



HAL
open science

Economic responses to fiscal incentives of owners of small corporations: evidence from the Netherlands. *

Francois Lafont

► To cite this version:

Francois Lafont. Economic responses to fiscal incentives of owners of small corporations: evidence from the Netherlands. *. 2018. halshs-04504190

HAL Id: halshs-04504190

<https://shs.hal.science/halshs-04504190>

Preprint submitted on 14 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Economic responses to fiscal incentives of owners of small corporations: evidence from the Netherlands.*

Francois Lafont 11386347

January 6, 2018

Abstract

This paper investigates fiscal incentives for owners of small corporations, *Directeur Grootaandeel-houders* (DGAs), inherent to the Dutch tax code and how DGAs react to them. We find evidence of inter and intra-temporal income shifting, and of bunching at the kink point. These behavioural responses correspond to an elasticity of corporate income with respect to the net-of-tax rate of 0.72. By comparing estimates for different subsamples of the DGA population, it is shown that this elasticity is composed of both a real-economic component and an income shifting component, principally via investment deductions.

*I am grateful to Hessel Oosterbeek for his supervision. I thank the CPB Netherlands Bureau for Economic Policy Analysis (CPB) for the helpful and stimulating work environment. I am especially grateful to Nicole Bosch for her close scrutiny of my work and to Arjan Lejour, Marteen van't Riet and Leon Bettendorf for their valuable comments. I thank Statistics Netherlands for access to their microdata (project nr 7458).

Statement of Originality

This document is written by Student Francois Lafont who declares to take full responsibility for the contents of this document.

I declare that the text and the work presented in this document are original and that no sources other than those mentioned in the text and its references have been used in creating it.

The Faculty of Economics and Business is responsible solely for the supervision of completion of the work, not for the contents.

Contents

1	Introduction	4
2	Institutional Background and Data	6
2.1	Data	7
2.2	Taxation of the DGAs	8
2.3	Income shifting among DGAs	10
3	Theoretical framework	13
3.1	Intuition of the kink analysis	13
3.2	A model of bunching	14
4	Empirical Strategy	15
5	Results	19
5.1	Results at the €200,000 threshold	19
5.1.1	Whole sample	20
5.1.2	Possible channels of fiscal optimisation	23
5.2	Results at the €0 threshold	26
6	Discussion and Conclusion	27
A	Technical Appendix	32
A.1	Estimates at the €0 threshold	32
A.2	Formal version of model of bunching	34
A.3	Box 1 and CTI rates for the 2007-2014 period	35
A.4	Precisions on Income shifting	36
A.4.1	The effect of income shifting on ETRs of DGAs	36
A.4.2	Evidence of income shifting	38

1 Introduction

In the Netherlands, the current coalition has revealed plans to decrease the top corporate income tax (CIT) rate from 25% to 21% (and to increase taxation on dividends). Moreover some forms of business organisation growing in popularity, such as that of the *Directeur-Grotaandeelhouders* (DGAs), grant their owners¹ the ability to shift income inter and intra-temporally. The purpose of this article is to investigate how the Dutch tax code creates possibilities and incentives for DGAs to reduce their fiscal burden and to estimate empirically consequent behavioural responses. I use comprehensive administrative data to reveal fiscal optimisation patterns for the different types of income perceived by DGAs. I then estimate a corresponding elasticity of corporate taxable income (ECTI) with respect to business taxation.

Taxation of corporate income occupies an important place in current European economic debates. Amidst vague calls for greater fiscal harmony for the Eurozone, some economists have exposed the existence of a “the race to the bottom”, referring to a downward trend of tax rates on businesses due to fiscal competition (Mooij & Nicodème, 2008). In the Netherlands, top CIT rates have already been lowered from 35% to 25% between 2000 and 2007. The French and the US governments have projects to lower their corporate taxation² and there is no sign of fiscal competition faltering in the near future. There is much uncertainty regarding its effects. The data suggest that it has not resulted in lower corporate tax revenues. Mooij & Nicodème (2008) however provide some evidence that this may be due to income shifting. In other words it could be that as the fiscal burden on businesses increased, the CIT base grew at the detriment of the personal tax base. It appears that part of the current uncertainty comes from the fact that there is little empirical evidence on corporate taxable income (CTI) responses to changes in tax rates.

Indeed while there is ample literature on the effects of taxation on business behaviour³, only three (to the best of the author’s knowledge) published

¹DGAs own at least 5% of the shares, thus owners does not necessarily mean sole-owners

²Official plans to bring the CIT rates from 33,34% down to 25% by 2022 in the former and from 35% to 20% in the latter.

³See for instance Golsbee (1998, 2004) or de Mooij and Nicodème (2008) for the influence of taxes on the choice of organisational forms; Graham (2003) for a review of the literature on changes of firms’ financing policy; or Hasset and Hubbard (2002) for a review on the literature on investment choices; or Chetty and Saez (2005) on dividend payouts.

papers have attempted to measure ECTIs. Gruber and Rauh (2007) do so using accounting data, which unfortunately means that tax liabilities are estimated rather than recorded. They estimate ECTIs around 0.2 using constructed marginal effective tax rates ($1 -$ marginal net return on investment). Dwenger and Steiner (2012), use German tax administration data around the time of a large decrease on the tax rate of distributed profits. They estimate a counter-factual average tax rates faced by firms absent any behavioural response and derive elasticities from the difference with the observed ones. Resulting elasticities are large (0.5) but likely to represent income shifting (on the form of losses brought forward pre-reform) rather than real responses.

The kinked structure of the CIT provides a means to measure ECTIs in the absence of tax changes using the so-called bunching method pioneered by Saez (2010) and Chetty et al. (2011). Saez (2010) demonstrated that the excess mass around kink points in the US income tax schedule is proportional to the elasticity of income. Chetty et al. (2011), using Danish tax records, corrected an inconsistency between the observed empirical distribution and its counterfactual. Kleven and Waseem (2013) then extended⁴ the method and showed that when kink points are located at round numbers, estimates might be overestimated⁵.

A recent paper by Devereux et al. (2014) and a working-paper by Patel et al. (2016) make use of the bunching method. The former uses UK corporate tax returns data to investigate tax base variations at two different kink points (one low at £10,000, one high at £300,000). They find elasticities between 0.13 and 0.17 for the high income kink point and between 0.53 and 0.56 for the lower one although these last ones are at least partially explained by income shifting and probably by evasion. Unfortunately Devereux et al. (2014) do not assess this possibility empirically and use a model which does not allow for this parametric relationship. Specifically, in their model, costs reported by firms are only an increasing function of output. The consequences are that a movement in the tax base implies a larger movement of output from the firm, which is not observed in the data. Patel et al. (2016) address this by augmenting the firm's model with a cost-reporting parameter. Using US administrative data, they find high ECTIs around 0.5 at the first (and only) kink point- at zero USD- of the US corporate tax schedule. Most

⁴The paper adapted the bunching techniques to notch points (discontinuities in the tax schedule).

⁵This is particularly relevant in the case of self-reported data (such as tax reports) due to a marked tendency to report income in round numbers.

of the observed movement in taxable income can be attributed to reporting responses.

This paper makes use of a rich dataset to investigate the source of tax base movements of DGAs. Specifically, data on the different deductible cost items reported by DGAs allow to check to what extent these are used to optimise the fiscal burden. The estimated ECTI is just 0.072 but drops to 0.047 when we control for cost-reporting responses via investment expenditures.

Feldstein (1995, 1999) showed elasticities of taxable income to be accurate measures of deadweight loss incurred from taxation and Saez (2001) then established their usefulness to estimate optimal tax rates. In some situations⁶ such as with income shifting, however, it may not be a sufficient parameter. Thus this paper makes two separate contributions. On the one hand it adds to the still-scarce evidence⁷ on ECTIs and their composition. On the other, it sheds qualitative light on fiscal optimisation by DGAs, who combine income shifting and cost reporting behaviour. Together these highlight difficulties of estimating societal cost using simply ECTIs.

The remainder of the paper proceeds as follows. The next section describes the fiscal context in the Netherlands and presents the data. Section 3 discusses the theory behind the empirical analysis used, which is elaborated upon in section 4. Section 5 presents the main results and section 6 provides a concluding discussion.

2 Institutional Background and Data

The next two sections serve two main purposes. The first one is to clarify the context of the present research. This is done in section 2.1 and 2.2 respectively, by presenting the data and decomposing the tax system faced by DGAs. The second purpose is to show that the CTI of DGAs is the most useful point of analysis to understand how taxation affects their economic performance. This is done in section 2.3, by showing how DGAs can manipulate their income tax bases.

⁶See Saez et al. 2012 for a review

⁷There is another forthcoming paper on the same topic by Lediga, Riedel & Strohmaier (2017) who use South African tax data.

2.1 Data

The paper uses a combination of merged administrative sets from Statistics Netherlands (CBS) covering the years 2007-2014. One is a new dataset from the CBS that contains comprehensive information on all registered small-businesses in the Netherlands, including detailed profits, costs, losses and investments declarations (as well as number of employees, date of creation of the business etc.). It is merged and linked to tax receipts of DGAs as well as that of other self-employed individuals in the case of hybrids (individuals registered as both self-employed and DGAs). The data also includes some socio-demographic information on DGAs. There are close to 13 million observations in the dataset, of which around 1.8 million are registered as DGAs and 346 thousands are separate entities.

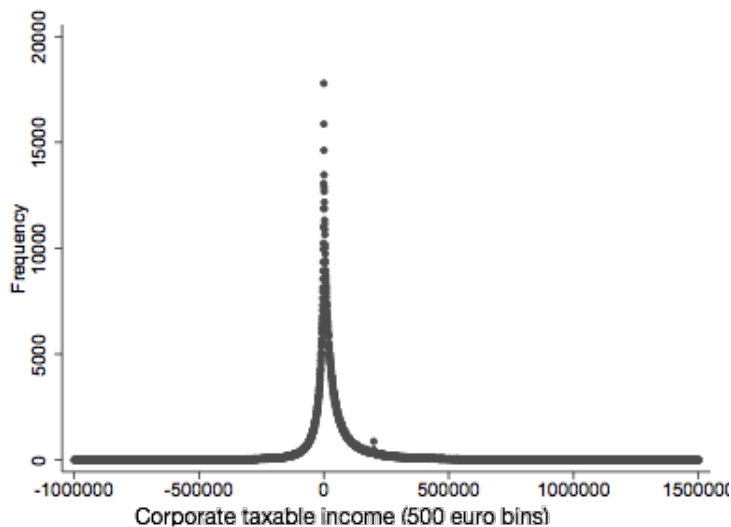
Administrative data is well-suited for fiscal analysis, as it is in general more reliable than similar survey data. Errors or incomplete information exist however and once we exclude observations without information on profits, there are 1,728,000 observations (309,920 separate entities) left. Due to the novelty of the data, some other uncertainties remain. For instance it is likely that there is a substantial number of dormant or inactive firms at the zero kink, which prevents assertive conclusions from the evidence at that point⁸.

There is comprehensive information on investment costs and losses. For instance the dataset contains separate figures for investment spending in R&D or in energy and environment related expenditures. However individual aggregate investment and losses figures are used to guarantee a sufficient number of observations. Table 1 provides descriptive statistics of the data for the pooled years. Given the tax schedule in 2007, it is excluded for most of the empirical analysis. Instead, the data for this year is used in order to verify the validity of a causal interpretation of the estimates.

Figure 1 shows that the majority of DGAs are concentrated at low levels of income, with an obvious peak in the distribution around the zero profit threshold. Importantly the kink of the CIT schedule is located at a relatively high level of profit, i.e. where there is a lower concentration of firms. There are 9852 observations with reported CTI between €195,000 and €205,000.

⁸Lediga et al. (2017) report substantially and systematically more than half of firms being either dormant or inactive at the zero kink in their data.

Figure 1: CTI distribution of the full population of DGAs 2007-2014



Note: This figure shows the full distribution of CTI. Outliers are excluded on both sides of the distribution so this figures actually shows the distribution of 99.5% of the data, collapsed into bins of 500 euros.

2.2 Taxation of the DGAs

Not uncommonly, different types of incomes are taxed at different rates in the Dutch tax system. Income received by people can be taxed in two different boxes, while income received in the form of corporate profits is taxed under the CIT. Box 1 concerns income received from a professional activity or from running a business and Box 2 is reserved to capital income gains, which include dividends and sale of assets. Together, they constitute the personal income tax (PIT).

The Corporate Income Tax- The Dutch tax code de facto implements some sort of differentiation between small and larger businesses, which are taxed at a different rate when it comes to corporate taxation. The ruling factor is the amount of profit declared by the firm⁹. The marginal rate is 20% in the first bracket. There is a 5% marginal increase at the €200,000 threshold that delimits the second bracket. The Dutch corporate tax system thus features a sizeable kink at this threshold. Table 3 in appendix A.1 contains the precise

⁹Rather than for instance its number of employees or operating revenues as in other countries (OECD, 2005).

Table 1: Descriptive Statistics

	<u>Figures</u>	<u>Number of obs.</u>
<i>Average age of business</i>		
DGAs	5.8 years	1,883,672
All categories	4.6 years	12,991,971
<i>Average profit level</i>		
very young businesses (< 5 yrs)	31,960	333,845
mature businesses (>20 yrs)	67,990	20,294
	<u>% of total</u>	<u>Number of obs.</u>
<i>DGAs</i>	100	1,883,672
<i>Per separate entity</i>	18.3	345,827
ZPs (and MPs)	56.3 (43.7)	1,061,481 (822,181)
Women (Men)	21.3 (78.7)	378,305 (1,397,775)
Around the kink	24.9 (75.1)	2,414 (7,079)
Businesses with dividends	14.9%	264,773
<i>Per separate entity</i>	14.1%	44,800
Businesses with investment	28%	514,000
<i>Around the kink</i>	48%	4,700
Businesses with pension contributions	38.5%	683,000
Average wage/profit ratio	16.9%	

Notes: Figures for the share of women, businesses with dividends, with investment and with pension contributions are done with a slightly different sub-sample: figures with missing information on profits were already excluded, which is why the number of observations add to a slightly different total.

Average wage/profit ratio is a weighted average of box 1 revenue declared over the profit of the firm, where the firms with profits systematically negative or above €2 million are excluded.

Business with dividends (per entities) correspond to observations (entities) with positive dividends. The 14.1% figure means that 14.1% declared positive dividends at least once. Around the kink refers to the [-5000, 5000] interval on either side of the kink.

rates and thresholds for all relevant years.

*The Personal Income Tax*¹⁰- Income in box 1 is subject to progressive taxation across four different brackets. The top marginal rate of 52% begins around 55K-60K depending on the year. Box 2 taxation on capital gains consists of a flat tax of 25% (see table 2 in appendix A.1 for details).

Figure 2 below shows the fiscal treatment of DGAs. Profits can be retained or distributed, either as a wage, in which case it is exempt from the CIT and taxed solely in box 1, or as capital gains. For the latter, the box 2 rate is applied after the corporate income tax. Each euro taken out in this form is thus in essence doubly taxed at the marginal rate of either 40%, before the kink, or 43.75% after the kink. The years 2007 and 2014 are notable exceptions where dividend taken out was taxed at 22% until a 250K threshold.

2.3 Income shifting among DGAs

It is important to take from section 4.1 that DGAs enjoy a certain degree of flexibility regarding how to distribute their profits. On the one hand they can choose whether to distribute or retain the profits within the company, which corresponds to inter-temporal income shifting. On the other they can choose to distribute profits between box 1 and box 2, which we call intra-temporal income shifting. The choice of box 1 income directly affects the CTI: one euro taken out in box 1 is one less euro taxed as business income. Thus it is vital to our analysis that box 1 income is properly reported by DGAs. It is possible to verify the accuracy of declared box 1 income in our data by checking for the presence of patterns reflecting fiscal incentives.

In essence, it is fiscally optimal for DGAs to minimise the amount of profits taken out in box 1 as soon as the second bracket is reached¹¹. The Dutch system features some rules concerning the wage of the DGA that somewhat restrict the income shifting possibilities. Although the degree of monitoring and enforcement of these rules is unclear there is a perceived minimum le-

¹⁰There is also a wealth component of personal taxation, however it does not apply to taxation of companies and is ignored here. See for instance Bettendorf et al. (2016), section 2, for a more detailed description.

¹¹To preserve clarity the differences in ETRs are not shown in any details in what follows. The interested reader can find a more complete depiction of the impact of income shifting on the fiscal burden of DGAs in appendix A.4.

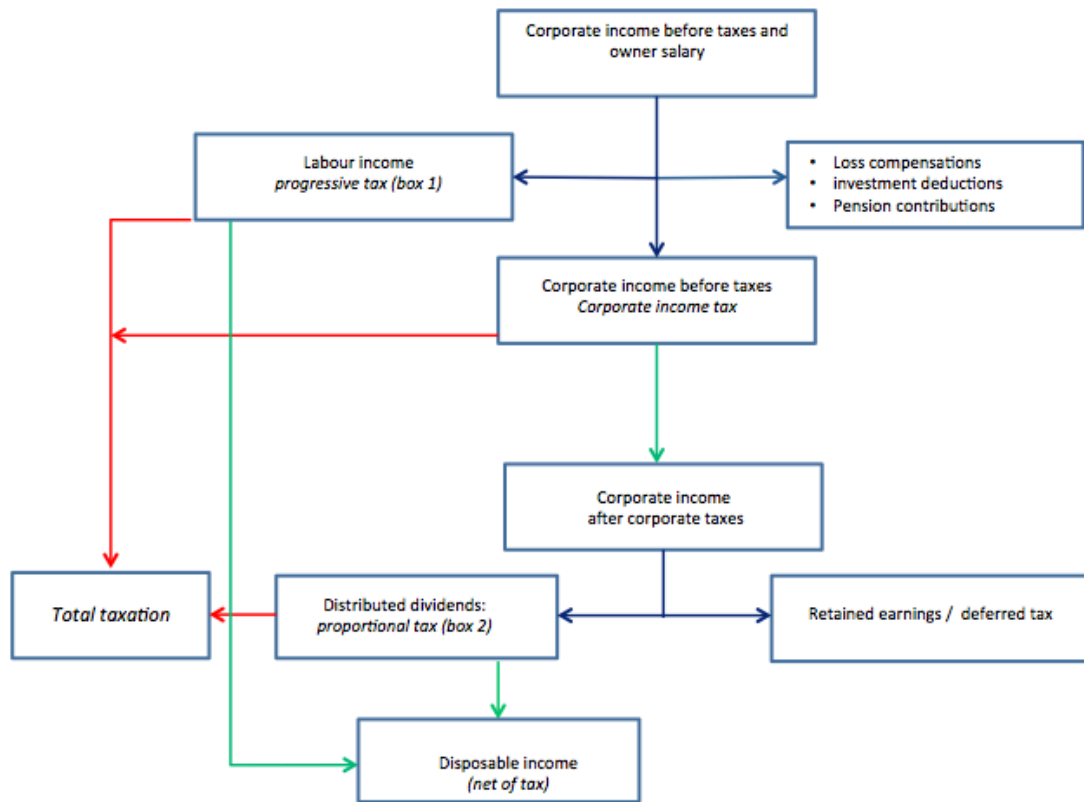
gal requirement in terms of wages for DGAs, that ranges from €40,000 to €44,000 across the sample period. In reality this is the threshold after which the responsibility on justifying the wage perceived by the DGA falls upon the tax authorities and not on the DGA (Bettendorf et al., 2016). What is observable in the data (see figure 13 in appendix A.4.2) is that a disproportionate amount of DGAs choose to pay themselves this minimum, possibly to avoid the costs of justifying their case to the tax authorities. Bettendorf et al. (2016) document this more extensively and also find significant bunching at this threshold for the 2007 and 2011 period.

On the other hand, due to inter-temporal income shifting, box 2 income for a given year is not necessarily related to the CTI. Since DGAs have total freedom in how much dividend they choose to distribute with regards to the tax authorities it is optimal to retain profits in wait for potential tax rebates¹². If this is the case, it would mean that dividends are not well suited to observe the economic performance of a given DGA, at least over a short-period, or how these are affected by a change in tax rates. Nonetheless observing fiscal optimisation patterns in box 2 income would reveal salience of the tax code, which is relevant for this analysis as we discuss further below. There is strong evidence that profit retention is a used tool for fiscal optimisation (see figure 14 in appendix A.4.2). DGAs substantially increased the amount of dividends declared in 2007 and 2014. It is possible that the announced reform to increase box 2 taxation to 28% will exacerbate the incentive to do so, especially if tax filer expect other box 2 tax holidays.

Let us recapitulate. Optimisation patterns in figure 13 and 14 testify for the accuracy of the data on wages and on dividends. It also suggests that, as far as possible, DGAs will try to minimise box 1 income, and that due to inter-temporal shifting, the flow of dividends may not depend on the CTI. Thus focusing on the CTI rather than on box 1 and box 2 income is arguably a better way to study the impact of small tax changes on the economic activity of DGAs. For these reasons it is useful to dispose of quantitative estimations of the various response of agents to changes in the CIT which are captured by ECTIs.

¹²Given the generous rebates on the taxation of bequests, it is actually quite optimal to retain profits until death. Such extreme scenario aside, it is to be expected that firm owners wait for potential temporary tax cuts to take out dividends.

Figure 2: the fiscal treatment of DGAs



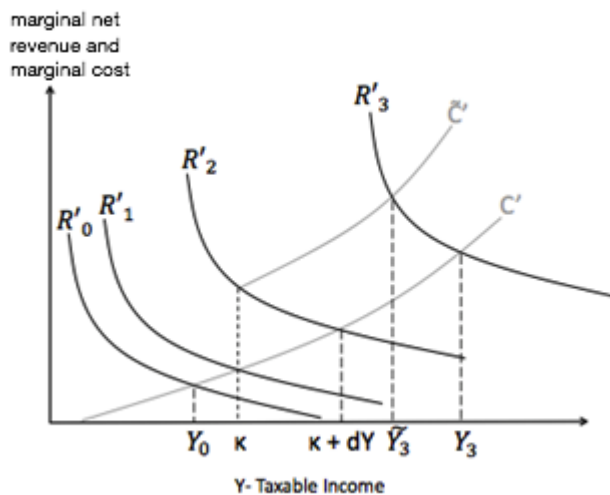
Note: Figure 3 is a simplified graph of the fiscal treatment of DGAs. Components not relevant for this paper are excluded such as for instance the general tax credit or the earned income tax credit. The arrows indicate the direction of the money flow. Red arrows indicate tax liabilities, and green arrows indicate net of tax income.

3 Theoretical framework

3.1 Intuition of the kink analysis

It is conventional in taxation theory that a smooth distribution of heterogeneously productive firms with heterogeneous preferences for corporate governance, subject to a linear tax schedule results in a smooth distribution of taxable income¹³. The intuition behind the bunching method is best exposed graphically with firms' marginal net revenue $R'_i(Y)$ and marginal cost $C'(Y)$ ¹⁴ as a function of gross business income. There is a kink in the tax schedule at gross income level k , which increases the marginal CIT rate from τ to $\tau + d\tau$ and shifts the marginal cost functions up to $\tilde{C}'(Y)$ for all $Y > k$. Firms choose an optimal level of taxable income, Y , at the point where marginal profit equals marginal cost.

Figure 3: bunching at a kink point



source: Patel et al. (2016). Firms maximise where marginal net revenue $R'_i(Y)$ equals marginal cost $C'(Y)$. Due to the non-linearity of the tax schedule, the optimal choice of taxable income of some firms is lower than their full capacity. Here, this effect is strongest for the firm at the margin, firm 2, who reduces her taxable income by dY .

¹³This is analogous to the case of personal income taxation, with a smooth distribution of heterogeneous ability and preferences.

¹⁴Costs here thus include the tax burden.

From Figure 3, we consider separate cases. Firm 0 and 1 are not affected by the kink because their optimal choice is such that $Y^* \leq k$. Firm 2's marginal revenue equals the post-shift marginal cost function $\tilde{C}'(Y)$ precisely at $Y^* = k$, thus its optimal choice of taxable income is at the kink. On a linear tax schedule (i.e. without a kink) however, firm 2 would instead choose $Y^* = k + dY$, where dY is the change in CTI for the firm responding to the kink at the margin. This is generalisable to all firms between Firm 1 and Firm 2 with $R'_i(k) \in [C'(k), \tilde{C}'(k)]$, which adjust their CTI by an amount between 0 and dY to bunch at the kink. Such response generates an excess concentration of firms at this point relative to the distribution when the tax schedule is linear, which is the counterfactual distribution. Note that firms with $R'_i(Y) > R'_2(Y)$, such as Firm 3 on the picture, also reduce their gross income levels although k represents a suboptimal level.

Intuitively, a larger number of firms bunching results in a larger aggregate change in the tax base. From the definition of the elasticity e of earnings with respect to the net of tax rate (in our case this is analogous to $1 - C'(Y)$), we have for a small change in the marginal tax rate

$$\frac{dY}{k} = e \frac{(1 - C'(Y))}{d(1 - C'(Y))}. \quad (1)$$

Saez (2010) provides the insight that the total number of tax filers bunching at k is equal to the product of the density of the counterfactual distribution of incomes and the change in income dY . As per the equation above, dY is proportional to the elasticity of earnings and it follows that greater excess mass corresponds to a larger elasticity. Thus it is possible to recover elasticities once excess mass is estimated. Importantly, the assumption of homogenous elasticities implicit in equation (1) is not necessary as bunching is proportional to the local average elasticity of a population with heterogeneous elasticities¹⁵.

3.2 A model of bunching

To make sense of the different reduced form estimates obtained when comparing the different subsamples of DGAs, and interpret them as structural parameters, it is necessary to model a parametric relationship between the

¹⁵See Saez (2010: 6) for a simple formal explanation.

data and the theory. The model by Patel et al. (2016) allows a distinction between reporting and real economic responses. This feature offers a good fit to the data. The intuition of the model is presented below and a formal version can be found in appendix A.2.

Consider a two-period setup. Firm i begins period 1 with an amount of retained earnings. Its choice of dividends and equity issuance defines its level of capital in period 2, which is the sum of its retained earnings and equity issue net of dividend payments. In period 2, the firm produces output using its stock of capital and its own technology described by a firm-specific productivity factor. Firm i can then lower its taxable income by shifting profit out of the CTI base at a cost c . The cost function is convex in the percent of CTI shifted, ρ . Intuitively this cost function incorporates potential inefficiencies linked to income shifting (in the case tax avoidance) and expected costs linked to monitoring and sanctions (in the case of tax evasion e.g. via false reporting), hence the convexity of the function in ρ .

Firm i maximises its value to shareholders, V , which is net profits minus the cost of shifting and the revenues foregone from government bonds. Following the standard definition, the ECTI with respect to the net of tax-rates can then be derived as

$$e_{y,\alpha} = \frac{dY}{d\alpha} \frac{\alpha}{Y} = \frac{(1-\rho)\alpha}{\alpha + \frac{\rho\tau}{2}} \epsilon + \frac{\rho\alpha}{(1-\alpha)(1-\rho)}. \quad (2)$$

The key intuition in equation (2) is the following. When there is no shifting (for instance if the cost of doing so is large) i.e. when $\rho = 0$, the ECTI is equal to ϵ and captures the real economic responses induced by the tax code. When $\rho > 0$ the estimated elasticity also captures CTI movements due to income shifting and the estimated elasticity corresponds to $e_y = \epsilon + e_s$, where e_s is elasticity due to income shifting. Importantly, the movement of the CTI base is in this case more than proportional to real economic responses, which seems to be the consensus of the literature.

4 Empirical Strategy

For clarity it is useful to follow Chetty et al.(2011) and to rewrite equation (1) as,

$$e(t_1, t_2) \simeq \frac{b(t_1, t_2)}{k \cdot \ln\left(\frac{1-t_1}{1-t_2}\right)} \quad (3)$$

where b is the excess mass. In this form, the equation clearly states that the estimated ECTI, \hat{e} , is proportional to the bunching around k . Determining what part of the observed mass is in “excess” requires a comparison with the counterfactual distribution. The plotted distribution of the DGAs in the Netherlands between 2007 and 2014 bares a spike right at the kink (red dot in Figure 4), whereas it is otherwise smooth and monotonically decreasing. Typically the data displays some noise around the kink due to optimisation frictions and it makes the precise identification of the counterfactual distribution more complex. The now standard method pioneered by Chetty et al. (2011) consists of estimating it by fitting a polynomial through the plotted data, excluding the bunching region¹⁶, as in the following regression

$$D_j = \sum_{i=0}^q \beta_i (Z_j)^i + \sum_{i=Z_-}^{Z_+} \gamma_i \mathbf{I}[Z_j = i] + \epsilon_j \quad (4)$$

where D_j is the number of DGAs in CTI bin j , Z_j is the income level in that bin, $[Z_-, Z_+]$ delimits the excluded interval containing dummies for the bunching region, and q is the order of the polynomial. The excess number of individuals located at the kink \hat{B}_N is then equal to $D_j - \hat{D}_j^{cf}$, the difference between the observed and the estimated counterfactual distribution in the bunching region. Equation (4) is useful to convey the intuition but the resulting estimate does not account for the fact that the individuals bunching come from beyond the threshold (i.e. the density function does not add up to 1), thus that \hat{B}_N is overestimated. The counterfactual distribution has to be shifted upward to the right of the kink via an iteration process until the integration condition is met¹⁷. It is then possible to estimate b , the ratio of the observed mass relative to the average density of the estimated counterfactual distribution over the interval $[Z_-, Z_+]$. It is derived as

$$\hat{b} = \frac{\hat{B}_N^*}{\sum_{Z_-}^{Z_+} \hat{D}_j / (Z_- + Z_+ + 1)} \quad (5)$$

¹⁶There are several ways to determine how far this bunching region extends. Commonly the bunching region is determined by visual inspection of the data near the kink, but some work has been done on identifying it endogenously; see for instance Bosch et al. (2016). Visual inspection is used for this paper.

¹⁷See Chetty et al. (2011:21-22 for details).

where \hat{B}_N^* is the corrected version of \hat{B}_N . The iteration convergence to correctly estimate the counterfactual distribution has consequences for inference as normal standard errors do not account for it. Specifically they correspond to the standard errors of the final iteration of the shifted data and are unaffected by the amount of shifting that has taken place. A common approach in the bunching literature is to estimate the standard errors of the excess mass via a bootstrap procedure. For all the results presented in this paper, the procedure is repeated 10,000 times. The standard errors for the elasticities can then be calculated with the delta method.

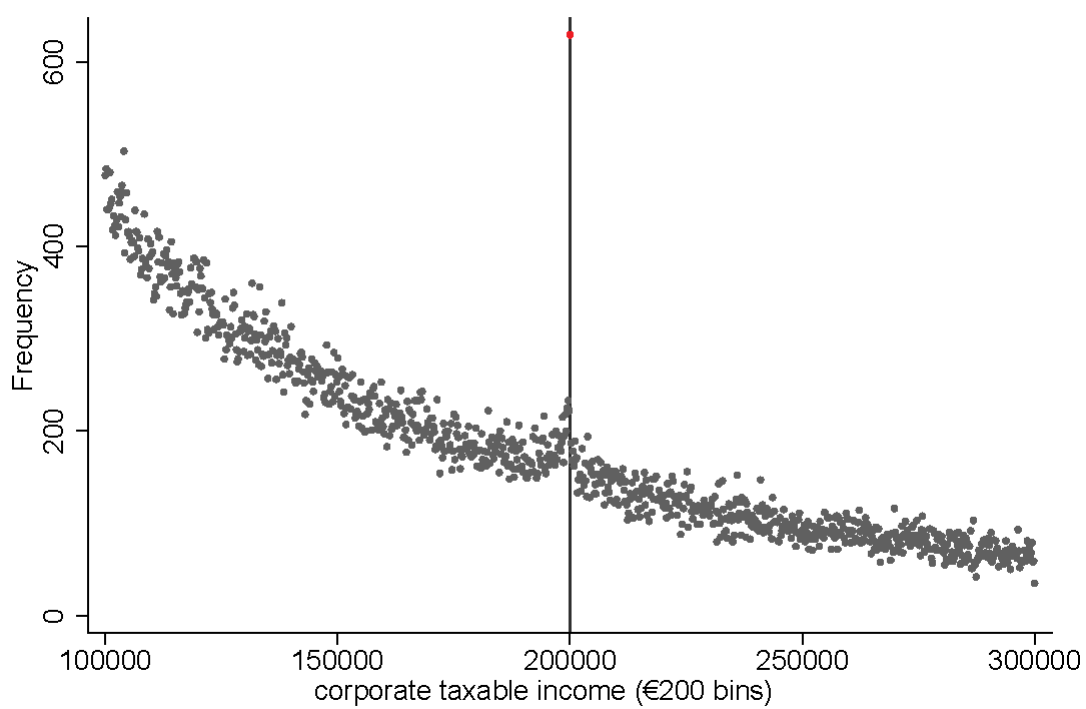
The bunching method often requires case-specific precautions. There is a tendency, documented by Kleven and Waseem (2013) for individuals to report income in round figures¹⁸. Indeed it is possible that DGAs round up their declared corporate income for ease of computation (figure 14 shows that this is the case for dividends for instance). If this is the case, there will always be some bunching at certain round figures of the CTI distribution. Given that the kink is located at a salient and convenient €200,000 the amount of bunching as a response to the tax increase would be exaggerated, ultimately yielding biased ECTI estimates. Under such circumstances, Kleven and Waseem (2013) show that equation (4) can be easily augmented to account for the potential bias. Conveniently, there does not appear to be significant bunching at other round numbers in the data. The yearly average number of firms at round points (every €5,000) between €5,000 and €100,000 is 245 in total, and less than ten¹⁹ at round points in increments of 50,000 at the exception of €200,000 of course. Another potential threat to identification is the existence of other policies located at the kinkpoint that also affect CTI (Kleven 2016), but this is not an issue in the Dutch fiscal setup.

To estimate the excess mass for 2008 and the subsequent years together, a “distance from the kink-point” variable is constructed as the difference between the CTI and the threshold. The year 2007 on the other hand is excluded as the CIT schedule at the time is ill suited to the bunching method and not comparable to that of subsequent years. Judging from the distribution of CTI in Figure 4 the bunching region seems quite narrow. Observa-

¹⁸Such behaviour is also observed for other monetary aggregates, such as housing prices for instance as observed by Kleven and Best (2016).

¹⁹According to the CBS’ confidentiality rules, precise numbers of observations below 10 cannot be reported.

Figure 4: bunching of firms at the kink



Note: Figure 4 shows the distribution of CTIs between €100,000 and €300,000. Observations for 2007 are included in this figure to give a comprehensive depiction but are omitted in later analysis.

tions are grouped into 200 euro bins in order to delimit it as accurately as possible²⁰. For each estimation it is selected by looking at the number of observations per bins in the distribution and from the reported standard errors with the different specification. It turns out that an asymmetric bunching window is better suited to the data in all cases.

In theory DGAs will only bunch until a set level of income and that level should define the total number of bins considered on either side of the kink. There is no obvious way to decide the CTI upper-bound for bunching i.e. how far along the CTI distribution Firm 2 from the example in figure 1 is located. If the number of bins included in the analysis is too small, bunching firms will be overlooked, yielding conservative estimates of the excess mass. On the other hand extending the analysis too far along the distribution of firms decreases precision as more noise is included and the polynomial is harder to fit. Successive analyses with respectively, 5,000, 30,000 and 50,000 on either side of the kink confirm this intuition. ECTI estimates are smallest for the case of the $\pm 5,000$ interval, and least precise with the $\pm 50,000$ interval. There is no statistically significant difference between the estimates with 30,000 and 50,000 euros interval however. Given that the threshold is €200,000, it seems likely that a [195,000;205,000] interval is too restrictive, therefore the estimates with 150 bins, i.e. 30,000 euros interval, are selected²¹.

5 Results

5.1 Results at the €200,000 threshold

The section that follows provides empirical estimates of the local behavioural response of DGAs to fiscal incentives in the form of the corporate tax schedule. The excess mass of tax filers at the kink point translate into statistically significant elasticities, consistent with previous findings such as those of Devereux et al. (2014). Some interesting additional features of the behavioural response will be highlighted by means of a year-by-year inquiry.

²⁰Bins of 100 euro would have been preferable but some subgroups had too few observations around the kink to allow this. Thus rather than change the size of the bins according to the subgroups or the year, 200 was chosen for consistency. Such small difference has no quantitative impact, nor changes the interpretation of the results.

²¹As a reference, Devereux et al. (2014) consider firms with CTI between £250,000 and £350,000 for a £300,000 threshold. The present choice is quantitatively similar.

The analysis is repeated for subsamples of the DGAs population in order to investigate the role of cost reporting.

5.1.1 Whole sample

The local estimates in figure 5 correspond to a large excess mass of about, $b \simeq 8$ (i.e. eight times superior to that of the counterfactual), which following equation (8) corresponds to an elasticity slightly over 0.07²², both statistically significant²³. The implied quantitative effect of a tax change is rather small: a 10% increase in the marginal rate of tax is associated with a 0.72% decrease in the declared business income. In the case of the announced 4 percentage points decrease in the CIT rate, i.e. a 16% decrease, it should increase reported income by 1,15%.

It is possible to test the causality link between the observed bunching and the tax schedule by looking separately at the results for the year 2007, 2008 and the subsequent years. In 2007 the tax schedule included two small kinks much earlier in the income distribution (the second one was at €60,000) and in 2008 the system was changed with one large kink at €275,000, which is at €200,000 since 2009. If the bunching is caused by the kink there should not be any excess mass in 2007, but there should be some in 2008 and after. Two additional complications must however be noted. First, the change in the corporate tax schedule in 2008 took place in December of that year²⁴ and thus if tax-filers react to tax changes with a lag (during the knowledge spread for instance) or optimise their taxes throughout the year, it is likely that no bunching will be observed²⁵. The lag explanation is also consistent with the ECTI estimate for 2009 being slightly lower than for the subsequent years. Second, due to the tax holiday in 2007, DGAs have incentives to take out dividends i.e. not to bunch (see figure 6). It is possible to verify that an

²²The denominator must be itself divided by the size of the income bins, in our case 200. Also note that because of 2008 where the kink is at €275,000, the average threshold when the years 2008-2014 are pooled together is 219,000.

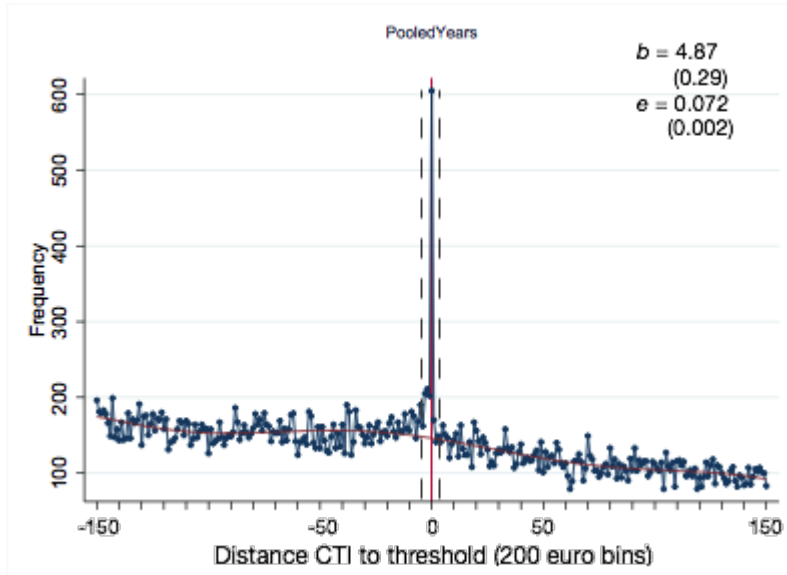
²³As a comparison, Devereux et al. (2014) find larger ECTIs -albeit of similar magnitudes, between 0.11-0.15- for a kink twice as large, consistent with the intuition from Chetty et al. (2011:25) that larger kinks yield larger behavioural responses as they represent greater incentives.

²⁴More precisely a plan of reform was announced in September but was only finalised in December.

²⁵The intuition for this is similar to the one about the knowledge effect from the salience of a kink mentioned below in 7.2.

absence of bunching is not caused by this by comparing it to the year 2014, when the same tax holiday was implemented. The statistically significant (albeit reduced) excess mass for 2014 strongly suggests causality.

Figure 5: bunching at €200,000

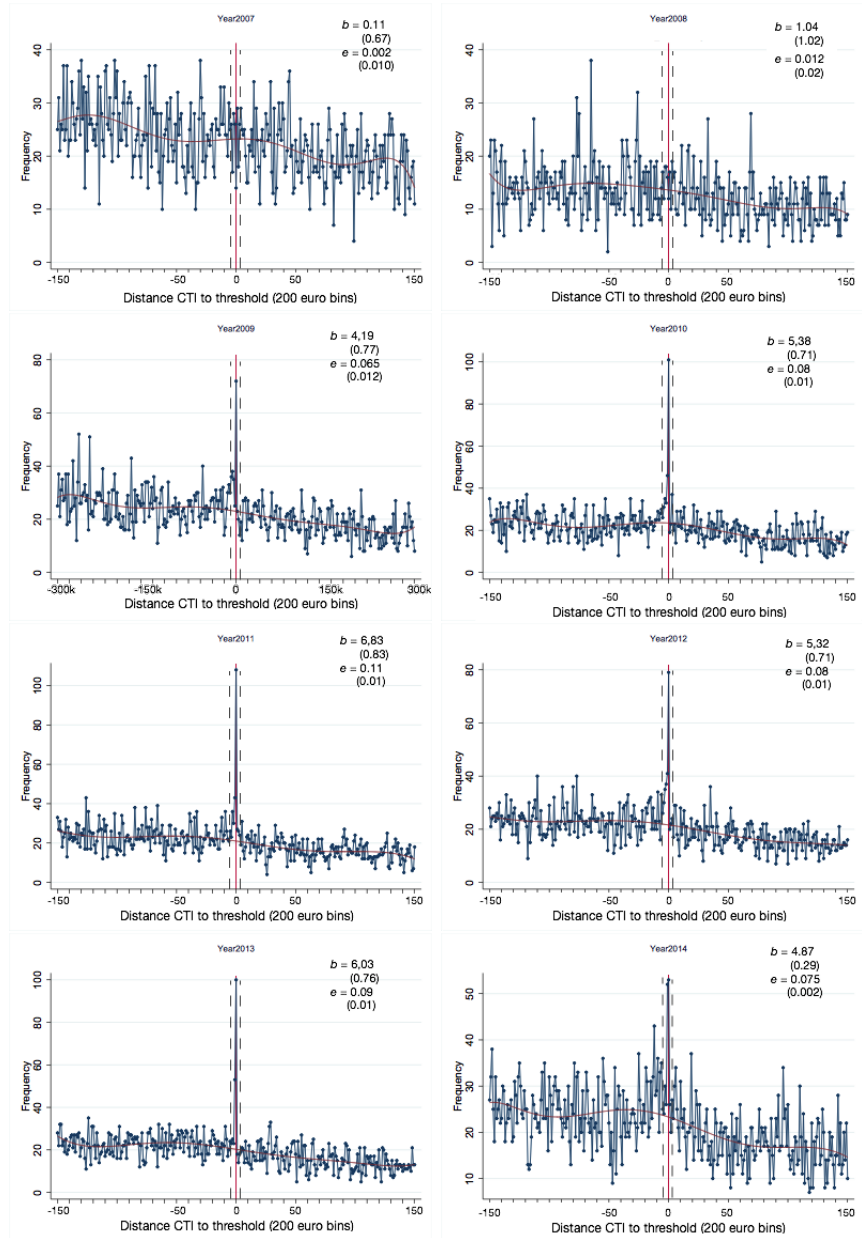


Note: Figure 5 shows the observed distribution (dotted line) and the estimated counterfactual distribution (solid smooth line) of CTI between 2008 and 2014 in 150 bins of 200 euros on either side of the kink (i.e. between €-30,000 and €30,000) represented by the vertical line. The counter-factual is a 7th order polynomial estimated as per equation (11). In brackets are bootstrapped standard errors for the excess mass and the ECTIs. The selected bunching region, delimited by the dashed lines, is 5 bins to the left of the kink and one to the right (i.e. within €1000 to the left and €200 to the right). Excess mass and elasticity figures are rounded up to second decimal.

As expected, figure 6 below shows that there is virtually no excess mass for the years 2007 and 2008. The period 2009-2014 displays statistically significant excess mass for each separate year. The excess mass for 2014 is smaller than for the previous years, probably due to the box 2 tax holiday. There are some small differences in the elasticities across years, $e_{2010} = 0.065$ and $e_{2013} = 0.09$ but overall the estimates are consistent except for the aforementioned years (2007, 2008, 2014 and to a lesser extent 2009) and a causal interpretation seems coherent.

Another striking feature is the precision of the observed bunching relative

Figure 6: bunching at €200,00 for the years 2007-2014



Note: Analogously to figure 5, Figure 6 shows the elasticities for each year between 2007 and 2014. The selected bunching region is 5 bins to the left of the kink and one to the right.

to that of previous papers²⁶. The bunching literature consistently points out optimisation frictions that impede the precision of bunching. Intuitively one can think of fixed contracts or orders over a period of time for companies or for relatively rigid amounts of money²⁷. In that light such a narrowly located excess mass suggests that bunching may be due to cost reporting, which, monitoring issues aside, is much more flexible since there is a wide variety of possible expenses to add as investment or costs.

5.1.2 Possible channels of fiscal optimisation

In what follows, the analysis is repeated for the following subsamples: DGAs not reporting any losses, only DGAs reporting losses, DGAs not reporting any investments costs, and only DGAs reporting investment costs. If losses and investment deductions play a role in fiscal optimisation, one expects bunching to be less pronounced for the subsamples without costs or losses declared, and vice versa.

In the vicinity of the kink, 88% of DGAs do not report losses. However, it is possible that those who do make use of the possibility to carry losses forward and backward to optimise their tax liability. For subsample without losses, the elasticity reported in figure 7 is $e = 0.068$, slightly smaller than for the full sample, although the difference is not statistically significant. The reported elasticity for the subsample reporting losses is much larger, $e = 0.20$, but given the small number of observations these results must be taken with caution. This is suggestive evidence that loss reporting may be used for bunching. Rigorously there is no way to assert causality here. It is good to note that in this part of the CTI distribution, close to half the firms report investment deductions and more than a third report pension contributions. Thus it is unlikely that the difference in bunching observed is linked to an involuntary exclusion of firms with a specific investment or pension behaviour.

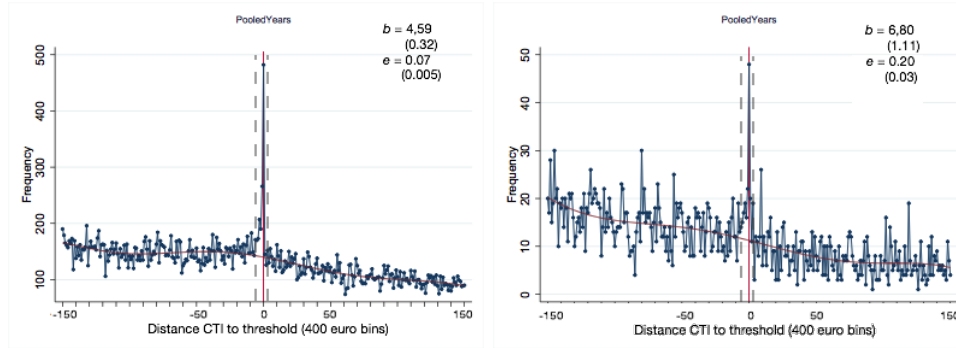
The shares of firms reporting (or not) investment costs are more even, which makes the comparison more trustworthy. Figure 8 shows a significant difference in the excess mass observed among businesses that do not report investment expenditures and among those that do, corresponding to respec-

²⁶For instance Devereux et al. (2014) specify an excluded region of £10,000 to the left of the kink and £4,000 to its right.

²⁷The intuition is the same for labour income (binding contracts, wages etc. See Kleven (2016) section 5.1.1 for a review.

tive elasticities of 0.047 and 0.094, i.e. a statistically significant difference of 0.047.

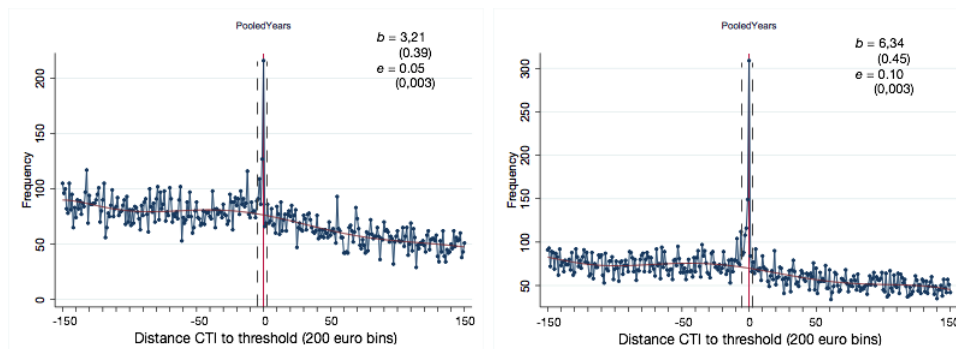
Figure 7: bunching at €200,000 DGAs with and without any losses declared



Note: Figure 7 shows the elasticities for DGAs with (right) and without (left) losses declared for the 2008-2014 period. The counter-factual is a 7th order polynomial. The selected bunching region is 5 bins to the left and 1 one to the right. Due to small number of observations for the sub-group with declared losses, the data is grouped in 400 euro bins.

A closer look at the data for the years 2008-2014 reveals additional evidence. For instance the share of firms with positive investment costs (48%) is substantially higher in the interval [195K , 205K] than it is in the interval [190K-195K , 205K-210K]. The figures of reported investment expenditures around the bunching region are also significantly higher. In the interval [170K , 230K] there are 544 observations with investment higher than 250K, and 31 observations with investment over 2,5M. In the interval [230K , 290K] there are less than ten observations over the with investment above 250,000 and none above 2,5M. Average investment spending in the bunching region considered in Figure 8 is €23,680, whereas if one extends that region to 10 bins on either side of the kink, that average falls to €16,308. Intuitively one would expect that firms with higher profits would have more to spend on investment, or that firms having invested more would be more profitable. This is strong evidence of cost reporting as tool for fiscal optimisation. Figure 9 shows this evidence graphically.

Figure 8: bunching at €200,000 DGAs with/without any investment costs declared



Note: Figure 8 shows the elasticities for DGAs without investments (left) and with investments (right) declared. The counter-factual is a 7th order polynomial. The selected bunching region is 5 bins to the left and 1 one to the right.

Figure 9: Investment spending by DGAs around the kink

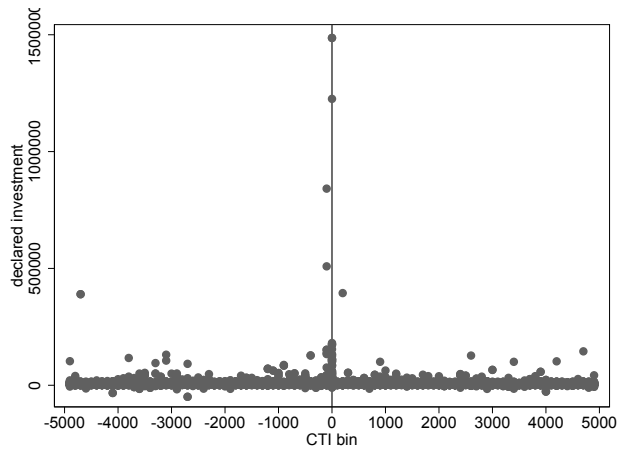


Figure 9 shows investment spending by firms in bins of 100 around the kink for the years 2008-2014. Firms located at the kink declare massive investment expenditures compared to others.

5.2 Results at the €0 threshold

The face of the distribution around that point evokes bunching, or at least the amount of observations around that point seems excessive. The estimated excess mass is just above 3.5 and statistically significant (see figure 10 in appendix A.1). However the bunching region appears very large and not clearly defined. In theory it is possible that given the size of the kink (the MTR jumps from 0% to 20%) there are many individuals bunching (see figure 1). Other factors may also affect the location of the bunching around this kink. DGAs with different box 1 revenues declared get different corresponding tax rebates (the aforementioned labour and general tax rebates) thus moving their “own” kink. Under perfect information, multiple personal kinks would then spread the tax filers who are bunching. It could also be that additional complexities such as these lower the salience of the kink²⁸.

It is also possible that larger firms manage to bunch around the zero threshold if they experienced large losses some years before and have carried them forward to smooth their fiscal burden. The percentage of firms reporting losses is also high in the (-5000, 5000) interval: 25% compared to the average of 15,5% of the rest of the sample. Taken together, these factors point in the direction of losses being used as an optimisation tool at this threshold.

On the other hand there is no guarantee of the precision of the counterfactual estimation at this point given the width of what seems to be the bunching area and the shape of the distribution at that point²⁹. There could for instance be many dormant/inactive firms at that point³⁰. Of course it also seems logical that small firms declaring CTI around €0 would on average suffer more losses than those with profits of several hundreds of thousand.

Figure 10 shows bunching at €0 for the sub-sample of DGAs that do not declare any losses. The difference with the rest of the DGAs is only marginal and not statistically significant. A last hurdle is the impossibility to compare bunching between years with and without a kink in the tax schedule at that

²⁸It is intuitive why less information about the kink would lead to a less precise and smaller aggregate behavioural response. For a more thorough theoretical explanation see for instance Chetty et al. (2009), and Chetty et al. (2013) for empirical evidence.

²⁹Indeed, with the possibility of losses i.e. of negative revenues, one can expect a peak at the zero threshold in a normal distribution of income, which makes excess mass harder to distinguish than in the case of an expected smooth distribution.

³⁰Lediga et al. (2017) for instance find many dormant firms at the zero threshold (27% of the total population).

point, since it is present in every year of the data. In sum, results hint that bunching at the zero kink is likely and perhaps achieved via loss-carrying. However the data does not offer enough certitude to make strong claims about fiscal behaviour of agents around that point of the CTI distribution and thus ECTIs are not calculated at that threshold.

On a more positive note, the consistency of the excess mass estimated also across the years 2007 and 2008 (see figure 11 in appendix A.1) strengthens the evidence for a causal interpretation of the excess mass at the €200,000 kink and the credibility of the resulting ECTI. In 2007 and 2008, changes occurred in the tax schedule at the €200,000 kink, and the data shows variations in the observed bunching for those years. At the €0 threshold, there were no changes in the tax schedule and the observed density of tax-filers shows no significant change over that period.

6 Discussion and Conclusion

There are several main findings in this paper. First, the estimated ECTIs are statistically significant and rather low. There is strong evidence for causality between the tax code and the observed bunching. In other words a change of the marginal rate on business income will bring a proportional change in tax revenue from the DGA population unlikely to be offset by significant behavioural response. This is true at least for DGAs around this part of the distribution. The paper also provides strong evidence of the use of investment costs by firms to optimise their fiscal burden. This emphasises the point made above about the expected loss of tax revenue due to a decrease in the CTI in the sense that the observed behavioural reactions to tax changes do not correspond fully to real economic responses. Importantly however, the data does not allow to verify whether the reported expenditures correspond to real investment or just to reclassifying other costs.

More generally this begs the question of the effectiveness of ECTIs to capture societal costs of taxation. Saez et al. (2012) show that in the presence of certain types of externalities, the estimation of elasticities must be accompanied by that of other parameters in order to accurately capture deadweight loss. When income shifting occurs for instance, it is crucial to also estimate the degree to which the income of one tax base shifts to another tax base, as well as the rate at which it is then taxed. When such shifting occurs through time however, it becomes harder to quantify the societal cost (Saez

et al. (2012:11). In the Dutch case for instance, the retained income may be taxed at a normal box 2 rate, or not depending on tax holidays such as in 2007 and 2014. In the case of DGAs, their tendency reported in the paper to shift income between boxes and through time would make the attempt to estimate deadweight loss a tentative endeavour at best. Taking the international dimension into account also add another layer of complexity, and it is useful not to see these results purely in the Dutch context but in the light of international fiscal competition too.

Classical externalities may be even more difficult to quantify. Even if the data allowed to distinguish between true and reported investments, the marginal social benefit would also depend on many other factors. For instance two firms may have different marginal returns to capital and benefit differently from additional investment. Or one may be investing in a green industry whether the other invests in polluting businesses. With data on a longer period, it may be possible to fill one of those caveats with future research. One could assess the returns on these investments by looking at the subsequent growth of firms with high investment expenditures relative to others for instance.

In light of the information that the paper can and cannot provide it is possible to highlight two policy aspects. It is important to monitor DGAs to ensure that reported investment expenditures correspond to what is intended by the fiscal scheme. Also it may be useful to provide incentives to channel investments in areas with positive externalities.

References

- [1] Stuart Adam, Helen Miller, and Thomas Pope. *Tax, legal form and the gig economy*. Feb 2017.
- [2] Michael Carlos Best and Henrik Jacobsen Kleven. Housing market responses to transaction taxes: Evidence from notches and stimulus in the uk. 2016.
- [3] Leon Bettendorf, Arjan Lejour, and Maarten van't Riet. Tax bunching by owners of small corporations. *De Economist*, 165(4):411–438, 2017.
- [4] Raj Chetty. The simple economics of salience and taxation. Technical report, National Bureau of Economic Research, 2009.
- [5] Raj Chetty, John N Friedman, Tore Olsen, and Luigi Pistaferri. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The quarterly journal of economics*, 126(2):749–804, 2011.
- [6] Raj Chetty, John N Friedman, and Emmanuel Saez. Using differences in knowledge across neighborhoods to uncover the impacts of the eitc on earnings. *The American Economic Review*, 103(7):2683–2721, 2013.
- [7] Raj Chetty and Emmanuel Saez. Dividend taxes and corporate behavior: Evidence from the 2003 dividend tax cut. *The Quarterly Journal of Economics*, 120(3):791–833, 2005.
- [8] Raj Chetty and Emmanuel Saez. Teaching the tax code: Earnings responses to an experiment with eitc recipients. *American Economic Journal: Applied Economics*, 5(1):1–31, 2013.
- [9] Ruud A De Mooij and Gaëtan Nicodème. Corporate tax policy and incorporation in the eu. *International Tax and Public Finance*, 15(4):478–498, 2008.
- [10] Vincent Dekker, Kristina Strohmaier, and Nicole Bosch. A data-driven procedure to determine the bunching window: An application to the netherlands. 2016.

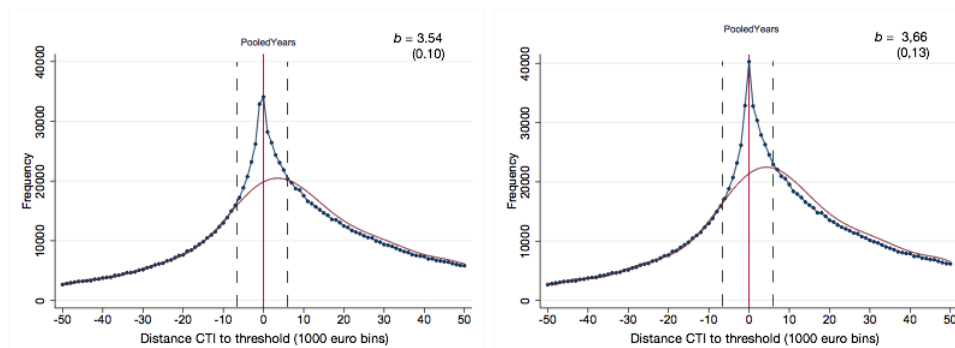
- [11] Michael P Devereux, Li Liu, and Simon Loretz. The elasticity of corporate taxable income: New evidence from uk tax records. *American Economic Journal: Economic Policy*, 6(2):19–53, 2014.
- [12] Peter Diamond and James Mirrlees. Optimal taxation and public production: I–production efficiency. *American Economic Review*, 61(1):8–27, 1971.
- [13] Nadja Dwenger and Viktor Steiner. Profit taxation and the elasticity of the corporate income tax base: Evidence from german corporate tax return data. *National Tax Journal*, 65(1):118–50, 2012.
- [14] Martin Feldstein. The effect of marginal tax rates on taxable income: a panel study of the 1986 tax reform act. *Journal of Political Economy*, 103(3):551–572, 1995.
- [15] Martin Feldstein. Tax avoidance and the deadweight loss of the income tax. *The Review of Economics and Statistics*, 81(4):674–680, 1999.
- [16] Organisation for Economic Co-operation and Development. *OECD SME and entrepreneurship outlook*. OECD, 2005.
- [17] Austan Goolsbee. Taxes, organizational form, and the deadweight loss of the corporate income tax. *Journal of Public Economics*, 69(1):143–152, 1998.
- [18] Austan Goolsbee. The impact of the corporate income tax: evidence from state organizational form data. *Journal of Public Economics*, 88(11):2283–2299, 2004.
- [19] John R Graham. Taxes and corporate finance: A review. *The Review of Financial Studies*, 16(4):1075–1129, 2003.
- [20] Jonathan Gruber and Joshua Rauh. How elastic is the corporate income tax base? *Taxing corporate income in the 21st century*, pages 140–163, 2007.
- [21] Kevin A Hassett and R Glenn Hubbard. Tax policy and business investment. *Handbook of public economics*, 3:1293–1343, 2002.

- [22] Henrik J Kleven and Mazhar Waseem. Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723, 2013.
- [23] Henrik Jacobsen Kleven. Bunching. *Annual Review of Economics*, 8:435–464, 2016.
- [24] Collen Lediga, Nadine Riedel, and Kristina Strohmaier. Corporate taxes and firm behavior-evidence from south africa. *Unpublished manuscript, University of Bochum*, 2016.
- [25] Ministry of Finance. Bruto-nettotrajecten, bijlage 2 bij Brief over de winstbox. Den Haag, 3 juli 2014.
- [26] J. A. Mirrlees. An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208, 1971.
- [27] Elena Patel, Nathan Seegert, and Matthew Grady Smith. At a loss: The real and reporting elasticity of corporate taxable income. 2016.
- [28] Emmanuel Saez. Using elasticities to derive optimal income tax rates. *The review of economic studies*, 68(1):205–229, 2001.
- [29] Emmanuel Saez. Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, 2(3):180–212, 2010.
- [30] Emmanuel Saez, Joel Slemrod, and Seth H Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, 50(1):3–50, 2012.

A Technical Appendix

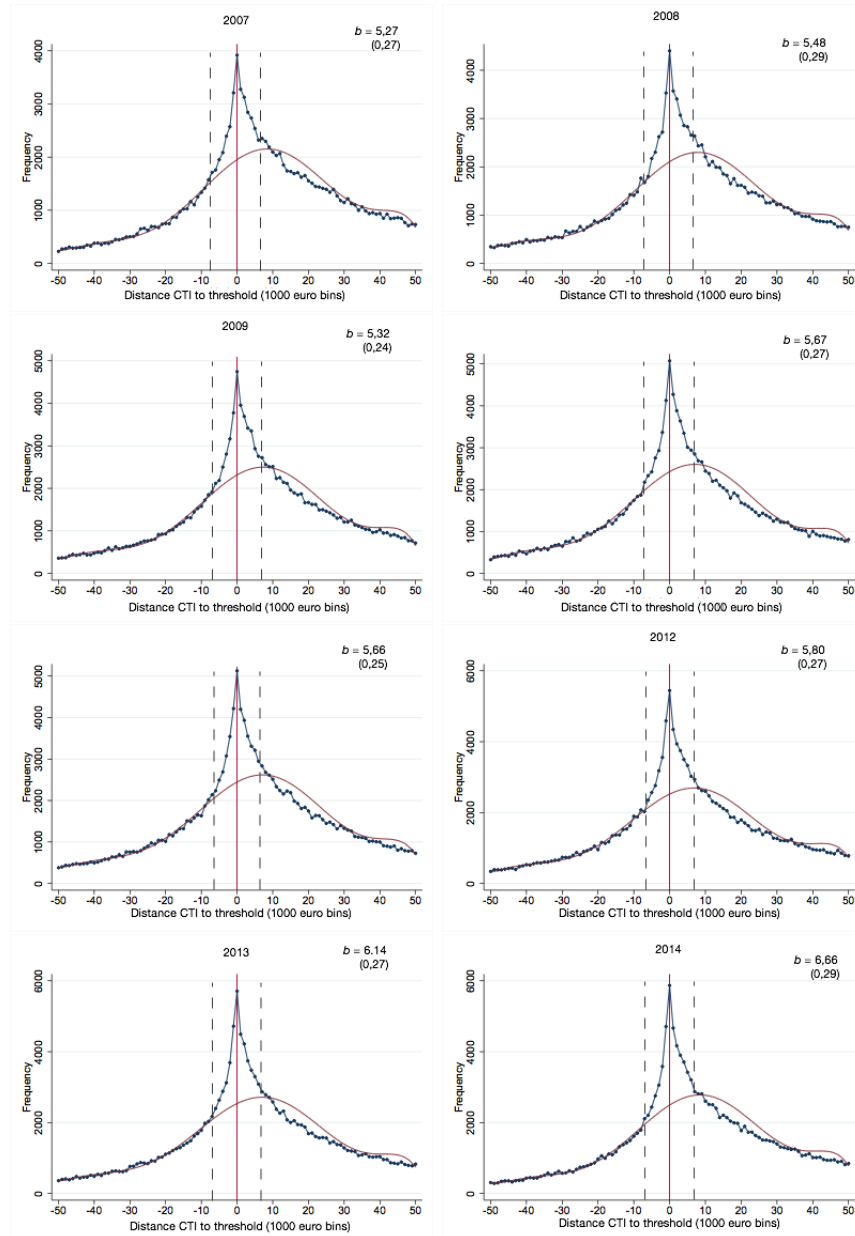
A.1 Estimates at the €0 threshold

Figure 10: bunching at €0 for firms with and without any declared losses



Note: Figure 10 shows the actual distribution and an estimated counterfactual distribution of the CTI between 2007 and 2014 in fifty bins of 1000 on either side of the kink (i.e. between €-50,000 and €50,000). On the left is the subsample of firms not reporting any losses and on the right all the firms. The counter-factual is a 7th order polynomial. The selected bunching region is 7 bins on each side of the kink.

Figure 11: bunching at €0 for the years 2007-2014



Note: Figure 11 shows the actual distribution and an estimated counterfactual distribution of the CTI for each separate year between 2007 and 2014. CTI data is grouped into bins of 1000 euros. The counter-factual is a 7th order polynomial. The selected bunching region is 7 bins on each side of the kink.

A.2 Formal version of model of bunching

In a two-period setup Firm i begins period 1 with E retained earnings and K capital in period 2, such that $K = E + I$ where I is equity issue net of dividend payments. Firm i produces profits following

$$\Pi(K) = \frac{1+e}{e} A_i K^{\frac{e}{1+e}} \quad (6)$$

where A is a productivity factor heterogeneously distributed across firms. Firm i chooses taxable income $Y = (1 - \rho)\Pi(K)$ by shifting profit out of the CTI base at a cost $c(\rho, \Pi(K)) = \rho^2 \Pi(K)/(2\phi)$, convex in ρ , where ρ is percent of CTI shifted and ϕ captures the relative cost of shifting. The firm maximises its value to shareholders V

$$\max_{K, \rho} V = -rK + (1 - (\tau(1 - \rho)Y) - c(\rho, \Pi(K))), \quad (7)$$

where rK is revenues foregone from government bonds yielding an interest of r . The first order conditions yield respectively

$$\Pi'(K) = \frac{r}{1 - \tau(1 - \rho) - \frac{\rho^2}{2\phi}} \Leftrightarrow K^* = r^{-(1+e)} [A(1 - \tau(1 - \rho) - \frac{\rho^2}{2\phi})]^{1+e} \quad (8)$$

$$\tau \Pi(K) = \frac{\rho}{\phi} \Pi(K) \Leftrightarrow \rho = \phi \tau. \quad (9)$$

Combining the two we obtain the equilibrium CTI in the second period

$$Y^* = (1 - \rho) \frac{1+e}{e} A^{1+e} r^{-e} (\alpha + (1 - \alpha)^2) \frac{\phi}{2} \quad (10)$$

where α is the net-of-tax rate $1 - \tau$. Following the standard definition, the ECTI with respect to the net of tax-rates can then be derived as

$$e_{y, \alpha} = \frac{dY}{d\alpha} \frac{\alpha}{Y} = \frac{(1 - \rho)\alpha}{\alpha + \frac{\rho\tau}{2}} \epsilon + \frac{\rho\alpha}{(1 - \alpha)(1 - \rho)}. \quad (11)$$

A.3 Box 1 and CTI rates for the 2007-2014 period

Table 2: Box 1 tax rates and thresholds

<u>Year</u>	<u>Bracket 1</u>	<u>Threshold</u>	<u>Bracket 2</u>	<u>Threshold</u>	<u>Bracket 3</u>	<u>Threshold</u>	<u>Bracket 4</u>
2007	33.65	17319	41.40	31122	42.00	53064	52.00
2008	33.60	17579	41.85	31589	42.00	53860	52.00
2009	33.50	17878	42.00	32127	41.00	54776	52.00
2010	33.45	18218	41.95	32738	42.00	54367	52.00
2011	33.00	18628	41.95	33436	42.00	55694	52.00
2012	33.10	18945	41.95	33863	42.00	56491	52.00
2013	37.00	19645	42.00	33363	42.00	55991	52.00
2014	36.25	19645	42.00	33363	42.00	56531	52.00

Note: The table reads as follows: in 2007 the first bracket starts at €0 and ends at €17319, and the applied marginal rate is 33.65%. Bracket 2 starts at €17320 and ends at €31122 for a rate of 41.40%, etc.

Table 3: Corporate income tax rates and thresholds

<u>Year</u>	<u>Bracket 1</u>	<u>Threshold</u>	<u>Bracket 2</u>	<u>Threshold</u>	<u>Bracket 3</u>
2007	20	25000	23,5	60000	25,5
2008	20	275000	25,5	-	-
2009	20	200000	25,5	-	-
2010	20	200000	25,5	-	-
2011	20	200000	25	-	-
2012	20	200000	25	-	-
2013	20	200000	25	-	-
2014	20	200000	25	-	-

Note: The table reads as follows: in 2007 the first bracket starts at €0 and ends at €25,000 the applied marginal rate is 20%. Bracket 2 starts at €25,001 and ends at €60,000 for a rate of 25,5% etc.

A.4 Precisions on Income shifting

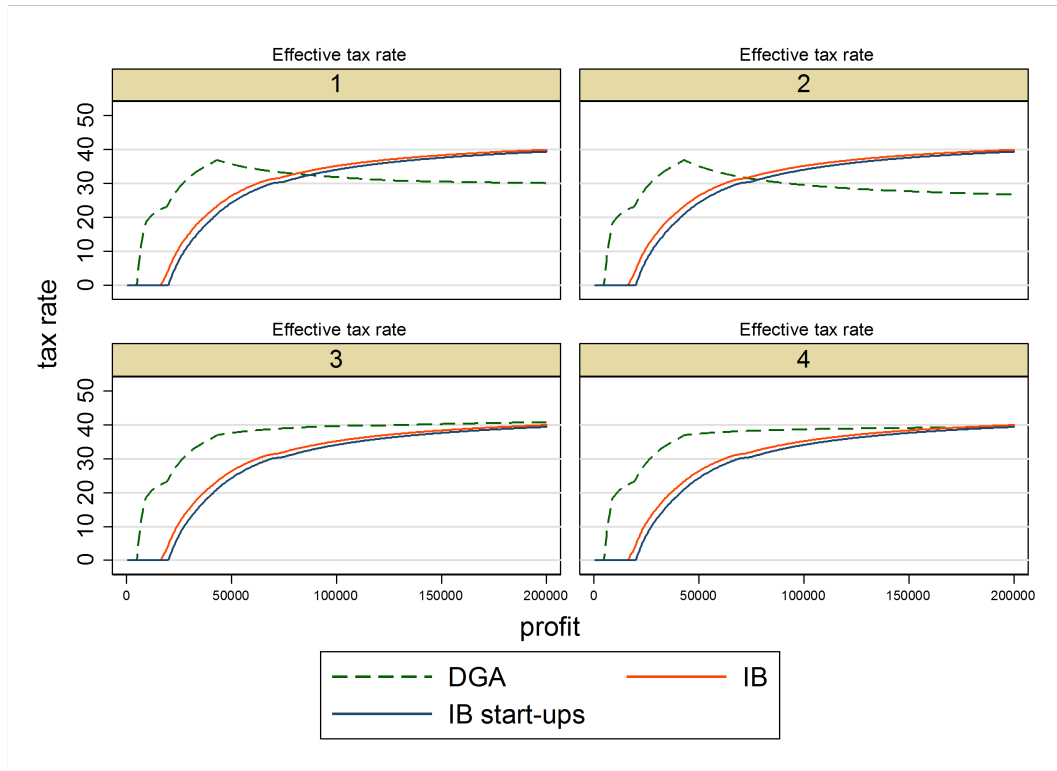
A.4.1 The effect of income shifting on ETRs of DGAs

This section shows the mechanisms of ETR optimisation via inter-temporal and intra-temporal income shifting. I compare below the ETRs of four different scenarios, where the share of profit retention and of income received as wages vary. These scenarios are largely theoretical and in reality income shifting possibilities may be more or less constrained. For instance, DGAs in firms with more than one shareholder may be engaged in more or less binding in-house agreements with the other owners. In the absolute, whether on average, DGAs retain 100%, 80% or 50% of the profit within the firm does not change the argument that profit retention is an effective way to reduce to reduce one's ETR by lowering the box 2 tax base.

As a ground for comparison, the ETR faced by DGAs are plotted against that of an IB-entrepreneur, another form of organisation often used for small businesses. The profits of IB-entrepreneurs are taxed entirely under box 1 after their various deductions are applied. The DGA tax burden is calculated following the same methodology as the Dutch ministry of finance. I use tax rates presented in the tables above. The calculations include the different tax credits, but exclude pensions or treatments that are the same across both forms of organisation (DGA or IB). For the scenarios where box 1 revenue is not capped at €44,000, it is computed as a function of profit. More precisely it is equal to profit until the cap (set by the minimum perceived requirement) plus 15% of each additional euro after that. This choice is somewhat arbitrary but as shown in Table 1, the average wage/profit ratio observed is 16,9%. Thus defining box 1 as 100% of the profit and then 15% seems a decent approximation.

The ETRs displayed are for the year 2013. The year 2013 was chosen because of the profit exemption for IB-entrepreneurs was already at 14%, which makes the importance of inter-temporal income shifting for the fiscal viability of owning a business as a DGA even more salient. Additionally the year 2014 is a special case due to the tax holiday. There is no qualitative difference during the years in the sense that profit retention decreases substantially the ETR of DGAs and profit retention allows flexibility for fiscal optimisation.

Figure 12: ETRs for all four scenarios, 2013

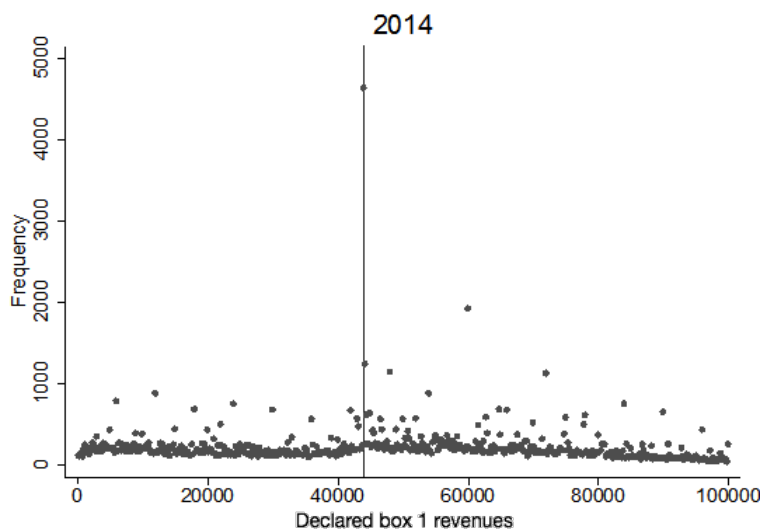


Note: Scenarios 1 and 2 have 80% of the profit retained whereas all profits are taken out in scenarios 3 and 4. In scenario 1 and 3, box 1 income is a function of profit, whereas in scenarios 2 and 4 it is fixed at the €44,000 level. IB start-ups designates those active for less than 3 years and serve to illustrate the effect of tax credits on the ETR.

DGAs enjoy much larger flexibility regarding their fiscal burdens than IBs do. Depending on the scenario, the ETR of DGAs will converge either below 30% or around 45% as income rises but it will never converge below the 25% of the CIT. This difference is primarily due to inter-temporal income shifting. The impact of shifting income between box 1 and 2 is not negligible (around 2 percentage points difference between scenarios 1-2, and 3-4), but it is substantially less important.

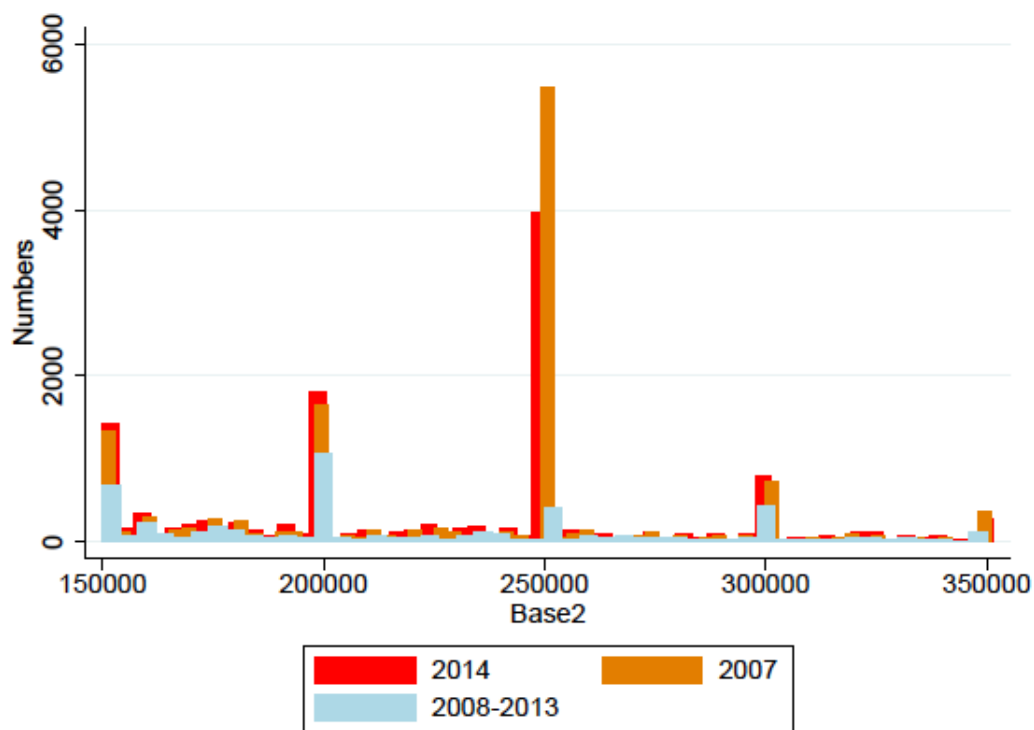
A.4.2 Evidence of income shifting

Figure 13: Wages declared by DGAs



Note: Figure 13 shows the histogram of wages for DGAs with a CTI between €0 and €100,000 for 2014. The bunching at that “legal” threshold is observed for all the years available in our dataset. Interestingly we can also notice large number of tax filers declaring wages around €60,000 which is about where the marginal tax rate in box 1 increases from 42% to 52% (the threshold of the 3rd bracket that year is actually 56,531, but once the labour tax credit and the general tax credit are taken into account, it is closer to €60,000).

Figure 14: declared Box 2 income, 2007-2014



Note: Figure 14 displays the number of firms per amount of dividends declared, for each year between 2007 and 2014 (only 2007 and 2014 are in different colour for better visibility). In 2007 and 2014, the number of firms distributing dividends, €250,000 especially, exploded. This is likely due to the tax holiday, a 22% marginal tax rate on dividends instead of 25% until €250,000, in 2007 and 2014.