information and the one available in this dataset. Each month, firms have to declare the wage they pay to which employee in order to pay the social contributions on each corresponding wage. Information available in these type of data are thought of as less prone to misreporting since they have a direct impact on employees social protection since these data are used by the Unemployment office for instance. Thus it allows me to check and compare the two reportings. The data are available exhaustively from 1994 onward such that it covers the reform but also 7 years prior to it.

Financial linkages data I also rely on financial linkages data (LiFi). It is a dataset that follows groups of companies and tries to track which firms are owned by which firm. From 1985 to 2011, it is only a survey, but from 2012 it is exhaustive. Unfortunately, I don't have the data for firms that are not part of a group. Thus it is not possible to really track which firms is effectively treated or not. However, I make the assumption that firms I will study are usually small enough that they are detained in majority by natural persons. However, for the most recent years, this dataset allows me to observe the discontinuity at the ratio of natural person and thus perform a RD analysis.

Companies' registry data Finally, in order to have the correct year of creation of each firm, I use the companies' registry data (Stocks entreprises and Stocks Entreprises et Etablissements).

4.2 Descriptive statistics

Tables 1 and 2 provide some descriptive statistics of RN firms in 2000, the year prior to the reform. The first one gives figure of firms based on their taxable income while the second based on their revenue. Firms around the kink before its implementation are smaller than the average firm. They have an operating revenue of less than 1m€, around 980.000€ and an operating cost of 940.000€. Note that their accounting income is smaller than their range of taxable income showing that on average, a firm needs to add back more costs than what they are allowed to deduct. Firms just below the notch are bigger than the average firm. They have on average a fixed asset of 2.4m€, around 3 times more than the average firm. They also have an accounting income of 165.000€ on average, higher than the kink suggesting that their taxable income is also way larger than the kink.

5 How do firms learn the tax code

In this section, I use the accounting data reported by firms especally the variable *hk* presented in section 4. First I will show how firms learn the tax code and how much do they take. Then I will study which firms learn and which don't. Finally, I will study the channels of learning.

5.1 The how

In this subsection I will only focus on firms below 80.000€ of taxable income and for now I only use RN firms. The results will be extended soon.

In 1999 and 2000, there was no change in the tax schedule. Thus it is possible to observe whether the reported amount of due taxes from firms follow the then tax schedule. In figure 3, the plain lines show the true tax schedule, and the dotted lines show the average reported amount of due taxes from firms by taxable income. In 1999 and 2000, the dotted lines closely follow the green plain line. Firms, before the 2001 reform, were reporting correctly in accordance with the true tax schedule. In 2001 however, the blue dotted line does not map the first blue plain line. There is a discrepancy between the average reported by firms and the true tax schedule. In 2002, when the reduced rate is decreased from 25 to 15%, there is still a discrepancy. The dotted line is in between the two blue plain lines, such that, an adjustment has started but is not yet achieved. From 2003 until 2005, there is a convergence of the average reported due taxes to the true tax schedule, which is the lowest blue plain line.

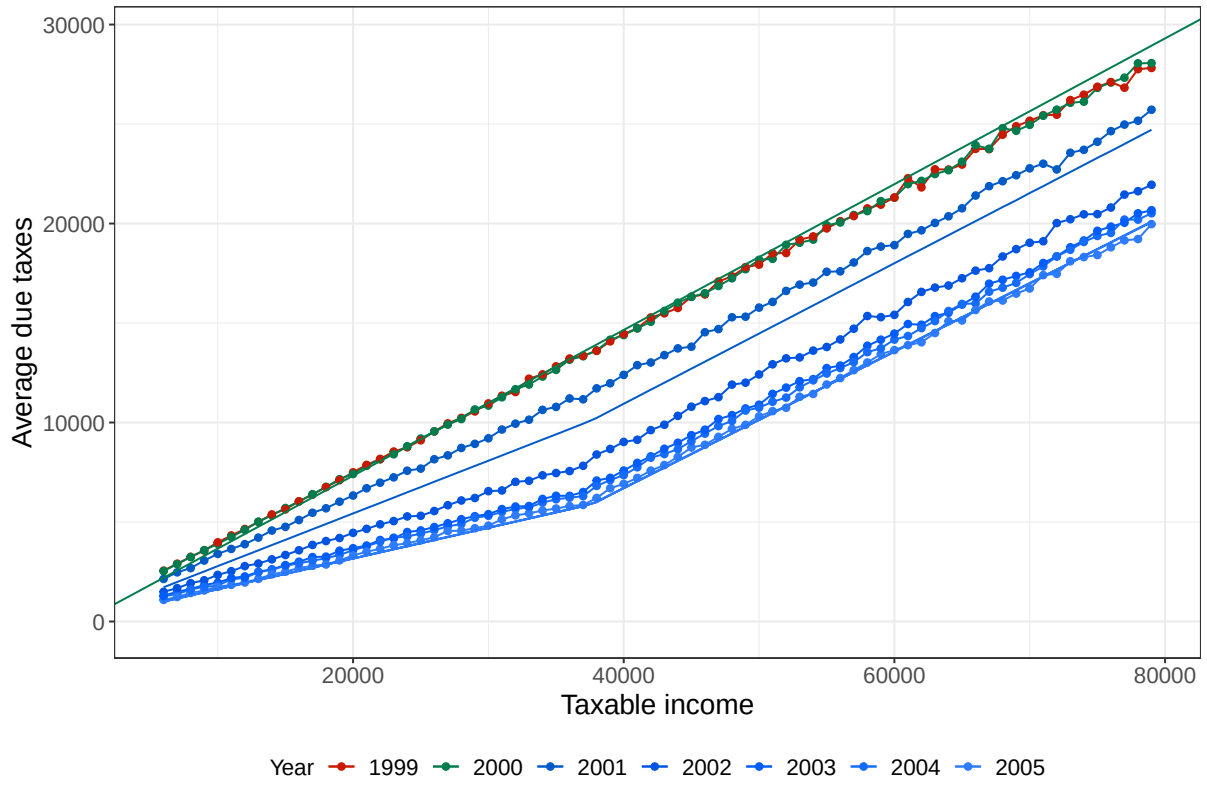

Figure 3: Average expected tax liability from firms, by bins of taxable income across years

Note: In 2001 (the first dark blue lines) the average expected tax liability as reported by the firm (dotted line) to the tax administration differs from the true tax schedule (plain line). In 2002, the reduced rate is further reduced from 25% to 15%, there is still a discrepancy. However, as we move forward to 2005, the knowledge of firms converges to the true tax schedule.

In order to understand how firms learn, I compute the counterfactual amount of due taxes in the absence of the reform. To do so, I apply the standard rate of 33.3% to the taxable income. In such a way, I can observe which firms continue, year after year, to report the due taxes as if it was the 2000 tax schedule.

In figure 11, the number of firms reporting using the 2000 tax schedule reaches 0 only in 2005. It means that for up to 5 years, a part of firms were using the pre-reform tax schedule. Note that this could also be partly explained by firms detained by more than 25% of legal persons that modify their shareholders structure so as to benefit from the reduced rate. On the same figure, the green dotted line shows the number of firms using the 2001 schedule. There was still a large amount of firms using this schedule in 2002. This suggests that either a part of firms switched from using the tax code of 2000 to the one of 2001 and still used it in 2002 or that some firms still using the 2000's one in 2001 did their adjustment in 2002.

In order to understand how firms understand changes in the tax code, I take the difference between the reported amount minus the counterfactual amount and divide the difference by the reported one. I call this ratio the normalized ratio of reported taxes. For firms below the kink, a 0 means that a firm

reports using the 2000 tax schedule (counterfactual no-reform), a 0.33 means that a firm reports using the 2001 tax code and 1.22 using the 2002 tax code. If the ratio is different from these three values, it means that firms report with other kind of rates in mind.

Figure 12 depicts the distribution of this normalized ratio of reported taxes. There are only very few other values than the three ones presented. This means that firms either do or don't learn but do not have other rates in mind than those that are present in the legislation. The learning process is thus a binary one.

5.2 Who's learning, who's not

In order to understand which firms are learning and which are not, I compute what firms should have reported applying the current and correct tax code to the reported taxable income.

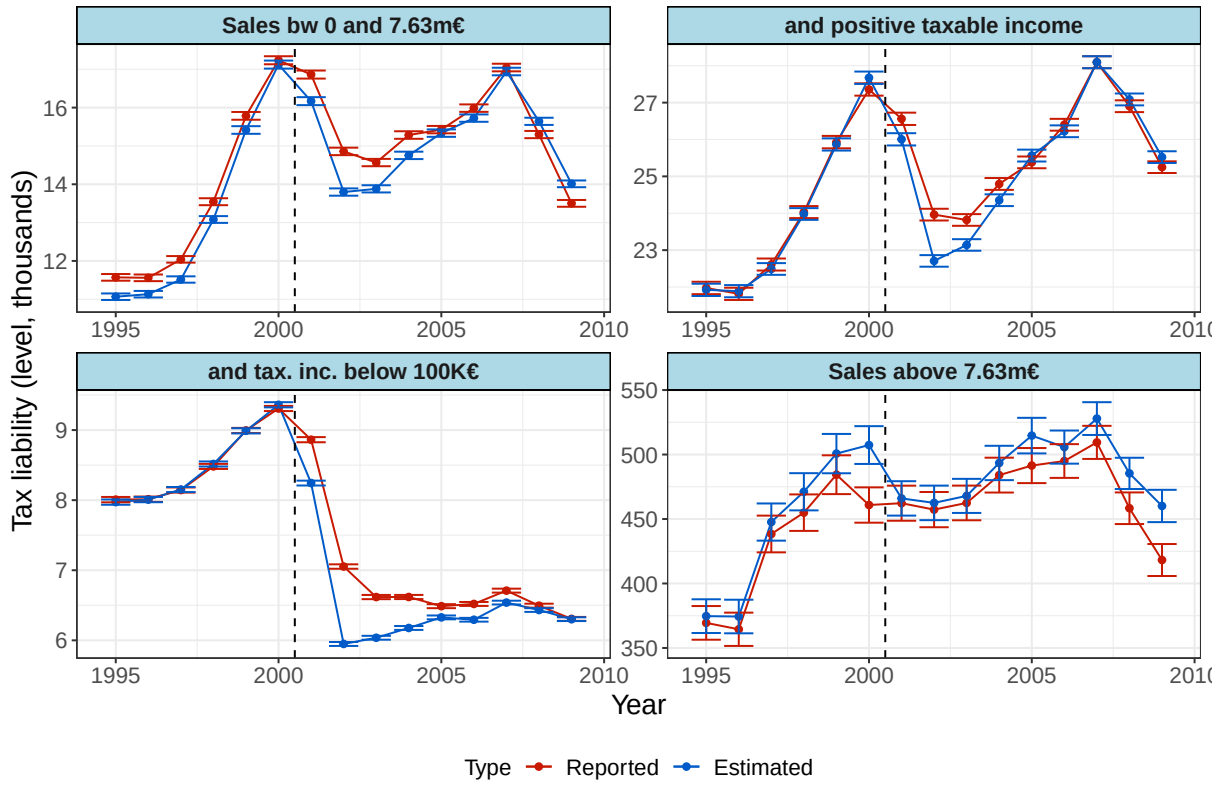

Figure 4: Average reported and computed tax liability by size across years

Note:

Figure 4 presents my computation both with the average reported amout of taxes by size across years. Before the reform, my computation closely maps the reported average which acts as a sort of *pre-trends* test. The last group, firms having sales above the notch, do not exhibit a discrepancy between the two. This is reassuring since these firms are not affected by the reform. However, the three first groups display substantial discrepancies between the reported and computed amount of

taxes. Though the averages are in level, the smaller the firms, the larger the difference in relative terms. In 2002, for firms below the notch and below 100.000€ of taxable income there is a 1000€ difference for a reported amount of 7000€. This translates into a 14,3% difference with respect to the reported amount.

Turning to the breakdown of this averages across time by number of employees, figure 16 shows that the smaller the companies in term of size, the longer they take to adjust and understand the tax code. Indeed, firms with no employees (top right panel) took many years to finally reach the correct tax schedule. However, firms with more than 10 employees (bottom right panel) were correct by 2005. For firms with more than 50 employees, the discrepancy is lower from the beginning and seems to be insignificant by 2005 too.

Figure 13 depicts both averages across years by the age of firms. Before the reform, for all age groups, the estimated and reported due amount closely follow each other. As before, after the reform, we observe a large shift of the two. However, it is surprisingly similar across age.

If we turn to the breakdown by industries, there are substantial variations across them (figure 14 and 15). Sundry professional, scientific and technical activities as well as Information and communication and Transportation and storage are the three industries for which the shift is the largest and takes more time to recover. Note that it could be also the industries in which firms are less likely to satisfy the 75% threshold on the ratio of natural person.

When it comes to sales, figure 17 shows the shift in the reported amount of due taxes by sales as percentage of the notch.

To continue study which companies learn,

Finally, random forests

5.3 The channels of learning

6 Behavioral responses

This section starts with graphical evidence of the bunching at the different discontinuities of the policy. Then it goes into the dynamic of the behavioral responses. After that, the Elasticity of corporate taxable income to the net-of-tax rate is estimated and the different channels of the Behavioral responses are considered. Finally, these parameters are used to compute the welfare effects of the reform.

6.1 Bunching evidence

As presented in section 2, a kink is introduced in the CIT code in 2001. Figure 5 shows the distribution of companies by taxable income around the kink pooled from 2001-2007. The distribution exhibits substantial bunching at the kink.

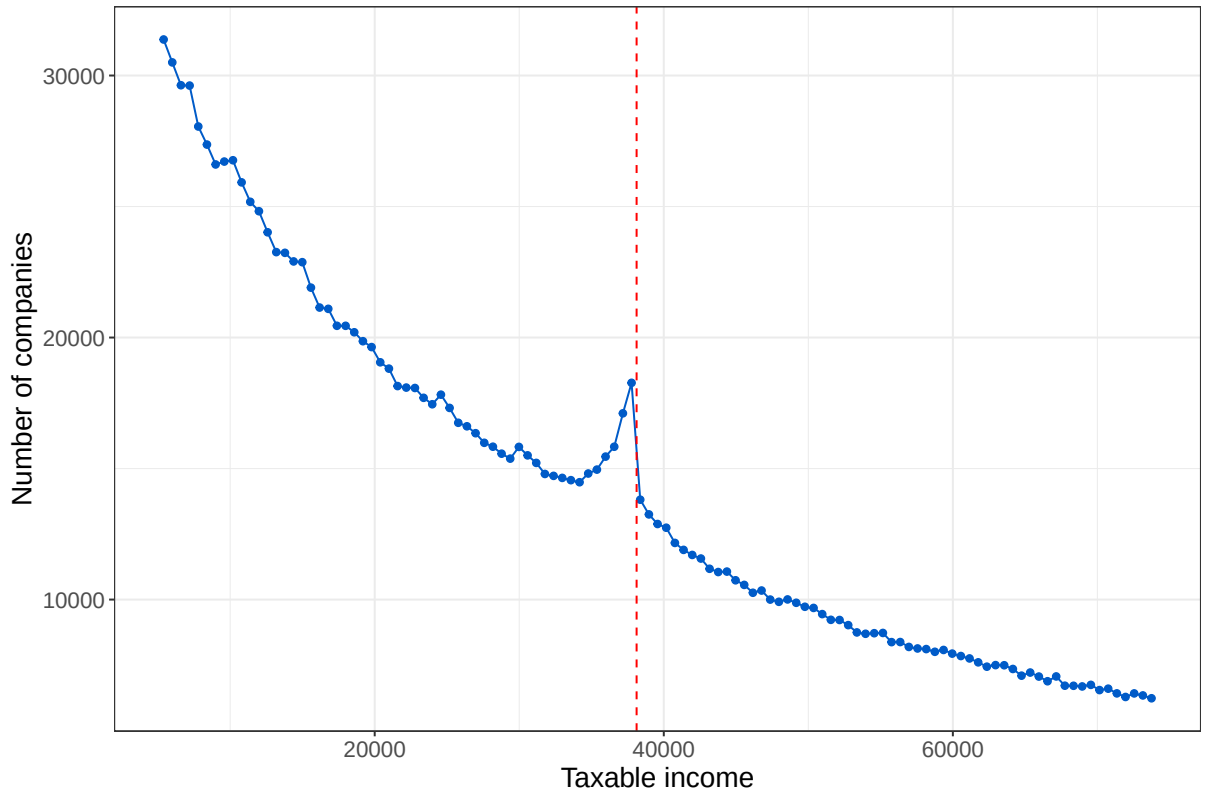

Figure 5: Distribution of companies by taxable income (2001-2007)

Note: Distribution of companies by taxable income around the kink pooled from 2001 to 2007. The red dashed line shows the location of the kink.

Figure 18 shows distribution of companies by total sales between 2001 and 2007 around the notch. The distribution does exhibit a small bunching mass. However, it is smaller than the kink. This can be explained by the small incentive to bunch for the type of firms around the notch. Firms close to the kink have an accounting income above 165.000€ (see table 2 in section 4.2) such that the incentive is less than 4% of their accounting income. As the incentive is quite low, there is only little bunching.

Finally, turning to the notch at the natural person ratio threshold figure 19 shows the distribution of companies by such a ratio for years 2006-2007 pooled. There are two extreme points, 0% and 100%. Removing these two points yield figure 20 in which we see that there are large masses at different ratios. Masses are located at very standard points such as 50%, 25%, 33.3% or any multiple of 10 which represent firms held by few shareholders and thus have those standard ratios. The 75% percentage does not seem to have a larger mass than other similar points. Though it seems that there

are more firms 1pp before than 25% for instance. This would suggest bunching behavior although of a relatively small amount.

6.2 Dynamic learning and learning cost

The figures of the previous section were all static and pooled multiple years. It is possible to learn something about the behaviors of firms by studying the dynamic of behavioral responses.

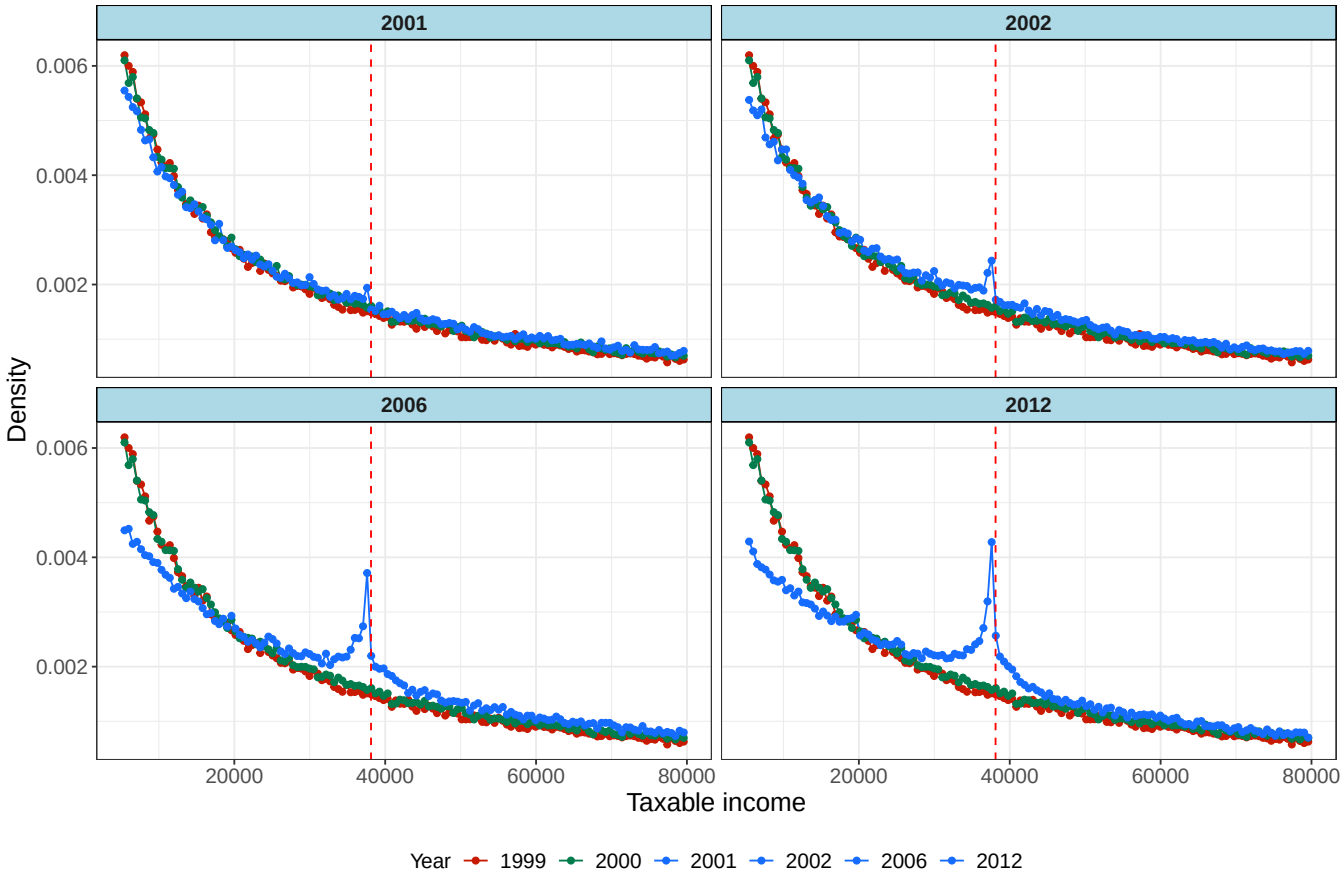

Figure 6: Distribution of companies by taxable income across years (2001,2002,2006 and 2012)

Note: Distribution of companies by taxable income around the kink for years 2001,2002,2006 and 2012. The red dashed line shows the location of the kink.

Figure 6 depicts the distribution of firms by taxable income around the kink, but this time broken-downed by years. In 2001, the year the reform is implemented, there is only a little bunching though the difference of marginal rates is 8pp. In 2002, the reduced rate is decreased from 25% to 15%. This time the bunching mass increases compare to 2001 which is consistent with the increased incentive. From 2002 to 2006, the tax schedule doesn't change, the bunching mass does however. This increase continues up until 2012 in which year the bunching mass reaches its maximum (see figure 21 for detail per year). This dynamic is very consistent with the learning process of the tax code explained in section 5. Indeed, firms get to know the tax code by repeatedly having to deal with it. Bunching

behaviors also display such a learning process. Using the tagging done in section 5.2, figure 23 shows the differential bunching based on the learning status. The bunching mass and its dynamic is entirely driven by firms that are correct in computing their due amount of taxes (right panel). Firms that are not correct do no contribute to the bunching mass until 2004. However, from 2005 onward incorrect firms do contribute to the mass (see figure 24). It is surprising that firms tagged as incorrect do bunch. It would mean that they understood where the kink is but maybe not the marginal rate change. Another explaination is my tagging starts to be inaccurate from 2005 which seems reasonable too. Indeed, as explained in section 5.2 my main assumption is that my computation is accurate enough so as to capture the learning process close the reform year. After 5 or 6 years, the threshold used to capture correctness about the computation of the due taxes becomes random in a sense since there is no large variation anymore. In order to really capture correctness away from large variations one would need a very accuracte micro-simulation. However, incorrect firms exhibit frictions, as the bunching is far more sharp for correct firms than incorrect ones. Though one might consider that the tagging is less accurate, these frictions still say something about these firms. They might still be confused about the tax code or be less strict on where they want to bunch and may really on other heuristics. This would relate to Rees-Jones and Taubinsky [2020].

Figure 25 shows the dynamic around the notch of total sales. The distribution exhibits a little dynamic though much smaller than around the kink. This is consistent both with learning and the fact that the incentive is quite small for firms around this notch.

Now that we have shown that the dynamic is led by learning, it is possible to recover meaningful information on the size of the cost of learning. Indeed if firms are bunching, it means that their optimal choice is to bunch. Contrary to all other level of taxable income, which would either be the optimal decision of an unknown maximization program or a suboptimal decision due to unknown frictions, when firms are bunching, it is their realized optimal decision. Such that if a firm locates multiple time at a discontinuity during a period of time, it means that being at such a discontinuity is its optimal deicison. I consider repeated bunchers in this identification framework to be certain that it is a bunching behavior that I capture. Because I am able to know whether firms understood the tax code or not, I can use the bunching decision in addition with the learning status to recover these cost parameters. Relying on the panel dimension, if a firm bunches multiple times within a period but starts strictly after understanding the tax code (learning happens in year N and bunching in N+1), it means that there are other types of costs that prevent them from doing so. I will call these costs pure adjustment costs. If a firm bunches multiple times within a period and starts strictly the year they learn, it means that what prevented such a move was a learning cost. In either case, I take the difference between the taxable income the year prior bunching and the kink for firms in each sets

defined previously.

Here, I am going to use the bunching at the kink, though of course it would be more convenient in term of identification to use the notch of sales. Unfortunately for the identification the bunching mass is small, the number of firms around the kink even without notch is small too and the estimation would suffer from a lack of power. This is the reason why I rely on the kink. The parameter I will be able to estimate is not an average but a bound, in this case an upper-bound.

I use the period from 2001 to 2007 and take the average by pooling all years. This yield the results shown in table 3. On average, the cost of learning is 7000€ which represents 24% of the original taxable income. This cost is substantial. It means that the year they learn, firms are able to increase their taxable income by a quarter. When we exclude learning costs from the adjustment costs, the average pure adjustment costs amounts to 3900€. This average amounts to 12% of the taxable income which is substantial though much smaller.

6.3 Elasticity estimation

Bunching is a form of behavioral responses from which it is possible to estimate elasticites Saez [2010]. In this section, I rely on the method developed by Bertanha et al. [2023]. The previous methods of estimating elasticites through bunching were relying on z restrictive assumption on the underlying and unobservable distribution of the parameter that generates bunching, usually thought as ability or productivity. Bertanha et al. [2023] and Pollinger [2023] propose methods that rely on less restrictive smoothness assumptions: the Lipschitz continuity with constant M of the underlying distribution or that such a distribution has a convergent power series representation, respectively. I estimate the elasticity cross-sectionally using the Stata package of Bertanha et al. [2022].

Figure 7 shows these cross-sectional elasticites. The evolution mirrors the one of the bunching mass in the previous section. The average elasticity is the one using the trapezoidal rule of Saez [2010] and the error-bars are the bounds given by theorem 2 of Bertanha et al. [2023] for the largest M for which the set has an upper bound. The elasticity goes from 0.05 in 2001 up to 0.4 in 2012. These values seem reasonable in comparison with Devereux et al. [2014] but below Bachas and Soto [2021], Coles et al. [2022], Bukovina et al. [Forthcoming]. One reason is that these elasticities are computed only using normal-schemed firms (RN) though the extent of bunching seems quite similar for simplified-schemed firms (RSI). These estimates are also consistent with the multi-country analysis done by Agostini et al. [2024] with elasticities ranging from 0.04 to 1.9 and a mean of 0.59.

If we think about the maximum value reached in 2012 as the structural elasticity, it means that it takes 12 years for the elasticity to reach its structural value. 12 years is substantial especially

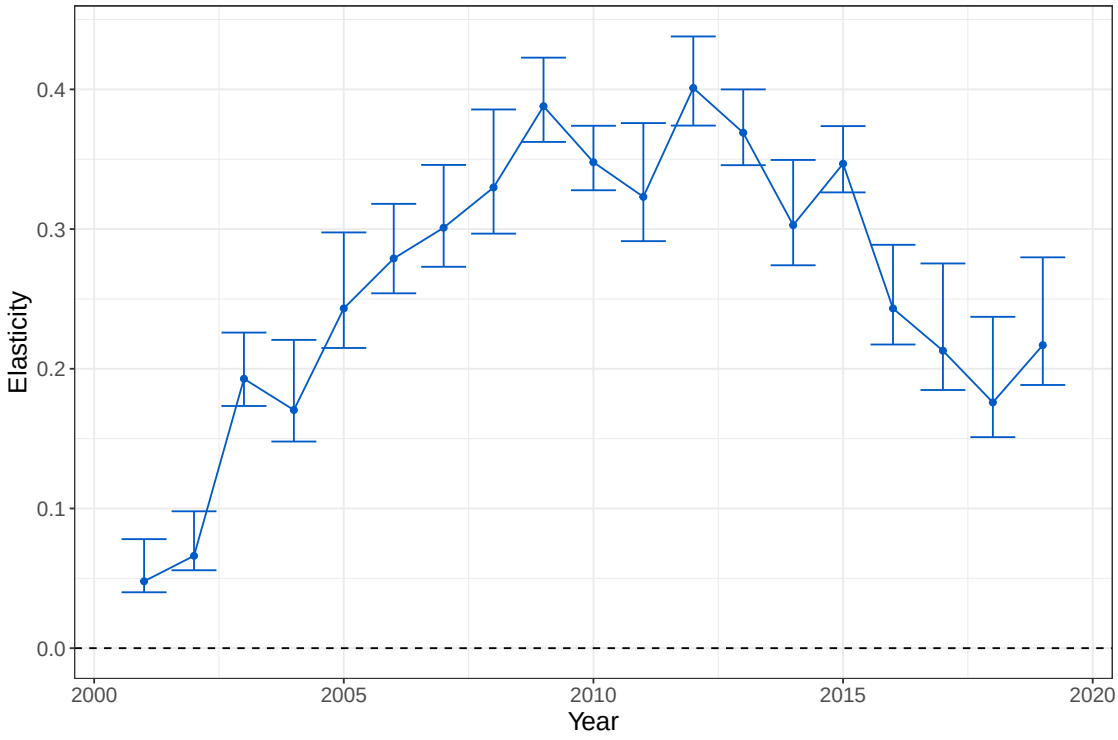

Figure 7: Cross-sectional elasticities using the Bertanha et al. [2023] method

since between 2006 and 2012 the tax code doesn't change. This might raise concerns about what is a structural elasticity [Slemrod and Kopczuk, 2002] and whether this conceptual parameter really makes sense or not. Note also that the development of a market for evasion can dynamically modify the elasticity [Alstadsæter et al., 2019, Bustos et al., 2022, Laffitte, 2024]. I argue that the elasticity is fundamentally a dynamic parameter.

Finally, relying on the learning status defined in the previous section, I can measure the contribution to the elasticity of new learners every year. Define perfect learners as firms being correct in their tax reporting in year $t-1$ and in year t. Define imperfect learners as firms being incorrect in year $t-1$ but correct in t. Conceptually, firms that are imperfect learners in year t should be perfect learners in year $t+1$ unless they make a mistake above 5% of their reported taxable income. Define $P_{it} = 1$ if a firm i is a perfect learner in year t, $P_{it} = 0$ if this firm is an imperfect learner in year t. Note that this indicator function is not a partition of firms as there are two types of mistakers: firms that continuously make mistakes and those that start to mistakes. To measure the contribution of each group, I consider the number of firms N_t at the kink (in bin b_k, hence $N_t(b_k)$) with respect to two counterfactual bins ($b_{c1} < b_k < b_{c2}$, and $N_t(b_{c1}), N_t(b_{c2})$), and define the excess mass as: $EM_t(b_k) = N_t(b_k) - \frac{N_t(b_{c1})+N_t(b_{c2})}{2}$. I do the same for each subgroups, thus conditioning on the

indicator P_{it}. Finally, I take the ratio by groups and define the contribution of learners $C_{it}(P_{it})$ as:

$$C_{it}(P_{it}) = \frac{EM_t(b_k \mid P_{it})}{EM_t(b_k)} = \frac{N_t(b_k \mid P_{it}) - \frac{N_t(b_{c1}|P_{it}) + N_t(b_{c2}|P_{it})}{2}}{N_t(b_k) - \frac{N_t(b_{c1}) + N_t(b_{c2})}{2}}$$

Figure 27 displays such contributions. In 2001, only perfect learners contribute to the bunching mass and thus to the elasticity. In 2002, imperfect learners account for 10% of the elasticity. This suggests that even after at least one year and a half, it is still mainly the perfect learners that contribute to the elasticity. 2003 is the year in which the adjustment takes place as 40% of the elasticity is explained by imperfect learners, it means that the year they learn, they join the kink to bunch. This substantial amount mimics the learning costs measured in the previous section. By 2004 onward, the imperfect learners joined the perfect learners group such that there is a residual but small adjustment. It takes then 3 years for firms to adjust their Behavioral responses.

6.4 Digging into the different channels

From the hierarchy in the timing of behavioral responses of Slemrod [2001], bunching behaviors can be driven either by avoidance (including evasion) motives or real responses. If one considers the black-box aspect of the elasticity, one does not need to investigate the origins of such behaviors. However, if one is interested in the welfare effects of the policy, one needs to.

If the running variable on which the discontinuity applies cannot be misreported, the behavioral responses at such a discontinuity will be due to real responses. However, if the running variable can be misreported, it is usely difficult to distinguish the two channels. In order to do so papers in the literature [Saez, 2010, Garbinti et al., 2023, Lobel et al., 2024, Londoño-Vélez and Avila-Mahecha, 2024] have tried to study hard-to-misreport-type of assets and to study differential bunching with respect to the composition of the running variable of such assets. If economic agents for which the running variable is mainly composed of third-party-reported assets bunch only a little, but those for which it is the opposite bunch extensively, it is reasonable to think that the source of bunching is avoidance. Other papers [Askenazy et al., 2022, Lapeyre, 2024] have directly matched other sources of data that are thought of as less prone to misreporting, like employer-employee data for instance. This section aims at studying the sources of behavioral responses using different strategies. Since it is difficult to have one clearcut result using reported data, I am relying on several strategies. If all the strategies point to the same direction, it would be plausible to say that the main source of these behavioral responses are due to the corresponding channel.

First of all, as explained in section 3, there is some anecdotal evidence that some business owners use the wages they pay themselves to locate their taxable income below the kink. It is not possible to observe these very wages, as it is mixed with other employees wages in the tax administrative data. However, business owners that pay themselves wages need to report their social contributions such that it is possible to use them as a proxy for wages. Figure 28 pictures the average owner-managers' social contributions by taxable income. There is a large increase at the kink which seems to be the empirical counterpart of the anecdotal evidence. For the years 2008-2013 the average goes from around 8.000€ just before the kink to around 10.000€ at the kink and goes down to a little bit more than 8.000€ after. The response would be of 25%, which is substantial but only appears at the kink.

To go further, as papers cited above, I match French linked employer-employee data (*DADS*) to the accouting tax administrative one to track the payroll declared in each data source. Figure 29 displays the average reported payroll by taxable income for firms reporting no employees in the employer-employee data. Here again, the increase is substantial at the kink going from around 48.000€ to around 62.000€ and down to 46.000€ after the kink. This yield a response of 29% at the kink. The fact that business owners' social contributions and wages seem to grow substantially at the kink while the average payroll for all the companies does not react much at the kink (see figure 30). Indeed, it is difficult for owner-managers to adjust through employees' wages but pretty easy through their own ones.

In the same spirit as studying third-party-reported assets, one can study the differential bunching by legal status. If firms in legal status for which it is easier to adjust are bunching more intensively and firms in legal status for which it is harder to do bunch less, it is again a weak signal of tax avoidance response. Figure 31 shows the differential bunching by legal status among commercial entities in level while 32 does it in percentage within groups. Professions Limited liability company (*Société d'exercice libéral à responsabilité limitée*) is a legal status for physicians, lawyers or architects that choose a LLC form. These companies would traditionally be levied at the PIT but with this form they can decide to opt for the CIT. The same goes for Individual LLC. These firms being the ones which bunch the most, with traditional LLC, might show the will of companies to benefit from legal status that are levied at the CIT rather than the PIT. By doing so, they benefit from the reduced rate and can adjust their wages such as to minimize the overall taxes.

Another way to study the channels underlying the behavioral responses is to use variations in the tax code implemented at specific dates by which only one of the different possible channels are available to react. In 2022, the government announced the modification of the kink: for the fiscal years that end from December 31st, 2022 it would be 42.500€ instead of 38.120€. Figure 8 shows the reaction of bunching companies to this change.

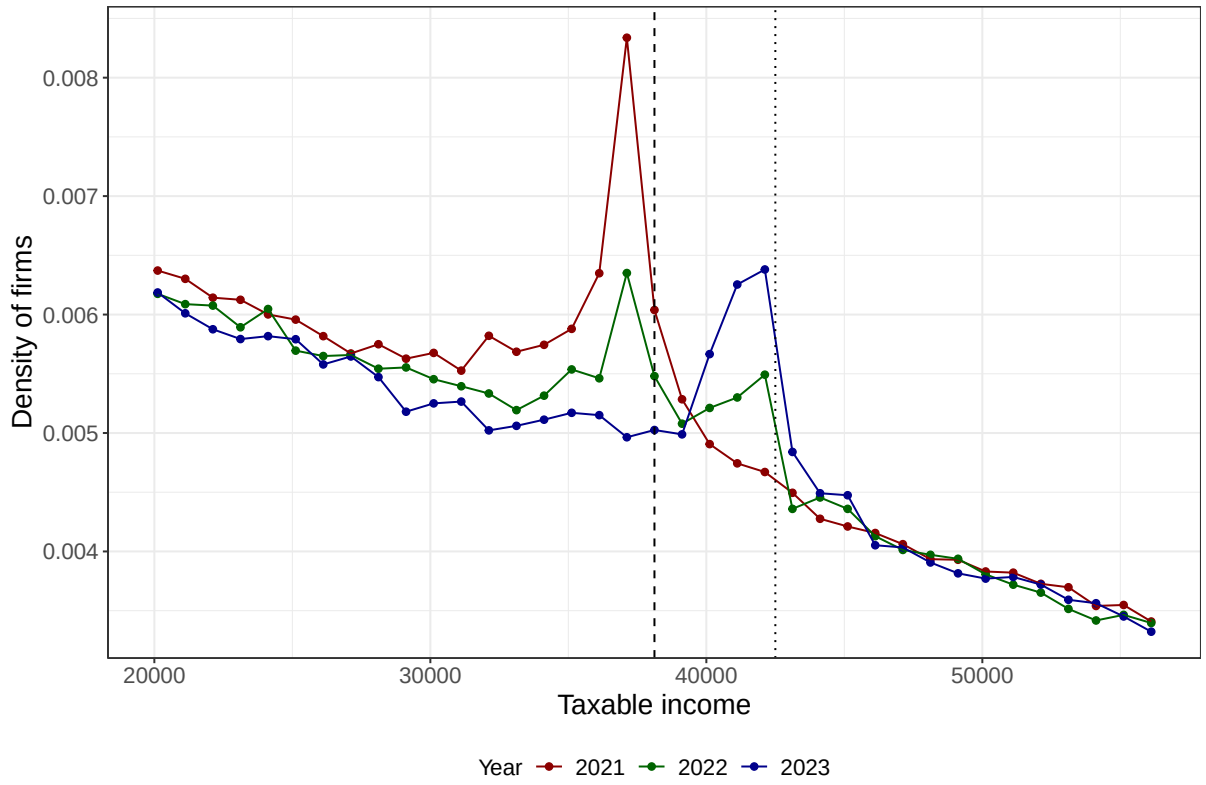

Figure 8: Distribution of companies by taxable income across years (2021-2023)

Note:

Since there were misunderstandings about the implementation date of this reform (see setting in section 2.2), the first adjustment of a part of the bunching mass for the fiscal year 2022 tells something about the underlying channels. First, only sophisticated entrepreneurs were fast enough to react by 2022, the rest did in 2023. Second, firms are reporting their accounting information in year N+1, once everything has been realized: the outputs and inputs can not be modified by real responses. Thus, the adjustment of the bunching mass for the fiscal year of 2022 conveys the idea that it has been made through changes in the reporting only. One could argue that sophisticated entrepreneurs could just be aware of all the legislative debates over the CIT rates and that this very change was discussed before allowing them to adjust. There is no mention of this modification in the original draft of the budget bill of 2023[22]. This amendment was introduced on December 8th, 2022[23], was voted with the law on December 15th and promulgated on December 30th, 2022. It seems quite reasonable to assume that a vast majority of firms were not aware of this change sufficiently in advance to react through real responses, and that a substantial part, if not all, of this 2022 reaction is due to tax avoidance.

Figure 33 exhibits the number of creation by taxable income across years. More precisely it is the location of firms during their first year. There is bunching at the kink for businesses creation. There

[22] Projet de loi de finances pour 2023, n° 273, déposé le lundi 26 septembre 2022.
[23] Amendement n°452 Déposé le jeudi 8 décembre 2022

might be a small effect of the reduced rate on firm creation as the numbers seem to be higher below the kink rather than above. However, the bunching mass at the kink shows that these new firms were created to maximize the amount of taxable income under the kink since on their first year they already bunch. This is consistent with firm splitting, another channel of tax avoidance.

As Brockmeyer [2014], Boonzaaier et al. [2019], Massenz [2024] I study the persistence of firms at the kink. Figure 34 shows the comparative persistence of firms at the kink and firms around it, in the area between 31.120€ and 45.120€ without the are 36.120-40.120€. Firms at the kink strictly remain more at the kink than firms at other level of taxable income, as the blue curve is always above the red one. The number of firms who stay for 2,3 or 4 consecutive years is always more than twofold the same number for taxable income around the kink. This suggests strong capacity to adjust for bunching firms, if not to manipulate their reporting.

Finally, and related, figure 35 shows the intra-firm volatility, the standard error of the taxable income of a firm computed between 2001 and 2016 then averaged by bins of taxable income, shown in 2010. From 5000€ the relation is surprisingly increasing and linear with taxable income. There is a drop at the kink, suggesting that indeed firms are capable of remaining multiple years at the kink as shown just above.

6.5 Welfare effects?

Chetty [2009] has shown that the Feldstein [1999] result of the sufficiency of the elasticity of taxable income to compute the deadweight loss of taxation was relying on assumptions regarding the cost of tax avoidance. If one relaxes these assumptions, one needs to consider whether the elasticity is driven by avoidance or real responses. That is what Devereux et al. [2014] do by distinguishing the elasticity of total income and the elasticity of the share of income taken as profit. Relying on the findings of the previous section, the purpose of this section is to rely on the conceptual framework to give bounds of the welfare effects of this policy.

7 The effects of lowering the CIT rate on small companies

The purpose of this section is to measure the real responses of the reform more broadly than close to the kink. The vast majority of firms were impacted by this reform since the majority of firms have their total sales below the notch of 7.63m€. Thus, compared to a no-reform situation, these firms paid 6800€ less in taxes per year (see section 2.2). This represent between 2.5bn€ to 4bn€ of foregone taxes per year [CPO, 2016]. As the policy is in place since 2001, this amounts to a substantial total.

In order to study whether this reform was efficient or not, I will rely on the discontinuities introduced by the reform and the difference of learning among groups to try and recover causal effects of the policy.

In this section, D_{it} will denote the treatment status equal to 1 if a firm i is treated in year t, 0 otherwise. The outcome will be denoted Y_{it}, any covariate (or vector of covariates) X_{it} and the error term ε_{it}. Finally, I introduce the learning status L_{it} which is equal to 1 if a firm i in year t understood the change in the tax code. Because of the timing of the reporting, it is difficult to know precisely the year in which the adjustment take place. For example, a firm could understand the new tax policy on January 1st, N+1 and report a correct tax amount on its N's taxable income but as the inputs and outputs of year N are realized, the adjustment would only take place in year N+1, reported in May N+2 (see section 2 for the timing of report). In section 7.2 we will consider how to overcome this limitation.

7.1 Using the discontinuities of the policy

Discontinuities in policies can sometimes be used to measure their efficiency through Regression-Discontinuity Design (RDD, Angrist and Lavy [1999], Hahn et al. [2001], Imbens and Lemieux [2008], Lee [2008]). One main assumption behind such a tool is that the running variable can not be manipulated. In our case, section 6 showed that there were substantial behavioral responses preventing from using such a tool. However, there is a literature trying to overcome this limitation, especially one focusing on so-called *Donut-hole* RDD [Barreca et al., 2011, 2016]. The idea is to exclude the area close to the discontinuity where the behavioral responses appear and to compare the units on each side of the threshold without this excluded window. In this section, I rely on this technique as well as on a modified RKD [Card et al., 2012, Landais, 2015] with such an excluded range to study the kink.

7.1.1 RDD at the notch

The RDD relies on the assumption that, within a small window, the location of the threshold is sufficiently random such that the units on each side of it are comparable and that the treatment is as-good-as-random. However, the overlap assumption is not met by design: the treatment depends on a running variable with a strict threshold [Imbens and Lemieux, 2008]. The probability is either 0 or 1 in the case of a sharp design or 0 and strictly positive in the case of a fuzzy one, such that the probability to be treated is not strictly positive and below 1 for each value of the running variable. Thus one requires the extrapolation of the outcome and a continuity assumption.

Define $\tau_{t,RDD} = \mathbb{E}[Y_{it}(1) - Y_{it}(0) \mid X_{it} = N]$ the average treatment effect at N, the notch on sales, X_{it} here being a unique covariate: sales.

Assumption 7.1.1 (Continuity of conditional regression functions)**.** $\mathbb{E}[Y_{it}(d) \mid X_{it} = x]$ *is continuous in* x *for each* $d \in \{0,1\}$

Under this assumption it is possible to recover the ATE as

$$\tau_{t,RDD} = \lim_{x \to N^+} \mathbb{E}[Y_{it} \mid X_{it} = x] - \lim_{x \to N^-} \mathbb{E}[Y_{it} \mid X_{it} = x]$$

It is possible to estimate this quantity by taking the difference between two linear regressions, one for each side of the notch. However, since the running variable can be manipulated, either through avoidance or real responses, the ATE is going to be biased if we include the bunching region. One solution is to exclude such a region. The drawback is that one looses the comparability of units across the threshold. Assumption 7.1.1 is violated in our setting, but equipped with the correct assumptions and setting it is possible to recover the causal effect of the policy.

Because the link function between Y and X is linear, we make the following assumption:

Assumption 7.1.2 (Linearity of the link function)**.** *Let* p *be a strictly positive integer,* δ *a positive parameter, for any* $x \in Supp(X) \setminus [N-\delta, N+\delta]$,

$$\mathbb{E}[Y_{it}(0) \mid X_{it} = x] = \alpha + \sum_{j=1}^{p} \beta_j \times x^j$$

In our setting, and within a strict subset of $Supp(X)$ around N, it is plausible to assume that $p = 1$ (see figures **??**). If we are ready to assume the **SUTVA** (Stable Unit Treatment Value Assumption), it is possible to write $Y_{it} = Y_{it}(0) + (Y_{it}(1) - Y_{it}(0)) \times \mathbb{1}(X_{it} \leq N)$. Let's define $\Delta_{it} = Y_{it}(1) - Y_{it}(0)$ the individual-time treatment effect, such that $\tau_{t,RDD} = \mathbb{E}[\Delta_{it} \mid X_{it} = N]$. Thus $Y_{it}(1) = Y_{it}(0) + \Delta_{it}$ and $\mathbb{E}[Y_{it}(1) \mid X_{it} = x] = \alpha + \sum_{j=1}^{p} \beta_j \times x^j + \tau_{t,RDD}$. Consider a parameter δ the size of the one-sided exclusion window around the notch, we are interested in the following quantity:

$$\begin{aligned}
\tau_t(N, \delta) &= \mathbb{E}[Y_{it}(1) \mid X_{it} = N - \delta] - \mathbb{E}[Y_{it}(0) \mid X_{it} = N + \delta] \\
&= \alpha + \sum_{j=1}^{p} \beta_j \times (N-\delta)^j + \tau_{t,RDD} - \alpha - \sum_{j=1}^{p} \beta_j \times (N+\delta)^j \\
&= \sum_{j=1}^{p} \beta_j \times \left[(N-\delta)^j - (N+\delta)^j\right] + \tau_{t,RDD} \\
&= -2\beta_1 \delta + \sum_{j=2}^{p} \beta_j \times \left[(N-\delta)^j - (N+\delta)^j\right] + \tau_{t,RDD}
\end{aligned}$$

If we are ready to assume that $p = 1$ then $\tau_t(N, \delta) = -2\beta_1\delta + \tau_{t,RDD}$. In order to recover $\tau_{t,RDD}$ one needs to set δ in a way that assumption 7.1.2 holds and then estimate β_1 on untreated units. Thus,

$$\begin{aligned}\tau_{t,RDD} &= \mathbb{E}\left[Y_{it}(1) \mid X_{it} = N - \delta\right] - \mathbb{E}\left[Y_{it}(0) \mid X_{it} = N + \delta\right] + 2\beta_1\delta \\ &= \mathbb{E}\left[Y_{it} \mid X_{it} = N - \delta\right] - \mathbb{E}\left[Y_{it} \mid X_{it} = N + \delta\right] + 2\beta_1\delta\end{aligned}$$

Where the last equality holds because of treatment allocation.

7.1.2 RDD at the 75% natural person threshold

7.1.3 RKD at the kink

Here instead of the causal effect of decreasing the average tax rate or the total amount of taxes by 6800 €, the object under study is the effect of changing the marginal tax rate. When crossing the kink, firms go from a marginal rate of 15% to 33.3%, the rate more than doubles. Understanding the effect of such a substantial increase of this marginal rate is the purpose of this section.

The setting here is more a fuzzy RK design since one is not able to observe which firms are below or above the ratio of shareholders. Though, at this level of taxable income, it is reasonable to think that a majority of firms are detained by natural persons only. Before turning to the framework, figure 36 and 37 show the average current assets by taxable income around the kink with or without the bunchig region, respectively, for year 2001-2007 pooled. Focusing first on the second figure, one can see how linear is the link function between the average outcome and the taxable income. When the bunching region is removed, the two lines seem parallel, suggesting no difference between the two slopes. One including the bunching region, the first figure shows a visible difference in slopes, the line is pulled down at the kink by bunching behaviors. Because we know that there is substantial bunching and that at least a part of it is due to avoidance, it seems reasonable to interpret this difference as the effect of avoidance on the link function. If it was the proper effect of the change in marginal rate, the slopes would be different even when excluding the bunching region. Here however, the two lines are strikingly similar. This result extends to pretty much all the variables, see for instance figures 38 and 39 for total operating costs and 40 and 41 for total sales.

The RKD estimand is composed of two parts: the numerator and the denominator. The latter one is deterministic and is the difference in the slope of the tax schedule in our case. Using the framework of Card et al. [2012], denoting x the taxable income and using the other notations as before: $\lim_{x_0 \to k^+} T'(x) \mid_{x=x_0} - \lim_{x_0 \to k^-} T'(x) \mid_{x=x_0} = 0.333 - 0.15 = 0.183$. The numerator however

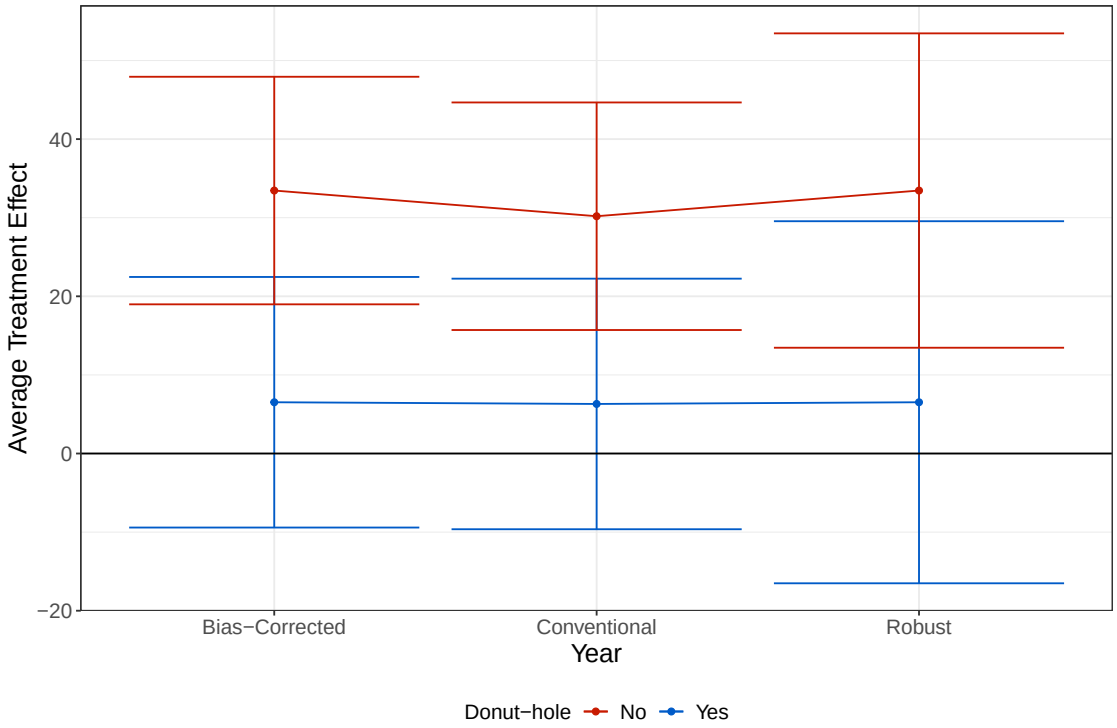

Figure 9: RKD estimation of current assets with and without the bunching region (2001-2007 pooled)

Note:

needs to be estimated and, denoting X the taxable income as random variable, is:

$$\lim_{x_0 \to k^+} \frac{\mathbb{E}[Y|X=x]}{dx}\bigg|_{x=x_0} - \lim_{x_0 \to k^-} \frac{\mathbb{E}[Y|X=x]}{dx}\bigg|_{x=x_0}$$

This link functions are estimated through local polynomial regressions, and the difference of slopes is then the difference of the slopes of these regressions. Since the link function seems very linear, I rely on local linear regressions. In order to compare the two differences, with and without the bunching region, for each variable I compute two times the local linear regression on each side: one with and one without the bunching area. I use the R package from Calonico et al. [2015] for the estimation.

Figure 9, 42 and 43 display the estimates for current assets, total sales and total operating costs respectively. The red estimates show the ones with the bunching region (without using the donut-hole technique, hence the "No" in the legend), the blue ones without the bunching region. For these three variables, without the bunching region, there is no effect of the change in the marginal rate on current assets as all the estimates are not significant. When including the bunching region the estimates all become significant and more surprisingly positive. One would expect that increasing the marginal tax rate would reduce the observed outcome. Including the bunching region shows the opposite. This suggets that there is no effect of increasing the marginal tax rate, which may suggest that in any case

firms are more interested in the average tax rate to base their decisions. Including the bunching area implies finding counter-intuitive results that would suggest that what is happening is due more to avoidance than to real responses.

7.2 Using difference in learning

Firms do not learn at the same path as section 5 have shown. In the spirit of Chetty et al. [2013], one can use differences in learning in order to uncover the effect of the reform. Companies can not adjust to something they are not aware of even though they are treated at the same time. So much so that the timing of the reaction, if any, will be different. One can then use the staggered Diff-in-diff à la Callaway and Sant'Anna [2021] to recover the policy's effects. Of course, before any causal claim, one should be cautious about the parallel trend assumption (PTA). In my case, the Callaway and Sant'Anna [2021] framework seems more appropriate than the Athey and Imbens [2022] one. Indeed, the latter relies on the assumption that the treatment timing probability is random which is a very implausible assumption in my setting. There is no variation in the treatment timing *per se*, only the understanding timing does vary and it seems non-random as shown in section 5.2. Thus only the Callaway and Sant'Anna [2021] framework is of use here.

However, because the timing by which companies understand the shift is non-random, there are concerns about the PTA. Though a proper check is not feasible, it is possible to check for pre-trends and try to figure out what would be the reasons for the different groups to break the PTA.

The framework relies on several assumptions. First of all, the irreversibility of treatment which means that if $D_{it-1} = 1$ then $D_{it} = 1$. Units can not leave the treatment once they have experienced it for the first time. This assumption is valid in our case if firms do not reduce their ratio of natural persons below 75% or grow such as to have their total sales above the notch. In order to satisfy this assumption, it is possible to focus on firms sufficiently below the notch on sales, it is trickier for the ratio of shareholders however. For many companies, this ratio is not observable (see section 4.1). This means that among the firms we will study some might leave the treatment. However, section 6 showed that the bunching response to this notch was very limited suggesting that this issue is not a large threat to the identification. Furthermore, as what we are studying here is L_{it} rather than D_{it}, for firms with a lower threshold of natural persons, L_{it} equal to 0 means D_{it} equal to 0 too. They would be in the group of the "never-treated". Finally, the change from treated to not treated through a change of a composition of shareholders seems less plausible since the firm would directly pay more taxes. Turning to the same assumption concerning L_{it}, it is difficult to imagine a firm forgetting the policy after having understood it. Thus, this assumption seems quite plausible. The second assumption is

random sampling and implies that we are considering panel data, which is the case here.

Assumption 7.2.1 (Irreversibility of treatment). $D_{i1} = 0$ *for all* i. *For* $t = 2, ..., T$, $D_{it-1} = 1$ *then* $D_{it} = 1$

Assumption 7.2.2 (Random sampling). $\{Y_{i1}, ..., Y_{iT}, X_i, D_{i1}, ..., D_{iT}\}_{i=1,...,N}$ *is i.i.d.*

Before going further with the assumptions, let's introduce the framework more in depth. Denote G the group variable, defined by the timing at which units understand the treatment. G_g being the indicator of firms understanding in period g, $G_g = \mathbb{1}(G_i = g)$, and $G = \infty$ for firms never treated. Denote time periods $t = 1, ..., T$. Denote $Y_{it}(0)$ the untreated potential outcome of firm i in period t for units that remain untreated during all periods. Denote $Y_{it}(g)$ the potential outcome of firm i in period t if they were to first experience the treatment by period g. One can then write the outcome as:

$$Y_{it} = Y_{it}(0) + \sum_{g=2}^{T} (Y_{it}(g) - Y_{it}(0)) \times G_{ig}$$

We consider the following quantity, called the *group-time average treatment effect*:

$$ATT(g,t) = \mathbb{E}\left[Y_{it}(g) - Y_{it}(0) \mid G_{ig} = 1\right]$$

In order to identify this quantity, one needs 3 more assumptions. The first one is *limited treatment anticipation*. Limited because the anticipation is parametrized by a quantity δ:

Assumption 7.2.3 (Limited treatment anticipation). $\forall g < \infty, t \in \{1, ..., T\} \mid t < g - \delta$

$$\mathbb{E}\left[Y_{it}(g) \mid X, G_{ig} = 1\right] = \mathbb{E}\left[Y_{it}(0) \mid X, G_{ig} = 1\right]$$

If $\delta = 0$, there is strict no ancitipation, if $\delta = 1$, there is one period of anticipation. In our case, the reduced rate was voted in December 2000, and already announced in the draft of the budget bill in September of the same year[24]. There might be some anticipation. One way to check for that is to measure the effect in pre-reform years. The second assumption is the *conditional parallel trends based on Not-yet-treated groups*:

Assumption 7.2.4 (Conditional parallel trends based on Not-yet-treated groups). $\forall g < \infty, (s,t) \in \{2, ..., T\} \times \{2, ..., T\} \mid t \geq g - \delta$ *and* $t + \delta \geq s < T$

$$\mathbb{E}\left[Y_{it}(0) - Y_{it}(0) \mid X, G_g = 1\right] = \mathbb{E}\left[Y_{it}(0) - Y_{it}(0) \mid X, D_s = 0, G_g = 0\right]$$

[24] The different steps of the 2000 Budget bill

This is the assumption 5 in Callaway and Sant'Anna [2021]. They warn that favoring this one over the one relying on *never treated* has implications in terms of restrictions of observed pre-treatment trends [Marcus and Sant'Anna, 2021]. The last assumption is the *overlap* one:

Assumption 7.2.5 (Overlap)**.** *For each* $t \in \{2,...,T\}$, $g < \infty$, *there exist some* $\varepsilon > 0$ *such that* $\mathbb{P}(G_g = 1) > \varepsilon$ *and* $\mathbb{P}(G_g = 1 \mid X, G_g + (1-D_s)(1-G_g) = 1) < 1-\varepsilon$.

Assumption 7.2.5 states that at each period g a positive fraction of the total units starts the treatment and that the generalized propensity score is bounded away from one. $\mathbb{P}(G_g = 1 \mid X, G_g + (1-D_s)(1-G_g)$ is the probability that a unit is in group g conditional on X and on either being indeed of group g or not-yet-treated by time s.

Using Theorem 1 of Callaway and Sant'Anna [2021] we can rely on the outcome regression, inverse-probability weighting or doubly-robust approach to recover the $ATT(g,t)$.

8 Conclusion

References

P. Aghion, M. Gravoueille, M. Lequien, and S. Stantcheva. Tax simplicity or simplicity of evasion? evidence from self-employment taxes in france. Technical report, 2024.

C. Agostini, G. Bernier, M. Bertanha, K. Bilicka, J. Bukovina, Y. He, E. Koumanakos, T. Lichard, J. Palguta, E. Patel, et al. The elasticity of taxable income across countries. 2024.

A. Alstadsæter, M. Jacob, and R. Michaely. Do dividend taxes affect corporate investment? Journal of Public Economics, 151:74–83, 2017.

A. Alstadsæter, N. Johannesen, and G. Zucman. Tax evasion and inequality. American Economic Review, 109(6):2073–2103, 2019.

J. D. Angrist and V. Lavy. Using maimonides' rule to estimate the effect of class size on scholastic achievement. The Quarterly journal of economics, 114(2):533–575, 1999.

P. Askenazy, T. Breda, and V. Pecheu. Under-reporting of firm size around size-dependent regulation thresholds: Evidence from france. 2022.

S. Athey and G. W. Imbens. Design-based analysis in difference-in-differences settings with staggered adoption. Journal of Econometrics, 226(1):62–79, 2022.

L. Bach, A. Bozio, B. Fabre, and A. Malgouyres. Follow the money! why dividends overreact to flat-tax reforms. Technical report, PSE Working Papers, 2021.

P. Bachas and M. Soto. Corporate taxation under weak enforcement. American Economic Journal: Economic Policy, 13(4):36–71, 2021.

A. I. Barreca, M. Guldi, J. M. Lindo, and G. R. Waddell. Saving babies? revisiting the effect of very low birth weight classification. The quarterly journal of economics, 126(4):2117–2123, 2011.

A. I. Barreca, J. M. Lindo, and G. R. Waddell. Heaping-induced bias in regression-discontinuity designs. Economic inquiry, 54(1):268–293, 2016.

A. Bauer and M. Rotemberg. Tax avoidance in firms. 2018.

Y. Benzarti. How taxing is tax filing? using revealed preferences to estimate compliance costs. American Economic Journal: Economic Policy, 12(4):38–57, 2020.

Y. Benzarti and L. Wallossek. Rising income tax complexity. National Tax Journal, 77(1):143–173, 2024.

M. Bertanha, A. H. McCallum, A. Payne, and N. Seegert. Bunching estimation of elasticities using stata. The Stata Journal, 22(3):597–624, 2022.

M. Bertanha, A. H. McCallum, and N. Seegert. Better bunching, nicer notching. Journal of Econometrics, 237(2):105512, 2023.

M. C. Best, A. Brockmeyer, H. J. Kleven, J. Spinnewijn, and M. Waseem. Production versus revenue efficiency with limited tax capacity: theory and evidence from pakistan. Journal of political Economy, 123(6):1311–1355, 2015.

W. Boonzaaier, J. Harju, T. Matikka, and J. Pirttilä. How do small firms respond to tax schedule discontinuities? evidence from south african tax registers. International Tax and Public Finance, 26:1104–1136, 2019.

A. Brockmeyer. The investment effect of taxation: Evidence from a corporate tax kink. Fiscal Studies, 35(4):477–509, 2014.

J. Bukovina, T. Lichard, J. Palguta, and B. Žúdel. Corporate minimum tax and the elasticity of taxable income: Evidence from administrative tax records. American Economic Journal: Economic Policy, Forthcoming.

S. Bustos, D. Pomeranz, J. C. S. Serrato, J. Vila-Belda, and G. Zucman. The race between tax enforcement and tax planning: Evidence from a natural experiment in chile. Technical report, National Bureau of Economic Research, 2022.

B. Callaway and P. H. Sant'Anna. Difference-in-differences with multiple time periods. Journal of econometrics, 225(2):200–230, 2021.

S. Calonico, M. D. Cattaneo, and R. Titiunik. Rdrobust: an r package for robust nonparametric inference in regression-discontinuity designs. R J., 7(1):38, 2015.

C. Carbonnier, C. Foffano, C. Malgouyres, L. Py, and C. Urvoy. Évaluation interdisciplinaire des impacts du CICE en matière d'emplois et de salaires. PhD thesis, Laboratoire interdisciplinaire d'évaluation des politiques publiques, 2018.

C. Carbonnier, C. Malgouyres, L. Py, and C. Urvoy. Who benefits from tax incentives? the heterogeneous wage incidence of a tax credit. Journal of Public Economics, 206:104577, 2022.

D. Card, D. Lee, Z. Pei, and A. Weber. Nonlinear policy rules and the identification and estimation of causal effects in a generalized regression kink design. Technical report, National Bureau of Economic Research, 2012.

Z. Chen, Z. Liu, J. C. Suárez Serrato, and D. Y. Xu. Notching r&d investment with corporate income tax cuts in china. American Economic Review, 111(7):2065–2100, 2021.

R. Chetty. Is the taxable income elasticity sufficient to calculate deadweight loss? the implications of evasion and avoidance. American Economic Journal: Economic Policy, 1(2):31–52, 2009.

R. Chetty and E. Saez. Teaching the tax code: Earnings responses to an experiment with eitc recipients. American Economic Journal: Applied Economics, 5(1):1–31, 2013.

R. Chetty, J. N. Friedman, and E. Saez. Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. American Economic Review, 103(7):2683–2721, December 2013. doi: 10.1257/aer.103.7.2683. URL https://www.aeaweb.org/articles?id=10.1257/aer.103.7.2683.

G. Chodorow-Reich, M. Smith, O. M. Zidar, and E. Zwick. Tax policy and investment in a global economy. Technical report, National Bureau of Economic Research, 2024.

J. L. Coles, E. Patel, N. Seegert, and M. Smith. How do firms respond to corporate taxes? Journal of Accounting Research, 60(3):965–1006, 2022.

CPO. Adapter l'impôt sur les sociétés à une économie ouverte. Technical report, Conseil des prélèvements obligatoires, 2016.

W. Cui, M. Wei, W. Xie, and J. Xing. Corporate tax cuts for small firms: What do firms do? Available at SSRN 3950973, 2021.

M. P. Devereux, L. Liu, and S. Loretz. The elasticity of corporate taxable income: New evidence from uk tax records. American Economic Journal: Economic Policy, 6(2):19–53, 2014.

I. Di Marzio, S. Moceeti, and E. Rubolino. Market externalities of tax evasion*. 2023.

N. E. Feldman, P. Katuščák, and L. Kawano. Taxpayer confusion: Evidence from the child tax credit. American Economic Review, 106(3):807–35, March 2016. doi: 10.1257/aer.20131189. URL https://www.aeaweb.org/articles?id=10.1257/aer.20131189.

M. Feldstein. Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics, 81(4):674–680, 1999.

B. Garbinti, J. Goupille-Lebret, M. Munoz, S. Stefanie, and G. Zucman. Tax design, information, and elasticities: Evidence from the french wealth tax. Working Paper, 2023.

I. Guceri and L. Liu. Effectiveness of fiscal incentives for r&d: Quasi-experimental evidence. American Economic Journal: Economic Policy, 11(1):266–291, 2019.

J. Hahn, P. Todd, and W. Van der Klaauw. Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica, 69(1):201–209, 2001.

J. Harju, A. Koivisto, and T. Matikka. The effects of corporate taxes on small firms. Journal of Public Economics, 212:104704, 2022.

C. L. House and M. D. Shapiro. Temporary investment tax incentives: Theory with evidence from bonus depreciation. American Economic Review, 98(3):737–768, 2008.

G. W. Imbens and T. Lemieux. Regression discontinuity designs: A guide to practice. Journal of econometrics, 142(2):615–635, 2008.

S. Laffitte. The market for tax havens. 2024.

C. Landais. Assessing the welfare effects of unemployment benefits using the regression kink design. American Economic Journal: Economic Policy, 7(4):243–278, 2015.

A. Lapeyre. Firm moral hazard in short-time work. 2024.

D. S. Lee. Randomized experiments from non-random selection in us house elections. Journal of Econometrics, 142(2):675–697, 2008.

F. Lobel, T. Scot, and P. Zúniga. Corporate taxation and evasion responses: Evidence from a minimum tax in honduras. American Economic Journal: Economic Policy, 16(1):482–517, 2024.

J. Londoño-Vélez and J. Avila-Mahecha. Behavioral responses to wealth taxation: Evidence from colombia. Technical report, National Bureau of Economic Research, 2024.

M. Marcus and P. H. Sant'Anna. The role of parallel trends in event study settings: An application to environmental economics. Journal of the Association of Environmental and Resource Economists, 8(2):235–275, 2021.

G. Massenz. Heterogeneity and persistence in tax responsiveness: Evidence from owner-managed companies. Technical report, IFN Working Paper, 2024.

S. Pollinger. Kinks know more: Policy evaluation beyond bunching with an application to solar subsidies. 2023.

A. Rees-Jones and D. Taubinsky. Measuring "schmeduling". The Review of Economic Studies, 87 (5):2399–2438, 2020.

E. Saez. Do taxpayers bunch at kink points? American economic Journal: economic policy, 2(3): 180–212, 2010.

J. Slemrod. A general model of the behavioral response to taxation. International Tax and Public Finance, 8:119–128, 2001.

J. Slemrod and W. Kopczuk. The optimal elasticity of taxable income. Journal of Public Economics, 84(1):91–112, 2002.

D. Yagan. Capital tax reform and the real economy: The effects of the 2003 dividend tax cut. American Economic Review, 105(12):3531–63, 2015.

E. Zwick. The costs of corporate tax complexity. American Economic Journal: Economic Policy, 13 (2):467–500, 2021.

E. Zwick and J. Mahon. Tax policy and heterogeneous investment behavior. American Economic Review, 107(1):217–48, 2017.

Figures and Tables

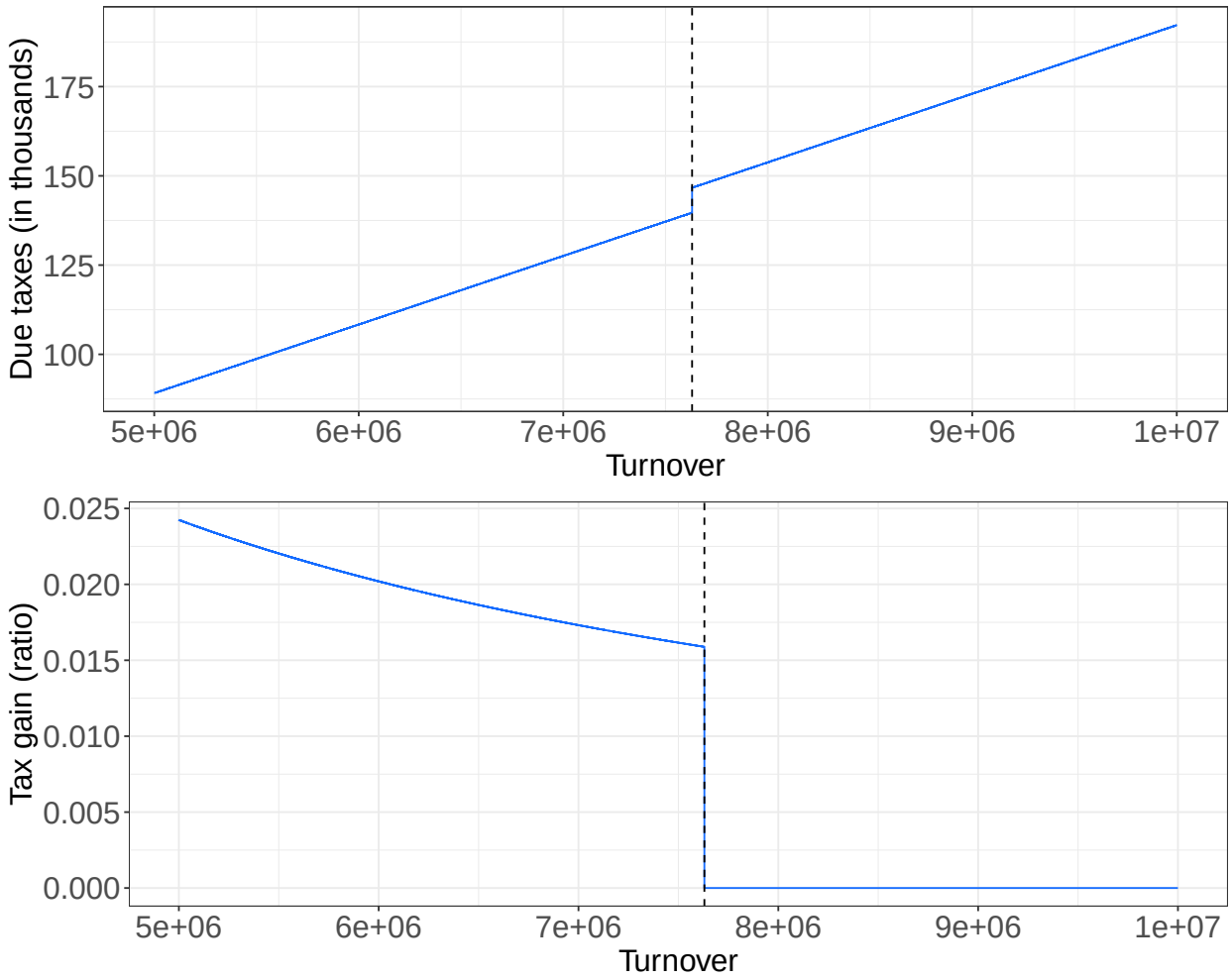

Figure 10: Representation of the tax schedule at the turnover notch

Note: For simplicity and graphical reason, an assumption of linearity between turnover and profit has been made.

Range of taxable income	Fixed asset	Current asset	Capital stock	Equity	Debts	Operating revenues	Operating cost	Accounting income
0,36k	192244	250162	37285	104387	248061	591880	574973	11222
36k,40k	324513	429349	63984	195831	398683	984207	942340	27357
40k,70k	378173	513392	76329	245635	464851	1135405	1080298	37282
>70k	3483419	3515186	628369	2012645	3477185	5471450	5118624	321022
>0	805564	962688	150294	503525	884822	1735154	1626542	81770
All	821595	812157	178034	415683	880906	1408133	1360294	43877

Table 1: Descriptive statistics of firms based on their taxable income in year 2000 (normal scheme only)

Note: In year 2000, firms with taxable income between 0 and 36.000 € had a fixed asset of 192.244 € on average. Overall, in year 2000 firms had a fixed asset of 821595 €. Firms between 36.000 € and 40.000 € in 2000, around the future kink, are smaller than the average firm. For example their capital stock if almost 3 times smaller.

Range of revenue	Fixed asset	Current asset	Capital stock	Equity	Debts	Operating revenues	Operating cost	Accounting income
0,3m	427411	361368	97585	211734	403051	624118	602164	22505
3m,6.63m	1338789	1945287	266366	803444	1745657	4337343	4152415	115984
6.63m,7.63m	2473625	3239542	484205	1303998	3094475	7249455	6936925	165286
>7.63m	52511890	31836385	5530544	15560969	43518202	56491366	54408172	1626368
All	821595	812157	178034	415683	880906	1408133	1360294	43877

Table 2: Descriptive statistics of firms based on their revenue in year 2000 (normal scheme only)

Note: In year 2000, firms with revenue between 0 and 3m € had a fixed asset of 427.411 € on average. Overall, in year 2000 firms had a fixed asset of 821595 €. Firms between 6.63m € and 7.63m € in 2000, just below the future notch, are bigger than the average firm. For example, their capital stock is more than 2 times higher.

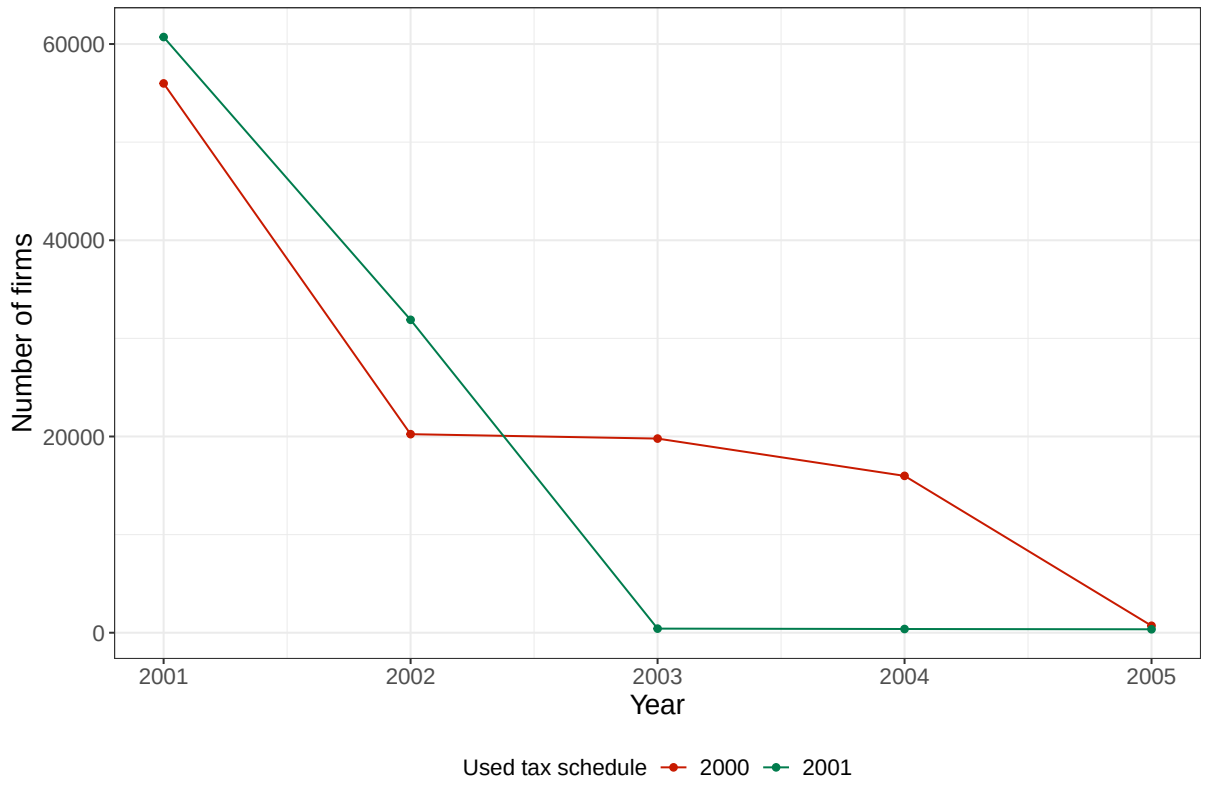

Figure 11: Number of firms reporting using the 2000 or 2001 tax schedule by year

Note:

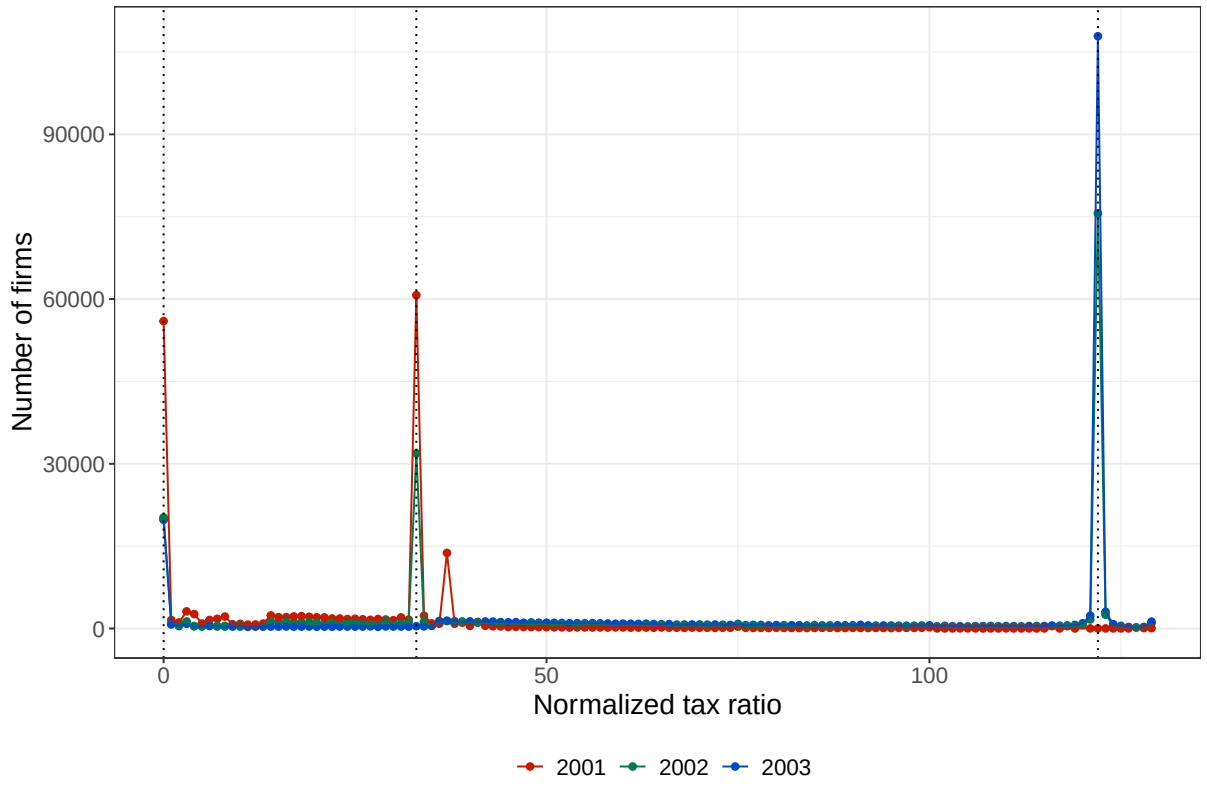

Figure 12: Distribution of the normalized ratio of reported taxes

Note:

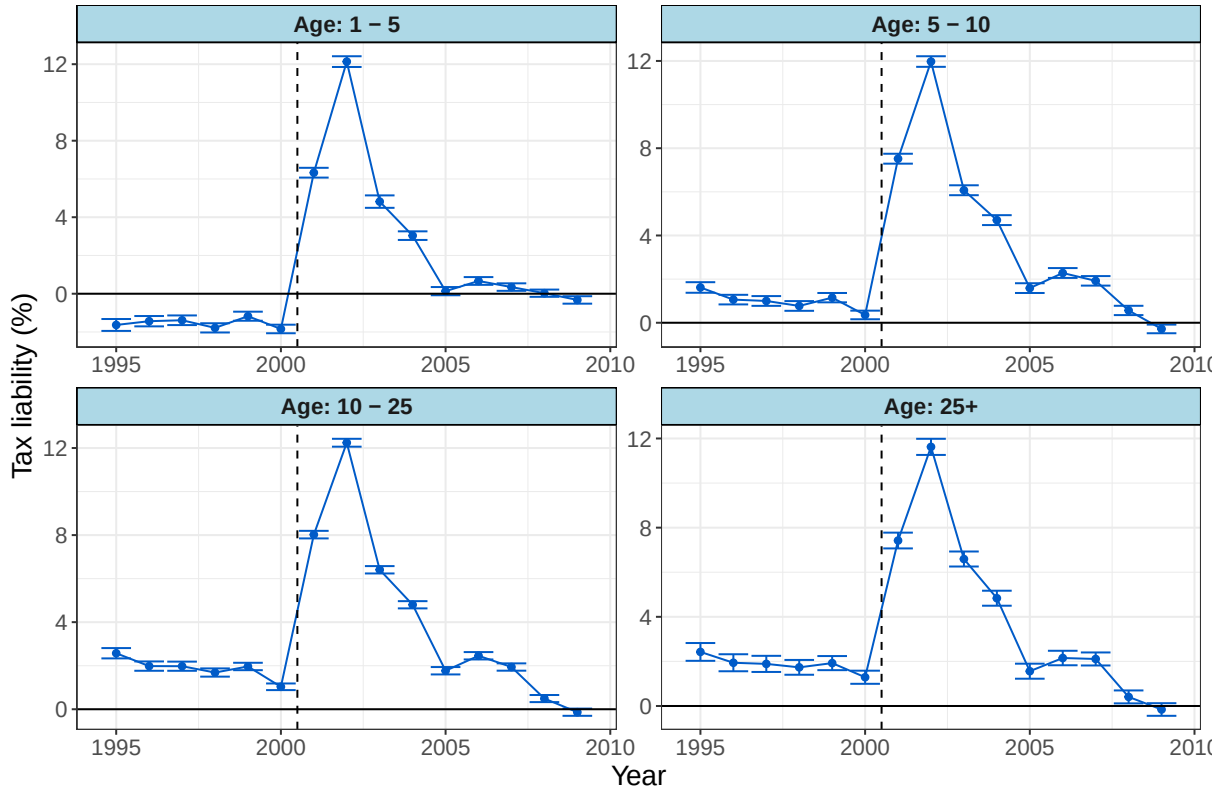

Figure 13: Average reported and computed tax liability by age across years

Note:

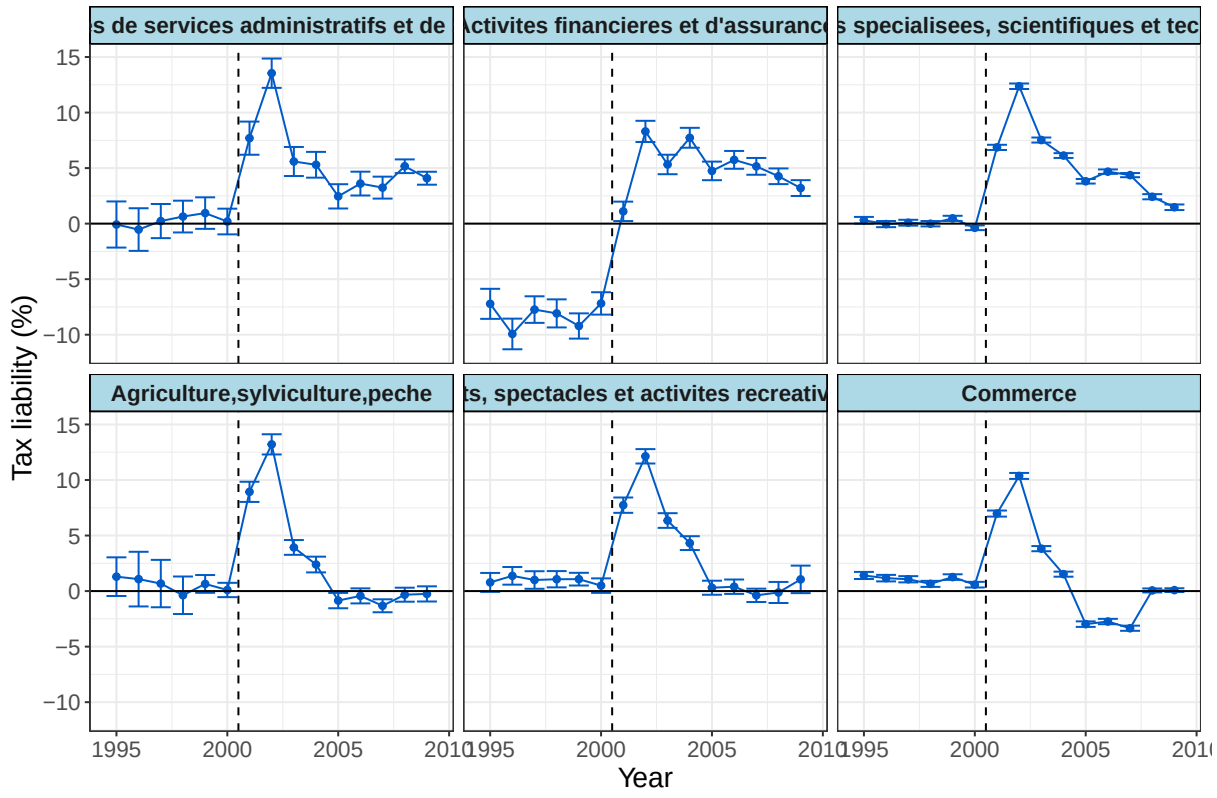

Figure 14: Average reported and computed tax liability by industry (I) across years

Note:

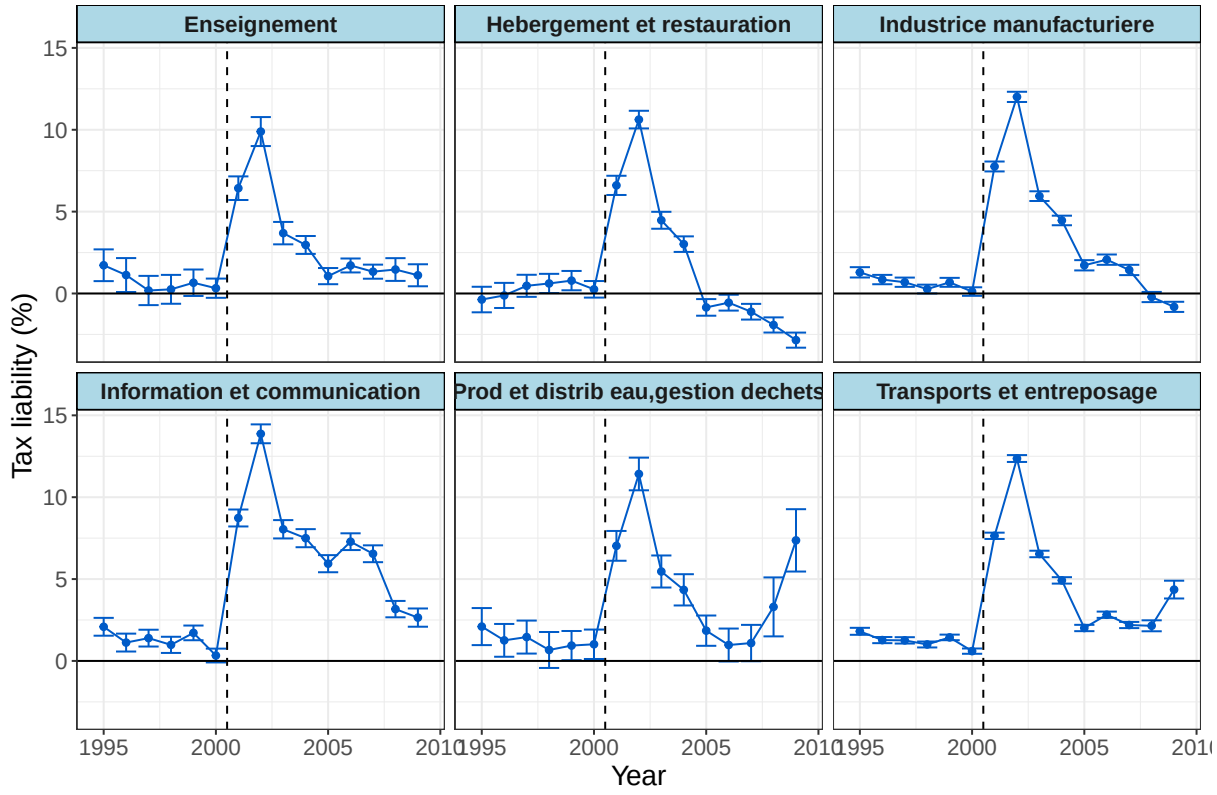

Figure 15: Average reported and computed tax liability by industry (II) across years

Note:

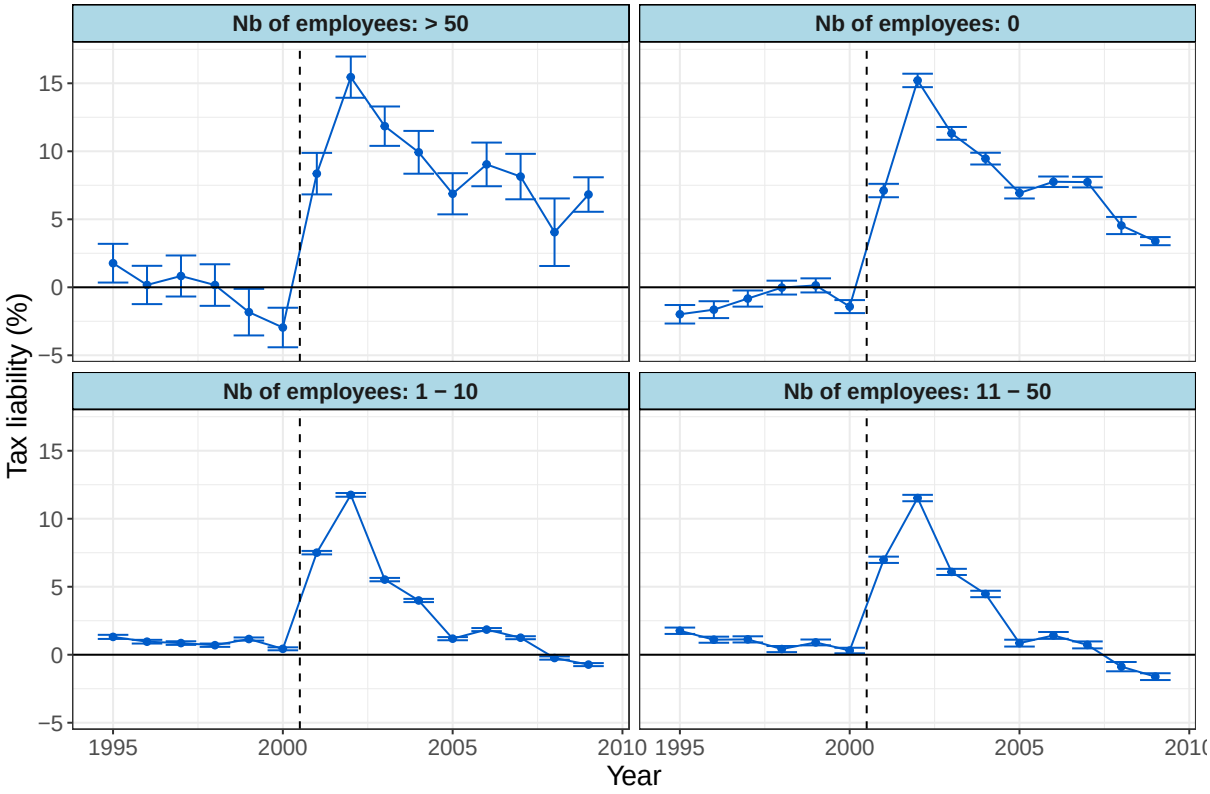

Figure 16: Average reported and computed tax liability by number of employees across years

Note:

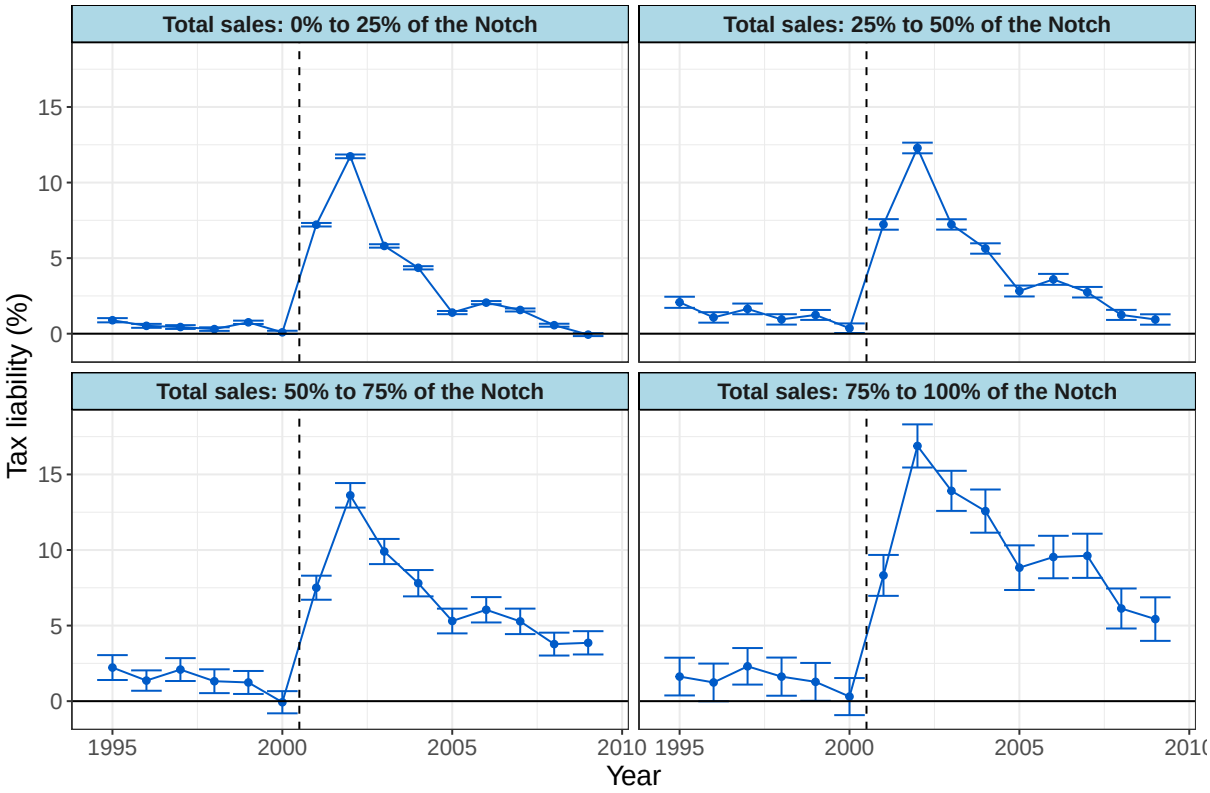

Figure 17: Average reported and computed tax liability by sales of employees across years

Note:

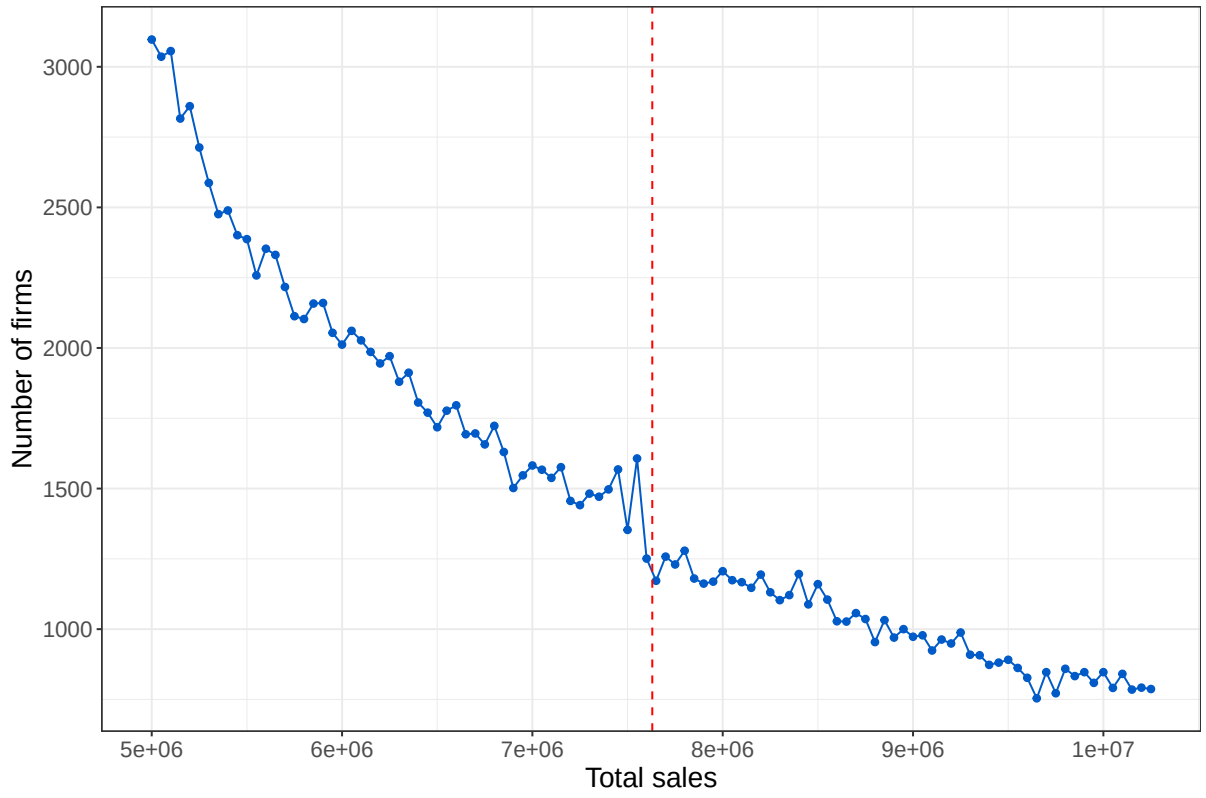

Figure 18: Distribution of companies by total sales (2001-2007)

Note: Distribution of companies by total sales around the notch pooled from 2001 to 2007. The red dashed line shows the location of the kink.

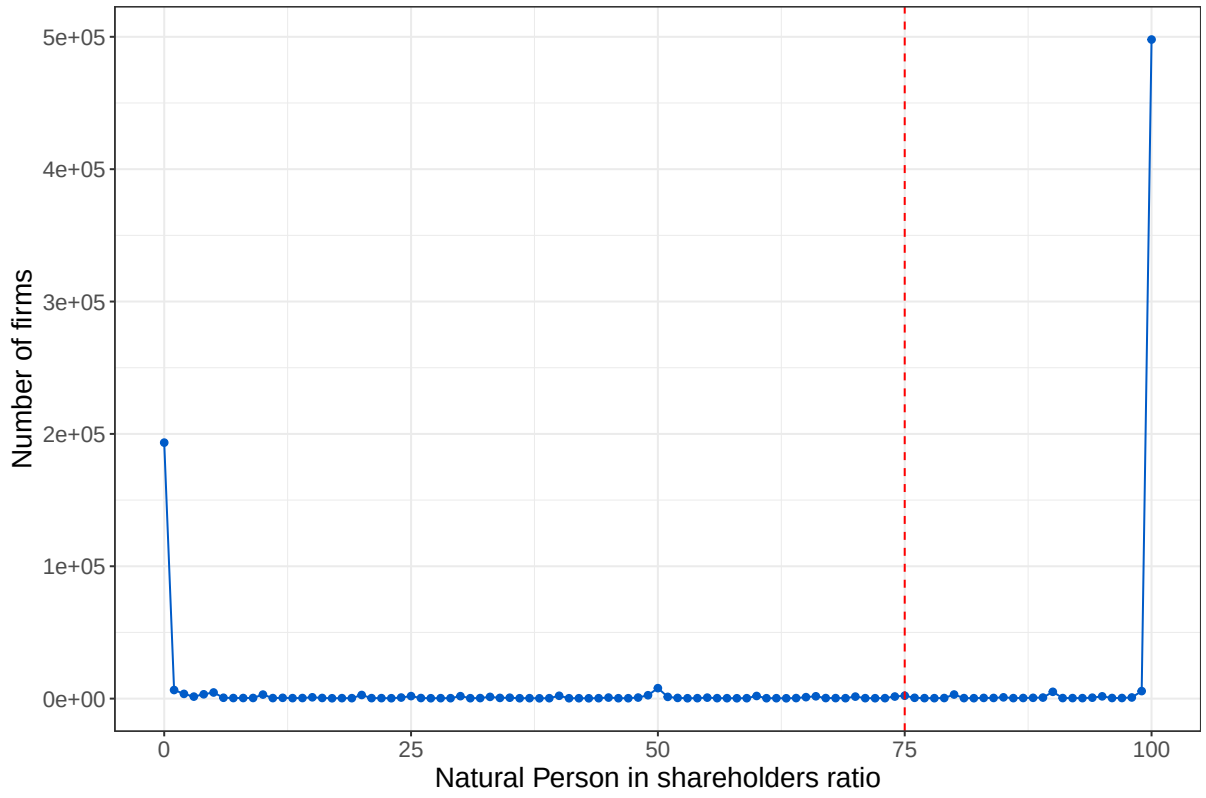

Figure 19

Note:

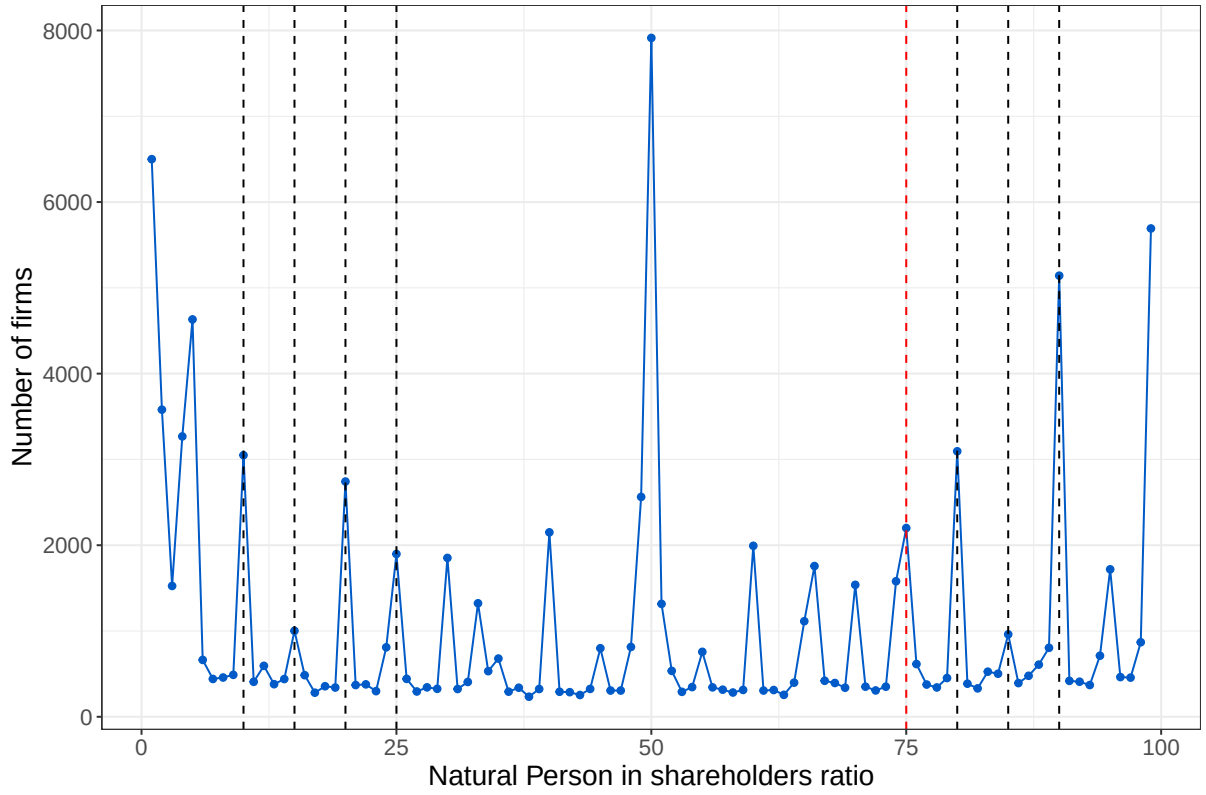

Figure 20

Note:

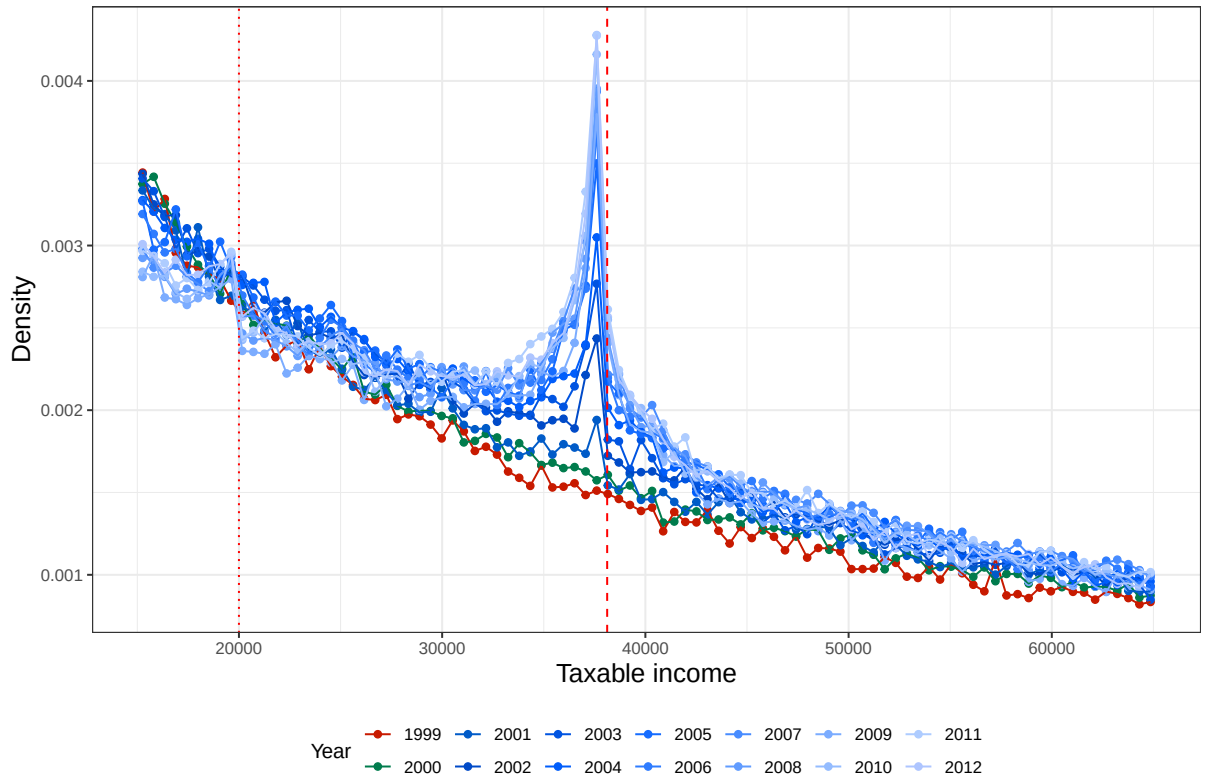

Figure 21: Distribution of companies by taxable income across years (1999-2012)

Note: Distribution of companies by taxable income around the kink across years. The red dashed line shows the location of the kink.

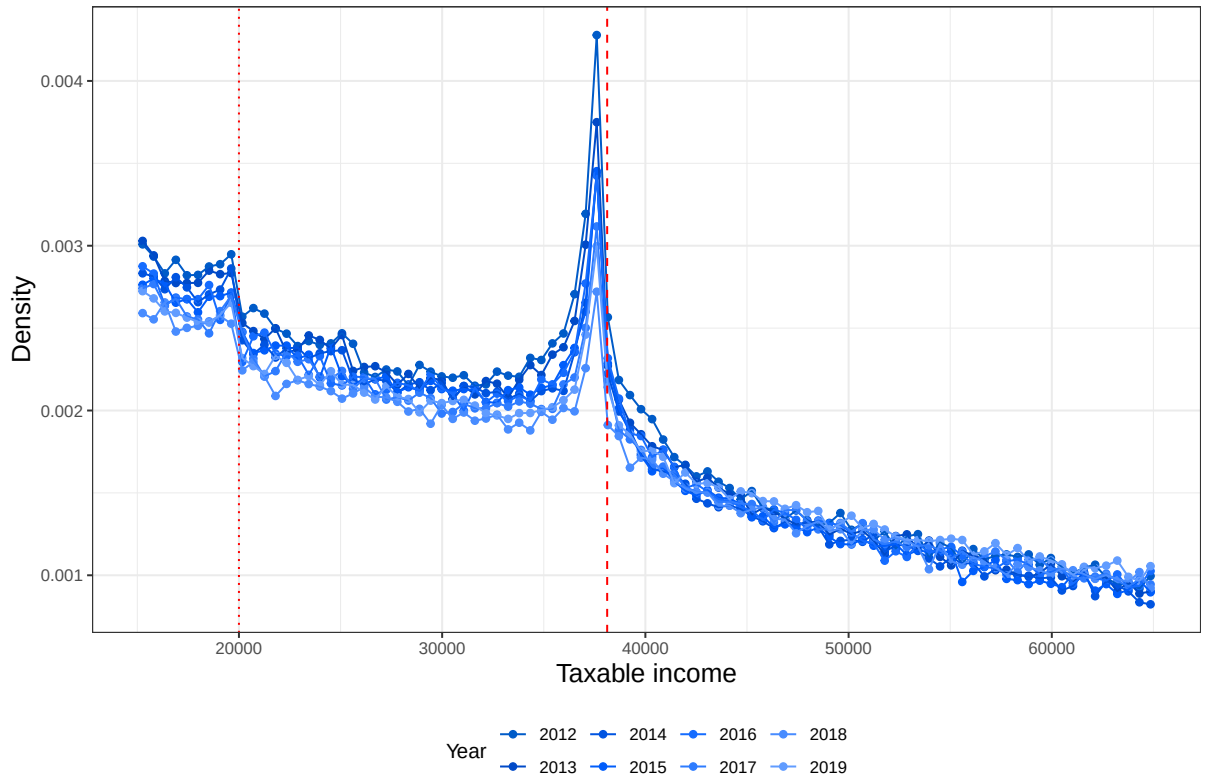

Figure 22: Distribution of companies by taxable income across years (2012-2019)

Note: Distribution of companies by taxable income around the kink pooled across years. The red dashed line shows the location of the kink.

	Learning cost	**Pure Adj. Costs**
Average	7032.4	3943.7
Stand. Dev.	22489.4	22086.4
N. of firms	518	16867
Conf. Interval	[5095.7, 8969.1]	[3610.4, 4277.1]
% of TI	24.0%	11.9%

Table 3: Estimation of adjustment costs

Note:

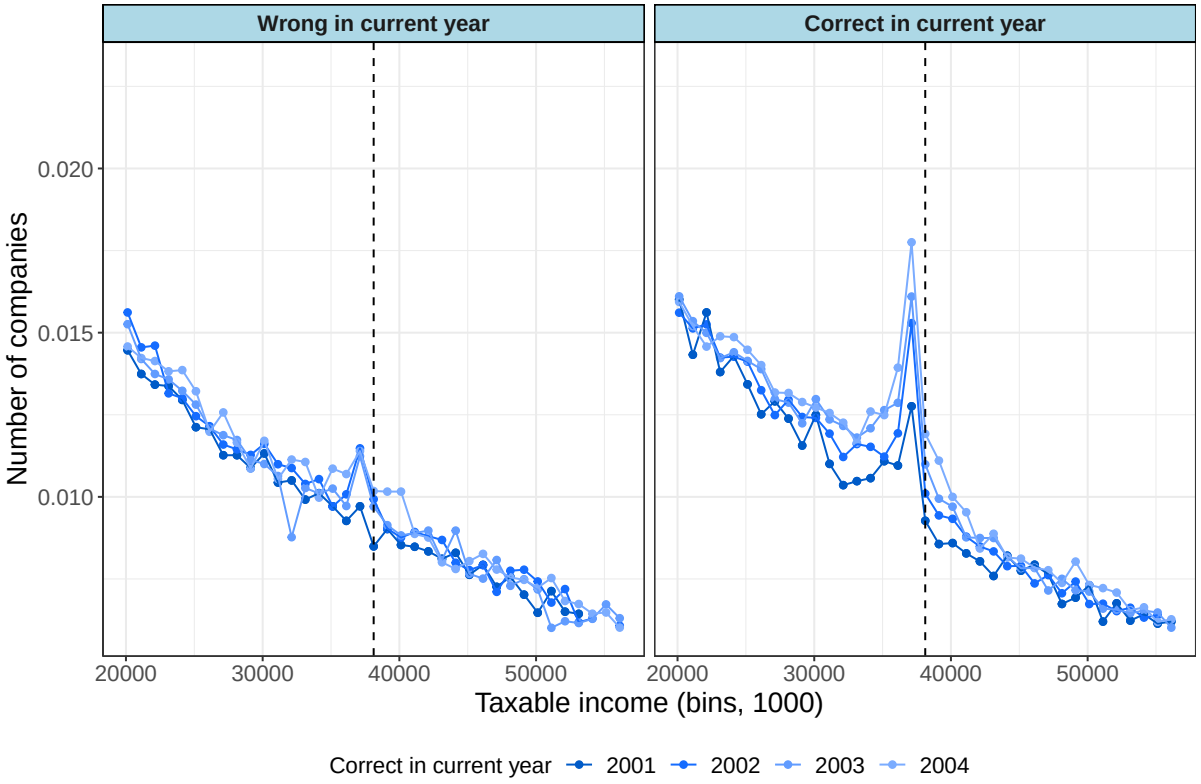

Figure 23: Differential bunching by learning status (2001-2004)

Note:

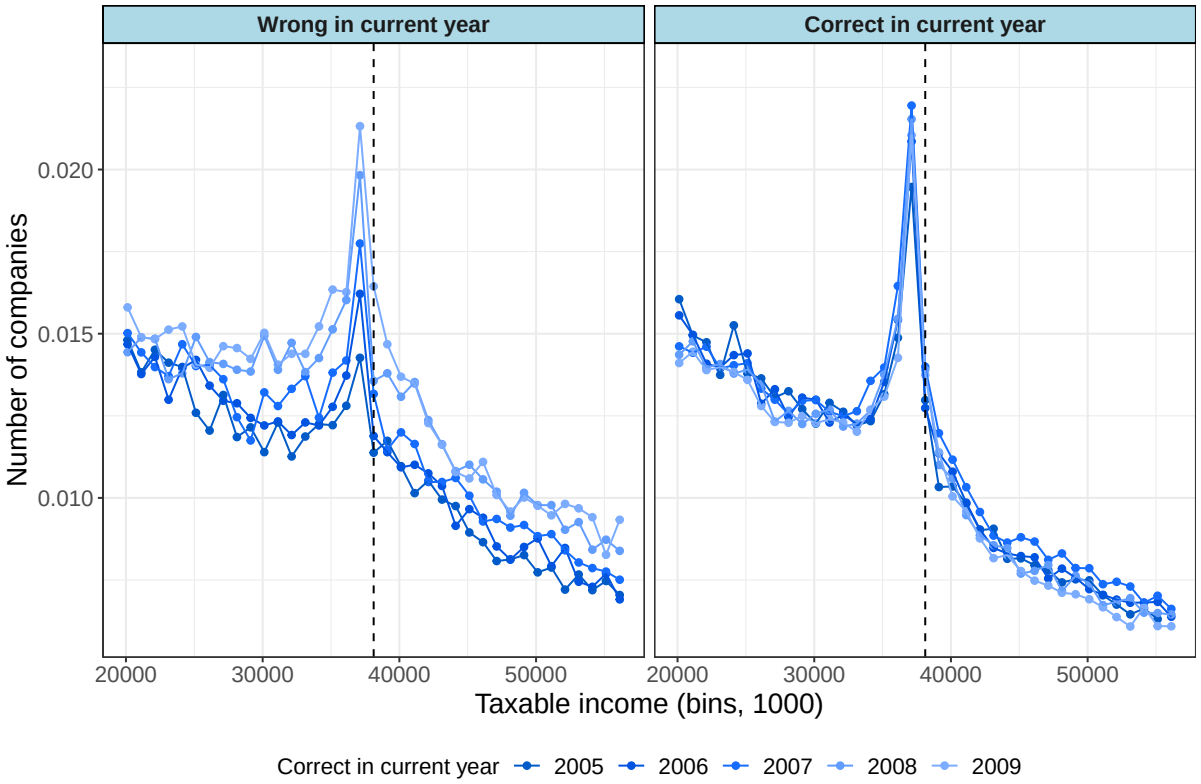

Figure 24: Differential bunching by learning status (2005-2009)

Note:

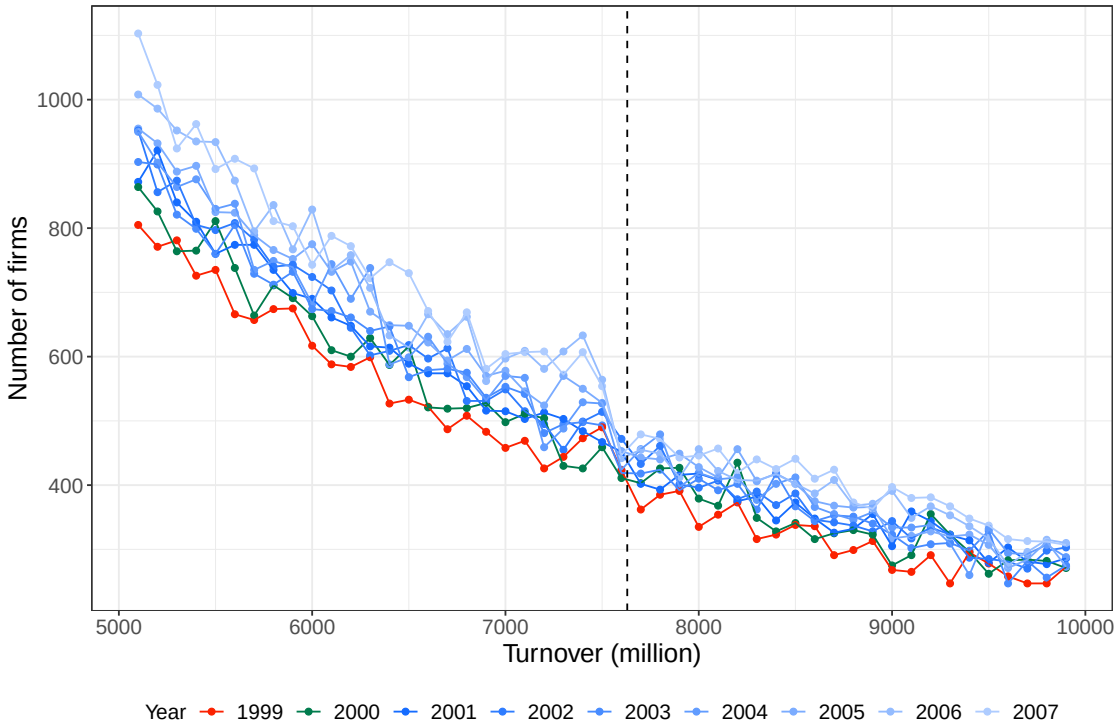

Figure 25: Bunching at the turnover threshold

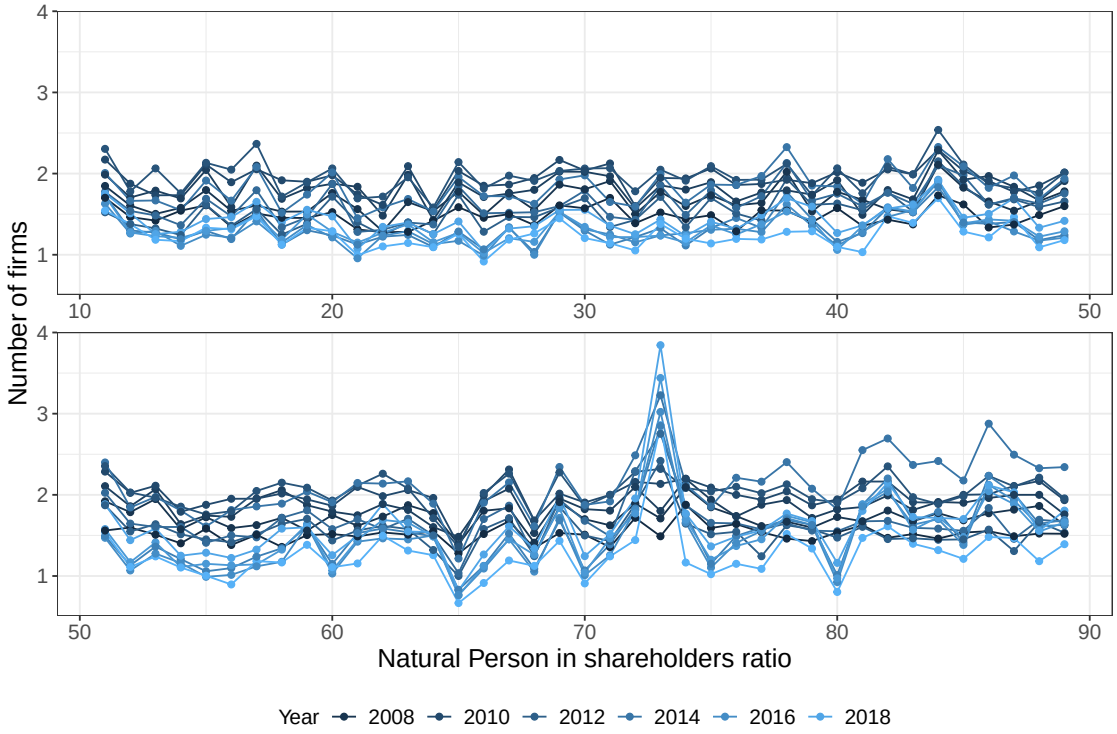

Figure 26: Bunching at the threshold of natural persons

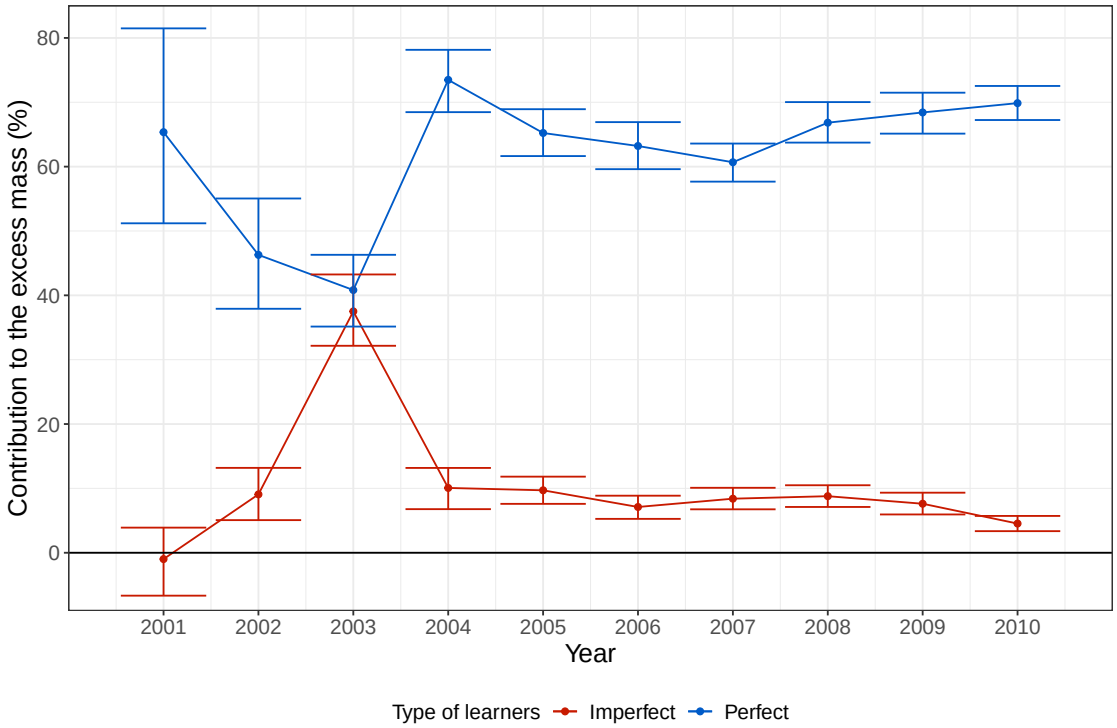

Figure 27: Contribution of perfect and imperfect learners to the elasticity across years

Note:

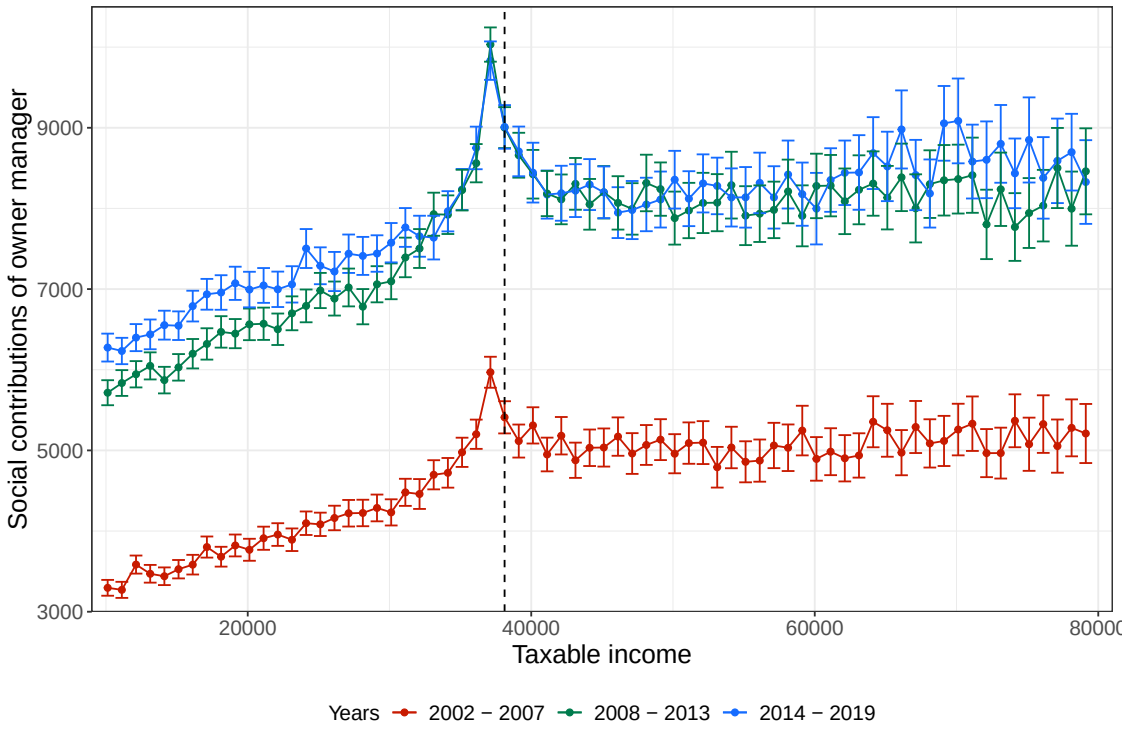

Figure 28: Average owner-managers' social contributions by taxable income

Note:

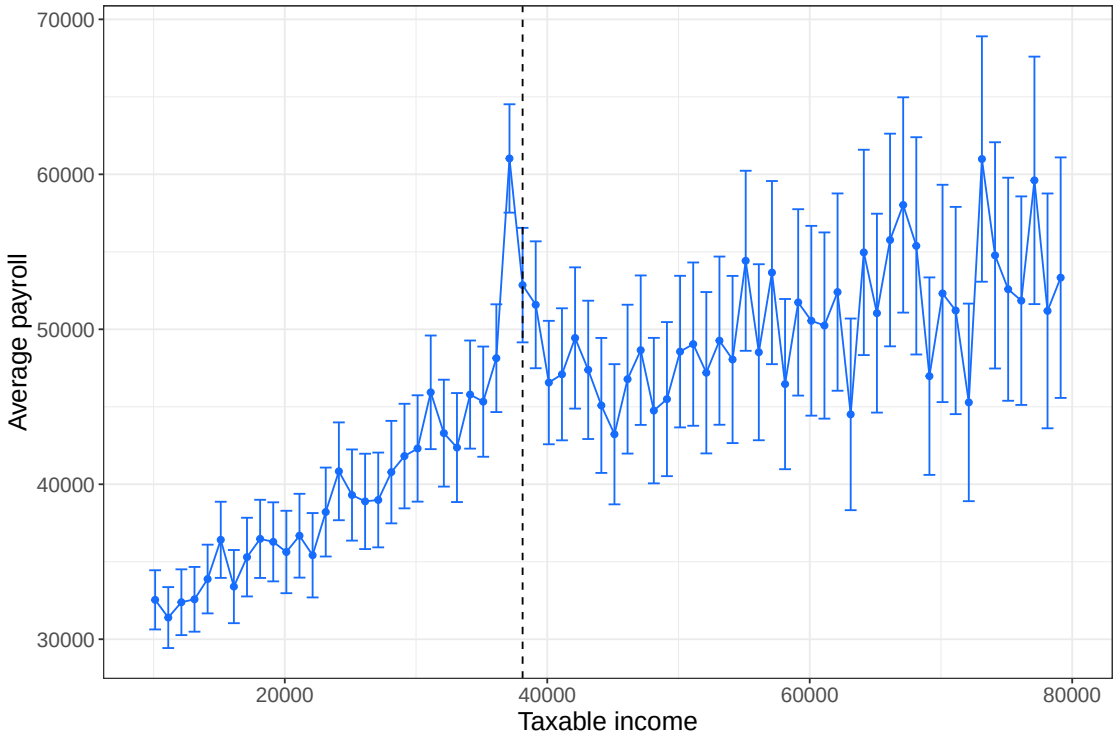

Figure 29: Average payroll for firms not declaring any employees in the DADS by taxable income

Note:

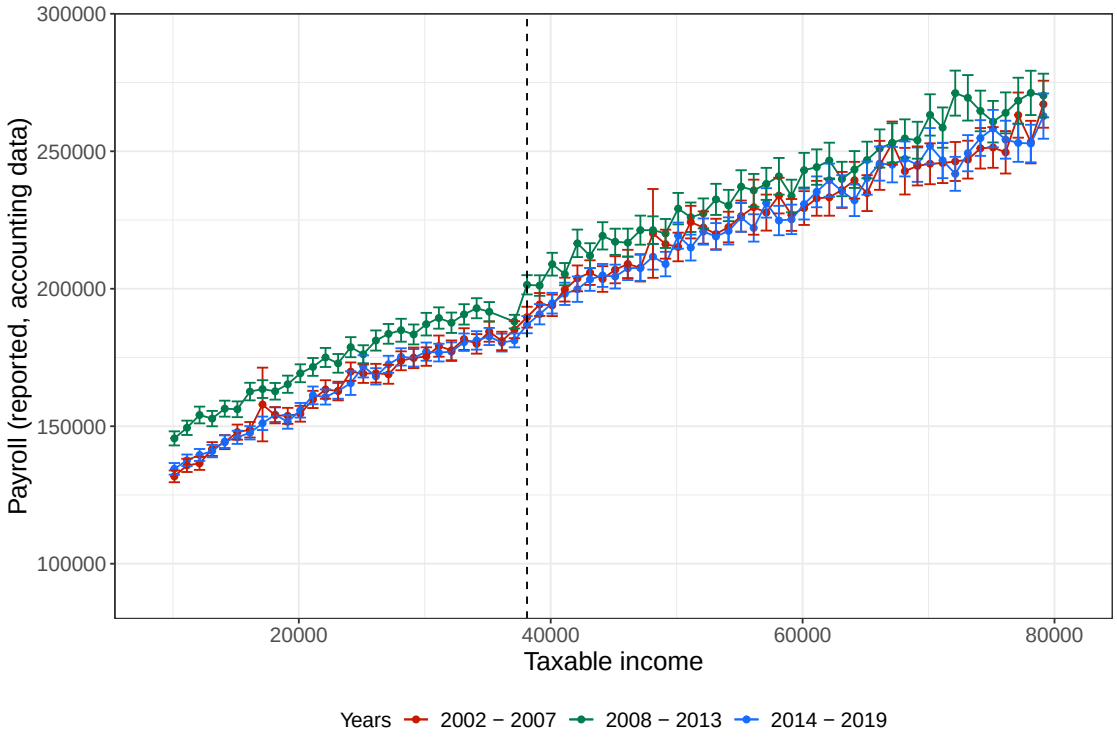

Figure 30: Average payroll by taxable income

Note:

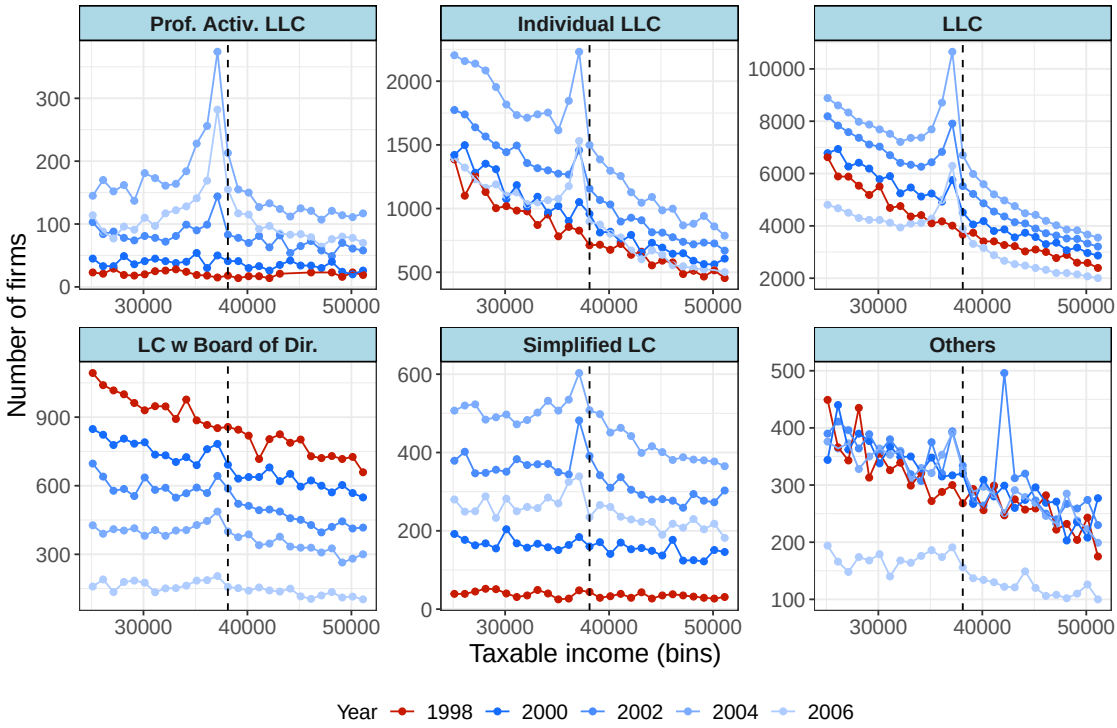

Figure 31: Distribution of firms by taxable income across years and legal status

Note:

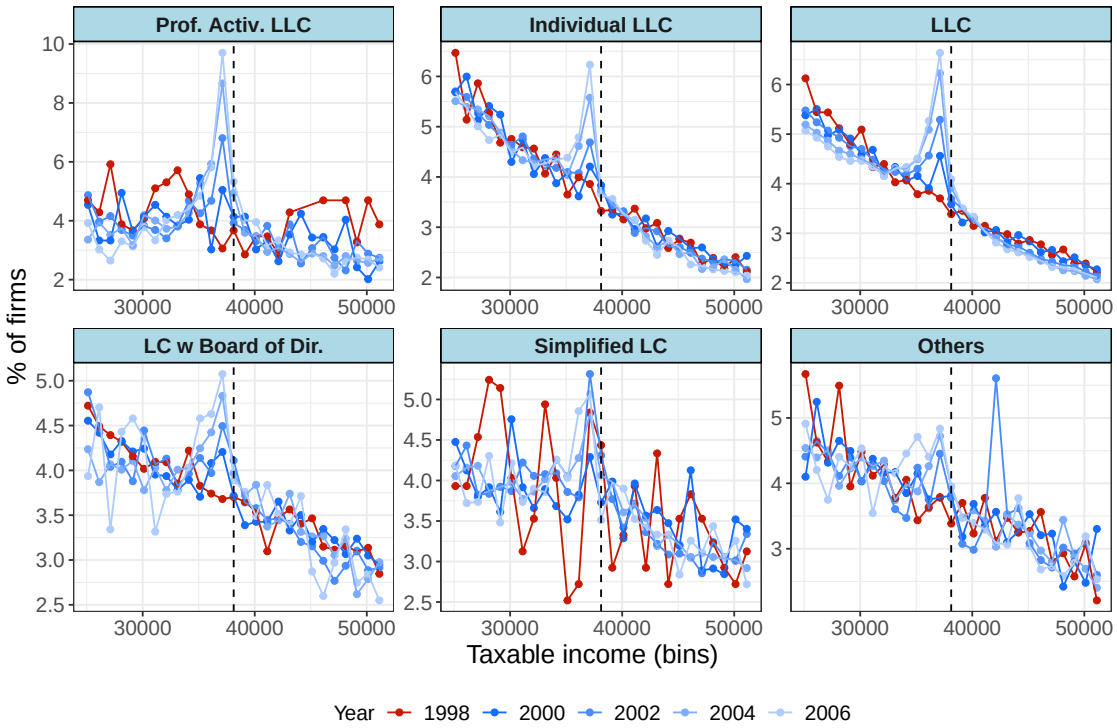

Figure 32: Distribution of firms (%) by taxable income across years and legal status

Note:

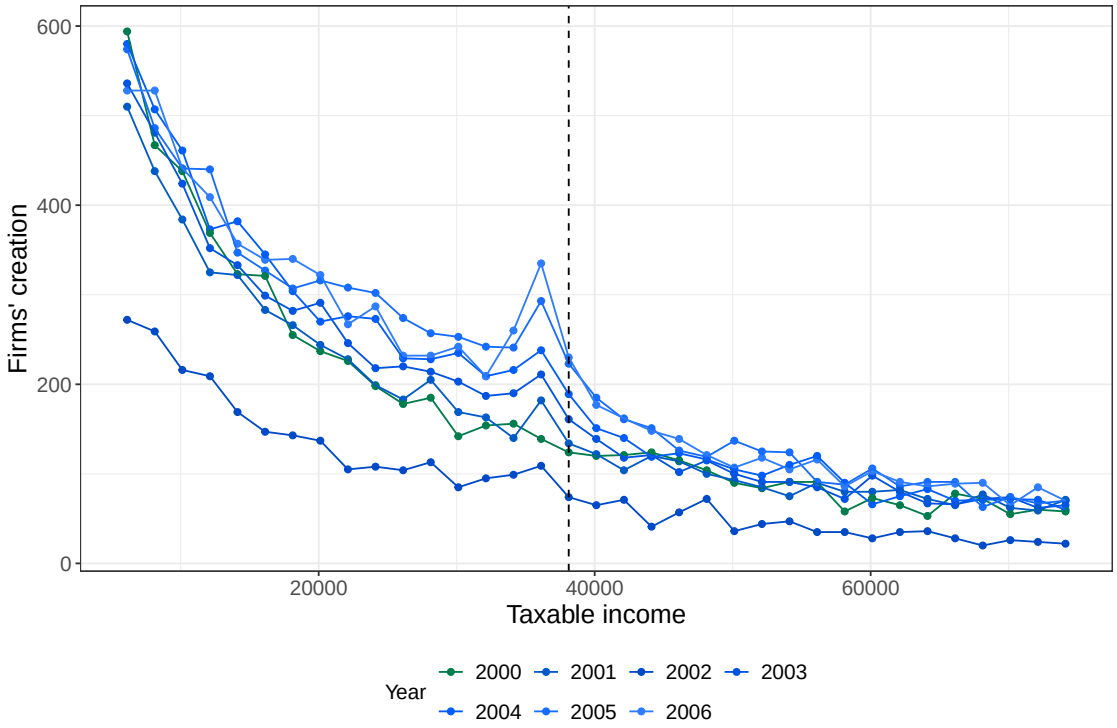

Figure 33: Number of business creations by taxable income across years

Note:

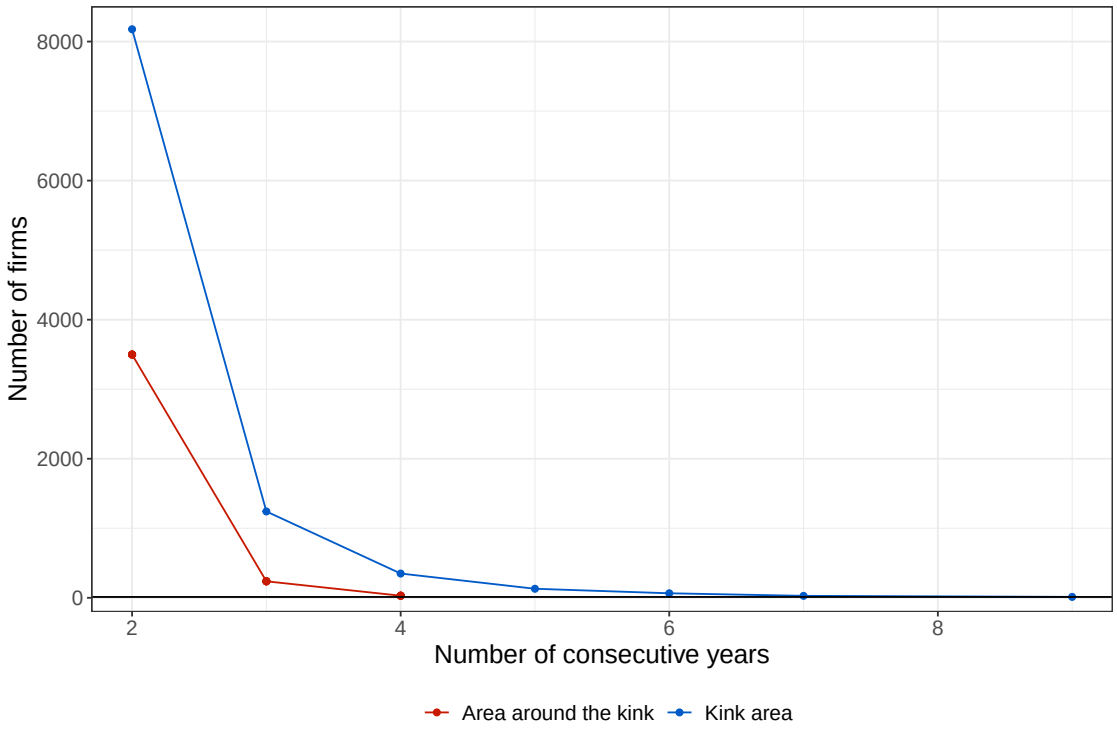

Figure 34: Number of firms consecutively reporting the same taxable income

Note:

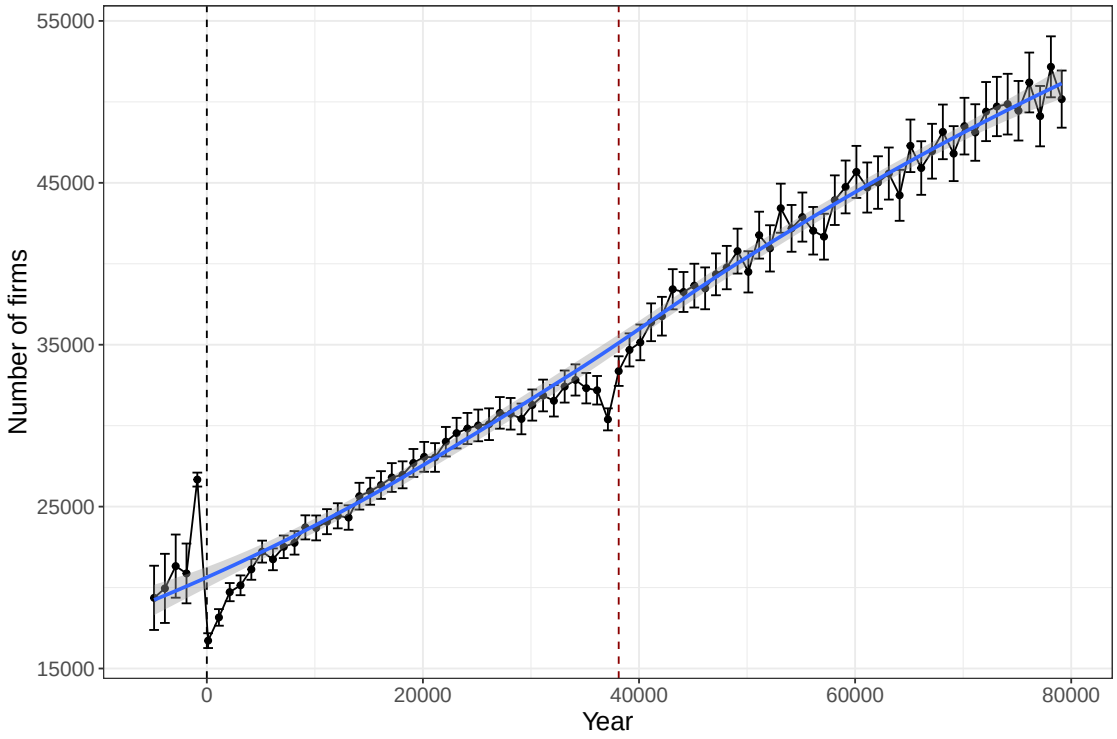

Figure 35: Intra-firm volatility between 2001-2016 in 2010

Note:

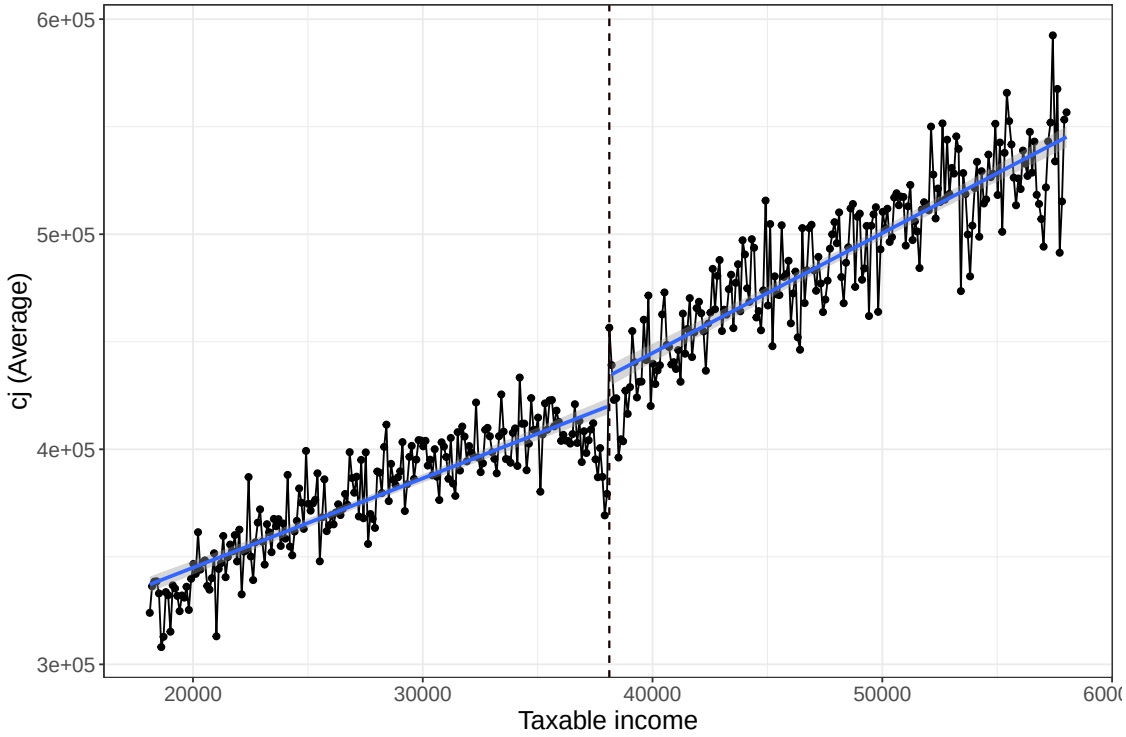

Figure 36: Average current assets by taxable income around the kink (2001-2007 pooled)

Note:

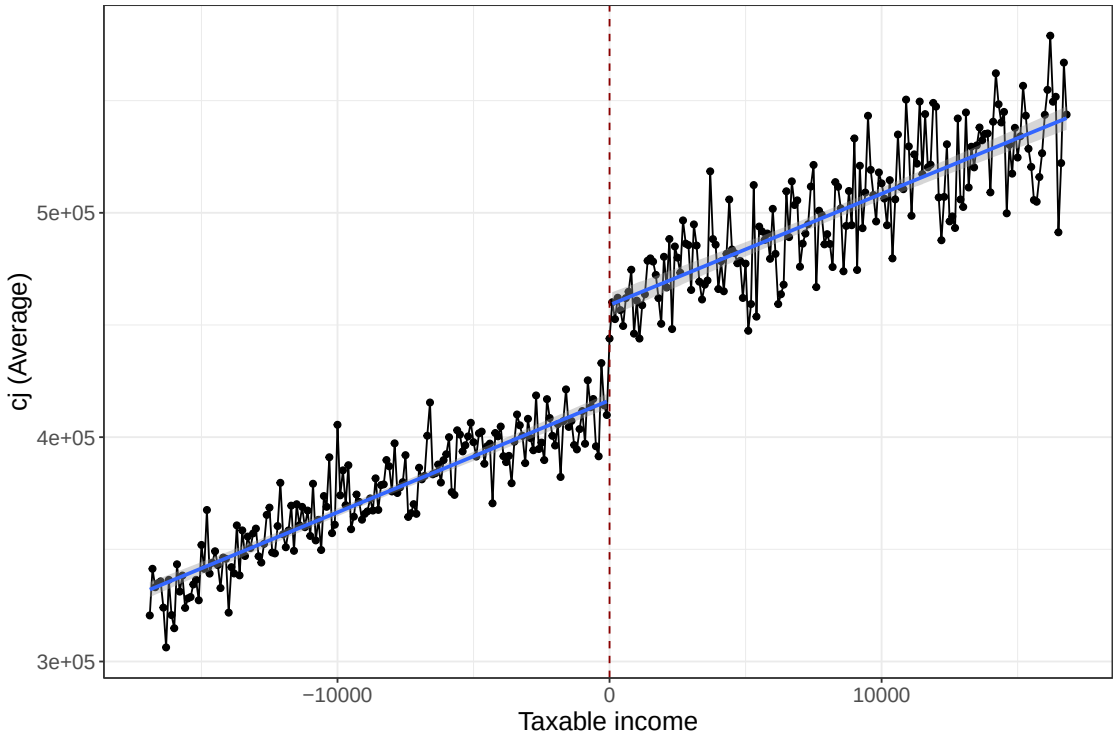

Figure 37: Average current assets by taxable income around the kink without the bunching region (2001-2007 pooled)

Note:

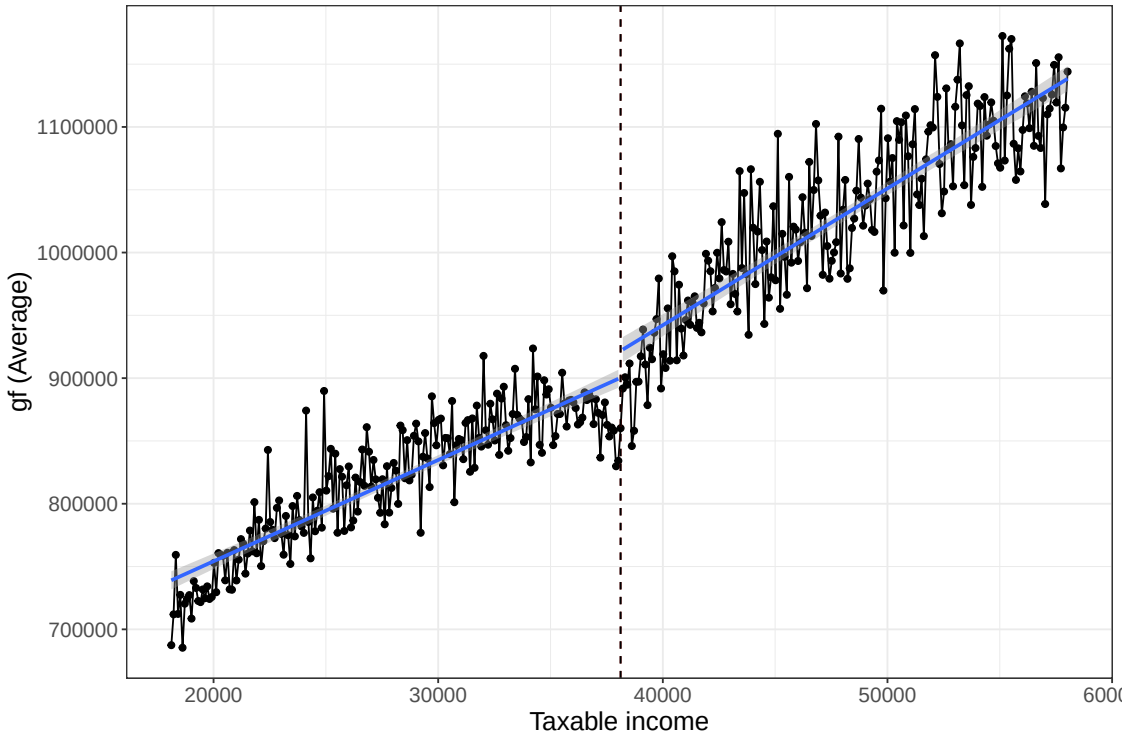

Figure 38: Average total operating costs by taxable income around the kink (2001-2007 pooled)

Note:

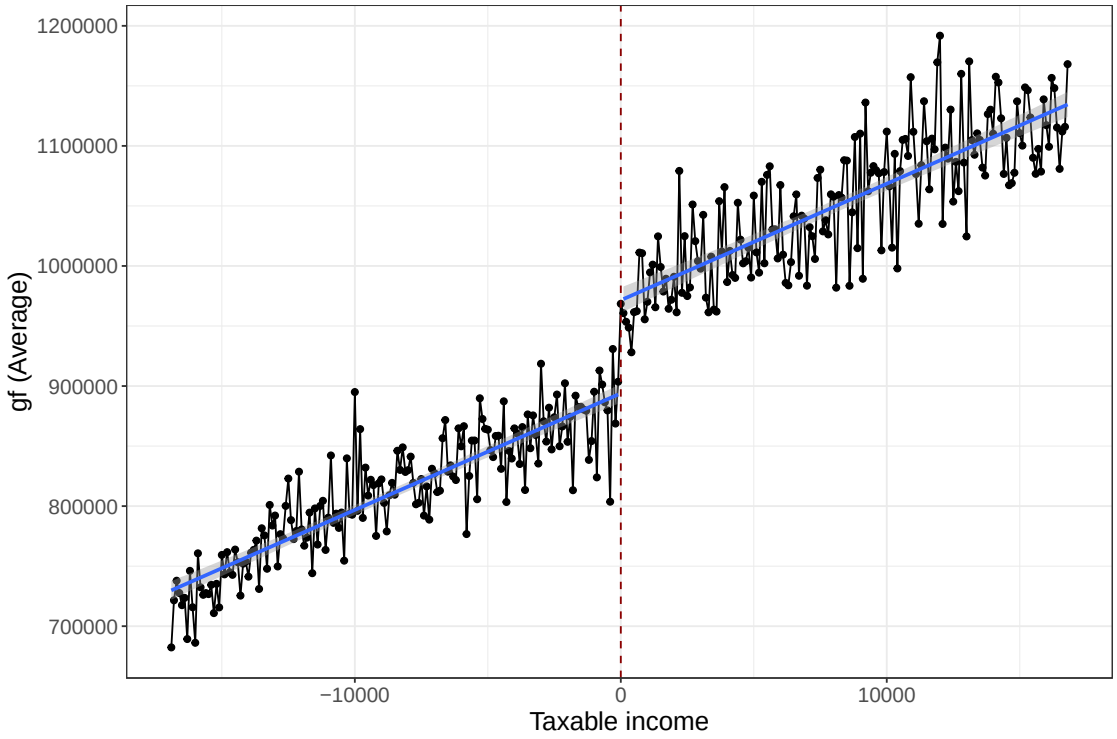

Figure 39: Average total operating costs by taxable income around the kink without the bunching region (2001-2007 pooled)

Note:

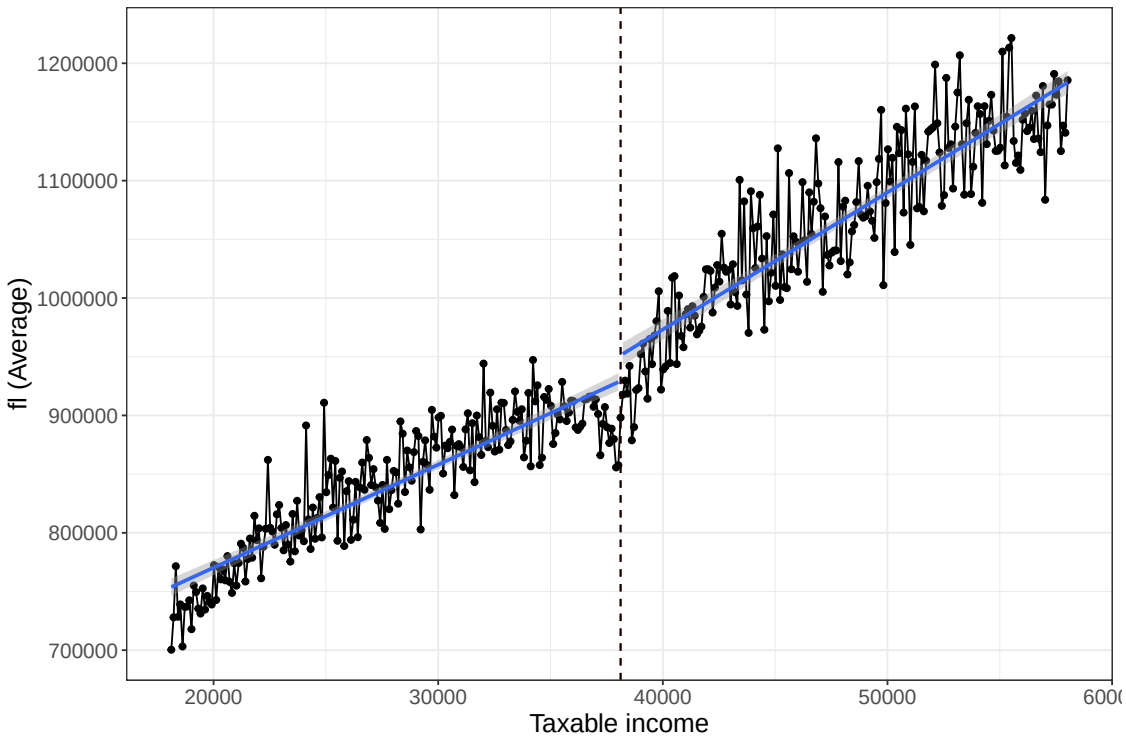

Figure 40: Average total sales costs by taxable income around the kink (2001-2007 pooled)

Note:

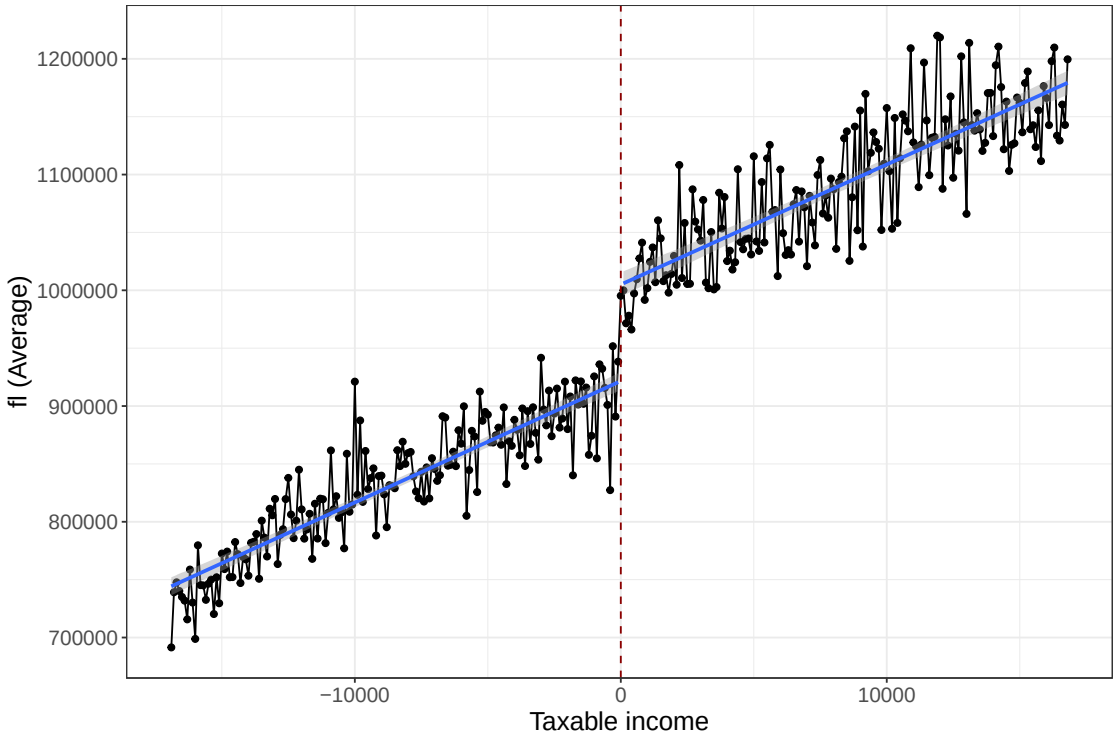

Figure 41: Average total sales by taxable income around the kink without the bunching region (2001-2007 pooled)

Note:

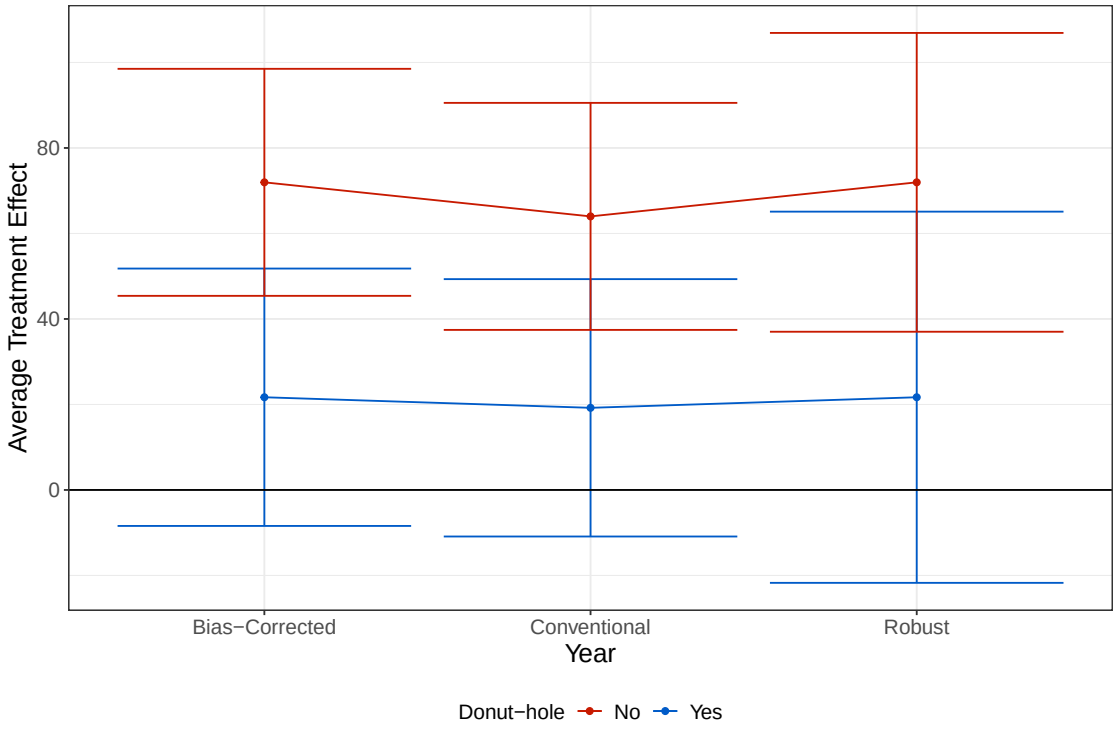

Figure 42: RKD estimation of total sales with and without the bunching region (2001-2007 pooled)

Note:

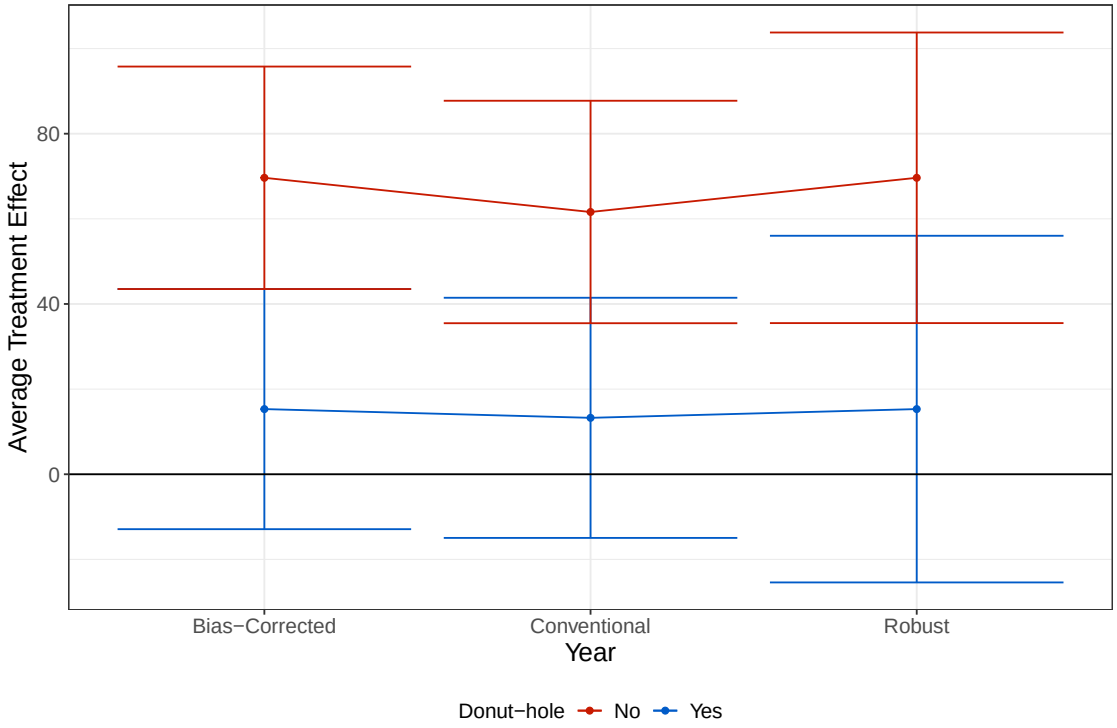

Figure 43: RKD estimation of total operating costs with and without the bunching region (2001-2007 pooled)

Note: